

EGOCENTRIC AND ALLOCENTRIC SCHEMES FOR MATERIAL UNDERSTANDING

A Thesis

Presented to the Faculty of the Graduate School
of Cornell University

in Partial Fulfillment of the Requirements for the Degree of
Master of Science

by

Jimmy Briggs
(James E. Briggs)

December 2019

© 2019 Jimmy Briggs

(James E. Briggs)

ALL RIGHTS RESERVED

ABSTRACT

Allocentric motion, in which the user moves with respect to an external subject's coordinate system, is commonly used in the context of rendering and material design, but little scientific work exists to support this scheme over egocentric motion, in which a user works according to their own coordinate system. We examine the effectiveness of these two navigation schemes with a user study spanning a standard desktop workstation, "fish tank" virtual reality, and a head-mounted display. Our results do not indicate a difference in performance between these schemes, a result we analyze in the discussion section.

BIOGRAPHICAL SKETCH

Jimmy Briggs began his undergraduate degree at Boston University studying Pure and Applied Mathematics. He dropped out during his junior year and worked full-time at an engineering firm while applying for transfer. He was admitted to Cornell University, where he would complete his Bachelor's degree in Computer Science before continuing into a Master's program in the same department.

Jimmy has won multiple awards for his contributions as a teaching assistant at Cornell, published in a flagship conference during his second year of graduate school, and served as an advocate for LGBTQ+ rights on campus. He is responsible in part for reforms in Computing and Information Science at Cornell, including the addition of an all-gender restroom to Gates Hall.

In solidarity with Cornell Graduate Students United.

ACKNOWLEDGEMENTS

I am obliged to thank my faculty committee: Professors François Guimbretière, Steve Marschner, and Alex Townsend for their guidance in this work. I am grateful to Lauren Kilgour, Huaishu Peng, and Matt Bierbaum for piloting my study; Michelle Carpenter, Anmol Kabra, and Vince Tseng for demonstrating the hardware configurations in Chapter 3; and Jean Costa and Matt Law for helping me to learn the statistical tools used in Chapter 4.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Tables	viii
List of Figures	ix
1 Introduction	1
1.1 Materials	2
1.2 Egocentric and Allocentric Schemes	3
1.3 Contribution	5
2 Related Work	6
2.1 Material Perception	6
2.2 Egocentricity and Allocentricity	6
2.3 Material Design Interfaces	8
3 Experimental Design	10
3.1 Overview	10
3.2 Participants	12
3.3 Tasks	13
3.3.1 Comparison	13
3.3.2 Tuning	16
3.3.3 Scanning	18
3.4 Apparatus	20
3.4.1 Viewport Configuration	21
3.4.2 Fish Tank Configuration	22
3.4.3 HMD Configuration	24
4 Results	26
4.1 Dependent Variables	26
4.2 Speed	27
4.2.1 Comparison	28
4.2.2 Tuning	30
4.2.3 Scanning	31
4.3 Accuracy	32
4.3.1 Comparison	33
4.3.2 Tuning	34
4.3.3 Scanning	36
4.4 User Perception	37
4.4.1 Comfort	38
4.4.2 Intuitiveness	39

4.4.3	Difficulty	41
4.4.4	Confidence	42
4.5	Summary	43
5	Discussion	44
5.1	Intuitiveness: Egocentricity or Confounding Factor?	44
5.2	Observed Behaviors May Explain Lack of Effect	45
5.2.1	Users Do Not Adjust Viewpoint During Evaluation	46
5.2.2	Users Explore the View Space Only Once	47
5.3	Are Single-Peak Reflectance Functions the Culprit?	48
5.3.1	BRDF(s)	48
5.3.2	Single-Peak BRDF(s)	49
5.3.3	Multi-Peak BRDFs Exist and Can Be Tested	52
6	Future work	53
7	Conclusion	55
A	Fish Tank VR with the Oculus Rift	56
	Bibliography	58

LIST OF TABLES

3.1	The order in which each user utilized the Viewport (V), Fish Tank (F), and Head-Mounted Display (H) configurations.	12
3.2	Each row shows a single block in three lighting conditions. The material rendered in each row (except the first) differs from the default material in a single shading parameter.	15
4.1	Comparison Task Response Time: Within Subjects Effects	28
4.2	Tuning Task Response Time: Within Subjects Effects	30
4.3	Scanning Task Response Time: Within Subjects Effects	31
4.4	Comparison Task Accuracy: Within Subjects Effects	33
4.5	Tuning Task Accuracy: Within Subjects Effects	34
4.6	Scanning Task Accuracy: Within Subjects Effects	36
4.7	Level of Comfort: Within Subjects Effects	38
4.8	Assessment of Intuitiveness: Within Subjects Effects	40
4.9	Post-Hoc t-Test for Intuitiveness	41
4.10	Level of Difficulty: Within Subjects Effects	41
4.11	User-Reported Confidence: Within Subjects Effects	42

LIST OF FIGURES

1.1	The Stanford Bunny assigned (from left to right) metallic, isotropic, anisotropic, and brick-textured materials.	2
1.2	Left to right: egocentric, hybrid, and allocentric descriptions of the position of a hare.	4
1.3	Many tasks can be arranged along a continuum between mostly egocentric and mostly allocentric.	5
2.1	A screenshot of Maya’s Hypershade material designer, reprinted from public documentation [2].	8
3.1	The virtual scene in the Comparison Task populated with a pair of identical blocks.	13
3.2	Iridescent shields for the Tuning Task.	16
3.3	A hidden star in the Scanning Task.	18
3.4	A non-participant volunteer examines a wood sample using the Viewport Configuration.	21
3.5	A lab-mate models the tracked helmet employed in the Fish Tank Configuration.	22
3.6	A non-participant volunteer examines a scene in head-mounted display configuration.	24
4.1	Combinatorial plot of mean response time for the Comparison Task.	29
4.2	Combinatorial plot of mean response time in the Tuning Task	31
4.3	Combinatorial plot of Scanning Task response times.	32
4.4	Combinatorial plot of Comparison Task accuracy.	33
4.5	Combinatorial plot of Tuning Task error.	35
4.6	Combinatorial plot of Scanning Task accuracy.	37
4.7	Combinatorial plot of participant comfort.	39
4.8	Combinatorial plot of intuitiveness as reported by participants.	40
4.9	Combinatorial plot of difficulty experienced by participants.	41
4.10	Combinatorial plot of user confidence as reported by participants.	43
5.1	Cross-sectional diagrams of simple BRDFs (blue) evaluated for many ω_o with fixed ω_i (yellow). In the ideal specular diagram, a normal is annotated for clarity.	48
5.2	Plots of the physically accurate version of the Phong BRDF [24] for several fixed values of w_o , each labeled with a red circle. Each graph should be understood as the orthogonal projection of the hemisphere from which w_o is drawn, with high-intensity directions more yellow and low-intensity directions more blue.	50
5.3	BRDFs of the anisotropic model used in the Scanning Task for fixed values of w_i , each labeled with a red circle, as in Figure 5.2	51

5.4	Plots of reflected intensity over the incident hemisphere (projected to the equatorial disc) for selected surface points of the Marschner wood model. Reprinted from [26] with permission from the author.	52
-----	--	----

CHAPTER 1

INTRODUCTION

A common occupation in the animation industry is that of the materials designer. These artists spend most of their time in the process of *material understanding*: either by tuning reflectance models to match design references, by evaluating the quality of virtual materials, or by scanning a surface visually to detect optical abnormalities. Professional material design tools like Maya’s Hypershade [2] employ an allocentric scheme for user interaction, in which the designer orbits the object of interest to better observe a material.

Our goal in this work is to interrogate this widespread usage of allocentric interfaces in the performance of material understanding tasks. In this thesis, we present a user study in which the material understanding tasks of comparison, tuning, and scanning were performed in three hardware configurations differing in exocentricity. We also provide a statistical analysis of the results and a qualitative discussion of our observations during the trials.

We begin in this chapter with an explanation of materials followed by a discussion of egocentric and allocentric schemes for navigation. In Chapter 2 we consider our work in the context of the existing literature in the fields of Computer Graphics, Human-Computer Interaction, and Cognitive Science. Chapter 3 details our experimental design and execution. We present our results with analysis in Chapters 4 and 5, respectively, before rendering our suggestions for future work in Chapter 6, and stating our conclusions in Chapter 7

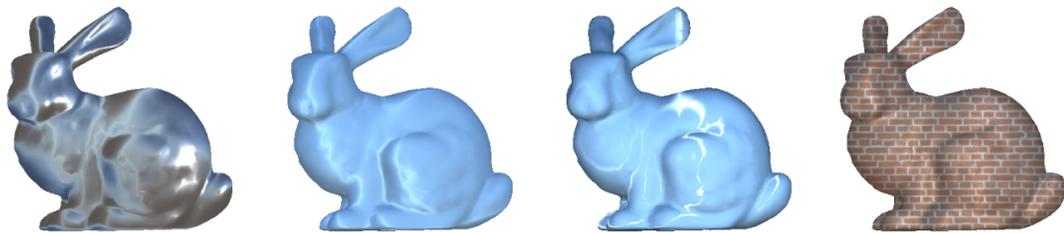


Figure 1.1: The Stanford Bunny assigned (from left to right) metallic, isotropic, anisotropic, and brick-textured materials.

1.1 Materials

Materials are an abstraction of surface reflectance and volumetric scattering models which characterize the way that virtual objects interact with light in a rendering environment. Materials approximate the microscopic structure of an object, not its larger geometric features [39]. Thus two virtual objects with identical shape can be modeled with different materials, as shown in Figure 1.1. In the context of photorealistic image synthesis, materials are designed to approximate the true optical behavior of physical references. Even in non-photorealistic contexts, materials artists spend a great deal of time approximating a complex physical materials like wood, rust, and hair.

The simplest materials are uniform across a surface or volume and interact with light in clear, predictable ways. The reflective behavior of matte wall paint can be approximated by an ideal diffuse material. By contrast, an ideal specular material mimics a polished silver mirror [33].

More complicated materials may vary across the surface of a virtual object. This document, for instance is black where there is text and white elsewhere. Thus color varies across the surface of the page. This spatial variation in re-

flectance is sometimes called *texture*. Color is not the only reflectance feature that may vary spatially across a surface. Roughness, anisotropy, iridescence, and myriad other optical properties contribute to texture.

Real-world materials tend to have texture. Wood exhibits interlocking patterns of light and dark often called “grain.” A rusty pipe ranges from glossy silver to matte orange in a pattern determined by wear on different parts of the metal. The feathers of a pigeon vary from dull grey at the tail to iridescent green and purple at the neck. These complex patterns of reflectance are difficult to replicate and evaluate for the designers of virtual materials working in the film and video gaming industries. This is partially due to the fact that the evaluation of a surface reflectance model requires a practitioner to observe it from many different directions and in many different lighting conditions to understand the full behavior of the texture.

Our user study uses the challenge of evaluating materials with complex textures as a performance test for several navigation schemes. Our goal is to determine the relationship between success in this realm and the choice of navigation scheme.

1.2 Egocentric and Allocentric Schemes

A spatial reference frame is a mental model which allows person to describe the positions of objects in space. *Egocentric* reference frames place objects with respect to the observer. *Allocentric* reference frames consider the spatial relationships between objects and some landmark external to the observer [23]. It is worth noting that hybrid schemes involving the simultaneous use of egocentric

and allocentric reference frames are also possible and perhaps the mode [5].

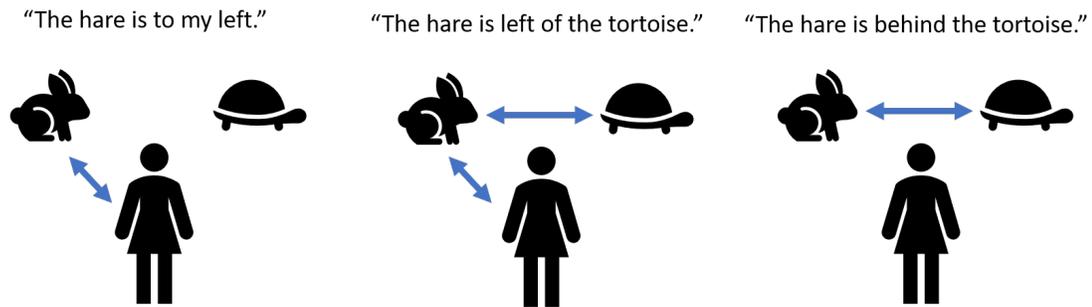


Figure 1.2: Left to right: egocentric, hybrid, and allocentric descriptions of the position of a hare.

Figure 1.2 illustrates spatial thinking by way of egocentric, hybrid, and allocentric reference frames. (We refer to these as egocentric, hybrid, and allocentric reasoning, respectively.) The scenes in Figure 1.2 are identical physically but their descriptions demonstrate the speaker's spatial reference frame. In the egocentric example, the hare is positioned relative only to the speaker. In the allocentric example, position is given only in relation to an external object – the tortoise. In the hybrid example, the hare is triangulated with information from the speaker's heading and the tortoise's position.

While some tasks are clearly egocentric or allocentric, many hybrid tasks rely on both egocentric and allocentric reasoning in varying amounts. One way to model this is a spectrum of exocentricity ranging from purely egocentric to purely allocentric, with most activities falling in between [11]. This spectrum is visualized in Figure 1.3.

Material understanding tasks have most commonly been solved with purely allocentric *motion models* – the navigation schemes designers use to interact with materials. The dominant paradigm in material design software requires users to orbit their view around a fixed object whose material they would like to eval-

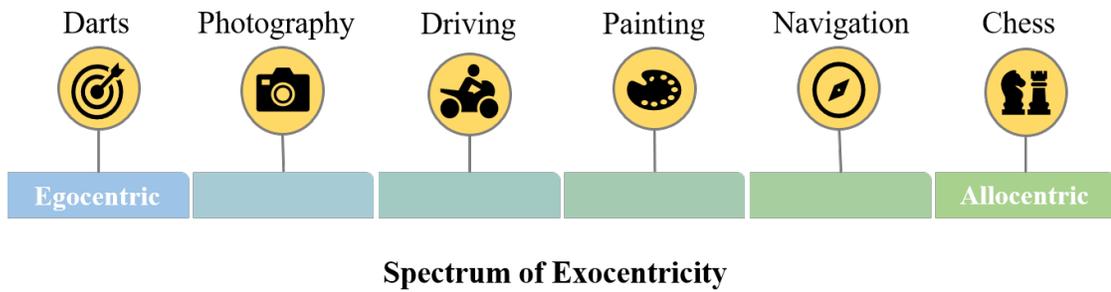


Figure 1.3: Many tasks can be arranged along a continuum between mostly egocentric and mostly allocentric.

uate. Our user study examines whether or not this choice of allocentric motion model affects material understanding in such a design tool.

1.3 Contribution

To evaluate the effectiveness of egocentric and allocentric navigation schemes that humans may use to understand the reflective properties of materials, we conducted a user study in which human subjects' performance of material understanding tasks was evaluated in several paradigms representing different points on the spectrum of exocentricity. Our analysis shows no significant effect of the exocentricity of an interface on the material understanding that it enables, indicating that little is to be gained by breaking with material design tradition.

CHAPTER 2

RELATED WORK

This thesis explores the performance of users completing several common material understanding tasks across several hardware configurations meant to represent various points on the spectrum of exocentricity from egocentric to allocentric. In this chapter, we briefly visit the spaces of material perception and material design user studies to show the novelty of our experiment therein.

2.1 Material Perception

The unconscious processes by which humans visually assess the physical, textural, and functional properties of materials is both much-studied and poorly understood [14]. Although our work does not directly involve neurological or psychological models of material perception, the adjacency of our work to this rich field merits mention. We direct the reader to [15] for its excellent top-down overview of visual perception. Alternatively, those with a Computer Graphics background are encouraged to read [43], especially Section III, which focuses solely on the relevant sub-field of material perception.

2.2 Egocentricity and Allocentricity

Our interest in discovering whether egocentric, allocentric, or hybrid motion models affect material understanding is not unprecedented. Researchers have investigated similar questions of exocentricity in various domains for decades, although not directly toward material understanding tasks.

There is a body of evidence suggesting that egocentric and allocentric reasoning are neurologically distinct phenomena. For instance, Gramman et al. observed that subjects displaying egocentric and allocentric preferences exhibited divergent alpha-blocking patterns as measured by EEG while simulating turns in a virtual tunnel [18], an observation corroborated by subsequent correlate studies [34, 10]. Other work has indicated that rotations are better understood egocentrically while translations are most intuitive in an allocentric scheme [20], and that human subjects have individual biases toward egocentric or allocentric navigation modes [17].

The observation of neurological divergence in egocentric and allocentric reasoning has inspired a broad array of research in human-computer interface design. While some research posits that humans are predisposed toward a preference for egocentric control schemes [30, 36, 6, 5], the majority of studies indicate that differing levels of exocentricity are better suited to different tasks. One line of research, for example, showed that a novel allocentric scheme empowered drone pilots in indoor spaces [12], while another showed improved spatial localization in a virtual tennis game in the egocentric mode [1]. Another study in virtual navigation indicated that scene geometry determines the efficacy of egocentric and allocentric navigation schemes [48]. Perhaps most similar to this thesis is the work of Qi et al., who determined that the egocentric motion afforded by a head-mounted display was inferior to the hybrid navigation mode of fish tank VR for volume visualization [35].

It has been shown that egocentric and allocentric motion models affect the brains of users in distinct ways and that some tasks are better suited to differing levels of exocentricity. Our study is the first to apply this line of thought to the

topic of material understanding tasks.

2.3 Material Design Interfaces

While material understanding has not been codified in the HCI literature, its application to material design interfaces has been a topic of study. We discuss several interesting material designers in this section to identify the typical treatment of material understanding tasks. For clarity and brevity, we will refer to these design interface as *material designers*, and we will refer to the human consumers of the interfaces as *artists* or *users*.

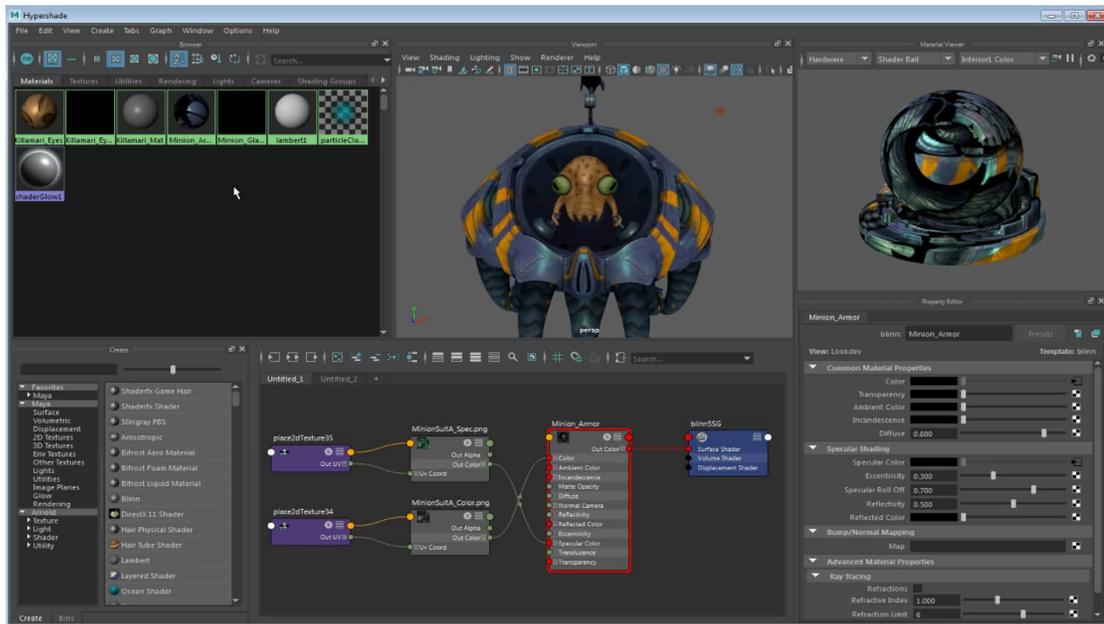


Figure 2.1: A screenshot of Maya's Hypershade material designer, reprinted from public documentation [2].

The standard paradigm for a material designer is best illustrated by Maya's Hypershade tool [2], which presents an allocentric view of a shaded object as well as a flowchart and sliders for editing material parameters (Figure 2.1).

More exotic material designers have been studied, but have not seen commercial implementation. Kerr and Pellacini produced a material designer for novice users which offers image-based navigation through a material parameter space [22]. The same experiment measured no improvement between users who navigated through physical material parameters over users who navigated through images which corresponded to a perceptual shading model [31]. While novel, the parameter-space navigation featured only an allocentric view of the object, whereas our work evaluates egocentric, allocentric, and hybrid motion models for material understanding.

Other material designers give the user an understanding of structure at various scales during design. One system simultaneously displays a macroscopic view of the object of interest and a microscopic view of the material with which it is adorned [49, 50], controlled allocentrically. Another class of material designers focuses on manipulating illumination [41, 8, 37], but does not interrogate spatial navigation schemes.

Finally, it has been noted that the majority of material design user studies have been conducted on novice users, and that any conclusions drawn from these studies may have biases that make them inapplicable to real-world material design pipelines [38]. Our experiment too was conducted on novices, and thus may not generalize to experienced material design artists, although we think this risk is sufficiently small.

While a diverse collection of material design questions has been addressed by existing research, we are not aware of any studies which diverge from the allocentric navigation model in favor of an egocentric or hybrid one. We intend to address this question in the forthcoming thesis.

CHAPTER 3

EXPERIMENTAL DESIGN

3.1 Overview

Our experiment was designed to determine if the exocentricity of a display system affects user performance in material understanding tasks. The research question occurred to us after presenting several volunteers with wood samples, which they were asked to evaluate visually. We were interested by the fact that all of our volunteers moved their heads around the wooden surface while assessing it. From this, we hypothesized that an egocentric navigation scheme – in which users could move their heads to observe changes in reflection across the surfaces of virtual objects – would enable human subjects to better understand the reflective properties of virtual objects when compared to a more allocentric scheme.

We decided to test our hypothesis by measuring user performance in three distinct tasks, each of which mirrored one of the three primary material design skills: *comparing* materials to a reference, *tuning* reflectance model parameters, and *scanning* surfaces for optical anomalies. Users were asked to complete each task once on various hardware configurations. All three tasks featured interaction with common anisotropic materials. The focus on anisotropy was designed to force users to change their viewpoints frequently during each task, thereby engaging with the navigation model we provided. The tasks surveyed were:

1. Comparison of shaded wooden blocks to a reference.
2. Tuning of a diffraction grating until it matched an exemplar.

3. Scanning brushed metallic surfaces for changes in brush direction.

We discuss each of these tasks in greater detail in Section 3.3.

The three tasks above gave us a broad sampling of material design skills to test. In order to test our hypothesis that egocentric motion would assist in these skills, however, we needed multiple user interfaces which diverged in exocentricity. We chose to examine three separate hardware configurations which represented different locations on the spectrum of exocentricity. The configurations employed were:

1. A head-mounted display (HMD), in which virtual motion was encoded one-to-one by user movement.
2. A "fish-tank" virtual reality system, in which users could both move their heads to change their views and orbit objects allocentrically.
3. A viewport configuration in which users could only orbit allocentrically.

Due to the motion models allowed in each configuration, we hold that the head-mounted display is a primarily egocentric scheme, that the viewport configuration is a primarily allocentric scheme, and that the fish-tank configuration is a truly hybrid scheme. These configurations are detailed further in Section 3.4.

Our experiment was chosen to be *within-subject* for hardware configurations to minimize nuisance variance due differential user performance. This was deemed especially important given the small number of users in our study. Conversely, the experiment was designed to be *between-subjects* for material un-

Scanning				Comparison				Tuning			
User	1 st	2 nd	3 rd	User	1 st	2 nd	3 rd	User	1 st	2 nd	3 rd
0	V	F	H	6	V	F	H	12	V	F	H
1	H	V	F	7	H	V	F	13	H	V	F
2	F	H	V	8	F	H	V	14	F	H	V
3	V	H	F	9	V	H	F	15	V	H	F
4	H	F	V	10	H	F	V	16	H	F	V
5	F	V	H	11	F	V	H	17	F	V	H

Table 3.1: The order in which each user utilized the Viewport (V), Fish Tank (F), and Head-Mounted Display (H) configurations.

derstanding tasks in order to avoid overwhelming users with training on three separate tasks.

3.2 Participants

We recruited 19 participants and used software logging and observation to record them in each of the three experimental conditions: Viewport, Fish Tank VR, and HMD VR. Participants were paid twenty dollars for roughly thirty minutes of their time. All participants reported normal or corrected-to-normal vision. One participant had difficulty understanding the assigned task and produced erratic results. This participant’s data was removed from the dataset and replaced with that of a subsequent participant’s. Our resultant population of 18 participants (10 women, 8 men, ages 18-33) was organized into three groups of equal size. Each of the Comparison, Tuning, and Scanning tasks were completed by one such group of 6 participants. Subjects completed their tasks in the Viewport, Fish Tank, and HMD configurations in the partial-factorial ordering shown in Table 3.1.

3.3 Tasks

In order to understand the effect of egocentric, allocentric, and hybrid motion models on material understanding, we required tasks which represented common material understanding functions. We chose three specific material understanding tasks which corresponded to the primary material understanding activities most common in the animation industry. We refer to these user tasks as the *Comparison*, *Tuning*, and *Scanning* tasks, and explain them below.

3.3.1 Comparison

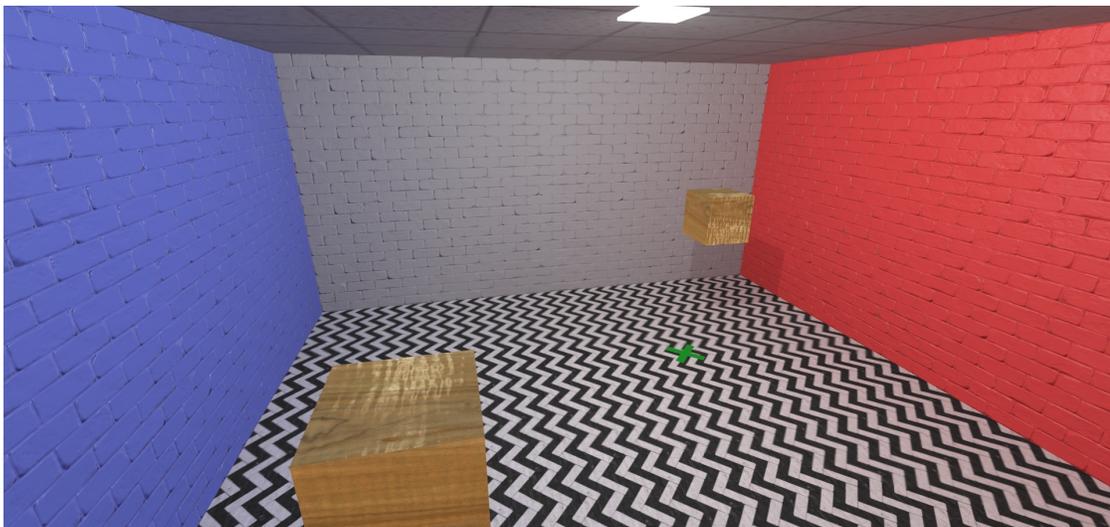


Figure 3.1: The virtual scene in the Comparison Task populated with a pair of identical blocks.

Holistic comparison of a virtual material to a reference is fundamental to material design; thus it served as a vital testbed for our evaluation of egocentric and allocentric motion schemes for material understanding. In the Comparison Task, participants were presented with a several pairs of virtual wooden boxes

and asked to identify whether or not they were the “same.”

The virtual setting of the Comparison Task – shown in Figure 3.1 – was modelled after the Cornell Box [16]. The design of this scene lends several advantages to the Comparison Task. Firstly, the scene contains a single light source, simplifying evaluation of the material. Secondly, the red and blue walls give obvious but non-intrusive landmarks to help users keep track of which block they are looking at. Finally, the simplicity of the scene made it easy to implement and interact with. To assist in depth perception, each surface in the scene was given a diffuse texture. Two green “X” markers were also added to the floor next to the boxes to suggest good starting positions to users.

We used a real-time implementation of the Marschner Finished Wood Model to simulate light transport on the blocks [26]. Aside from physical accuracy, the advantage of using this model is that it includes several tunable texture parameters, each with differing levels of visual salience and anisotropy. We presented users with pairs of wooden blocks differing in either wood fiber orientation, fiber highlight width, diffuse color, or fiber color. The differences in these texture parameters are not clear or even visible in some configurations of light and user position, as Table 3.2 demonstrates. This property of the shading model encouraged users to move around more when comparing the blocks, which forced them to engage with the given motion model.

In the course of the Comparison Task, users were presented with 16 pairs of blocks. Each pair was either identical or differed by exactly one shading component (as illustrated in Table 3.2). Some of the blocks were consistent with actual wood samples, with each of their shading components coming directly from a measurement of the same piece of wood. Other “chimera boxes” used param-

	Light Behind Camera	Light Behind Block	Light From Right
Default			
Fiber Direction			
Highlight Width			
Diffuse Color			
Fiber Color			

Table 3.2: Each row shows a single block in three lighting conditions. The material rendered in each row (except the first) differs from the default material in a single shading parameter.

eters from several different wood samples. Chimera boxes were included to nullify the potential bias that boxes with one modified texture parameter might look inexplicably “unnatural.” The accuracy of each user response was measured as a boolean reporting whether or not the user was correct. The time taken for each decision was measured as the duration between the time a new pair of wooden blocks was loaded into the virtual environment, and the moment the subject declared them the same or different verbally.

3.3.2 Tuning



Figure 3.2: Iridescent shields for the Tuning Task.

While holistic comparison is an important step in the overall execution of the material design process, local evaluation of small adjustments in a shading model is required for the minute parameter tweaks on which designers spend much of their time. The Tuning Task was designed to assess user performance of these small-scale changes under egocentric and allocentric navigation schemes. In this task, users were shown a pair of virtual iridescent shields and asked to

adjust a hidden parameter of one shield until it matched the other (Figure 3.2).

Iridescence was chosen for the Tuning Task for its dramatic dependence on viewpoint. In order to compare the tunable shield to the reference shield, users were forced to move back and forth between the pair in order to see them from congruent viewpoints. This encouraged users to make small changes to the tunable shield parameter and immediately return to the reference, mirroring the kind of small adjustments we were interested in observing.

The iridescence shader coloring the shields was implemented based on a diffraction grating model originally due to Jos Stam [13]. Users adjusted the grating spacing to alter the appearance of the shield. The diffraction grating model was selected because of its simple parameterization by grating distance and easily visible anisotropy.

As in the Comparison Task the virtual setting of the Tuning task was a modified Cornell Box. Several additional changes were made for the Tuning task to facilitate user evaluation. The two shields were placed beside each other a short distance away so that the user could move back and forth between the two quickly. For this reason, the one central light had to be replaced by two spotlights – one per shield – so that the spatial relationship between each shield and its source of illumination was the same. Finally, the room was darkened from the original to better show the iridescent materials of the shields.

Over the course of the Tuning Task, users were presented with eight pairs of virtual shields (fewer than the sixteen pairs of blocks in the Comparison Task, chosen to keep each session less than an hour in duration). Each pair was composed of a reference and a doppelganger. Users were asked to tune the diffrac-

tion grating distance of the doppelganger until it matched that of the reference. The diffraction grating model of the shields was not explained to users. They were instead asked to match the shields solely based on their own perception of visual similarity. Users were permitted to adjust the parameter until they were satisfied with its similarity. Once they reported satisfaction, a new pair of shields was loaded.

Timing information was measured as in the Comparison task. Accuracy was measured as the linear distance from the reference value of the tuning parameter to the value of the tuning parameter effected by the user.

3.3.3 Scanning

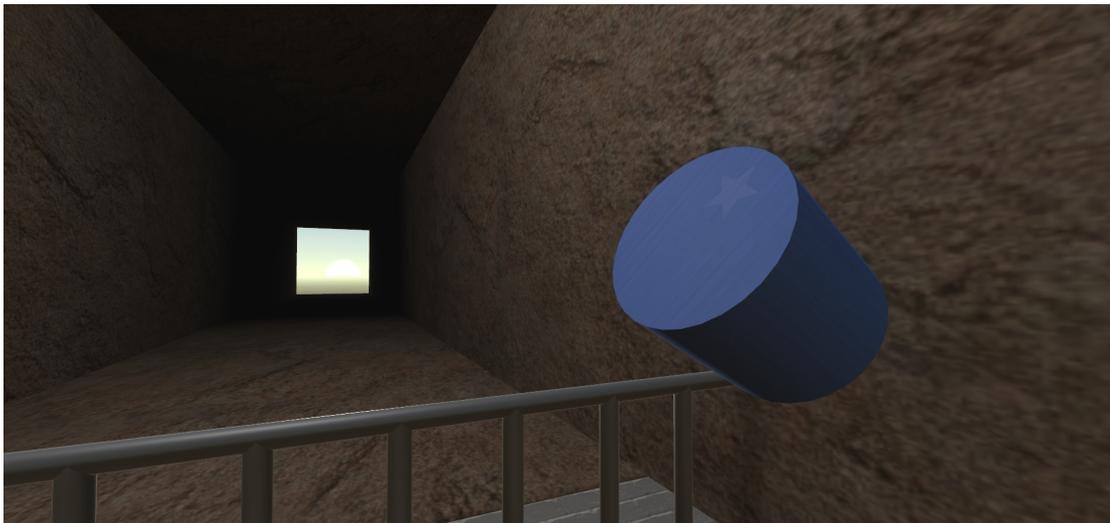


Figure 3.3: A hidden star in the Scanning Task.

Scanning for optical anomalies, the final material understanding task completed frequently by material designers, is important for avoiding visual disturbances in rendered animations. Whereas comparison and tuning require in-

teraction only with the material, scanning necessitates that users understand the interactions between surface, material, and illumination. The Scanning task was built to evaluate how egocentric and allocentric motion models contribute to success in this multifaceted material understanding problem. In the Scanning Task, users were asked to find a hidden shape on brushed metal cylinder.

The brushed metallic material was selected to force users to adjust the positions of both the cylinder and themselves with respect to a fixed light source. The anisotropic shader used to represent this material was based on a Bruce Walter's 2007 microfacet model [46], but simplified to run the GPU. Brushed metal has an unusual highlight behavior that is uniform along the direction of brushing. This highlight is only visible from certain configurations of light, camera, and surface. Therefore to see any disruptions in the highlight that might reveal a hidden shape, users needed to move themselves and the cylinder relative to the fixed light source. The cylinder itself was brushed in a uniform direction except for a triangle, square, or star (the hidden shape) on the surface which was brushed in a contrasting direction (Figure 3.3).

The virtual setting of the Scanning Task differed significantly from the Comparison and Tuning tasks. Rather than inside the Cornell Box, the user and object were placed at the end of a long virtual tunnel with one overhead light and a far-away directional light coming from the opposite end of the tunnel. The tunnel was designed to encourage users to find grazing angles of light off of the surface of the cylinder, as these angles were most likely to reveal the hidden shape. The overhead light was added to increase the number of configurations which could reveal the shape, since pilot experiments showed that a single light source made the task too difficult.

Users were given a total of twelve cylinders per configuration to evaluate (more than in the Comparison Task but less than in the Tuning, again to comply with time constraints). Each cylinder was given a random color to indicate that a new cylinder had been loaded between trials. This choice came as a result of feedback that users perceived the twelve cylinders to be a single object whose brush direction was changing over time, causing confusion and frustration. Timing was measured as the duration from the moment each new object was loaded into the scene to the moment the user called out the shape of the anomaly perceived on the surface of the object. Accuracy was measured as in the Comparison Task.

3.4 Apparatus

Users were asked to interact with virtual scenes constructed and simulated using the Unity3D Game Engine [42]. All shading models required for the tasks were implemented as Unity shaders in the Nvidia Cg shading language [28].

To interact with these virtual scenes, users employed three unique hardware configurations, each corresponding to a different level of exocentricity. In order from most allocentric to most egocentric, we refer to these UI modes as the *Viewport Configuration*, the *Fish Tank Configuration*, and the *Head-Mounted Display (HMD) Configuration*. Each participant completed a single task three times – once in each of the three hardware configurations. Trials were protected from analytical nuisance factors by partial combinatorial counterbalancing (as shown earlier in Figure 3.1).

All of the code used to run the experiment can be found in the author’s

github repository¹.

3.4.1 Viewport Configuration

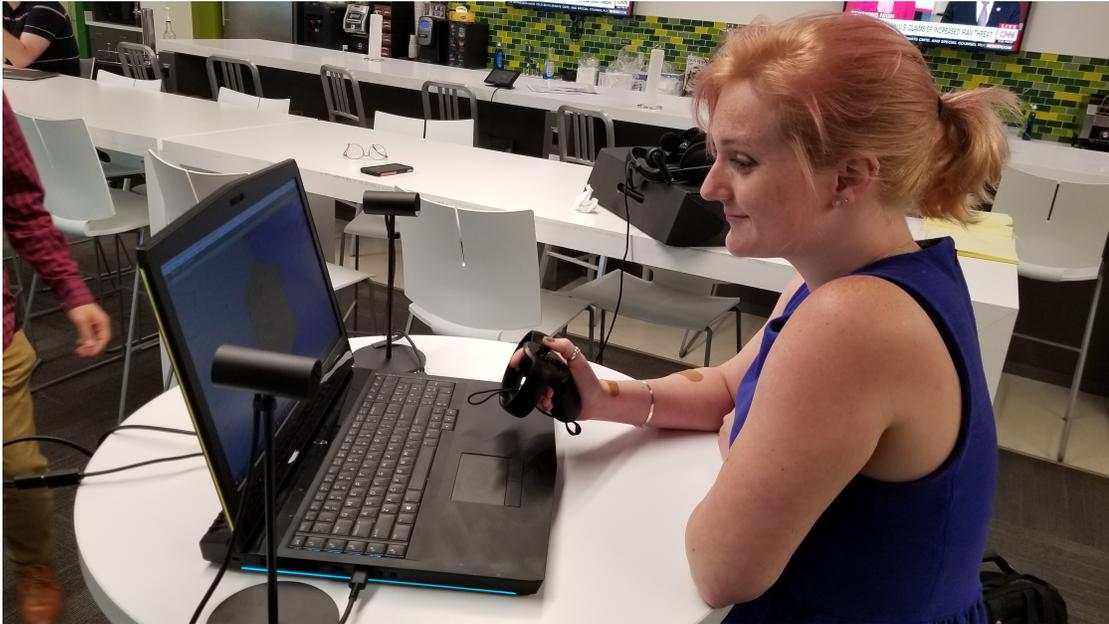


Figure 3.4: A non-participant volunteer examines a wood sample using the Viewport Configuration.

The Viewport Configuration is similar to the actual material design experience most common in the animation industry. In this configuration, users stand at a desk and interact with a virtual scene shown through a widescreen desktop monitor. A picture of the apparatus can be seen in Figure 3.4. In the Viewport Configuration, users interact with the virtual scene using a subject-orbiting control scheme similar to those available in Maya and Unity [2, 42]. This configuration is distinctly allocentric, as all navigation occurs with respect to a coordinate system internal to the computer program, namely that of the subject, rather than the natural coordinate system of the user. Unlike common material

¹<https://github.com/jeb482/morphomaterial>

design tools, the Viewport configuration does not use a mouse-and-keyboard for orbiting. Instead, the user operates an Oculus Touch controller to orbit and zoom. This decision was made to avoid using separate input devices for the three configurations. The Oculus controller can be used in virtual reality and desktop conditions, while a mouse-and-keyboard is difficult to use with a head-mounted display.

3.4.2 Fish Tank Configuration



Figure 3.5: A lab-mate models the tracked helmet employed in the Fish Tank Configuration.

The Fish Tank Configuration is inspired by the eponymous 1993 Interchi paper in which a desktop monitor was augmented with a head-tracking system so

that the image could be updated to match the viewer's position with respect to the screen [47]. This modification allows the display to replicate view frustum effects like parallax. The resultant system gives the viewer the impression that she is looking through a window into a 3D space, rather than at a screen containing a flat images. It also allows the user to move her head for new views of the subject of interest.

We used the Oculus Rift tracking system and Oculus Touch controllers to construct a fish tank VR installation. The decision to use Oculus tracking eliminates the nuisance factor of using multiple tracking systems between the Fish Tank and HMD configurations, and simplifies our system design, since Oculus tracking architecture is already in place. To provide head tracking without a head-mounted display, we affixed an Oculus Touch controller to the back of a bicycle helmet (Figure 3.5) and asked users to wear the helmet while interacting with the fish tank screen. Our initial tests showed that users experienced difficulty in using the head tracking while seated. This constraint and the desire to minimize experimental variance led us to require users to stand for all three hardware configurations. For more information on the setup and calibration of Fish Tank VR using the Oculus Rift, please refer to Appendix A.

Although the head-tracked motion of the Fish Tank display is intrinsically egocentric, users can only access a small subset of viewing directions by moving their heads. For instance, the user cannot walk behind the fish tank monitor to see the back of an object, because the screen is one-sided. Thus, the user must use another control scheme to access these additional views. We chose the same allocentric control scheme used in the Viewport Configuration to supplement head-tracking-based motion. The Fish Tank Configuration therefore is a hybrid

of egocentric fine motions and allocentric gross motions.

3.4.3 HMD Configuration

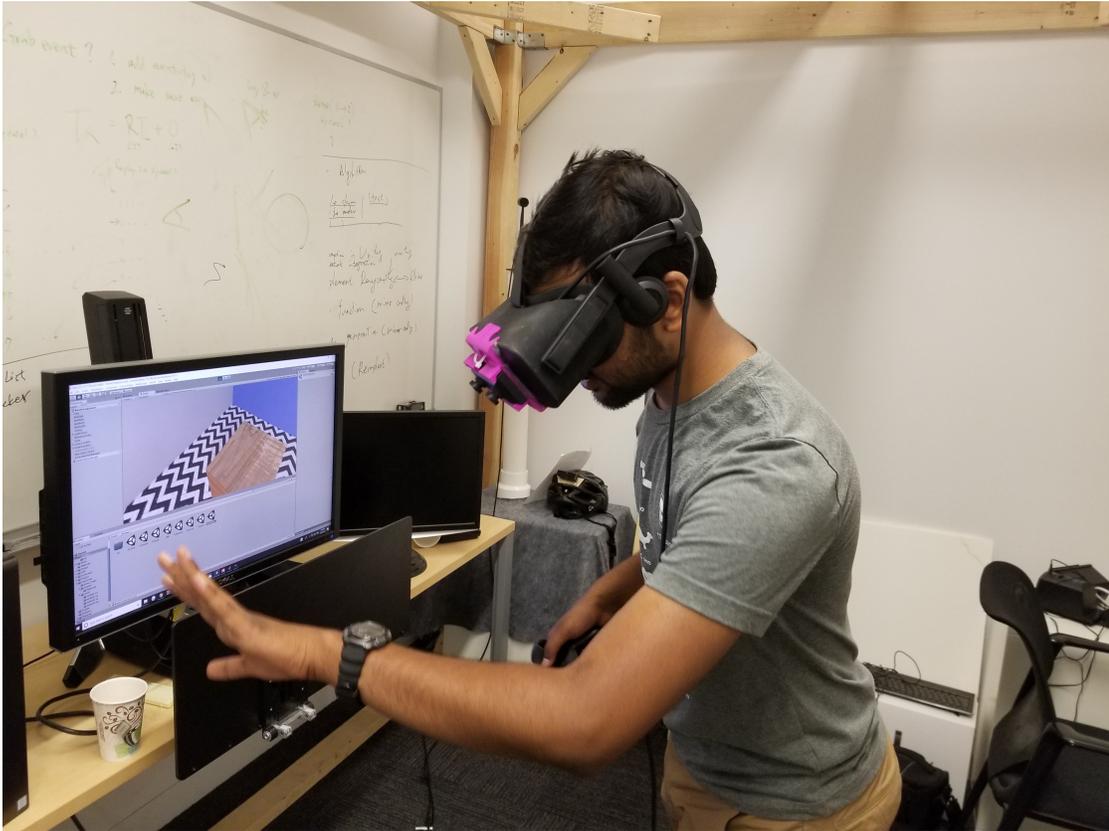


Figure 3.6: A non-participant volunteer examines a scene in head-mounted display configuration.

The Head-Mounted Display (HMD) Configuration employs the Oculus Rift and an Oculus Touch controller in their typical roles within a virtual reality system. Users were asked to wear the Oculus headset and interact with virtual scene by moving about the experiment room and inspecting or manipulating objects with their headset and controller (Figure 3.6). Due to the limited size of the experiment room, participants were also equipped with a “teleporter”

activated by pressing the joystick of the controller. The teleporter could be used to jump discontinuously between points in the virtual world. This scheme is common for virtual reality games and was chosen for our experiment because it is more user-friendly than other common forms of locomotion [4].

Whereas the Fish Tank Configuration includes both egocentric and allocentric motion patterns, the HMD Configuration is purely egocentric, as all motion – including teleportation – is handled with respect to the user's own position and orientation.

CHAPTER 4

RESULTS

In this chapter we explore the data acquired during our user study using the tools of statistical analysis. We begin by introducing the dependent variables broadly. We then answer questions around speed, accuracy, and user perception in terms of these variables in sections 4.2, 4.3, and 4.4, respectively. We finish with a summary of our statistical findings, leaving interpretation of the results to Chapter 5.

4.1 Dependent Variables

Our experiment was designed to ascertain whether performance in three material understanding tasks was affected by the usage of egocentric, allocentric, and hybrid navigation schemes. Performance was measured by the following dependent variables:

- Average time taken to complete each task.
- Percentage of correct responses for the Scanning and Comparison tasks.
- Average final error in shading parameter for the Tuning Task.
- Self-reported comfort during each task on a Likert-type scale [25].
- Self-reported difficulty during each task on a Likert-type scale.
- Self-reported intuitiveness during each task on a Likert-type scale.
- Self-reported confidence during each task on a Likert-type scale.

In our evaluation, a superior navigation scheme should reduce average time per task, tuning parameter error, and perceived difficulty, while increasing percentage of correct responses, comfort, intuitiveness, and user confidence.

Average time taken per task, percentage of correct responses, and average Tuning Task error were all measured by software logging. Timing data was measured as the duration from the moment a new trial was loaded to the moment the user gave a verbal response to the prompt. Response correctness was recorded with keyboard commands operated by an experimental observer. Tuning Task error was calculated as the difference in nanometers between the reference object's diffraction grating spacing, and the tuned object's diffraction grating spacing at the end of each Tuning trial. The remaining self-reported quantities were recorded by a written post-experiment survey.

In contrast to the above, task, user, hardware configuration, and recording procedure were all considered as independent variables in our analysis.

4.2 Speed

Each of the Comparison, Tuning, and Scanning tasks involved a series of individual trials for the user to complete. The participants' speed was measured by the time taken to perform each trial. In general, the distribution of user response times was highly right-skewed. We found that these samples closely matched a log-normal distribution in shape. The log-normal distribution has been previously proposed as an appropriate model of user response times [45], so we conducted our analysis under the assumption of log-normality.

One convenient property of the log-normal distribution is its duality with the normal distribution. Under the assumption of a log-normal distribution, when we apply the natural logarithm to each sample response time, we receive a collection of log-space response times which are normally distributed. The resultant data set is compatible with the repeated measures ANOVA available to us for hypothesis testing.

The one-way repeated measures ANOVA was a natural choice for analysis because each participant attempted the same task in three separate conditions. We cannot expect a user’s speed in one configuration to be independent from their speed in another, indicating that the repeated measures ANOVA is required. An additional benefit of the repeated measures ANOVA procedure is that it is generally more powerful than the independent ANOVA [40]. This is especially important to our design due to our small number of participants.

In the following subsections, we summarize our analysis of the timing data from each of the three tasks based on our assumption of log-normality.

4.2.1 Comparison

Table 4.1: Comparison Task Response Time: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	0.110	2	0.055	0.547	0.595	0.099	0.000
Residual	1.005	10	0.100				

As explained in the introduction to this section, we performed a logarithmic transformation on participant response time data collected during the Compar-

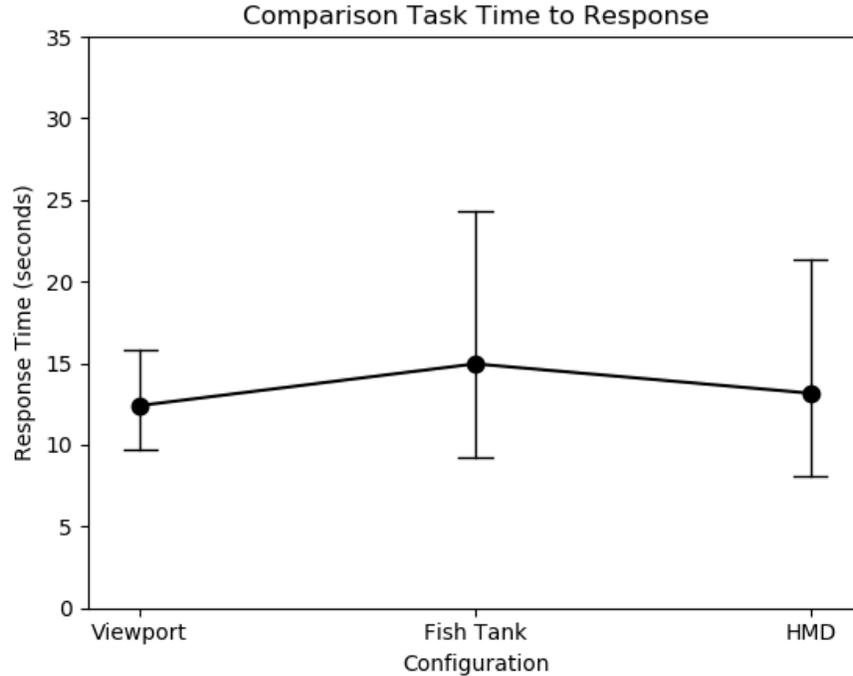


Figure 4.1: Combinatorial plot of mean response time for the Comparison Task.

ision Task. We averaged each participant’s logarithmically transformed response time and conducted a repeated measures ANOVA on the averages (Table 4.1). The ANOVA revealed no statistically significant relationship between log response time and hardware configuration ($p = 0.595$). Figure 4.1 shows a combinatorial plot of the response times across the three configurations with 95% confidence intervals superimposed over each datum. There is a large overlap between each pair of confidence intervals, indicating that no conclusion can be reached with the current data.

It is difficult to produce a valid power analysis from the results obtained from our repeated measures ANOVA. Although our η^2 (0.099) indicates a small or medium effect, this metric is known to overestimate effect size for small sam-

ple sizes [44]. ω^2 , known to be a less biased estimator of effect size for small N [29], indicates trivial or nonexistent effect size.

If we accept the risk of a power analysis based on η^2 , we can conclude that a pool of 330 participants should yield enough power (80%) to determine whether or not there is a significant interaction between exocentricity and speed of execution.

4.2.2 Tuning

Table 4.2: Tuning Task Response Time: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	0.209	2	0.105	1.478	0.274	0.228	0.011
Residual	0.707	10	0.071				

Just as we did for the Comparison Task, we performed a log transform of the response time data before beginning our analysis for the Tuning Task. A repeated measures ANOVA (Table 4.2) showed no significant effect of configuration on the average response time of users during the Tuning Task ($p = 0.274$).

Confidence intervals around the log-mean of each of the three configurations show a longer duration on average in the HMD configuration than in the other two, although this difference is not significant (Figure 4.2).

As mentioned in the previous section, we must be cautious in treating a power analysis based on η^2 as unbiased. With that warning in mind, such a power analysis indicates that to achieve 0.8 power at the 0.05 significance level, we would need a pool of 210 participants

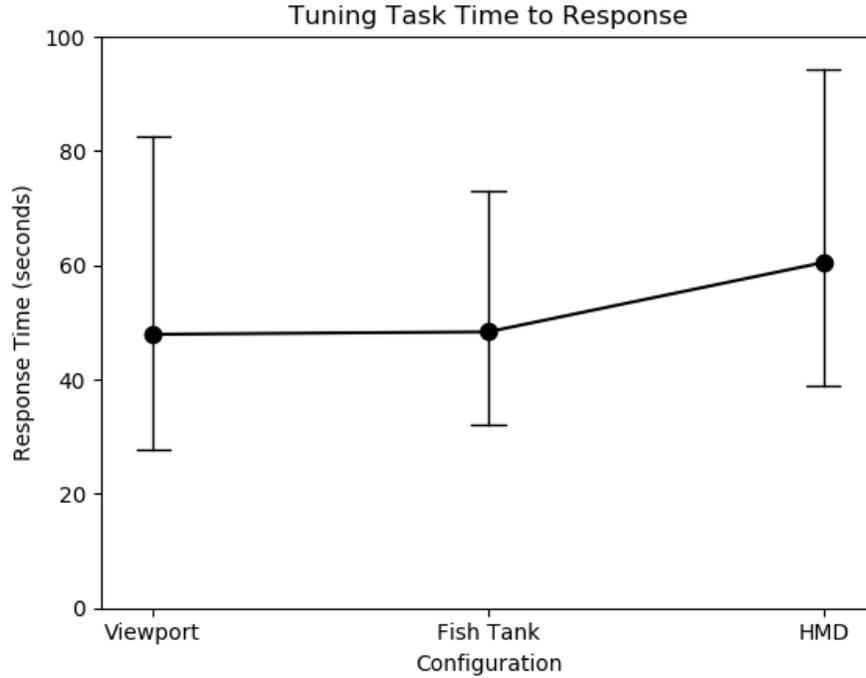


Figure 4.2: Combinatorial plot of mean response time in the Tuning Task .

4.2.3 Scanning

Table 4.3: Scanning Task Response Time: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	0.206	2	0.103	0.724	0.509	0.126	0.000
Residual	1.425	10	0.143				

The analysis of user response times in the Scanning Task proceeded identically to that of the Comparison and Tuning Tasks. A repeated measures ANOVA (Table 4.3) found no significant effect of hardware configuration on user response time in the Scanning Task ($p = 0.509$). Confidence intervals of mean response time, shown in Figure 4.3, overlap heavily. An η^2 power analysis, with the same caveats as before, shows that 80% power at the 0.05 significance level

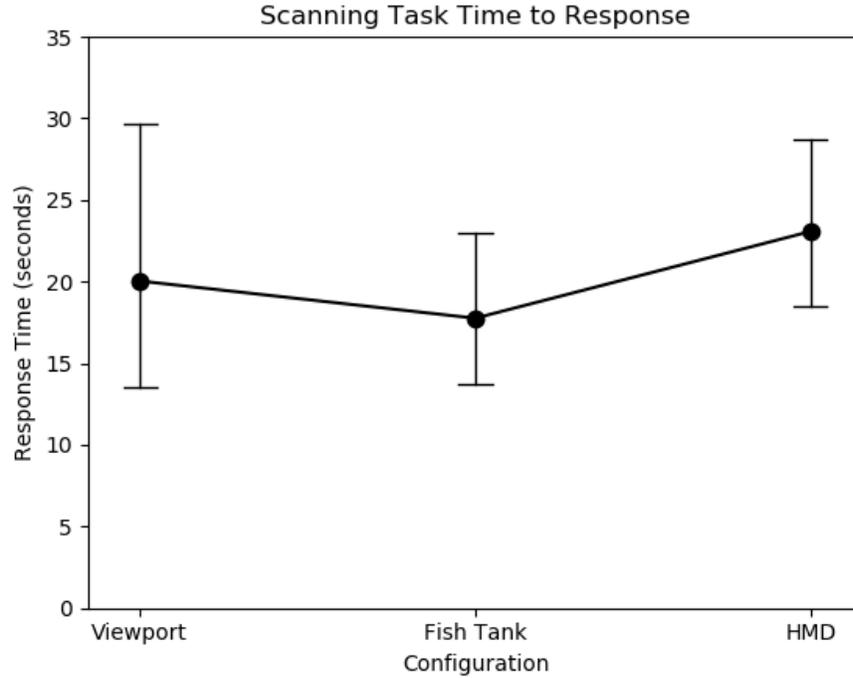


Figure 4.3: Combinatorial plot of Scanning Task response times.

could be achieved with 204 participants.

4.3 Accuracy

Another set of dependent variables we explored correspond to user accuracy while performing each task. For each of the Comparison, Tuning, and Scanning tasks, we analyzed an appropriate metric of user accuracy for dependence on exocentricity condition (i.e. hardware configuration). In this section, we explain our choice of metrics, present our findings, and analyze the results. We do this for each of the Comparison, Tuning, and Scanning tasks in that order.

4.3.1 Comparison

Table 4.4: Comparison Task Accuracy: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	0.008	2	0.004	0.514	0.613	0.093	0.000
Residual	0.080	10	0.008				

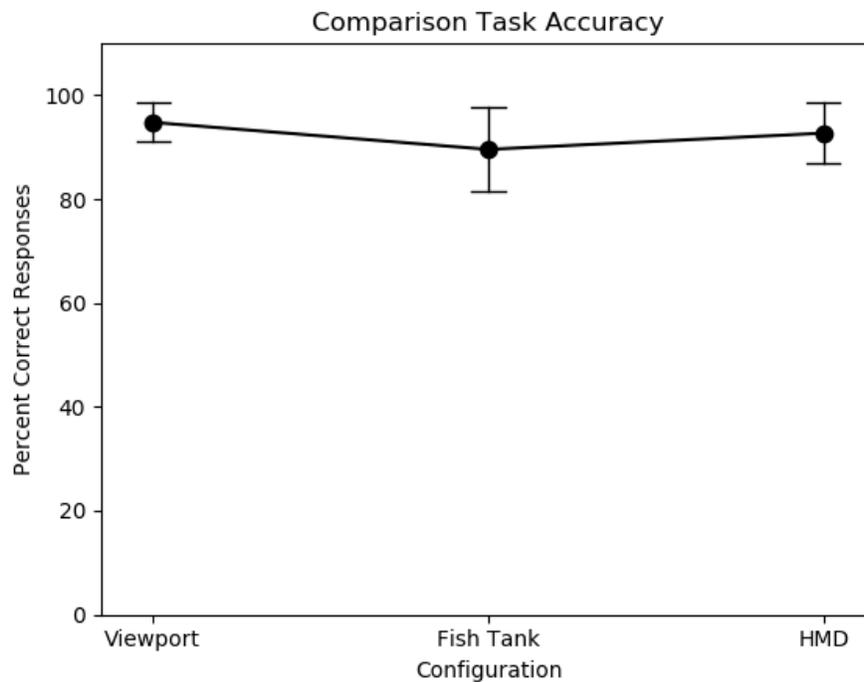


Figure 4.4: Combinatorial plot of Comparison Task accuracy.

The Comparison Task required that users give a boolean response to the question "are these two boxes made from the same material" sixteen times. The user response was considered correct in the following two cases:

1. The shading parameters were identical and the user responded "yes."
2. The shading parameters differed and the user responded "no."

Otherwise, the responses were considered incorrect. We defined a user’s Comparison Task accuracy as the percentage of correct responses given by the user during the Comparison Task. Accuracy was fairly high in each of the three configurations, with the lowest score for any user in any configuration at 75%,

A repeated measures ANOVA was run on the accuracy averages from the Comparison Task (Table 4.4). The ANOVA did not indicate that hardware configuration had a significant effect on accuracy ($p = 0.613$). Furthermore, 95% confidence intervals constructed around the per-subject average accuracy (Figure 4.4) showed large overlap.

As explained in the previous section, η^2 may overestimate effect size for small studies like ours. However, ω^2 is too small for many of our metrics, so we include an η^2 power analysis for completeness. Based on our η^2 , we should be able to determine if an effect on accuracy due to exocentricity exists with 0.8 power at the 0.05 confidence level given 378 participants.

4.3.2 Tuning

Table 4.5: Tuning Task Accuracy: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	127.4	2	63.70	0.277	0.763	0.053	0.000
Residual	2296.4	10	229.64				

The design of the Tuning Task allowed us to obtain a continuous error metric for each Tuning Task trial. During each trial, users adjusted a shading parameter to match a reference. We then used the linear difference between the user’s

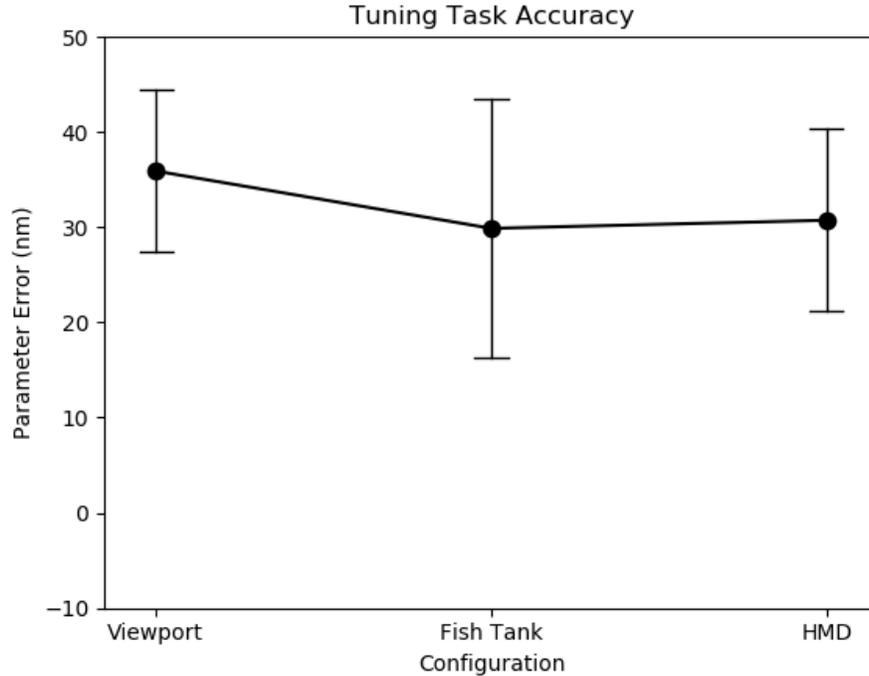


Figure 4.5: Combinatorial plot of Tuning Task error.

approximation of the shading parameter and the reference’s shading parameter as an error metric. Because the shading parameter was diffraction grating spacing, error in this task has units of nanometers.

We removed the trials whose error was more than three standard deviations above the mean error from our analysis. We believe that this removal is justified because these errors were caused by a misunderstanding of the task, rather than failure to evaluate the difference between materials. Our conclusion that these errors are due to misunderstanding is supported by the fact that they all tended to occur within the first few trials that a user attempted.

In order to get a per-user measure of accuracy, we averaged each user’s error (after removing outliers) per configuration. We ran a repeated measures

ANOVA on the resultant average errors (Table 4.5). The ANOVA did not reveal a significant effect ($p = 0.763$) of hardware configuration on error. Although it is interesting that Viewport error tended higher than the other configuration-errors, 95% confidence intervals show a heavy overlap between the mean errors for each configuration (Figure 4.5).

An η^2 power analysis, such as it is, indicates that a collection of 1146 users at could provide 80% power at the 0.05 significance level.

4.3.3 Scanning

Table 4.6: Scanning Task Accuracy: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	0.002	2	0.001	1.000	0.402	0.167	0.000
Residual	0.012	10	0.001				

Accuracy data for the Scanning Task was collected and processed in a similar manner to accuracy data from the Comparison Task. For each virtual object, users were asked to determine whether the shape inscribed on the object was a square, a star, or a circle. User responses were considered correct if and only if they matched the true shape inscribed on the virtual object. User averages were then taken and analyzed as in the Comparison Task. The average accuracies in the Scanning Task were quite high, with every user-configuration score about 90%.

A repeated measures ANOVA (Table 4.6) found no significant effect of configuration on average accuracy ($p = 0.402$), and heavily overlapping confidence

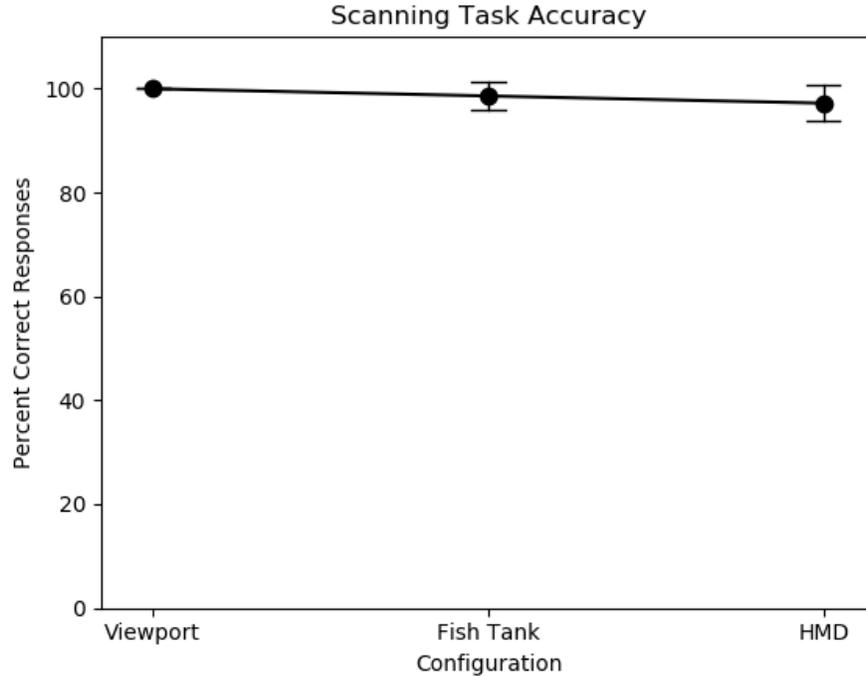


Figure 4.6: Combinatorial plot of Scanning Task accuracy.

intervals (Figure 4.6). An η^2 power analysis indicates that for 0.8 power at the 0.05 significance level would be achieved with 120 participants.

4.4 User Perception

We collected survey data from participants with respect to their perceptions about each of the hardware configurations they used. Users were asked to rate the comfort, intuitiveness, difficulty, and confidence they experienced in each configuration on a Likert scale ranging from 1 to 10. Traditionally, a Likert-style scale contains a neutral option, but there is precedent for an even scale, and even scales are indicated when users are likely to experience ambivalence between options [27]. Users completed the same task in each configuration so that only

configuration varied within subject. A Latin squares experimental design [19] was chosen to account for order-dependent nuisance effects.

Ordinal data obtained from the survey was distributed in a bell shape. We modeled each of the the underlying distributions as Gaussians. Following the assumption of normality, we ran repeated measures ANOVAs on each of the four user prompts. Our justifications for using this test mirror those given in Section 4.2 and 4.3. In the following sub-sections, we show the results of our analysis in order on comfort, intuitiveness, difficulty, and confidence — in that order — as reported by participants.

4.4.1 Comfort

Table 4.7: Level of Comfort: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	10.77	2	5.386	2.065	0.142	0.103	0.028
Residual	93.89	36	2.608				

Users were asked to rate their comfort on a scale between one (uncomfortable) and ten (comfortable) of each of the three configurations. A repeated measures ANOVA (Table 4.7) on the results of the survey did not show a significant difference in comfort between the three configurations ($p = 0.142$).

Figure 4.7 shows the mean level of comfort reported by users in each of the three configurations, along with a 95% confidence interval for each mean. While the sample mean for the Fish Tank Configuration is noticeably lower than the others, the confidence intervals all overlap, so no conclusion can be reached.

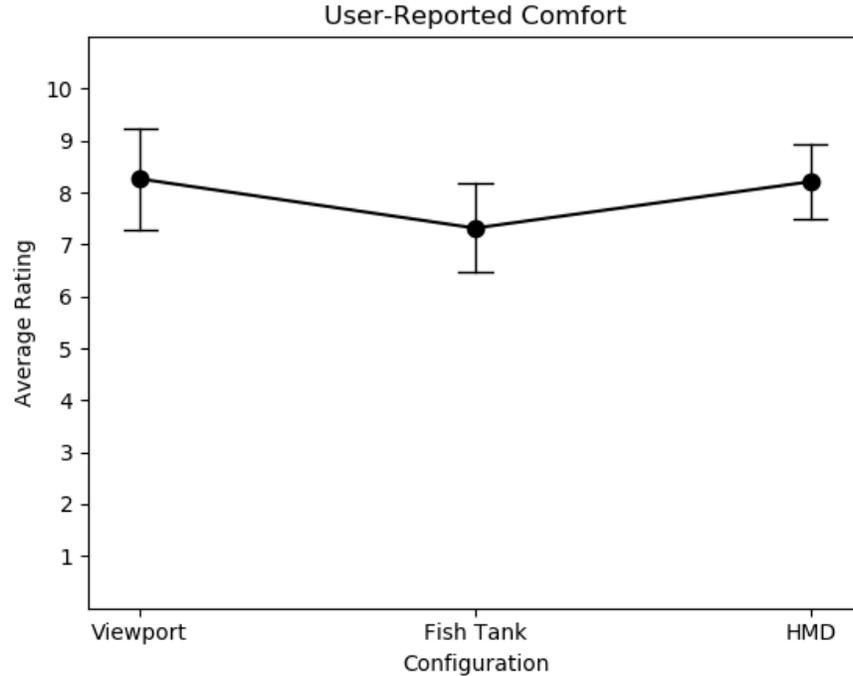


Figure 4.7: Combinatorial plot of participant comfort.

We have warned in previous sections that a power analysis based on η^2 may not be valid for such a small sample size, but we include it for completeness. Given a 0.05 significance level and our estimate of effect size, it would take 306 participants to determine with 0.8 power that a significant effect exists in this variable.

4.4.2 Intuitiveness

Participants were asked to report how intuitive each hardware configuration was on a scale from one (confusing) to ten (intuitive). A repeated measures ANOVA (Table 4.8) showed that hardware configuration had a significant effect on intuitiveness ($p = 0.041$), with the mean score of the HMD configuration

Table 4.8: Assessment of Intuitiveness: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	17.09	2	8.544	3.485	0.041	0.162	0.074
Residual	88.25	36	2.451				

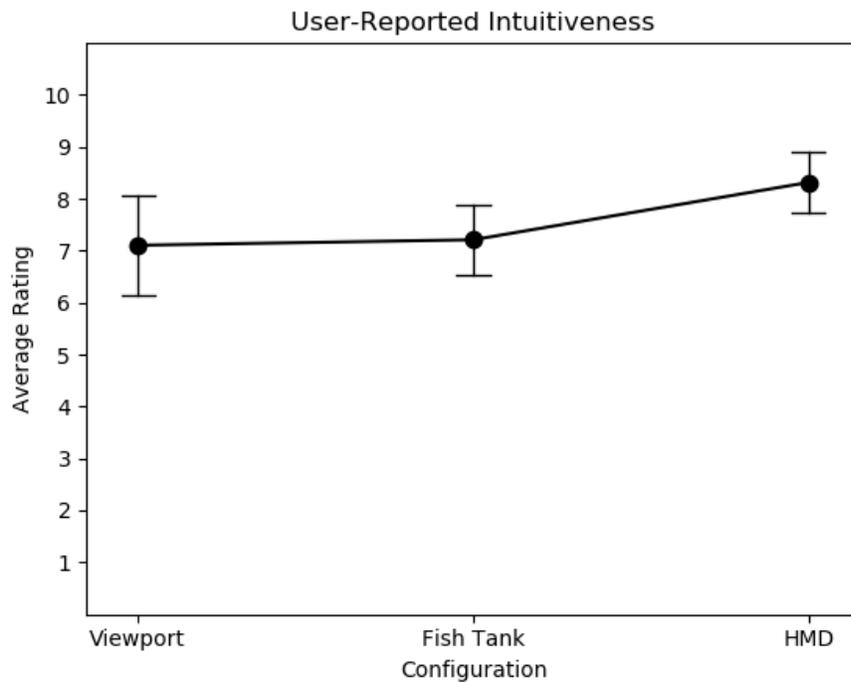


Figure 4.8: Combinatorial plot of intuitiveness as reported by participants.

higher than the mean scores of the other two, as is visible in Figure 4.8. The effect size as understood from η^2 (0.162) is moderate, but as understood from the less biased ω^2 (0.074) would be considered small [7].

Given that an effect was found with the ANOVA, we conducted a post-hoc Bonferroni-corrected t-test [3]. The results of the t-test, presented in Table 4.9, showed a significant difference in user-rated intuitiveness between the Fish Tank and Head-Mounted Display configurations ($p_{bonf} = 0.05$).

Table 4.9: Post-Hoc t-Test for Intuitiveness

Conditions	Mean Difference	Std. Error	t	p_{bonf}
Viewport vs. Fish Tank	-0.105	-0.582	-0.181	1.000
Viewport vs. HMD	-1.211	0.511	-2.371	0.087
Fish Tank vs. HMD	-1.105	0.418	-2.643	0.050

4.4.3 Difficulty

Table 4.10: Level of Difficulty: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	2.667	2	1.333	0.364	0.698	0.020	0.000
Residual	132.000	36	3.667				

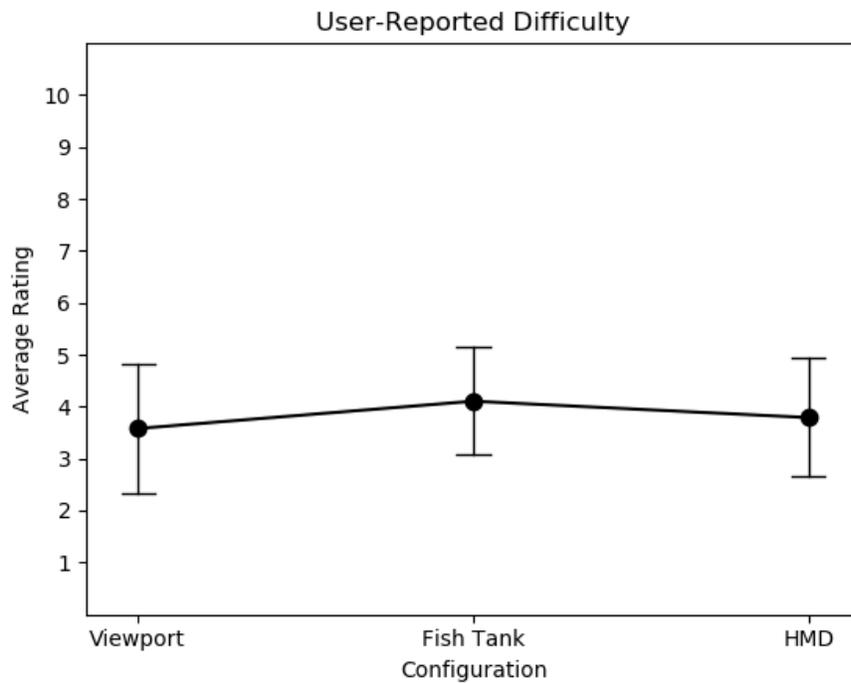


Figure 4.9: Combinatorial plot of difficulty experienced by participants.

We asked subjects to report the level of difficulty they experienced in each of the hardware configurations on a scale from one (easy) to ten (difficult). A repeated measures ANOVA (Table 4.10) revealed no significant effect of configuration on user-reported difficulty ($p = 0.698$). Although on average, the Fish Tank configuration was rated as the most difficult, large overlap can be seen in the confidence intervals for the three configurations (Figure 4.9).

With the usual caveats, a power analysis based on the η^2 statistic indicates that to receive a power of 0.8 at the 0.05 significance level, an impressive 8030 participants would need to be surveyed.

4.4.4 Confidence

Table 4.11: User-Reported Confidence: Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	ω^2
Configuration	1.719	2	0.860	0.436	0.650	0.024	0.000
Residual	70.947	36	1.971				

Users were asked to rate how confident they felt while using each of the three hardware configurations on a scale between one (unsure) and ten (confident). A repeated measures ANOVA (Table 4.11) showed no significant effect on confidence based on hardware configuration ($p = 0.650$). Confidence intervals for the means of user-reported scores showed significant overlap between the three hardware configurations (Figure 4.10). A power analysis based on η^2 indicates that 5580 participants would be required to obtain 0.8 power at the 0.05 significance level.

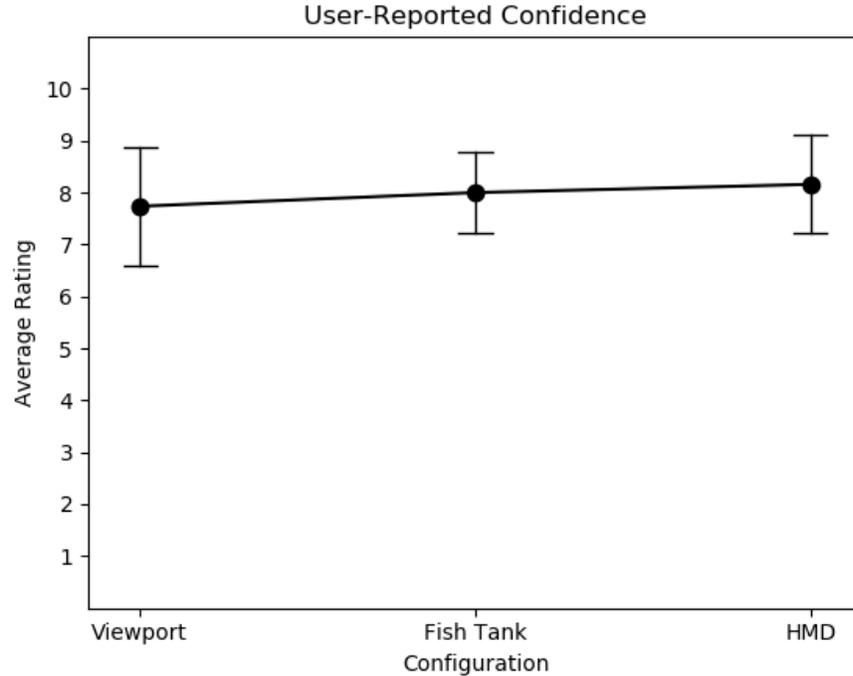


Figure 4.10: Combinatorial plot of user confidence as reported by participants.

4.5 Summary

In this chapter, we examined the data gathered from our user study in detail. Our analysis indicates that of the ten dependent variables we studied, only user perception of intuitiveness can be shown vary significantly between configurations. We further demonstrated through a set of optimistic power analyses that any additional conclusions about the remaining variables would require follow-up studies featuring many hundreds of participants. In the next chapter, we present our qualitative observations and use them to discuss the quantitative results from this chapter.

CHAPTER 5

DISCUSSION

In quantitatively assessing the speed, accuracy, and perception of users in material understanding tasks across three hardware configurations differing in exocentricity, we found only one significant result: that intuitiveness as reported by users varies between hardware configuration. We will explain below why we believe that this significant result is not very informative. With this one result removed, our experiments give us no reason to believe that egocentric, allocentric, and hybrid navigation schemes have a significant effect on the performance of these tasks. We hypothesized that an effect would be readily found, and are surprised that there is so little signal in our data to indicate that an effect exists. We discuss this further in the subsequent sections.

5.1 Intuitiveness: Egocentricity or Confounding Factor?

The only significant result in our experiment indicated that users found the HMD Configuration more intuitive than the Viewport and Fish Tank configurations. While it is tempting to interpret this as a preference for egocentric navigation schemes, we think that our experiment was slightly biased in this metric and that interpreting these data as a win for egocentrism would be premature.

Users of the Viewport and Fish Tank configurations had to use a joystick to change their viewing angles, while users of the HMD configuration could locomote as usual to achieve the same result. We suspect that the intuitiveness attributed to the HMD is influenced by the fact that users did not have to learn

such a control scheme to perform their tasks.

While we may have been more confident in this result if it corresponded to an increase in speed or accuracy, the absence of significant in those domains indicates to us that the effect is limited to the user's perception of the interface. Although this result could be an interesting data point for the advantages of VR as a medium, we think its implications for the intuitiveness of egocentricity should be interpreted conservatively. A more carefully controlled experiment would be necessary to discern between the factors contributing to this result.

5.2 Observed Behaviors May Explain Lack of Effect

Having hypothesized that an egocentric or hybrid navigation scheme would prove superior to an allocentric scheme, we were surprised to find so little indication that an effect existed at all. We returned to our qualitative observations of participants' behaviors to investigate what errors in our own reasoning may have led to this unsupported hypothesis. Informally, we found that user behaviors differed from our expectations during the trials. Specifically, while we expected users to interact constantly with the navigation model to evaluate each material, we found that:

1. Users do not adjust their viewpoints during evaluation.
2. Users explore the view space only once: at the beginning of their first trial.

We explain these assertions in the subsections below. In the next section, we present our hypothesis for why the behaviours we observed might dominate in

a study like our own. Please note that the design of our experiment does not allow for a numerical assessment of these results. We encourage the reader to follow our example by interpreting these observations cautiously.

5.2.1 Users Do Not Adjust Viewpoint During Evaluation

Although at the onset of the experiment we expected that subjects would frequently update their viewpoints while evaluating each material, we now believe that users prefer to rely on a single informative view. This pattern separates the processes of navigation and evaluation temporally, and likely uncouples any interactions between the two processes. We discuss below the observations that suggest this pattern to us.

In the Fish Tank configuration, users relied almost exclusively on the control stick to find views and almost never moved their heads while evaluating. Furthermore, in all three configurations and tasks, users would pause in one view while evaluating rather than moving about with the control stick, locomotion, or head movement. This pattern was evident not just in the Comparison and Tuning tasks, where users had to memorize one surface in order to compare to the other, but also in the Scanning task, where the entire surface had to be understood. In the Scanning task, all but one user relied primarily on rotating the surface of interest rather than by changing their viewpoints, although either method could give the same information to the user.

The fact that users did not adjust their viewpoints during evaluation may partially explain why the exocentricity of each control scheme had no significant effect on material understanding in our study. If users are not engaging

with the navigation scheme, there is likely no meaningful difference between the configurations.

5.2.2 Users Explore the View Space Only Once

We have mentioned that users would use fixed viewpoints, but moreover, users *always* returned to the same fixed viewpoint once they had found one that helped them to answer the experimental prompt. Users found grazing angles especially useful in the Scanning and Comparison tasks, and would always return to a grazing angle upon finding it. Participants in the Tuning Task preferred to look at the shields head-on, and would attempt to return to the same position in front of each shield to evaluate. One user in the tuning task went so far as to put her back to a wall and squat at a consistent height to replicate a viewpoint she liked for each shield.

It seems then that users not only avoided novel movement while evaluating, but in fact avoided novel motion altogether: further reducing their interaction with the motion model being tested. Unfortunately, we did not ask the users for their motivation in using only the first useful views they found. We conjecture that because these views continued to work between trials, inertia discouraged participants from attempting to find a better viewpoint. No conclusion can be drawn, however, without a follow-up study.

5.3 Are Single-Peak Reflectance Functions the Culprit?

The preliminary observation that participants appeared reluctant to employ multiple views in material understanding tasks surprised us. It is possible that if such a reluctance truly exists it is simply a product of human inertia. That is, if participants have no reason to change their successful strategies, why would they? That said, we found it interesting that each task had a small collection of viewing positions favored by most participants. We suspected there may have been a deeper reason for this, and began to investigate it further, hoping to reveal a new path for future research. We have a tentative explanation for why participants seemed to stick with a single view for each task. To summarize, we believe that very few views give good information for material understanding tasks. To explain more, we introduce the concept of the bidirectional reflectance distribution function (BRDF), a mainstay of the Computer Graphics literature.

5.3.1 BRDF(s)

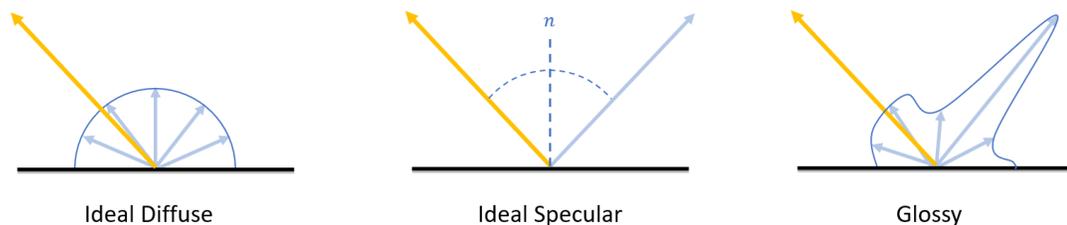


Figure 5.1: Cross-sectional diagrams of simple BRDFs (blue) evaluated for many ω_o with fixed ω_i (yellow). In the ideal specular diagram, a normal is annotated for clarity.

Bidirectional reflectance distribution functions (BRDFs) are the Graphics

community’s tool of choice for describing reflective surfaces. Formally, a BRDF is a function

$$f : \Omega_H \times \Omega_H \rightarrow [0, +\infty],$$

where Ω_H denotes the unit hemisphere normal to the surface at the point where reflection occurs. Given an incident direction ω_i and an outbound direction ω_o , $f(\omega_i, \omega_o)$ gives the proportion of all light entering the surface from ω_i which then leaves the surface toward ω_o . The physics of light transport places additional constraints on BRDFs which we will elide. We encourage the reader to refer to [33] or [39] for a full discussion thereof. A perfectly diffuse “Lambertian” BRDF must satisfy the equation

$$f_{\text{diffuse}}(\omega_i, \omega_o) = \frac{1}{\pi},$$

as light coming in toward the surface from any direction should exit the surface uniformly in every direction. In contrast, an ideal specular BRDF must satisfy

$$f_{\text{specular}}(\omega_i, \omega_o) = \begin{cases} +\infty & \omega_o \text{ reflected over the normal from } \omega_i \\ 0 & \text{otherwise,} \end{cases}$$

as all of the incident light from a single direction is reflected over the normal. Most common BRDFs (such as the Phong model in Figure 5.2) fall in between these two extremes, with some light exiting the surface everywhere, but more of it exiting at angles close to the reflected direction. Schematic diagrams of these simple BRDFs are shown in Figure 5.1.

5.3.2 Single-Peak BRDF(s)

We chose the materials for our study that we thought would present interesting visual effects for users who navigated through the view space. Unfortunately,

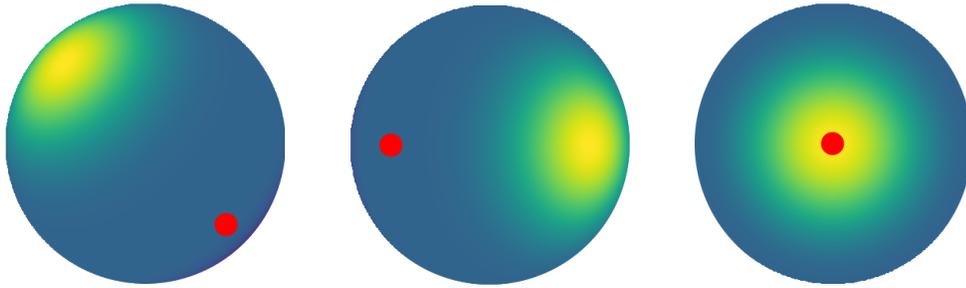


Figure 5.2: Plots of the physically accurate version of the Phong BRDF [24] for several fixed values of w_o , each labeled with a red circle. Each graph should be understood as the orthogonal projection of the hemisphere from which w_o is drawn, with high-intensity directions more yellow and low-intensity directions more blue.

we believe we ended up with materials whose BRDFs shared a structure discouraging the behavior we wanted. The BRDFs of the materials we chose seem to have a single, contiguous, smooth band of large values in the BRDF.

Peaks in the BRDF are interesting because there is a dramatic change in reflectance at their borders. We conjecture that reflective behavior at this boundary is highly visible to users. In the case of single-peak BRDFs, this would mean that there is a small range of viewpoints (the boundary of the sole peak) where users are most likely to detect reflectance features.

It is interesting to us that users seemed to be drawn to grazing angles in the Scanning and Comparison tasks of our study. If our conjecture — that users are attuned to the borders of peaks in the BRDF — were true, this observation could be explained readily. As shown in Figure 5.3, the peaks in the anisotropic BRDF used in the Scanning Task subtend a region of the hemisphere reflected over the normal from the incident light direction. This would indicate that grazing angles give users the most information about the shape of these peaks,

especially where anisotropy changes direction.

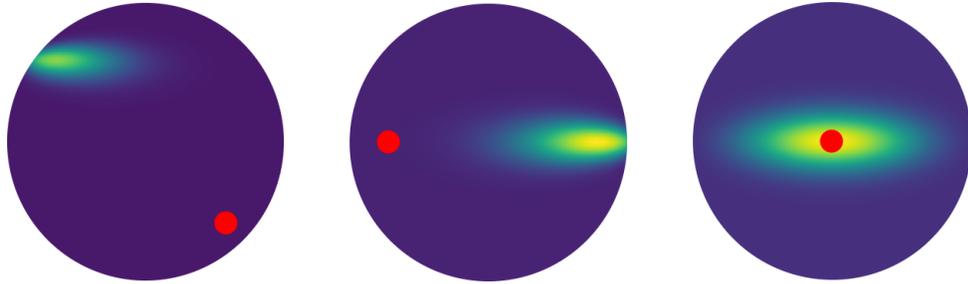


Figure 5.3: BRDFs of the anisotropic model used in the Scanning Task for fixed values of w_i , each labeled with a red circle, as in Figure 5.2

Grazing angles may give a similar view of the peaks of the wood reflectance model used in the Comparison task. This model is better described with a bidirectional subsurface scattering reflectance distribution function (BSSRDF), but there is reason to believe that a single-peak effect could dominate here. Figure 5.4 shows reflected intensities over the hemisphere from which ω_i is drawn for various points on the surface of wood blocks approximated by our shader. Notice that these functions also have a single broad peak.

While the single-peak BRDF model may help to explain why users did not seem to engage with the motion model in the Scanning and Comparison tasks, we do not think it explains the same behavior in the Tuning task. We will leave this assessment to future work.

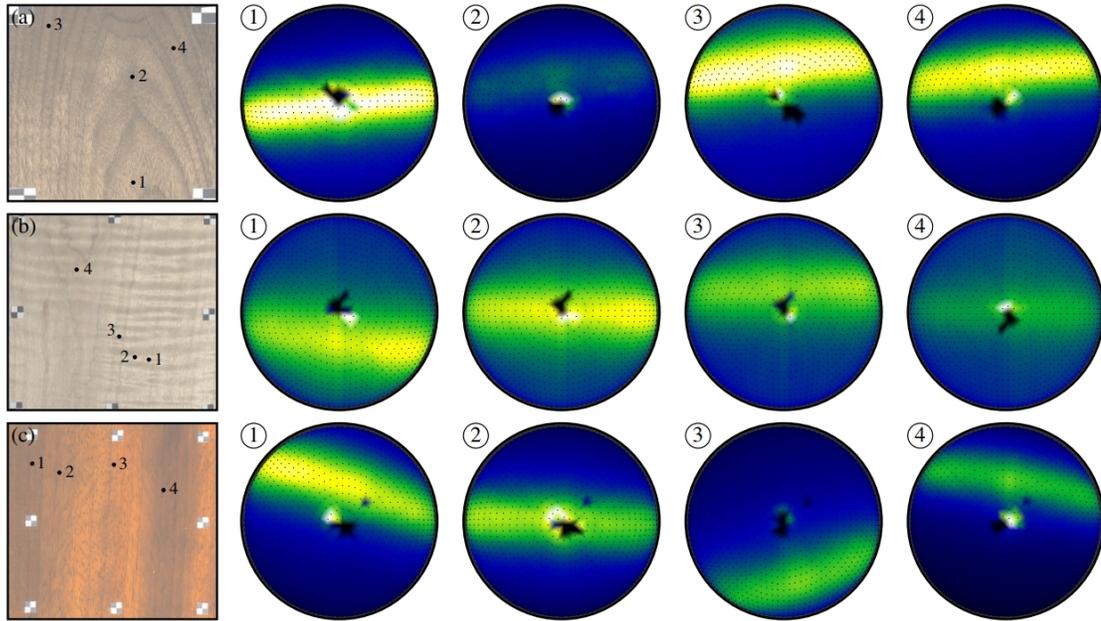


Figure 5.4: Plots of reflected intensity over the incident hemisphere (projected to the equatorial disc) for selected surface points of the Marschner wood model. Reprinted from [26] with permission from the author.

5.3.3 Multi-Peak BRDFs Exist and Can Be Tested

Our hypothesis that users chose not to vary their view because they had found the small region of views that were actually useful could be tested if we were to find BRDFs with more complicated peaks. Fortunately, several well-known models of glinty and glittery surfaces are available for testing [51, 21, 52, 53]. These more complicated models may encourage users to explore the view space more, although it is possible that their visual complexity will make the tasks more difficult overall.

CHAPTER 6

FUTURE WORK

We have found no evidence of a difference in performance in material understanding between users employing egocentric and allocentric motion models. That said, we have not acquired enough data to fully rule out the possibility that such a difference could exist. Future work may focus upon acquiring a larger sample of users, especially expert users, to empower statistical equivalence testing. It is also possible that the relative nascence of VR technology serves as a confounding factor in our experiment. Users more experienced in VR along with better VR technologies, such as holograms, may give better performance against the perfected desktop setup of the Viewport configuration.

We also highlight several observations in our discussion section which could serve as hypotheses for further study. We conjecture that the single-peak nature of the BRDFs used in our study may have made the motion models irrelevant to material understanding. Experiments with more complicated BRDFs, such as those found in glinty or glittery materials, may expose a utility in egocentric or allocentric motion modalities. It is unclear, however, how often these more complex shading models actually arise in a material design pipeline.

To test if users evaluate materials statically, researchers could track the head movement of participants as they inspect physical materials. This experiment could be framed as a semi-structured interview to give clues as to the subjects' thinking process while they evaluate the material. Additional software logging on a similar experiment could provide a quantitative examination of user gaze and motion planning, which may also give more fruitful results. Statistical testing on a larger set of users with head tracking could indicate whether or not

users do truly avoid exploration in the view space.

CHAPTER 7

CONCLUSION

We set out to learn whether an egocentric, allocentric, or hybrid motion model improved user performance in material understanding tasks. To this end we ran a user study on three common material understanding tasks across three hardware configurations, each representing a different level of exocentricity. Our preliminary study did not indicate an effect on speed, accuracy, or user perception of any task. Power analyses based on the η^2 statistic suggested that a study run on 200 - 1000 participants would detect these effects if they existed. We hypothesized that the single-peak nature of the BRDFs we used in the study may have stopped users from engaging with the motion controls after they had found a single good view. Finally, we proposed that a larger study on more complex materials could be the next step in this line of research.

APPENDIX A

FISH TANK VR WITH THE OCULUS RIFT

Calibrating a Fish Tank Virtual Reality installation is equivalent to constructing the viewing frustum defined by the user’s retina and the computer monitor for each frame. Deering’s 1992 paper gives an explicit form for the four-by-four projection matrix that produces this frustum [9]. However, this construction requires knowledge of the position of the retina with respect to computer monitor.

With a bit of linear algebra it is possible to track the retina and screen in the Oculus Rift’s global coordinate system (hereafter called “world space”). We can use the pose of a controller mounted on the user’s head with some fixed offset to approximate the position of the retina in world space for each frame. To track the screen, we need some way to assign world space coordinates to real-world positions.

We use a bootstrapping method developed for the Robotic Modeling Assistant [32], in which a needle is attached to the VR controller to function as a “pointer” into the real world. Given the controller’s four-by-four local-to-world matrix, P , we know there is some constant offset \hat{o} such that the position of the tip in world space can be expressed in homogeneous coordinates as $P\hat{o}$.

We solve for o_{tip} by holding the needle tip against a fixed point in physical space and recording the controller’s pose in at least four orientations this allows us to construct the linear system. With multiple measurements for the controller pose ($P^{(1)}, P^{(2)} \dots P^{(n)}$) and a fixed needle tip, we see that for any i, j ,

$$(P^{(i)} - P^{(j)})\hat{o} = 0.$$

Assuming that each $P^{(k)}$ is affine, this gives us a 3x3 matrix equation for each $i-j$

pair. Let $Q^{(ij)} = P^{(i)} - P^{(j)}$. Then we have: Equation A.1.

$$\begin{bmatrix} q_{11}^{(ij)} & q_{12}^{(ij)} & q_{13}^{(ij)} \\ q_{21}^{(ij)} & q_{22}^{(ij)} & q_{23}^{(ij)} \\ q_{31}^{(ij)} & q_{32}^{(ij)} & q_{33}^{(ij)} \end{bmatrix} \begin{bmatrix} o_x \\ o_y \\ o_z \end{bmatrix} = \begin{bmatrix} q_{14}^{(ij)} \\ q_{24}^{(ij)} \\ q_{34}^{(ij)} \end{bmatrix} \quad (\text{A.1})$$

Although it is possible that this matrix equation will be uniquely solvable for o after two measurements, it is more likely that there will be a circle of valid solutions, and we will need more measurements to determine a unique solution. To handle these additional measurements, we stack copies of Equation A.1 for each i - j pair into an over-determined system and execute a least squares solve. Our code does this with a QR factorization and gives convincing results.

Once we have a known offset from the controller origin to the needle point, we can record positions of real-world objects in the Oculus Rift's tracking space. We use this capability to record the corners of the screen we will use for Fish Tank VR. From here, we could apply Deering's view-projection frustum equation, but we find it more convenient to construct a projection matrix with `GLFrustum` and use our own view matrix so that we can choose where to position the camera within the virtual scene.

BIBLIOGRAPHY

- [1] A. V. Asutay, A. P. Indugula, and C.W. Borst. Virtual tennis: a hybrid distributed virtual reality environment with fishtank vs. hmd. In *IEEE International Symposium on Distributed Simulation and Real-Time Applications*, 2005.
- [2] Autodesk. Maya documentation, 2018.
- [3] Carlo Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8:3–62, 1936.
- [4] Evren Bozgeyikli, Andrew Rajj, Srinivas Katkoori, and Rajiv Dubey. Point & teleport locomotion technique for virtual reality. In *CHI PLAY*, 2016.
- [5] Neil Burgess. Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, 10(12):551 – 557, 2006.
- [6] A. Byagowi and Z. Moussavi. Design of a virtual reality navigational (vrn) experiment for assessment of egocentric spatial cognition. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012.
- [7] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, 1988.
- [8] Mark Christopher Colbert. *Appearance-driven material design*. University of Central Florida, 2008.
- [9] Michael F. Deering. High resolution virtual reality. In *SIGGRAPH*, pages 195–202, 1992.
- [10] Benedikt Ehinger, Petra Fischer, Anna Gert, Lilli Kaufhold, Felix Weber, Gordon Pipa, and Peter Knig. Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile eeg study. *Frontiers in Human Neuroscience*, 8:71, 2014.
- [11] Arne D. Ekstrom, Aiden E. G. F. Arnold, and Giuseppe Iaria. A critical review of the allocentric spatial representation and its neural underpinnings: toward a network-based perspective. *Frontiers in Human Neuroscience*, 8:803, 2014.

- [12] Okan Erat, Werner A. Isop, Denis Kalkofen, and Dieter Schmalstieg. Drone augmented human vision. In *IEEE Transactions on Visualization and Computer Graphics*, 2018.
- [13] Randima Fernando. *GPU Gems: Programming Techniques, Tips and Tricks for Real-Time Graphics*. Pearson Higher Education, 2004.
- [14] Roland W. Fleming. Material perception. *Annual Review of Vision Science*, 3(1):365–388, 2017.
- [15] E Bruce Goldstein. *Sensation and perception*. Cengage Learning, 2009.
- [16] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modeling the interaction of light between diffuse surfaces. *SIGGRAPH Comput. Graph.*, 18(3):213–222, January 1984.
- [17] Klaus Gramann. Embodiment of spatial reference frames and individual differences in reference frame proclivity. *Spatial Cognition & Computation*, 13(1):1–25, 2013.
- [18] Klaus Gramann, Julie Onton, Davide Riccobon, Hermann J. Mueller, Stanislav Bardins, and Scott Makeig. Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience*, 22(12):2836–2849, December 2010.
- [19] David A Grant. The latin square principle in the design and analysis of psychological experiments. *Psychological bulletin*, 45(5):427, 1948.
- [20] Paul Issartel, Lonni Besançon, Florimond Guéniat, Tobias Isenberg, and Mehdi Ammi. Preference between allocentric and egocentric 3d manipulation in a locally coupled configuration. 2016.
- [21] Wenzel Jakob, Miloš Hašan, Ling-Qi Yan, Jason Lawrence, Ravi Ramamoorthi, and Steve Marschner. Discrete stochastic microfacet models. *ACM Trans. Graph.*, 33(4):115:1–115:10, July 2014.
- [22] William B. Kerr and Fabio Pellacini. Toward evaluating material design interface paradigms for novice users. *ACM TOG*, 35(10):1–35, 2010.
- [23] Roberta L. Klatzky. *Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.

- [24] Eric P. Lafortune and Yves D. Willems. Using the modified phong reflectance model for physically based rendering. Technical report, 1994.
- [25] Rensis Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 140:1–55, 1932.
- [26] Stephen R. Marschner, Adam Arbree, and Jonathan T. Moon. Measuring and modeling the appearance of finished wood. In *SIGGRAPH*, 2005.
- [27] Stephen M. Nowlis, Barbara E. Kahn, and Ravi Dhar. Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research*, 29(3):319–334, 2002.
- [28] Nvidia. Cg 3.1 toolkit documentation, 2018.
- [29] Kensuke Okada. Is omega squared less biased? a comparison of three major effect size indices in one-way anova. *Behaviormetrika*, 40(2):129–147, 2013.
- [30] Thomas Pederson. Egocentric interaction. In *Workshop on What is the Next Generation of Human-Computer Interaction*, pages 22–23, 2006.
- [31] Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. Toward a psychophysically-based light reflection model for image synthesis. In *SIGGRAPH*, 2006.
- [32] Huaishu Peng, Jimmy Briggs, Cheng-Yao Wang, Kevin Guo, Joseph Kider, Stefanie Mueller, Patrick Baudisch, and François Guimbretière. Roma: Interactive fabrication with augmented reality and a robotic 3d printer. In *CHI*, 2018.
- [33] Matt Pharr and Greg Humphreys. *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2010.
- [34] Markus Plank, Hermann J Müller, Julie Onton, Scott Makeig, and Klaus Gramann. Human eeg correlates of spatial navigation within egocentric and allocentric reference frames. In *International Conference on Spatial Cognition*, pages 191–206. Springer, 2010.
- [35] Wen Qi, Russell M. Taylor, Christopher G. Healey, and Jean-Bernard Martens. A comparison of immersive hmd, fish tank vr and fish tank with

- haptics displays for volume visualization. In *Symposium on applied perception in graphics and visualization*, 2006.
- [36] Roman Rädle, Hans-Christian Jetter, Simon Butscher, and Harald Reiterer. The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, ITS '13*, pages 23–32, New York, NY, USA, 2013. ACM.
- [37] Iman Sadeghi, Heather Pritchett, Henrik Wann Jensen, and Rasmus Tamstorf. An artist friendly hair shading system. In *ACM SIGGRAPH 2010 Papers, SIGGRAPH '10*, pages 56:1–56:10, New York, NY, USA, 2010. ACM.
- [38] Thorsten-Walther Schmidt, Fabio Pellacini, Derek Nowrouzezahrai, Wojciech Jarosz, and Carsten Dachsbacher. State of the art in artistic editing of appearance, lighting and material. In *Computer Graphics Forum*, volume 35, pages 216–233. Wiley Online Library, 2016.
- [39] Peter Shirley and Steve Marschner. *Fundamentals of Computer Graphics*. A. K. Peters, Ltd., Natick, MA, USA, 3rd edition, 2009.
- [40] Laerd Statistics. Repeated measures anova, 2018.
- [41] Xin Sun, Kun Zhou, Yanyun Chen, Stephen Lin, Jiaoying Shi, and Bain-ing Guo. Interactive relighting with dynamic brdfs. *ACM Transactions on Graphics (TOG)*, 26(3):27, 2007.
- [42] Unity Technologies. Unity - manual: Unity manual, 2018.
- [43] William Thompson, Roland Fleming, Sarah Creem-Regehr, and Jeanine Kelly Stefanucci. *Visual perception from a computer graphics perspective*. AK Peters/CRC Press, 2016.
- [44] Susan Troncoso Skidmore and Bruce Thompson. Bias and precision of some classical anova effect sizes when assumptions are violated. *Behavior Research Methods*, 45(2):536–546, Jun 2013.
- [45] Wim J. van der Linden. A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics*, 31(2):181–204, June 2006.
- [46] Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. Microfacet models for refraction through rough surfaces. In *Proceed-*

ings of the 18th Eurographics Conference on Rendering Techniques, EGSR'07, pages 195–206, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

- [47] Colin Ware, Kevin Arthur, and Kellogg S. Booth. Fish tank virtual reality. In *Interchi*, 1993.
- [48] Colin Ware and Steven Osborne. Exploration and virtual camera control in virtual three dimensional environments. *ACM SIGGRAPH computer graphics*, 24(2):175–183, 1990.
- [49] Hongzhi Wu, Julie Dorsey, and Holly Rushmeier. Physically-based interactive bi-scale material design. *ACM Transactions on Graphics (TOG)*, 30(6):145, 2011.
- [50] Hongzhi Wu, Julie Dorsey, and Holly Rushmeier. Inverse bi-scale material design. *ACM Transactions on Graphics (TOG)*, 32(6):163, 2013.
- [51] Ling-Qi Yan, Miloš Hašan, Wenzel Jakob, Jason Lawrence, Steve Marschner, and Ravi Ramamoorthi. Rendering glints on high-resolution normal-mapped specular surfaces. *ACM Trans. Graph.*, 33(4):116:1–116:9, July 2014.
- [52] Ling-Qi Yan, Miloš Hašan, Steve Marschner, and Ravi Ramamoorthi. Position-normal distributions for efficient rendering of specular microstructure. *ACM Trans. Graph.*, 35(4):56:1–56:9, July 2016.
- [53] Ling-Qi Yan, Miloš Hašan, Bruce Walter, Steve Marschner, and Ravi Ramamoorthi. Rendering specular microgeometry with wave optics. *ACM Trans. Graph.*, 37(4):75:1–75:10, July 2018.