UNIVERSITY OF CALGARY

3D Tabletop Display Interaction

by

Mark Hancock

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

August, 2010

Abstract

Recent advances in digital technology, both in research prototypes and commercial products, have introduced a plethora of multitouch horizontal display surfaces. Perhaps because most of these devices are flat, initially most of the multitouch interfaces were two-dimensional (2D) in nature. However, on traditional tables, people frequently make use of the third dimension; they pick up, turn over, stack, build, and otherwise manipulate objects on physical tables. Furthermore, they frequently use the visual cues made available by the third dimension, such as viewing the different sides of an object or scene, or hiding something underneath another object. Successful interaction with three-dimensional (3D) objects on tabletops involves both manipulation and visual feedback. Thus, I simultaneously explore the research questions of both viewing and interacting with 3D virtual artifacts on digital tables.

The use of 3D virtual objects on a tabletop display introduces many research questions. For example, most applications that support 3D graphics do so by assuming a single viewpoint directly in front of the display. This assumption is no longer valid when using a large horizontal surface that affords many people working at different sides viewing the 3D virtual scene. It is an open question to what extent this discrepancy in viewing angle is problematic, and, when necessary, how it can be mitigated. Furthermore, the horizontal table imposes a physical barrier to the 3D virtual world, meaning that "touch" input will be within the 2D plane. Another open question is how this 2D information can be used to control 3D virtual objects "below" the table's surface and whether interacting through this surface can enable

the kinds of 3D abilities common to physical artifacts.

To address the research questions about the discrepancy in viewing angle between the different people around a table and the viewpoint used to render a scene, I empirically study this perceptual phenomenon. Results show that, in a tabletop display setting, viewing projected 3D virtual objects from multiple viewpoints is indeed problematic and becomes more problematic as the discrepancy in viewing angle increases. In this dissertation, I describe how to apply these results to 3D applications, either through an understanding of the compromises implied by each design or by using mitigating techniques to reduce the problem.

In this dissertation, I also build on previous work that explores manipulation of a virtual 3D object by introducing several techniques which use the 2D touch input provided by multiple fingers or contact points. Results of a comparative user study showed that both performance and preference increased as participants were provided with more touches to control the virtual objects. While this study only explored 2D movement and 3D rotation (the techniques did not allow lifting of virtual objects), the insight gained was used to create *sticky fingers* and *opposable thumbs*, which extend the three-touch technique to allow lifting (the sixth degree of freedom). By combining the power of this full control over any 3D virtual object with physically-based reactions of other virtual objects and interface components, *sticky tools* provides a framework for 3D tabletop interfaces that eliminates the need for specialized gestures or an abstract menu system.

This framework, together with insights gained from the exploration of viewpoint discrepancy, were applied to the practical application of enabling sandtray therapy, a form of art therapy for children, on a digital table. This application was cooperatively designed with therapists who use sandtray therapy in their regular practice. This application serves as a demonstration of how to apply the concepts in this dissertation to the design of 3D interaction on a tabletop display.

Publications

The materials, ideas, tables, figures, and videos in this dissertation have previously appeared in the following peer-reviewed conference publications. The chapters which make use of the material are noted.

Hancock, Mark, ten Cate, Thomas, Carpendale, Sheelagh, and Isenberg, Tobias (2010). Supporting sandtray therapy on an interactive tabletop. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2133–2142. New York, NY, USA: ACM Press. Material appears in chapter 7.

Hancock, Mark, Nacenta, Miguel, Gutwin, Carl, and Carpendale, Sheelagh (2009). The effects of changing projection geometry on the interpretation of 3D orientation on tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 175–182. New York, NY, USA: ACM Press. Material appears in chapter 3.

Hancock, Mark, ten Cate, Thomas, and Carpendale, Sheelagh (2009). Sticky tools: Full 6DOF force-based interaction for multi-touch tables. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 145–152. New York, NY, USA: ACM Press. Material appears in chapter 6.

Hancock, Mark and Carpendale, Sheelagh (2007). Supporting multiple off-axis viewpoints at a tabletop display. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*, 171–178. Los Alamitos, CA, USA: IEEE Computer Society. Material appears in chapter 3.

Hancock, Mark, Carpendale, Sheelagh, and Cockburn, Andy (2007). Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1147–1156. New York, NY, USA: ACM Press. Material appears in chapter 5.

Acknowledgements

There have been many people over the last several years who have helped to make this dissertation possible, and I wish I could thank them all, but that would add too many pages to an already large thesis. Instead I will highlight some of the most memorable and significant.

First I would like to thank my supervisor, Sheelagh Carpendale. It is hard to describe just how great of a supervisor Sheelagh has been, but more importantly, she has also grown to be one of my closest friends. Sheelagh takes great care to develop and nurture the very best qualities in all of her students, and this dissertation is an example of how skilled she is at that process. I look forward to continuing this friendship.

Thanks also go to my committee members Saul Greenberg, Faramarz Samavati, Gerald Hushlak, and Gordon Kurtenbach. Their feedback on my work and challenging questions have helped to strengthen the contributions.

I would also like to thank each and every member of the iLab during my time here. I have worked in many labs throughout my time as a student, and this lab far exceeds them all, and this is entirely because of its members. The people here have helped me both to produce good research and to maintain my sanity. I would especially like to thank: Petra Isenberg, who endured almost this entire process as my cubicle buddy, for always asking questions that made me think, answering questions about just about anything, and especially for reigning in my poor colour choices; Christopher Collins for being one of my "twins", almost always agreeing with my opinions, and for packing a whole bunch of cubicle-buddy goodness into an unreasonably short stay in the role; Jim Young for being another "twin", going to coffee with me almost every morning, great discussions and advice, and providing me with an endless supply of video games, reading material, and silly puns; Miguel Nacenta for always ensuring scientific rigour and for sharing a unique excitement about the work we were doing; Tobias Isenberg for making great hot chocolate and for always sending great students my way; Tony Tang for lifting my spirits at one of the lowest times in the process (sometimes shared agony does wonders). There are just too many great people in this lab to name them all by name, but for the rest: you know who you are.

I have had several internships during my time as a student, but the internship at Mitsubishi Electric Research Labs (MERL) and the people there had a significant impact on my research. Specifically I would like to thank Chia Shen, Kathy Ryall, and Clifton Forlines, who inspired a lot of thought about movement and rotation of objects just before I began my Ph.D. work.

The financial support from Alberta Ingenuity, iCORE, and the Natural Sciences and Engineering Research Council of Canada (NSERC) have been generous and have made it possible to both live and do research at the same time.

I would also like to thank my family: Kenneth (Dad), Tuula (Mom), Raija, Jeffery, Peter, Nicole, Danielle, Isaac, Kyle, Mackenzie, Jacob, and Brynn. Their support has made it possible to get through some of the hardest family times that we've had to endure and still manage to get my Ph.D.

Last, I would like to thank my partner, Kimberly Tee. There were many times during this past several years that have been particularly difficult emotionally, and she has been there for me through them all. On top of this companionship, she has also been great at editing whatever I send her way and knocking sense into me before I get into trouble with my (usually quite strong, though sometimes misguided) opinions.

To my mother

Tuula Hancock

Table of Contents

List of Tables

List of Figures

List of Acronyms

2D two-dimensional. iii, iv, 2, 4–6, 12, 13, 20–24, 26, 28, 29, 32–34, 38, 40–42, 44–47, 49, 51, 83, 85, 86, 94, 116–118, 120, 126, 127, 132, 153, 154, 170, 193, 198

3D three-dimensional. iii, iv, 1–14, 16, 18–30, 32–35, 38, 40–49, 51–53, 56, 58, 59, 61, 64, 67, 73, 75, 81–88, 92, 94, 96, 97, 108, 110–113, 115–121, 123–126, 131, 145–147, 149–154, 156, 163–166, 170, 174, 177, 190–198, 277

ANOVA analysis of variance. 58, 62, 64, 70, 72, 77, 80, 136, 139, 140

CAD computer automated design. 165

COP centre of projection. xvii, 8, 21–26, 44, 45, 47–53, 55–57, 62, 64, 65, 70, 71, 73, 77, 81–85, 94, 96, 131, 167, 194, 196, 277

CRT cathode ray tube. 14, 15, 28

DI diffuse illumination. 17, 98, 109, 110, 123

DLP digital light processing. 15

DOF degrees of freedom. 5, 29–35, 39, 41, 42, 97–99, 104–106, 112, 114–116, 133

DOFI degrees of freedom of the input. 105, 109–114, 123, 124, 126, 128, 141, 142, 146, 149

DOFO degrees of freedom of the output. 8, 105, 108, 110–115, 119, 120, 123, 124, 126, 128, 130, 131, 139, 144, 146, 147, 149, 150, 177, 194

DOP direction of projection. 22, 57

DRE differential rotation effect. 26, 77, 82, 83

FTIR frustrated total internal reflection. 17, 98, 109, 110, 116, 123, 176

GUI graphical user interface. 37

HCI human-computer interaction. 1, 4, 6, 8, 11–14, 26, 29, 30, 33, 35, 97, 98, 104, 105, 113, 164, 165, 192, 193, 195–197

IR infrared. 17, 110, 111

J joints. 154, 155

LA large-absent. 56–58, 63, 64, 70, 71, 73, 78, 80

LCD liquid crystal display. 15, 17

LED light emitting diode. 17, 41

LP large-present. 56–58, 63, 64, 70–73, 78, 80

MA medium-absent. 56–58, 62–64, 70–73, 78, 80, 82

MOF magnitude of freedom. 35, 98, 99, 104, 106, 112, 114, 115

MOFI magnitude of freedom of the input. 106, 107, 112–114

MOFO magnitude of freedom of the output. 106, 107, 109, 112, 114

NA none-absent. 56–58, 63, 70–73, 78, 80

NP none-present. 56–58, 63, 64, 70, 71, 73, 78, 80

P proxies. 154, 155

POV point of view. xvii, 21, 23–26, 44, 45, 47–50, 52–54, 56, 57, 65, 70, 73, 77, 81–84, 93, 94

RNT rotate n' translate. 32, 33, 39, 124, 126, 132, 146

SDG single-display groupware. 38

SDK software development kit. 38

SF sticky fingers. 154, 155

SF+OT sticky fingers with opposable thumbs. 154, 155

ST sticky tools. 154, 155

TCT task completion time. 61, 63, 64, 69, 72, 75, 78, 80, 81, 136, 140, 141, 146

TUI tangible user interface. 2, 43, 164, 196

WIMP windows, icons, menus, and pointers. 46

Introduction 1

The combination of rich three-dimensional (3D) interaction with rich 3D visuals to support co-located collaborative work is the focus of this dissertation. I concentrate specifically on digital tabletop display technology, since it affords many people standing or sitting around the display, and is a promising medium for co-located collaboration. Limiting the work to both 3D and tabletop display technology has allowed me to focus my exploration and develop a deeper, fuller understanding of this design space. This deeper understanding contributes to the larger field of human-computer interaction (HCI) and interaction design by opening the door to the possibility that the use of multitouch technology together with 3D visuals can improve the design of a co-located collaborative tabletop display application.

To introduce this work, I first motivate the use of multitouch technology, horizontal surfaces, and 3D interaction (section 1.1). I then describe how my research fits within the larger scope of HCI, graphics, and perception literature (section 1.2). Next I specify the issues that are addressed in this thesis (section 1.3) and the methods I use to address them (section 1.4). I then list the specific contributions made in this work (section 1.5) and present an overview of the structure of the thesis (section 1.6).

1.1 Motivation

The day-to-day lives of many people have been drastically changed by technology. While paper, books, pens, and other physical artifacts are still commonly used, people must also be familiar with their digital counterparts: digital documents, webpages, mice, and keyboards.

The digital world offers many advantages, such as the ability to easily modify, share, or publish information. Nonetheless, there are many aspects of physical artifacts that are lost in the current digital world.

A large part of early human development involves the discovery and mastery of how to interact with physical objects. At a young age, we learn to grasp, pick up, turn over, stack, build, and otherwise manipulate these objects. Thus, when we encounter physical artifacts in the real world, we bring a significant amount of prior knowledge about a great variety of rich interactions through which we can interact with these artifacts. In the digital world, we must relearn how virtual objects act and react.

One of the most common properties of the physical world that is broken in the digital realm is the property that to invoke change in our environment, we must apply some sort of force to a physical artifact. While this force can be made to propagate through some complex system (e. g., a system of pulleys or levers), there is always some form of initial contact or force through touching something. For this reason, digital technology that allows people to physically touch the digital display offers a compelling medium through which to explore leveraging some of the physical properties we learn about in our childhood. Similarly, devices that sense other forms of physical motion, such as Nintendo's Wii and tangible user interfaces (TUIs) [e. g., Underkoffler and Ishii, 1999], provide a means of achieving embodied interaction [Dourish, 2001] with a computer through physical touch and motion.

Another fundamental aspect of the physical artifacts with which we are familiar is that they are 3D. This three-dimensionality is precisely what enables some physical interactions, such as flipping and stacking. However, digital displays are typically flat and present information in a two-dimensional (2D) plane. Nonetheless, computer graphics technology can be used to mimic some of the 3D cues available in physical space so that people can perceive artifacts as having depth, despite being displayed in a 2D plane. Thus, the combination of rich physical interactions with rich 3D visuals has the potential to leverage familiar interactions from the physical world in the way that we interact with computer technology.

Familiarity with and frequent use of physical artifacts is ubiquitous, so much so that these artifacts can take on a variety of meanings and can become an integral part of both our understanding of the world and our ability to communicate with others in it. Indeed, a plausible definition for communication might include the common understanding of something in the world, whether it be a physical artifact or some abstract idea. People can even communicate implicitly through the use of objects when they are co-located, due to this common understanding. For example, a person looking at their watch may indicate that they are in a hurry, a person sipping the dregs of their drink may indicate that they are still thirsty, or a pilot flipping a switch on a panel may notify a co-pilot of a change in flight conditions. This implicit communication through artifacts can make collaboration far richer than when it is absent. Thus, the use of physically familiar interactions in the digital realm is a promising means to improve communication in a collaborative environment.

1.2 Scope

The specific area of research that this dissertation addresses is *3D tabletop display interaction*. Figure 1.1 shows how this research fits across the boundary of two broader areas: human-computer interaction and graphics & perception. The area inside human-computer interaction can be narrowed down from computer-supported cooperative work into co-located collaboration. More specifically, my focus is on tabletop display environments. The broad area of graphics & perception can be further narrowed to 3D graphics & depth cues. Narrowing the research in this way has enabled the in-depth exploration of 3D tabletop display interaction, which has led to a deeper understanding of how to create 3D tabletop display applications.

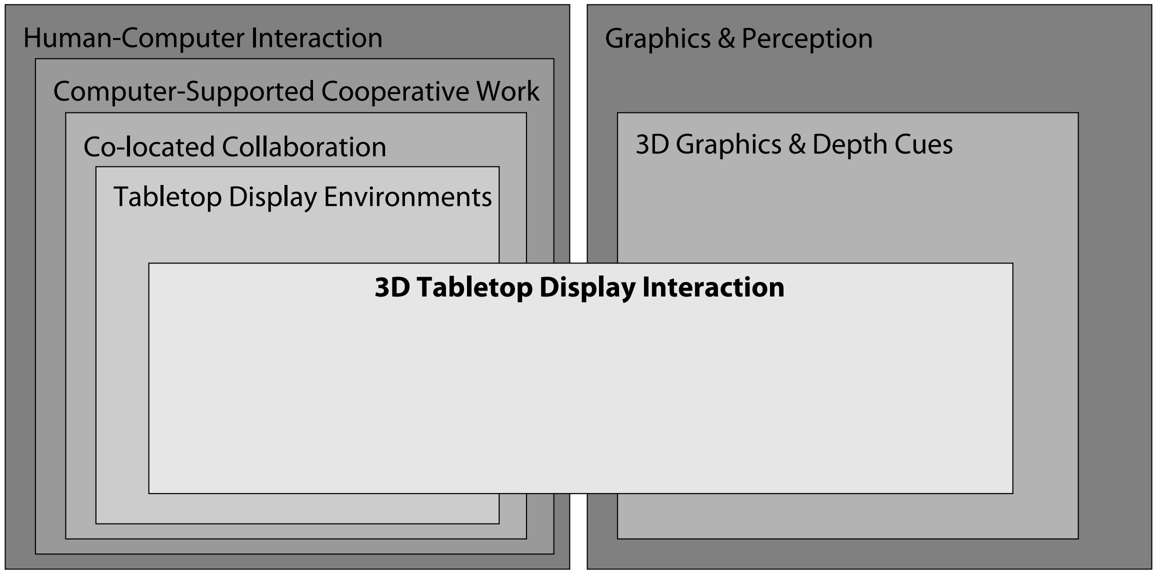


Figure 1.1: The context of my research.

1.3 Issues

Of the many interaction and perceptual issues that arise when designing and developing 3D interaction for co-located collaboration tabletop displays, in this thesis I focus on the following three issues.

Issue 1: **Perceiving 3D depth cues on a horizontal surface**. Combining tabletop displays with 3D virtual objects introduces a large variety of challenges in the areas of visual perception and HCI. In particular, large interactive surfaces afford many people collaborating at the display simultaneously, introducing several points of view of the same displayed information. However, the existing technology that enables 3D depth information in a 2D display assumes at a basic level that the information will only be viewed by one person. Thus, expectations from the physical world, such as people on opposite sides of a table seeing different sides of a physical artifact, do not apply in this environment. In this dissertation, I provide empirical evidence that using standard projection techniques will introduce errors in perception when viewed by multiple people (section 3.2). I also introduce several techniques that can be used to mitigate this discrepancy between multiple points of view (section 3.9).

Issue 2: **Interacting with 3D virtual objects on a horizontal multitouch surface**. Many of the existing techniques for manipulating 3D virtual objects were designed with the same assumption: that only one person would be manipulating these virtual objects. Those interactions that provide full control of all six degrees of freedom (DOF) (movement in x, y, and z; rotation about x, y, and z) typically involve the use of a physical device that is held separate from the display (e. g., a joystick or controller). These devices make it difficult for multiple people to remain aware of each other's actions. For example, information about the small movements required to move a small joystick might be easily missed by a collaborator, whereas the movement required to rotate a physical object would naturally be communicated as a consequence of the action itself. My approach to manipulating 3D virtual objects is to maintain a visual connection between a person and the object they are manipulating by using a direct touch interface — when the display space and touch input space are superimposed. Since display surfaces are typically 2D, a fundamental challenge is how to map this 2D input to 3D manipulation, which I address by using multiple simultaneous touches. In this dissertation, I describe the design, implementation, and evaluation of three techniques for manipulating 3D virtual objects, as well as several design refinements that resulted from the evaluation of these techniques (chapter 5).

Issue 3: **Providing a framework to design 3D tabletop display applications**. The use of direct touch together with 3D visuals results in an environment that has a noticeable appeal; participants in our studies stated that it felt like they could "pick objects up". However, while the ability to move and rotate a virtual object in a collaborative 3D environment has some clear connections to practical applications, such as supporting the examination of 3D medical scans, the collaborative analysis of 3D visualizations, and so on, it is not immediately obvious how to use this rich environment to support collaborative work in general. The computer applications with which we are most familiar are typically 2D and are controlled using familiar input devices, such as the mouse and keyboard. Many interactive tabletop researchers have proposed the use of menus and gestures to enable complex or abstract interactions, such as "printing" or "saving to disk". In this dissertation, I introduce a framework of force-based interaction, and describe how the combination of rich visuals with rich interaction can allow virtual objects to take on a wide spectrum of meaning. Through these different types of meanings, we can enable virtual objects to be repurposed as tools and can support more complex or abstract interactions, without the need for mice, keyboards, menus, or gestures (chapter 6). I then demonstrate how to apply this framework through the co-operative design of a virtual sandtray application to support a form of art therapy for children (chapter 7).

1.4 Method

The method used to address these three issues was an iterative design process familiar in the domain of HCI [Nielsen, 1993; Buxton and Sniderman, 1980]. The interplay between the perception of and interaction with 3D virtual objects in combination with the complexities of multiple people both viewing and interacting with these virtual objects on a digital table necessitated a complex folding of the design, implementation, and evaluation of these many aspects. Figure 1.2 shows these aspects both in the approximate temporal order of the research (from left to right) and the foundational relationship and order they are presented in this thesis (bottom to top). For all three issues, the research includes a complete cycle of design, implementation, and evaluation. However, this iteration varies in all three cases. For the first issue, the perceptual study forms the basis for the design and implementation of several projection alternatives. For the second issue, the design and implementation of three techniques forms the basis for the comparative evaluation. For the third issue, the results of the first two iterations, in combination with previous work, was used to develop a framework, which was then used to design and implement a real-world application. The co-operative

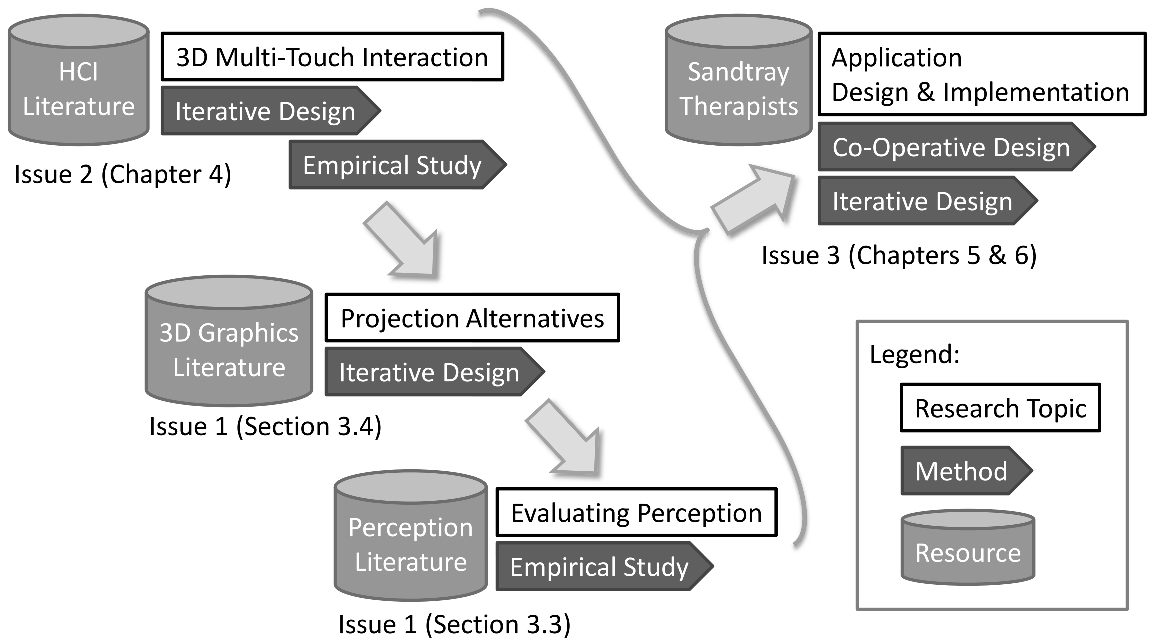


Figure 1.2: This diagram describes the method used to conduct the research in this thesis. The left-to-right progression and grey arrows are approximately correlated with the temporal order of each part. The bottom-to-top progression of the left three research topics indicates the relationship that this work feeds back into the previous research and together form a more solid foundation. These all fed into the final framework and application design.

design process of this application involved several iterations of design, implementation, and assessment by the sandtray therapists.

1.5 Contributions

The main contributions of this dissertation provide support for 3D tabletop display interaction. I contribute a better understanding of both the visuals and interaction in this medium, and show that this combination can ultimately be used in a practical application. I make the following specific contributions:

- A study providing evidence that, when projecting 3D onto a horizontal table using standard 3D graphics techniques, there is an established viewing location, and perception errors will increase as the viewer moves away from this location.

- This same study provides evidence that a parallel projection with a centre of projection (cop) directly above the table may reduce these perceptual errors.
- This same study also provides evidence that providing direct-touch interaction with the virtual artifacts being perceived will also reduce these perceptual errors.
- The design and implementation of a set of non-standard 3D projections that can be used to mitigate the problem of multiple viewpoints and viewpoint discrepancies.
- A study providing evidence that using more fingers to control the six degrees of freedom of the output (dofo) of manipulation improves performance and is preferred.
- The design, implementation, and evaluation of a set of interaction techniques that use multitouch to manipulate 3D virtual objects.
- A description of how to combine 3D interaction with 3D visuals to control 3D virtual tools, thus providing the ability to do more complex actions in a virtual world. This description provides a framework for how to create 3D tabletop applications.
- A case study of sandtray therapy to demonstrate the use of the framework described in chapter 6 to create an actual application.

1.6 Thesis Overview

In chapter 1, I first motivate the need for 3D tabletop display interaction, then provide the scope of the dissertation. I introduce the issues involved in designing a 3D tabletop display environment, my method for addressing these issues, and a list of the contributions that this research provides to the field of hci, to which I will refer throughout the document.

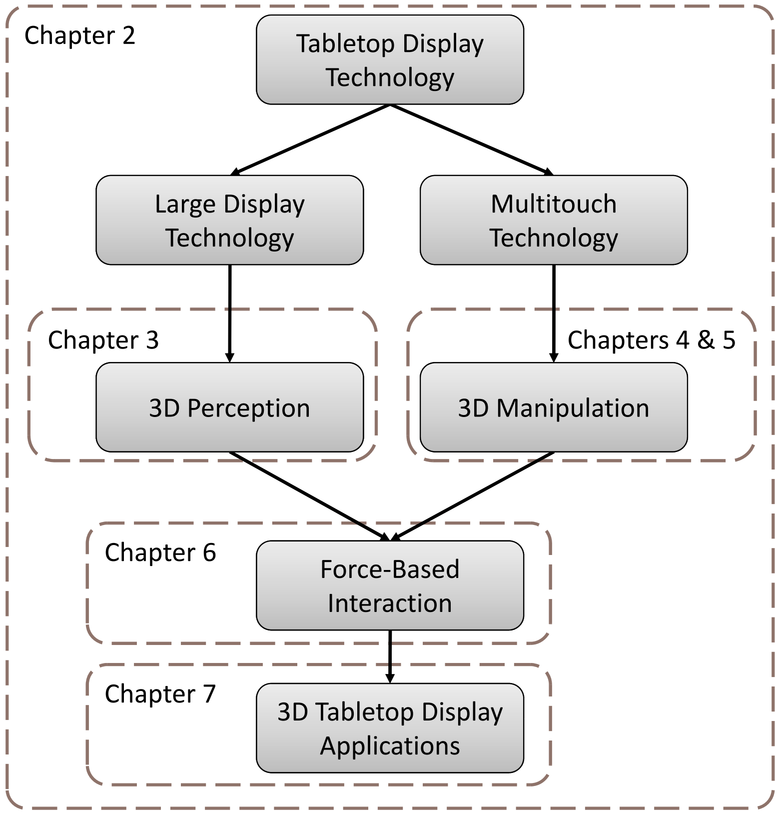


Figure 1.3: This diagram depicts the outline of this dissertation. Chapter 1 introduces the thesis, chapter 2 describes the related research in all of these areas, chapter 3 presents contributions in 3D perception on tabletop displays, chapters 4 and 5 present new interaction techniques for 3D manipulation on tabletop displays, chapter 6 describes a framework which unifies these contributions, chapter 7 presents the design of a real-world tabletop display application that makes use of this framework.

In chapter 2, I present the work most related to this dissertation using the same structure as the thesis as a whole. That is, I first briefly describe the state of the art in tabletop display technology, including both large display technology and multitouch technology. I then describe the research which examines the perception of 3D on these large displays, with particular attention to the issues which occur when they are used in a collaborative environment. This perception research largely informs the work described in chapter 3. I then discuss a parallel track of research which examines how people manipulate 3D artifacts, with particular attention to taxonomies of input devices and some specific techniques that make use of multitouch technology. This 3D manipulation research largely informs the work described in chapters 4 and 5. I then describe the relevant theoretical background research specific to tabletop display environments, as well as several notable designs of tabletop display applications. I also synthesize this research based on how it has made use of force-based metaphors, which helps to inform the framework described in chapter 6.

The remainder of the thesis builds upon this related research. In chapter 3, I present a series of experiments which empirically validate the presence of perceptual error when people view 3D information collaboratively on an interactive table. The results of this study provide clear guidelines about how a designer can avoid these perceptual errors when creating 3D tabletop display applications. I also provide a variety of techniques for mitigating this perceptual error when the constraints provided to the designer do not allow them to follow this advice. I demonstrate the application of the results of this chapter by following this same advice in the remainder of the dissertation.

In chapter 4, I turn my attention to the design of 3D manipulation techniques suitable for multitouch input devices on a tabletop display. Specifically, this chapter presents a formalism which demonstrates the thought process I used when designing the one-, two-, and three-touch 3D virtual object manipulation techniques described in chapter 5. This formalism provides a mechanism for describing how I mapped the input from the multitouch input device to the movement and rotation of the virtual objects being displayed on the large display technology.

In chapter 5, I describe the iterative design, implementation, and evaluation of a set of multitouch interaction techniques for manipulating 3D virtual objects on a tabletop display. The design and implementation makes use of the formalism described in chapter 4 and the evaluation provides empirical evidence that the use of up to three fingers for manipulating 3D virtual objects can improve both performance and preference.

In chapter 6, I describe a framework for force-based interaction which builds upon the synthesis of the literature provided in chapter 2 and provides a theoretical model for how to design 3D tabletop display applications. Specifically, this chapter introduces the notion of sticky tools and demonstrates how they can take on a variety of meaning, thus allowing 3D tabletop display applications to provide full functionality, without the need for additional gestures or menus.

In chapter 7, I present a case study that uses this framework to create a real-world 3D tabletop display application. This case study involved the participatory design of a virtual sandtray to be used in art therapy. I then conclude the dissertation in chapter 8 by describing how the contributions of this thesis apply to the larger field of HCI and by describing some extensions of this work as well as directions for future research.

2 Related Work

In this chapter I provide the context from the related literature to establish the necessary vocabulary and to position my research in the considerable amount of concurrent related work. All of the areas most related to my research—tabletop hardware, perception of 3D virtual objects, interaction with 3D virtual objects, and tabletop interaction theory and applications—have been rapidly changing throughout the years of my thesis research.

The availability of multitouch tabletop display hardware has recently expanded, providing the hardware basis for a tabletop research community within the field of HCI. The recent emergence of this hardware can be attributed to developments in two forms of technology: the display hardware and the input device hardware. The combination of these technologies is what makes tabletop display environments possible, and is particularly compelling because it provides the ability for people to interact with their hands and fingers in the same space that the information is being displayed. Furthermore, this hardware can be large enough that many people can gather around the same device and interact simultaneously. In parallel with the two forms of technology, two categories of research are particularly relevant to this community: the research exploring how we perceive information and the research exploring how we manipulate artifacts. In the same way that the combination of the display technology with the input device technology can lead to a new interactive environment, the advantages that can be gained from the direct visual link between a person's fingers and the virtual object they are manipulating are likely to result from understanding both how people perceive the object on the 2D display and how interaction can be enabled using the multitouch technology.

Because the display space and the input space are the same, there is a tendency for people to perceive that there is a direct connection between their fingers and the virtual artifacts on the display, much like in the physical world. However, a significant distinction between these environments and the physical world is that the table's surface is only 2D, whereas the physical world is 3D. In this dissertation, I explore tabletop display environments where the illusion of 3D information is provided on these displays and investigate how interactions can progress towards helping people move and rotate the virtual artifacts in the same dimensions that they can with physical artifacts. To provide the background for this research, I categorize the literature relevant to 3D tabletop display interaction into four main areas:

Tabletop Display Technology: A review is provided in section 2.1 of the existing technology that is capable of supporting the 3D interaction. This review is included to provide the background for the technology that is used and discussed throughout this dissertation. I separate my discussion of this technology into technology that can display information and technology that can sense multitouch input.

Perception of 3D: A review is provided in section 2.2 of the depth cues that make 3D perception possible and how these can be recreated in a 2D display. The perception literature is then reviewed on the potential perceptual errors that arise due to the viewing of 3D information on 2D displays. The HCI literature which investigates possible solutions to these perceptual errors is also discussed.

3D Manipulation: A review is provided in section 2.3 of the related work on manipulating digital artifacts, including a high-level discussion of a taxonomy of input devices and a low-level discussion of specific techniques for rotating and translating 2D and 3D virtual objects. The HCI literature that includes specific techniques for manipulating 3D virtual artifacts on a digital table is also reviewed.

Tabletop Display Theory & Design: A review is provided in section 2.4 of the existing HCI literature on the theory and design of tabletop display applications. The existing research is also synthesized by how many designs both implicitly and explicitly use force-based metaphors. Several specific applications for storytelling and therapy on tabletop displays, which relate directly to the design in chapter 7, are also described.

2.1 Tabletop Display Technology

Since the introduction of the DigitalDesk by Wellner [1993], there have been several significant advances that have led to the large variety of both large displays and multitouch technology that is available today. The combination of this display (output) technology with multitouch (input) technology provides the hardware that makes the work in this thesis on 3D tabletop display interaction possible. I include this brief discussion of hardware to set the context of my 3D interaction research within the varying advantages, challenges, and constraints of currently available multitouch tabletop technology.

2.1.1 Large Displays

This dissertation focuses on interacting with 3D information on a multitouch tabletop display. There are a variety of factors that determine the quality of the interactive image that is being displayed on these tables, including the size of the image, the display resolution (i. e., the number of pixels in both the horizontal and vertical dimensions[1]), the frame rate, the scan method (interlaced or progressive scan), and many more. This section provides only a high-level overview of the two factors that present the biggest challenges for the research described in this thesis: display size and display resolution.

Display technology has evolved significantly since the invention of computers. Desktop computers, at one point, would typically make use of a cathode ray tube (CRT) monitor

[1] This is actually a misnomer, as resolution should instead refer to the density of pixels (e. g., pixels per inch). I will adopt the common usage of the term (as described here) throughout this thesis.

to display information. More recently, liquid crystal display (LCD) technology has become more common for this purpose. However, practical hardware limitations make the creation of CRTs larger than 1 m in diagonal difficult. LCD technology does not have this limitation, and displays as large as 1.65 m are currently commercially available. Plasma displays can also be made quite thin (typically less than 10 cm) and are currently commercially available in much larger sizes (as large as 3.8 m). Projector technology can also be used to magnify a digital image onto a distant viewable screen. A variety of technologies can be used to create the initial image, including CRT, LCD, and digital light processing (DLP). The large image can also be projected from in front (i. e., the projector is separate from the surface) or from the rear (i. e., the projector can be contained in the display unit). For projector technology, the size of the display is determined by the lens(es) used to magnify the image, and so is not theoretically limited. However, the larger the size, the further away the projector must be, and so practically the size of the room will still limit the size of the display.

Besides the physical size, each of these hardware technologies will be limited in its display resolution. Some technologies, such as LCDs and plasma displays, have a fixed-pixel array, and so have a native resolution. Other technologies, such as the CRT, can provide a large variety of resolutions. Currently, commercially available displays are typically limited to 1920p (1920 × 1080 pixels) resolution. However, this limitation is largely due to standards and available media, and not inherent in the technology itself. Nonetheless, this commercial limitation makes it difficult to provide high pixel density (e. g., a 1.65 m 1920p display would have 34 pixels per inch). To achieve higher pixel densities, displays using any of the above-mentioned technologies can be tiled. However, this approach introduces complications, including the need to drive many displays at once with the same or multiple computers and to manage what happens at the bezels or seams of each tile [e. g., Bi et al., 2010].

One of the distinguishing factors of a tabletop display is that it is setup horizontally, like the Digital Desk [Wellner, 1993], or with a slight tilt, like a drafting table [Fitzmaurice et al., 1995; Buxton et al., 2000]. This setup means that the people using the table are close to the display and that the density of pixels must be higher than if the display were at a distance (e. g., a large wall display for presentations). So, the choice of a larger table size, often comes at the cost of a lower density of pixels. However, the choice of a small display has the implication that less people can gather around the digital work area.

In this dissertation, the research questions have required implementation of 3D visuals in a variety of hardware configurations. Because of the versatility of projector technology (size variability and choice of projection surface), all of the studies and demonstrations have been developed using either a projector or a set of tiled projectors.

2.1.2 Multitouch

Another defining factor of the tabletop displays described in this thesis is the ability for people to interact with them by directly touching the display. A variety of hardware can make direct touch a possibility on these devices. In this section, I provide an overview of the various possible technologies for achieving multitouch on a tabletop display. Note, however, that because this is a particularly active area for research, new technologies are constantly emerging and it is difficult to keep up to date.

2.1.2.1 Capacitance-based

One method for detecting touches is through capacitive coupling. The DiamondTouch [Dietz and Leigh, 2001] makes use of this technology, which has several advantages. By sitting or standing on a pad and touching the device, a person becomes capacitively coupled with an array of sensors that can be used to identify the horizontal (x) and vertical (y) position of the touch, as well as the area being touched. Although there are some ambiguities which increase with the number of touch points, the technology is capable of sensing multiple touches simultaneously. Furthermore, by sitting or standing on different pads, multiple people's touches can be identified as separate input. Because this technology uses an opaque surface, front-projection is necessary to create a superimposed display and input space.

Many commercial products also make use of capacitive touch technology, such as the iPhone [Apple Inc., 2007a], iPod Touch [Apple Inc., 2007b], iPad [Apple Inc., 2010], Zune [Microsoft Corp., 2009], Nexus One [Google, 2010], and more. These technologies benefit from a thin form-factor, because they can be used with LCD technology. However, these commercial products have thus far been limited in display size (the iPad is the largest of these and advertises a 25 cm diagonal display).

2.1.2.2 Camera-based

Another multitouch input device technology is the use of infrared (IR) light to create an image of what is touching the surface. Frustrated total internal reflection (FTIR) [Han, 2005] transmits IR light so that it reflects internally within the surface, and when interrupted by touching (and applying a small amount of pressure), the light gets transmitted perpendicular to the surface. An IR camera can then be placed so that it receives this light, and image processing can be used to translate these images into multitouch input. SMART Technologies [2008] and Perceptive Pixel [2010] create commercial multitouch tables that make use of FTIR technology.

IR light can similarly be used by simply directing the light so that its reflection will be aimed toward an IR camera that captures the reflected image in a technology called diffuse illumination (DI). The Microsoft Surface [2008] is an example of a commercial multitouch table that uses DI. The light source for DI can vary from IR light emitting diodes (LEDs) to laser light technology.

Camera-based techniques have the advantage that the image being captured by the IR cameras can be processed in a variety of ways. Blob detection can be used to identify contiguous forms in the image and blob tracking can be used to track those forms as they move across the screen [Lindeberg, 1994]. This method has the advantage that the size and shape of the forms can be used to develop interaction techniques that make use of not only touch points, but hand and arm shapes as well as physical objects.

2.1.2.3 3D Motion Tracking

Another technique for identifying the position of touch points on a surface is to track the 3D motion of a person's hand and fingers in the space on and above the table. The virtual reality community has explored a variety of technologies for tracking an object in 3D [Foxlin, 2002; Meyer et al., 1992; Richards, 1999; Welch and Foxlin, 2002; Bowman et al., 2005], and many of these can be applied to the tracking of hands and fingers.

3D tracking technology can be magnetic [e. g., Polhemus, 1994], optical [e. g., Vicon, 2006], acoustic, inertial, or mechanical [Bowman et al., 2005, chapter 4]. Typically, these systems require the use of markers or physical devices to be mounted to the hands and fingers in order to track their 3D positions in real-time. These 3D positions can be used to detect touches on a tabletop display by having information about the 3D geometry of the table itself, and therefore use the 3D position of the hands and fingers to detect when they have touched the table.

2.1.2.4 Other Technologies

There are a variety of other technologies that are capable of providing multitouch input, such as acoustic wave, strain gauge, optical imaging, dispersive signal technology, acoustic pulse recognition [Harrison et al., 2010], and interpolating force-sensing resistors [Rosenberg and Perlin, 2009]. While some of these technologies seem promising for use in multitouch tabletop display environments, most of these are very new (2009 & 2010) and the resulting construction of tables that use them and accompanying interactions are still to come, and so they will not be discussed in more detail here.

2.1.3 Summary of Tabletop Display Technology

When I started this research in 2004, the best available hardware was the multitouch DiamondTouch [Dietz and Leigh, 2001], which offered four identifiable touches, and the custom-built large tabletop display in our lab (a 146 cm × 110 cm table that used four tiled projectors for a total of 2800 × 2100 pixels), which was limited to two touches.

Currently, a great variety of multitouch surfaces are available but most are either handheld (e. g., iPod Touch [Apple Inc., 2007b]), intended for one person (e. g., Tablet PCs) or medium sized and limited in resolution (e. g., SMART Table 2008, which is 57 cm × 43 cm and uses a single 1024 × 768 pixel projector). While there has been research indicating that ideally for teams or small groups, tabletops would be large enough to accommodate 4 to 5 people comfortably (ĩ27 cm diagonal); have sufficiently high resolution for readability (in the order of 40 pixels / cm); and have identifiable multitouch (at least 10 touches per person), this is not yet available. However, interest and research productivity are both currently high and therefore it is reasonable to think that this will be achieved shortly. Because of the rapid changes in hardware over the last four years, I have conducted my research on a variety of displays, choosing (and sometimes augmenting) the best available at the time for the type of questions I was considering. For example, my perceptual empirical research was conducted on the display with the highest resolution, my interaction techniques were developed on a modified DiamondTouch device to provide multiple touches per person, and my final application was built for a fully multitouch display, albeit with a low resolution (1024 × 768).

While there has been a lot of work on engineering the physical devices that comprise a tabletop display, in this dissertation I focus on using these devices to interact with 3D virtual artifacts. Just as the physical hardware can be separated into the display output and the multitouch input, 3D interaction on these devices can be separated into the *perception of 3D* and the *manipulation of 3D artifacts* on these devices.

2.2 Perception of 3D

A significant aspect of 3D interaction is the visual feedback that is perceived by the person or people using the technology. There has been longstanding research into human 3D perception and, more recently, there has also been a wide range of research that applies this understanding to the perception of virtual artifacts. I start this discussion on perception by providing pointers to the significant literature about how people perceive *3D depth cues* (section 2.2.1). Artists have for centuries created 2D images that make use of these perceptual phenomena. The work of these artists, as well as much of the research in the field of computer graphics has informed how people *perceive 3D virtual artifacts in a 2D plane* (section 2.2.2). In the creation of these 2D images, on either a display or canvas, the typical assumption is that it will be viewed from a point of view in front of the display or canvas, and with a view angle that is not large. However, these assumptions will often not hold for the visuals presented on a tabletop display; there will typically be many viewers who are very close to the display (i. e., have a large view angle) and are looking from a variety of locations (i. e., anywhere other than in front of the display). There has been some work in the psychology literature which explores the *effect of off-axis viewing* of 3D virtual artifacts (section 2.2.3). There have also been a few *technologies for viewing 3D at a tabletop display* which take into account some of these perceptual phenomena (section 2.2.4).

2.2.1 3D Depth Cues

A variety of perception literature has surveyed the depth cues [Collett and Harkness, 1982; Zeil, 2000; Ware, 2004] that we use when we perceive 3D. Ware [2004] describes the depth cues listed in table 2.1. This list is separated into *dynamic* versus *static* cues, depending on whether or not the cue requires a moving picture, and *monocular* versus *binocular* cues, depending on whether the cue requires one or two eyes to perceive.

	Static	Dynamic
Monocular	Linear perspective Texture gradient Size gradient Occlusion Shape-from-shading Cast shadows Depth of focus	Kinetic depth Motion parallax
Binocular	Stereoscopic depth Eye convergence	

Table 2.1: A summary of the cues that provide depth perception in 3D.

Ware [2004] provides an in-depth description of each of these depth cues. Because the interest in these cues, for my research, is concerned primarily with how they manifest in 2D images, I limit my discussion of the extensive literature to provide brief descriptions in section 2.2.2 for each depth cue according to how it can be recreated in a 2D image.

2.2.2 Perception of 3D Virtual Artifacts in a 2D Plane

With knowledge of the depth cues that people perceive, it is possible to create the illusion of a 3D image within a 2D plane. This illusion is accomplished by fooling the viewer's eye into perceiving depth cues (section 2.2.1).

2.2.2.1 Recreating Static Monocular Cues

To introduce *linear perspective* as a depth cue in a 3D scene in a picture, straight lines (rays) that go from every point in the 3D scene to the centre of projection (COP) are created, and the intersection of these lines with the plane of the picture are drawn (figure 2.1, left). The picture is geometrically correct for the viewer if the centre of projection (COP) coincides with the point of view (POV)—the location from where the picture is observed. This method, usually called perspective (or pinhole) projection, has been used by artists for centuries [Jones and Hagen, 1978]. Currently, the same perspective projection fundamentals underlie most of 3D computer graphics and virtual reality. Other depth cues can be achieved using the same

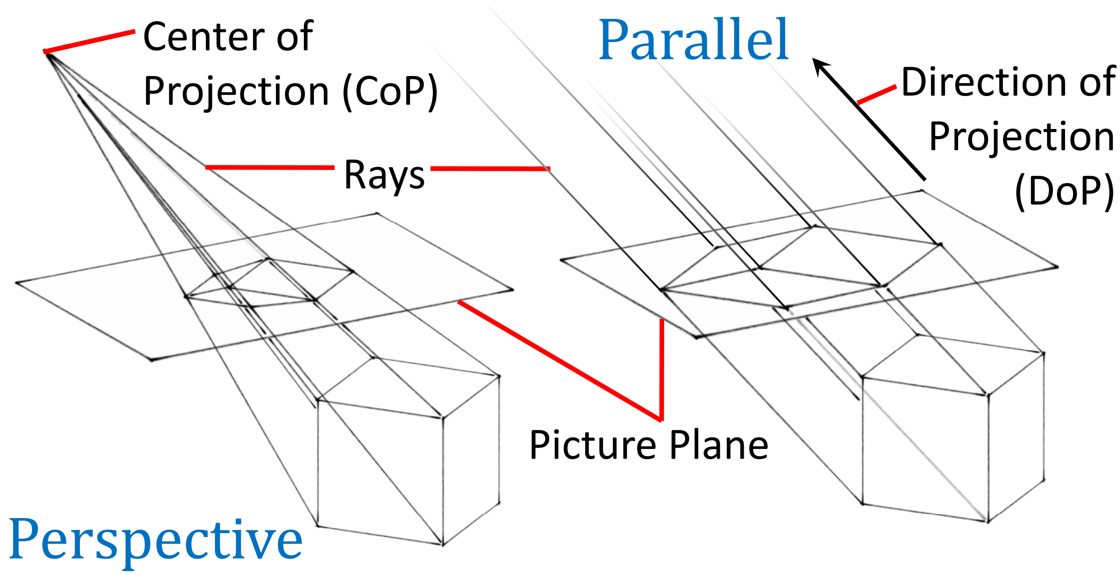


Figure 2.1: Projection geometries.

mathematical principles: *texture gradients* can be created by taking a 2D texture and rotating along an axis within the plane of the picture, and *size gradients* are achieved because 3D elements represented in a picture through perspective projection appear smaller if they are further away from the COP. A feature of linear perspective pictures is that most lines that are parallel in the 3D scene converge to vanishing points in the picture.

An alternative type of perspective geometry uses rays that are parallel to each other and project in a fixed direction instead of converging to a COP (figure 2.1, right). Parallel projections result in pictures where parallel lines in the 3D scene are preserved as parallel in the picture. Parallel projections do not have a COP because the projection lines do not converge, but instead the COP determines a direction of projection. Parallel projections are often used in architecture and engineering because they preserve parallel lines and because they make direct measurement easier; however, parallel projections cannot generate impressions in the viewer's retina that are equivalent to what the real scene projects on the viewer's eyes. That is, they will not be able to reproduce *linear perspective*, *texture gradients*, or *size gradients*.

In both perspective and parallel projections, *occlusion* is achieved by only drawing the first point that each ray from the cop hits. *Shape-from-shading* can also be approximated in both parallel and perspective geometries using, for example, Phong [1975] or Gouraud [1971] shading.

Cast shadows can also be created using computer graphics, and have also been used for centuries by artists. *Depth of focus* can also be created by blurring parts of the image, however, computer graphics would require knowledge of the viewer's eye gaze in order to adjust the image as their gaze changes.

2.2.2.2 Recreating Dynamic Cues

Dynamic depth cues require the use of a moving picture, and so need the use of technology such as a computer or movie camera to reproduce the animated image in a 2D display. Perhaps the easiest of the two dynamic cues to recreate is *kinetic depth*. This depth cue is provided whenever a viewer (themselves not moving) perceives an object in motion (e. g., a spinning object). This cue can be easily created using real-time computer graphics by rotating the model of the object being drawn.

In the physical world, as an observer moves left or right, near objects will move at a different rate than far objects, which provides another depth cue called *motion parallax*. Motion parallax is a depth cue that arises from the continuous change of the point of view (pov) with respect to the perceived objects, where many slightly different retinal images are composed to reconstruct the 3D structure of a scene. This depth cue will be missing from any static picture of a 3D scene. This cue is easily recreated in movies by simply moving the video camera. The resulting effect for the observer will be similar to the experience of driving in a moving vehicle, where the pov is changing without physical effort on the part of the observer. Fishtank virtual reality [Ware et al., 1993], where the observer's position is constantly monitored using motion tracking technology, also relies heavily upon motion parallax. In such a system, the image can be dynamically adjusted so that the cop coincides with the pov of the observer at each frame.

2.2.2.3 Recreating Binocular Cues

The first of the two binocular cues, *stereoscopic depth*, is perceived in the physical world because slightly different images are provided to each eye. The brain then fuses these two images to establish the 3D depth of various parts within the image. This cue is also heavily relied upon in fishtank virtual reality [Ware et al., 1993] and can be recreated by providing different images to each eye. In computer graphics, this can be achieved by generating two images with slightly different COPs, using an assumed distance between the viewer's eyes. These two images can be presented separately to each eye by using special glasses that, for example, display separate images in each lens or filter the light differently to each eye.

The second depth cue, *eye convergence*, is a result of a person's two eyes rotating differently to fixate at a particular depth. However, when a person is fixating on an image plane, the rotation of the eyes does not usually match the depth expected by the illusion. Technologies such as holography and volumetric displays [e. g., Ebert et al., 1999; Lucente, 1997] can recreate this depth cue, but these displays do not project onto a 2D image plane. Paradoxically, this cue can also be achieved if the viewer covers one eye and alters their gaze, as the covered eye will sometimes appropriately adjust its rotation to converge as if viewing a 3D scene as the uncovered eye examines the 2D scene.

2.2.3 Effect of Discrepancy between POV and COP

Although pictures represented on a 2D plane are able to create an impression of depth [Cutting, 1988, 1987; Sedgwick, 1993; Ellis et al., 1991; Saunders and Backus, 2006; Ware, 2004], flat projections of 3D space only create retinal images equivalent to the 3D scene if the POV of the observer is located at (or very close to) the COP used to generate the image [Sedgwick, 1993]. When POV and COP are at very different angles to the picture plane, or are at very different distances (e. g., when we look at photographs on a table, a painting from a lateral point of view, or a movie from the side aisle), the differences between correct and distorted retinal projections can be very large. If pictorial perception is dependent only on the geometry of the projected retinal image, this should result in the perception of a space that is deformed compared to the depicted space [Goldstein, 1987; Sedgwick, 1993; de la Gournerie, 1859].

Regardless of the distortion, observers are remarkably good at still perceiving a relatively accurate pictorial space [Vishwanath et al., 2005]. However, there is still controversy in the perception research community about the underlying processes that support correct space perception from geometrically incorrect retinal images (what is called space constancy). Some suggest that the visual system corrects distortions based on geometrical information from the represented scene (e. g., assuming certain angles are straight [Perkins, 1973], objects are rigid [Cutting, 1987], or certain converging lines on the picture are actually parallel in the real scene [Saunders and Backus, 2006]), and others propose that information about the correct COP can be recovered from perceptual information about the surface where the picture is projected (e. g., from accommodation and other 3D cues [Cutting, 1997], or from the shape of the frame of the picture [Koenderink et al., 2004]).

Although the perception of pictorial space is relatively stable regardless of the discrepancy between the locations of the COP and the POV, the relationship between the pictorial space and the physical space is not equally stable. In particular, the perceived orientations towards the physical space of elements within the picture plane can vary depending on the position of the observer [Goldstein, 1991; Cutting, 1988; Sedgwick, 1993]. This effect is best exemplified by the famous U.S. recruiting poster of Uncle Sam, in which he points directly at the observer regardless of how far she is or how oblique she stands to the plane of the poster. For elements within the picture that point perpendicular to the picture plane (e. g., Uncle Sam's finger), the perceived orientation always follows the observer regardless of its position, and therefore the perceived orientation of the object with respect to the plane of the image can vary almost 180°. For objects that do not point perpendicularly to the picture plane the possible variation in the perceived angle is reduced; at the extreme (objects that are aligned with the picture plane) geometrical accounts of orientation perception [de la Gournerie, 1859; Cutting, 1988] predict that the pointing direction will not vary with changes in the POV. This effect is referred to in the literature as the differential rotation effect (DRE) [Goldstein, 1987] or the *la Gournerie* effect [Cutting, 1988]. The DRE is also subject to scientific controversy; experiments have shown that the geometrical predictions do not necessarily fit all data, especially for very oblique POVs [Ellis et al., 1991]. The possible causes might be found among the cues that cause the pictorial space constancy discussed above (e. g., frame and perceptions of the picture surface through binocular cues [Vishwanath et al., 2005]).

The perception and manipulation of shapes in oblique displays has received some attention in the HCI literature. For example, Wigdor et al. [2007] studied how the slant of the surface affects the perception of several magnitudes (length, angle, area) for 2D data and Nacenta et al. [2007] studied how the correction of perspective distortion in oblique displays affects basic motor and cognitive processes. Finally, Grossman and Wigdor [2007] surveyed different 3D technologies for horizontal surfaces, including the 3D cues that they provide.

2.2.4 Multiple Viewers at a Tabletop Display

The problem of a discrepancy between COP and POV is particularly pertinent to 3D tabletop display interaction, because these physical tables afford many people gathering around and using them simultaneously. Since each person will be close to the table, if they are at different sides, their POVs will be different, and so there will likely be a discrepancy between the COP and POV for all but one person. For example, if the technology is displaying an image or virtual object that is pointing out of the display (e. g., the U.S. recruitment poster), the people around the table may all be perceiving that it is pointing at them. Thus, they would each have a different understanding of the represented scene, and so their shared knowledge may be compromised. This inconsistency in the shared knowledge may interfere with their collaboration.

There have been some attempts to correct for this discrepancy using technology. Three categories of solutions are relevant to 3D tabletop display interaction: *alternate perspective rendering*, *fish tank virtual reality*, and specific solutions for *3D tabletop display interfaces*.

2.2.4.1 Alternate Perspective Rendering

One technique for mitigating the perceptual error is to use a non-linear perspective rendering of the virtual objects on the 3D tabletop display. With these alternate renderings, parts of the image can appear "correct" at one side of the table and other parts can appear "correct" at another. Agrawala et al. [2000] present a way of providing multiple camera viewpoints for each object in a scene. Ryan [Coleman and Singh, 2004] allows a static image to be created by stitching together multiple viewpoints so even single objects can be distorted and viewed at multiple angles. Brosz et al. [2007] provide a unified flexible projection geometry that also supports complex and extended viewpoints that can look at all sides of an object or scene. Zorin and Barr [1995] describe many of the limitations of the use of standard perspective projection in the creation of static 3D images. Specifically, they argue that a picture of a 3D scene cannot simultaneously satisfy both properties (1) that straight lines should appear straight and (2) that objects should appear as if viewed directly. They present a formalism to manually balance these two properties. In the case of 3D tabletop display interaction, this problem is exacerbated due to (a) the desire to have a 3D image appear correct for *multiple* people viewing the image and (b) to have the "images" being viewed be interactive, and thus dynamically changing perspective.

2.2.4.2 Fish Tank Virtual Reality

Previous research also explores the correction of a 3D projection based on viewing angle [Deering, 1992; McKenna, 1992] and has been dubbed "Fish Tank Virtual Reality" [Ware et al., 1993]. These systems typically use one of two approaches: they either use stereoscopic (headtracked) goggles that project an image onto two surfaces that move with a single viewer's head motion, or they track the viewer's head position and correct the view for the measured eyepoint. These systems have also considered many variables not directly related to my research interest of the shared perception of 3D objects on a 2D display (including refraction and the effect of curvature of CRTS), and so is beyond the scope of this dissertation.

2.2.4.3 3D Tabletop Display Interfaces

The two-user responsive workbench [Agrawala et al., 1997] addresses the problem of different viewpoints at a table by providing correct stereoscopic 3D images to two different people. The IllusionHole can be used to integrate 2D and 3D [Nakashima et al., 2005] on a table by limiting the portion of the display presented at each viewing angle via a hole in the table's centre. These systems both provide stereoscopic cues via headgear and a tracking system. While these systems provide a means for accurately mitigating the problem of discrepancy, they come with the disadvantage of potentially interfering with the ability for people to make eye contact, requiring a large display while extremely limiting the usable display region, and they may otherwise interfere with the social expectations in a collaborative setting, due to the need to wear additional hardware.

2.2.5 Summary of Perception of 3D

While there has been a significant amount of work that investigates the human ability to perceive 3D, and applications of this theory to create the illusion of 3D on a 2D image plane, this work has previously not been considered in light of the unique properties of 3D tabletop display interaction. This dissertation extends the literature on 3D perception by investigating to what degree errors in perception might affect collaborative work with 3D virtual artifacts on an interactive table. This investigation corroborates the need for the existing mitigating techniques, and is also used to inform the design of new mitigating techniques that contribute to the existing work on technology which addresses this perceptual error. This investigation also informs the design of 3D interaction manipulation techniques and 3D tabletop display applications throughout the dissertation.

2.3 3D Manipulation

There are many actions which are relevant to a person's interactions with a 3D space. Bowman et al. [2005] discuss three particularly relevant examples in the consideration of 3D virtual environments: selection, manipulation, and navigation. However, the scope of this dissertation is limited to the selection and manipulation of 3D virtual artifacts. While navigation can be a useful action in many virtual environments, it is not clear what role it should play in a collaborative tabletop environment, where a person navigating may interfere with another's understanding of the space. 3D manipulation can also be further broken down into a variety of manipulations, but this dissertation considers only the movement and rotation of rigid virtual bodies, and not any change in shape (e. g., a pliable artifact).

There are a variety of relevant sources which discuss 3D manipulation, including significant research into *understanding how and why people manipulate* both physical and virtual artifacts (section 2.3.1). The intention of the HCI literature which explores this understanding is often to map the design space of input devices, which has resulted in a variety of relevant taxonomies. There have also been many *2D and 3D multitouch manipulation techniques* introduced (section 2.3.2) as well as some work conducted in parallel with this dissertation which explores *3D manipulation on tabletop displays* (section 2.3.3).

2.3.1 Understanding Manipulation

The human hand has many joints and muscles which provide a total of 29 DOF [Zhai, 1995]. From a very young age, human children learn to use these DOF to move, rotate, lift, assemble, and otherwise manipulate physical artifacts. Indeed, the word "manipulate" comes from the Latin "manipulus", meaning "handful", and so has its roots in the use of one's hands. Because a physical artifact exists in 3D space, there are three positional DOF (movement in x, y, and z) and three rotational DOF (rotation about x, y, and z) that can be manipulated. So, with little thought, an adult can use the 29 DOF of their hand to manipulate the 6 DOF of a rigid body (given that the object has certain properties, e. g., is light and easy to grab). However, in order to "manipulate" a virtual artifact being displayed on a computer screen, it is necessary to make use of some form of input device. Thus, much of the literature in HCI which aims to understand this manipulation also aims to understand input devices.

Early input device taxonomies categorized a device based on either its elementary virtual device components (e. g., button, locator) [Foley and Wallace, 1974] or the elementary tasks it could perform (e. g., select, move, rotate) [Foley et al., 1984]. However, these taxonomies had the disadvantage that two devices that look and feel very different (e. g., a trackpad and a mouse) could be classified in the same way. Buxton [1983] addressed this problem by providing a taxonomy based on the notions of DOF, what type of input is sensed, and whether the device is touch-sensitive or instead uses an intermediary between the person and what is sensed (e. g., an arm or knob). Mackinlay et al. [1990] further improved upon this taxonomy by separating movement and rotation as well as absolute and relative qualities of the devices. Figure 2.2 shows an example of how this taxonomy can be used to classify an input device.

This progression of taxonomies shows that their roots are in how physical objects can be manipulated. The later formulations separate the manipulation that is occurring in the physical world and being sensed by the device from the manipulations in the digital world (called the "task" in the earlier taxonomies). While this separation provides a useful classification system, later work suggests the additional need to consider the cognitive processes involved in the task at hand. Jacob and Sibert [1992] describe a theoretical model that considers, when a person composes multiple dimensions of action, whether the components are *integrated* or *separable*. For instance, movement in x, y, and z are considered to be integrated, since people frequently tend to move simultaneously in all three. However, it has been ar-

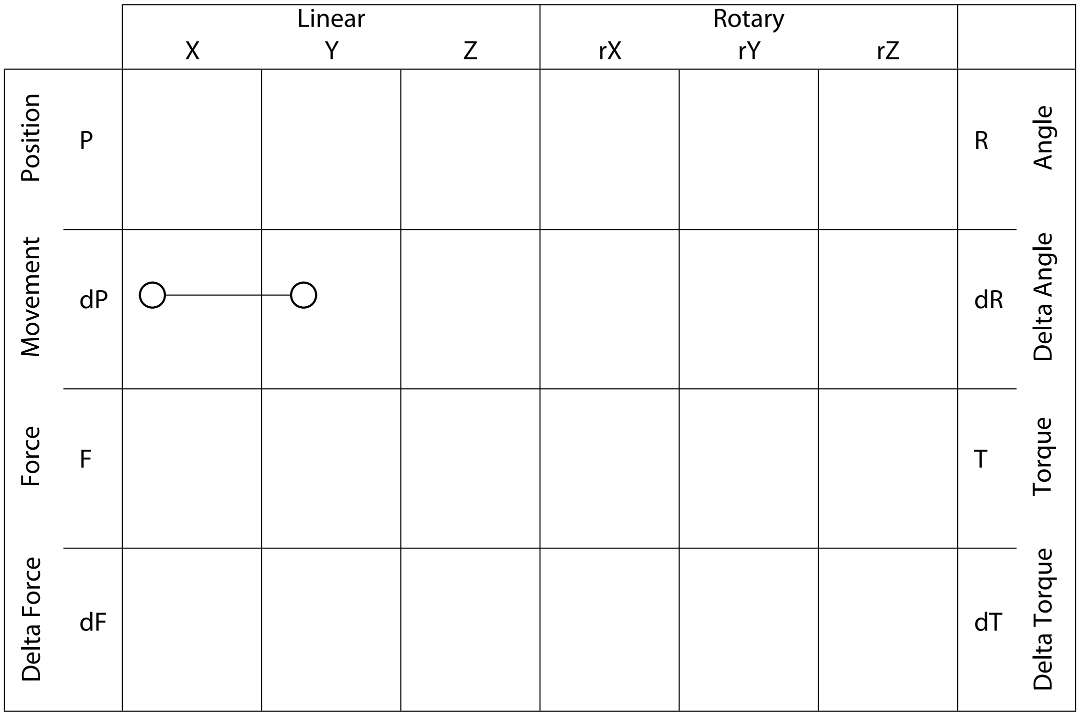


Figure 2.2: Mackinlay et al.'s [1990] input device taxonomy. A device can be classified by stating one or more boxes in which the device fits. E. g., a mouse is a relative input device that provides both X and Y linear DOF.

gued that rotation and translation are separable in the human mind [Masliah and Milgram, 2000]. This cognitive separability has led to the suggestion that these actions should involve separate movements in an input device, to match one's understanding.

Zhai [1995] further classifies input devices into *isotonic*—those which sense movement—and *isometric*—those which sense force or torque—and described how this input could map onto manipulation in 6DOF as the *transfer function*. Transfer functions that map the input device directly to the position of an object are referred to as *position control* or *zero order* transfer functions. Those that map this input to the velocity of an object are called *rate control* or *first order* transfer functions. Zhai [1995] showed empirically that isotonic devices were more compatible with position control, and isometric devices were more compatible with rate control. Another type of device along the isotonic-isometric scale is *elastic* input, where the amount of resistive force increases as the input device's position moves. Fröhlich et al. [2006] observed that isotonic input mapped well to the 3DOF of rotation, and elastic input mapped well to the 3DOF of translation, further justifying the separation of translation and rotation.

However, other researchers suggest that rotation and translation are *not* separable in the human mind [Jacob et al., 1994; Wang et al., 1998]. Studies of 2D interaction techniques for 2D tabletop interfaces, such as rotate n' translate (RNT) [Kruger et al., 2005], which require only 3DOF, tend to confirm that people typically ignore this difference and that integration of rotation and translation is in harmony with the need to support communication in a collaborative setting [Kruger et al., 2004, 2005; Hancock et al., 2006]. Although these two claims appear contradictory, this dissertation demonstrates that rotation and translation can be separated, but performed simultaneously, and that this provides an interface where it is possible to communicate without sacrificing performance. The techniques used in this dissertation achieve this feat by separating translation and rotation across both hands, and thus also contribute to the literature on bimanual interaction [e. g., Kurtenbach et al., 1997; Balakrishnan and Kurtenbach, 1999].

2.3.2 2D & 3D Multitouch Manipulation Techniques

While these input device taxonomies give significant insight into how one can manipulate a 3D virtual artifact *through* an input device, there has also been some work on how a person's hands can be used to directly manipulate a virtual object (both in 2D and in 3D). Westerman [1998] describes in detail how to use multitouch gestures (including a pinch gesture) to manipulate 2D virtual artifacts. The DiamondSpin toolkit [Shen et al., 2004], which was built upon the DiamondTouch multitouch input device [Dietz and Leigh, 2001] also describes methods for moving and rotating an object so that it orients itself toward the closest edge of a tabletop display; however, these techniques make use of only one touch point. Kruger et al. [2005] introduce the RNT technique, which provides another method for simultaneously rotating and translating a 2D virtual object with a single point of contact. Techniques such as TNT [Liu et al., 2006] have extended RNT by mapping the rotation of a hand or stylus to the orientation of the virtual object. Hancock et al. [2006] summarize a variety of techniques for manipulating 2D artifacts using a multitouch tabletop display.

A single point can similarly be used to rotate a virtual 3D object using a technique called ARCBALL [Shoemake, 1992]. This technique maps the change in position of the input device to an arc on a sphere encompassing the virtual 3D object. The object is then rotated about the centre of the sphere in the direction of the arc. While this technique was designed for mouse input, it can easily be used with a single touch on a multitouch device. Sturman [1992] also describes techniques for using hand and finger postures in mid-air to move and rotate 3D artifacts. Wu and Balakrishnan [2003] also used gestures and hand postures to control a room planning application. Other approaches combine posture and gesture interaction with speech input [Tse et al., 2007]. Even with relatively simple size-based recognition hardware [SMART Technologies, 2003] it is possible to use a set of hand postures to parametrize or control actions in an interface [Grubert et al., 2008]. Lepinski et al. [2010] also extend Kurtenbach and Buxton's [1993] marking menus to the use of multitouch input. Input techniques that provide more information on the shape of a touch, however, can be used to define postures inspired by physical interaction and to infer forces to be used in the interaction [Cao et al., 2008]. An approximated touching force can also be used to control the layering of 2D objects [Bers et al., 1998].

This dissertation contributes to this body of work by providing several interaction techniques together with an empirical evaluation which demonstrates that the consideration of multitouch input and the DOF it affords can lead to a combined technique for moving and rotating 3D virtual objects using the fingers on a person's hand. This work was introduced to the HCI community before the completion of this dissertation [Hancock et al., 2007], and since its introduction, there has been a significant trend toward 3D interaction techniques within the community. In the next section, I describe these techniques and how they relate to the work in this thesis.

2.3.3 3D Manipulation Techniques on a Tabletop Display

Prior to this dissertation, a few collaborative tabletop displays that included 3D effects had been proposed. Ståhl et al.'s [2002] tabletop 3D virtual pond floats items in use to the surface and allows items to sink when they are no longer in active use. The Lumisight table [Matsushita et al., 2004] and Nakashima et al.'s [2005] 3D table provide up to four people with a coherent view of a 3D image in the centre of the display. While these systems are capable of rich 3D visuals in a collaborative setting, they do not fully address the interaction with these 3D models. Furthermore, some of these systems require a very large tabletop to achieve a small central 3D display.

While interaction techniques that rely on input captured in 3D space have been investigated in detail in 3D stereoscopic environments such as the Responsive Workbench [Krüger and Fröhlich, 1994], the work in this dissertation concentrates on the input provided through the contact points of one, two, or three fingers on the surface of a horizontal display to interact with a 3D scene. Another metaphor to translate 2D input to interaction with 3D shapes in a horizontal interface, BumpTop, was introduced by Agarawala and Balakrishnan [2006] for mouse- or tablet-based interfaces. This interface is based on a physics simulation with 3D shapes, controlled through gestures as well as some menus. A similar interface that also relies on physical simulation was introduced by Wilson et al. [2008] who extend the interacting objects such as fingers virtually into a 3D environment in which objects are located. In contrast to the shallow-depth techniques introduced in [Hancock et al., 2007], objects in Wilson et al.'s [2008] system move in 6DOF (constrained by the physical simulation), but it is difficult to control them precisely, (e. g., to arrange them in a specific position and orientation), and movement in the z dimension is particularly problematic.

2.3.4 3D Manipulation Summary

The HCI community has spent a significant effort in understanding input devices and how they can be used to manipulate virtual artifacts. This work has led to a better understanding of how to classify devices according to a variety of qualities, including: number of DOF, types of properties sensed, whether it is direct or indirect [Buxton, 1983]; whether it is absolute or relative, translational or rotational [Mackinlay et al., 1990]; and whether it is isotonic, elastic, or isomorphic and what type of transfer function is being used (position or rate control) [Zhai, 1995]. This work distinguishes between the DOF of the input devices and the DOF of the virtual artifact being controlled, and also describes the need for a mapping or transfer function between the input and the output. However, these taxonomies were created prior to the increased availability of multitouch hardware. The work in this dissertation helps to elucidate some of the subtleties unique to this input technology and introduces the concept of magnitude of freedom (MOF) to help classify the different types of multitouch hardware (chapter 4). Chapter 5 then demonstrates how to apply this new model to the creation of novel multitouch interaction techniques for 3D virtual object manipulation.

2.4 Tabletop Display Theory & Design

The work in chapter 3 provides the basis for an understanding of how 3D visuals can be used in a collaborative tabletop setting and chapter 5 provides the basis for how a person can manipulate the 3D virtual artifacts that are being visualized. This work is brought together in chapter 6 to describe a framework for how to design 3D tabletop display interfaces. In this section, I review the current theory in *tabletop display interaction*, and then survey the presence of *force-based metaphors* throughout the design of tabletop display applications, both of which inform the framework in this dissertation. I also describe the specific *storytelling and therapeutic* tabletop display applications which are most closely related to the design presented in chapter 7.

2.4.1 Tabletop Display Interaction

At the turn of the millennium, there began to emerge several readily available multitouch input devices that could easily be coupled with a large display [e. g., Dietz and Leigh, 2001; SMART Technologies, 2003]. This emergence of the technology encouraged many to observe people using tabletop displays for collaboration. Scott et al. [2003] describe eight collaborative design guidelines for co-located collaborative work at a tabletop display: "(1) support interpersonal interaction, (2) support fluid transitions between activities, (3) support transitions between personal and group work, (4) support transitions between tabletop collaboration and external work, (5) support the use of physical objects, (6) provide shared access to physical and digital objects, (7) consider the appropriate arrangements of users, and (8) support simultaneous user actions." Scott et al. [2004] also developed a theory of territoriality based on observations of people using physical tables [see also Scott, 2005]. This theory suggests that people use three kinds of territories: personal, group, and storage territories. They further described the functionality and space requirements for each type of territory to provide guidelines for the design of tabletop applications.

Morris et al. [2004] also examine the use of coordination policies to mediate collaboration at an interactive tabletop display, rather than relying on social protocols to mediate this interaction [see also Morris et al., 2006a]. These policies are largely based on the notion of ownership and resolving conflicts based on who owns the virtual elements displayed on the screen. They separate these policies based on the type of conflict that occurs (global or whole-element conflicts) and what kind of initiative is taken (proactive, mixed-initiative, and reactive). Global and whole-element conflicts are those that arise based on questions of ownership over the entire environment or individual elements, respectively. Proactive and reactive initiatives are decisions about ownership which are decided before or after the conflict arises, respectively, and mixed-initiative are those where some prior ownership or hierarchy exists, but requires some contextual information to inform the decision. Morris et al. [2006b] later describe in depth how collaborative gestures can be used in a co-located tabletop display environment to coordinate conflicts such as these. Morris et al. [2006c] also performed a study comparing the use of centralized versus replicated controls in a tabletop environment and found that the use of many copies of each control may be detrimental to collaborative coordinated activity.

Tang et al. [2006] performed an observational study at a tabletop display to examine how people perform collaboration and discovered at least six varieties of collaborative coupling (from loosely coupled to closely coupled). Isenberg et al. [2008] observed people performing information analysis on a physical table and discovered that people frequently and rapidly switch between a variety of information analysis tasks. Both studies conclude that digital tables must support fluid, frequent, and rapid transitions between different modes of work and styles of collaboration.

There has also been some work that explores what types of interactive gestures are suitable for use in a tabletop display application. Wu and Balakrishnan [2003] describe several multi-finger and whole-hand gestures which can be used to invoke a variety of commands on a multitouch table. Freeman et al. [2009] extend this work by describing a technique to provide visual feedback during the learning phase of multitouch gestures. Wobbrock et al. [2009] also provide a method for generating a set of gestures by sampling a group of people who demonstrate what gestures they think correspond to the required actions. The intention is to create a gesture set that best matches a person's intuition, instead of determining a gesture set based on the intuition of a designer or researcher.

There have also been several toolkits provided for the creation of tabletop display applications. DiamondSpin [Shen et al., 2004] provides a means of using standard GUI widgets with a multitouch table (specifically the DiamondTouch table [Dietz and Leigh, 2001]).

These widgets provide support for multiple touches, orientation, and multiple people. The single-display groupware (SDG) Toolkit [Tse and Greenberg, 2004] also provides support for multitouch-capable hardware [e. g., Dietz and Leigh, 2001; SMART Technologies, 2003], as well as multiple people, and rotatable widgets. More recently, technologies such as the SMART Table [2008] and Microsoft Surface [2008] provide a software development kit (SDK) for programming their technology.

This dissertation makes use of this existing theoretical and practical work on tabletop display interaction. Specifically, the guidelines provided through observations of people using physical tables have helped to inform the design guidelines used to create the interaction techniques in chapter 5 and the sandtray therapy application in chapter 7. Chapter 3 also adds to this body of theory by empirically studying the effects of 3D perception specific to a co-located collaborative tabletop display environment. The results of this empirical work have direct implications on a designer's ability to follow the guidelines and theory laid out by this prior work. This theoretical and practical work has also directly informed the framework described in chapter 6, which integrates some of this theory with the consistent use of force-based interaction in many existing tabletop display designs, described in the next section.

2.4.2 Force-Based Interaction

Digital tables have used force-based interaction since they were introduced, both explicitly through metaphor and implicitly through 2D or 3D manipulation.

2.4.2.1 Force-Based Metaphors

Many tabletop display interfaces use force-based metaphors to create compelling new interactions. The Personal Digital Historian [Shen et al., 2003] uses the idea of a "Lazy Susan" to invoke the metaphor of spinning virtual objects to another side of the table. The Pond [Ståhl et al., 2002] uses the metaphor of a body of water where virtual objects can sink to the bottom over time. Interface currents [Hinrichs et al., 2006] demonstrate how the idea of flow can be applied to virtual objects; virtual objects can be placed in a dedicated area on the table that acts like a river, carrying the virtual objects to another part of the screen.

A more abstract property of force-based interaction is that local actions only cause local behaviour, though this behaviour can then propagate to have a larger area of influence. For example, dropping a stone in water initially affects a small area, and over time its ripples eventually affect the entire body of water. Isenberg et al. [2006] integrated this locality property into a framework for building tabletop display interfaces. With this framework, tabletop interfaces can be created where virtual objects adhere to this property.

2.4.2.2 2D Force-Based Interaction

A significant body of tabletop literature focuses on how to move and rotate virtual objects on a digital surface. One of the overarching results of studies [Kruger et al., 2003; Liu et al., 2006; Wilson et al., 2008] involving movement and rotation is that simulating (at least to some degree) how movement and rotation happen with physical forces typically results in both improved performance and a compelling feeling of being embodied with the virtual objects.

The RNT technique [Kruger et al., 2005] for moving and rotating objects uses the metaphor of an opposing force acting on a virtual object to make it rotate while moving. This technique has also been extended so that, when let go, an object will continue along its trajectory according to the current speed of movement. This extension produces the ability to "flick" or "toss" objects across the screen [Hinrichs et al., 2006; Isenberg et al., 2006]. The TNT techniques [Liu et al., 2006] use 3DOF to more directly simulate the movement observed in studies of moving and rotating paper on physical tables. With this method, a person can place their hand or a physical block on a virtual object, and the position and orientation of the hand or block controls the movement and rotation of the virtual object. On multitouch tables, two fingers are typically used for a combined movement, rotation, and scaling of a virtual object. The position of the first touch is used to determine the movement of the object and the position of the second touch relative to the first is used to determine the rotation and scale. This technique simulates how movement and rotation can occur with physical objects if frictional force between the fingers and objects is considered. The scaling aspect is an example of how this familiar force-based behaviour can invoke virtual behaviour not possible in the physical world (i. e., magically growing or shrinking objects).

ShapeTouch [Cao et al., 2008] provides force-based interactions on 2D virtual objects, such as pushing objects from the side, tossing them across the screen, peeling them back to place other objects underneath, and more. These techniques use the sensory data to invoke complex but physically familiar behaviour on the objects that are in direct contact with a person's hands and arms.

2.4.2.3 3D Force-Based Interaction

Force-based effects such as collisions, gravity, mass, and inertia can also be integrated into 3D environments through the use of a physics engine [e. g., BumpTop, Agarawala and Balakrishnan, 2006]. The image data provided through many multitouch input devices (see section 2.1) can be more directly integrated into such physics engines by creating physical bodies (either through proxies or particle proxies) that then can interact with the virtual objects through the physics engine [Wilson et al., 2008]. Because a person's hands and fingers (or even other physical objects) have a virtual representation in the physics engine, these can be used to push other virtual objects around.

The use of forces in general, and the use of physics-based forces in 3D virtual worlds in particular, have immense appeal as a basis for interaction on multitouch tables. However, while many appealing interactions have emerged, they fall short of the full functionality required for practical applications. For instance, BumpTop [Agarawala and Balakrishnan, 2006] resorts to a symbolic gestural language, which has an associated learning curve and the need for memory retention. Wilson et al. [2008] point the way to interactions that extend physical real-world responses into the virtual world, but fall short in that the realized virtual interactions provide only the ability to move invisible proxies, and not to spin, flip, or lift the virtual objects. In essence, this work provides no equivalent to an opposable thumb and has made a direct call for the ability to pick objects up and place them inside others—capabilities offered by the sticky tools approach presented in chapter 6. Another approach to manipulating 2D and 3D objects is to use the space in front of the display [Keijser et al., 2007; Malik et al., 2005] to extend interaction capabilities; however, this has only been accomplished through additional hardware such as markers and vision-based systems. Sticky tools achieves all 6DOF without additional hardware.

2.4.3 Storytelling and Therapeutic Applications

The specific application domain of the design presented in chapter 7, virtual sandtray therapy, is related to a number of approaches where modern touch and tangible technology is used to support work with children for storytelling or therapeutic purposes. Early examples include the use of robotic stuffed animals [Bers et al., 1998] to help young cardiac patients cope with their situation by encouraging them to talk about it, comparable to virtual sandtray therapy. Later work employed an interactive physical play mat (StoryMat) to record children's storytelling activities [Ryokai and Cassell, 1999; Cassell and Ryokai, 2001]. Li et al. [2008] developed a tangible tabletop game to support the therapy of children with cerebral palsy who need to train specific motor skills. The game combined tangible elements with a tabletop surface that was illuminated with coloured LEDs from below and was found to encourage children to train the desired therapeutic movements. Morris et al. [2006a] describe an interface designed for children with Asperger's syndrome. Here, the fact that tabletop displays afford collaboration is used in a game form to allow four children simultaneously to train social skills and collaboration. Similar to the motivation for the design in chapter 7, the authors name the children's affinity to technology as one of the criteria that makes tabletop technology well suited for such therapeutic applications.

Although not used to tell a story, Piper and Hollan [2008] describe the design of an interactive table used to facilitate communication between a doctor and a deaf patient. This design process has many similarities to the process used in chapter 7 and shows the benefit of tabletop technology for the deaf community. As would be expected, the work in this dissertation reinforces the idea that working closely with the people who professionally understand the application area can lead to successful tabletop applications.

2.5 Chapter Summary

In this chapter, I have summarized and synthesized the literature in four primary areas: *tabletop display technology*, *perception of 3D* , *3D manipulation*, and *tabletop display theory & design*. In this dissertation, I expand upon this body of work in several ways.

In chapter 3, I provide empirical evidence to demonstrate that the perceptual error introduced by viewing 3D information projected in 2D both exists for tabletop display applications and is likely to be both large and problematic when collaboratively viewing information. Furthermore, I add to this body of work by empirically demonstrating that a careful choice of projection geometry can mitigate this problem. This study expands the literature on the perception of 3D by examining these theories in the specific domain of collaborative tabletop display environments. This work also contributes to the theory and design of tabletop display applications. I then implement several techniques which distort the projection geometry to further mitigate this problem. These techniques contribute to the design of 3D tabletop display applications.

In chapter 4, I provide a mathematical description of DOF, and in chapter 5 demonstrate how to apply this theory to match multitouch input to 3D output. That is, I demonstrate empirically that people are both faster and prefer the use of multiple fingers to manipulate a 3D virtual object. This work contributes to the vast literature on the manipulation of 3D virtual objects and extends it to include manipulation techniques on a table's surface.

In chapter 6, I describe a framework for how to create 3D tabletop applications based on the evidence presented in the preceding chapters. This framework leverages the previous work that uses force-based metaphors in application design, and directly synthesizes the TUI literature with tabletop display literature. Chapter 7 then demonstrates how to apply this framework to the domain of sandtray therapy. This technology directly benefits both the tabletop display community, by demonstrating the application of this theory, and the psychotherapy community, by providing technology to improve their ability to practice art therapy.

Viewing 3D at a Tabletop Display 3

As discussed in chapter 2, there has been a significant amount of research into the perception of 3D information displayed in a 2D plane. While the research presented in this chapter was ongoing throughout my thesis, I discovered early on that it was essential to understand how people perceive 3D when designing multitouch interaction techniques on large 2D displays, and so I present this research first to set the stage for the interaction work presented next. This chapter builds upon the existing literature by presenting a series of experiments which evaluate a viewer's perception of this projection from 3D onto 2D on an interactive table. The purpose of these experiments was to empirically validate that some methods of projecting 3D onto 2D can lead to perceptual errors and that these errors are prevalent in a collaborative tabletop display environment. This chapter also builds upon the research which examines collaborative viewing of 3D information in these environments (section 2.2.4) by providing novel techniques for mitigating perceptual error that may be caused by a discrepancy between centre of projection (COP) and point of view (POV) (see section 2.2.2 for a detailed description of COP and POV).

There are a number of factors which may influence the perception of 3D virtual artifacts on a 2D display. The series of experiments presented here is intended to examine the factors which are most apparently different when using a tabletop display versus other common devices, such as wall displays, desktop computers, or mobile devices. Specifically, tabletop display environments afford multiple people gathering around the display and, on the other hand do not afford viewing the information from directly in front of the display. Thus, in each of these experiments, the discrepancy between the COP used to render the information and the POV of the viewer was the primary factor. A series of experiments was run to examine different perceptual tasks. Experiment 1 examines the task of estimating the angle of a virtual object within the space of the table, experiments 2 and 3 examine the task of matching the structure of a virtual object to others within the same space, and experiment 4 examines the task of estimating the orientation of a virtual object with respect to its physical surroundings. Furthermore, experiment 3 differs from experiment 2 in that the participants could interact with one of the virtual artifacts, which makes it possible to examine whether interactivity affects one's perception of the 3D virtual information.

In this chapter, I first motivate the need to study how people perceive 3D information at an interactive tabletop display, particularly when multiple people are simultaneously viewing the same information (section 3.1). I then present a series of experiments that demonstrates that a discrepancy between the COP and a person's POV can result in a real and significant amount of perceptual error (section 3.2) and then describe a series of techniques to mitigate the problem when designing 3D tabletop display interfaces for multiple viewers (section 3.9). I then summarize the major contributions of this chapter (section 3.10).

3.1 Collaborative Viewing of 3D on an Interactive Table

The causal relationship between the world and our perception is not entirely certain: either we think the world is 3D because we perceive things visually in 3D, or we perceive things visually in 3D because we think the world is 3D. In either case, it is notable that there seems to be something intrinsically 3D about both the world and our perception. However, throughout history, there have been attempts to represent the 3D world within a 2D surface (e. g., on cave walls, tablets, canvases, paper, and now digital displays). While these 2D surfaces were not always planar (e. g., The Sistine Chapel, a celestial globe), the 3D information was mapped (i. e., projected) into fewer dimensions. By projecting this data, some of the information from the

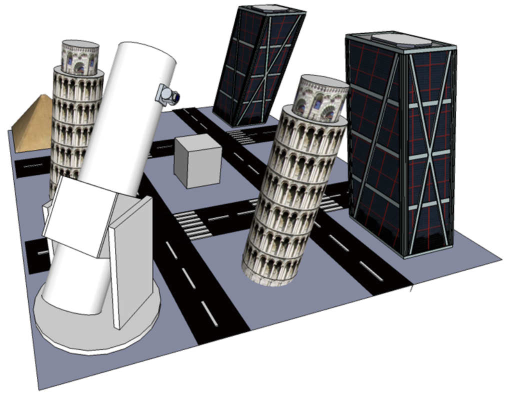


Figure 3.1: The model used to render the images in figures 3.2 to 3.5.

physical world being represented may be lost. There are, however, a variety of techniques to map 3D into 2D that preserve some of the lost information, creating the illusion that one is looking at something that is in 3D. This illusion can be created by simulating one or more of the *depth cues* that are naturally available to us when we perceive physical 3D objects. With the introduction of computer technology, many of these depth cues are far easier to simulate than ever before (see section 2.2.1 for a discussion of how they can be recreated digitally).

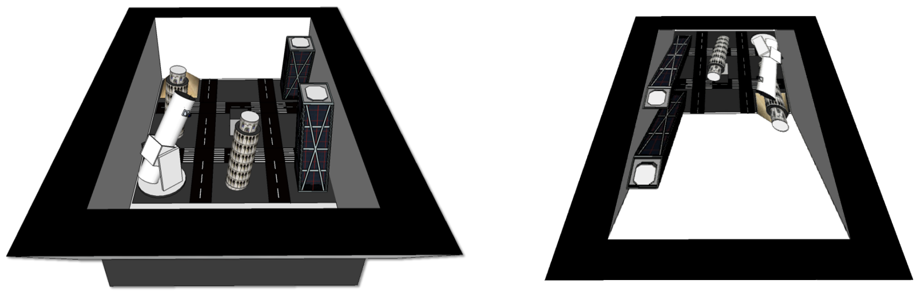
The ability to "draw pictures" led eventually to more abstract visual representations of the physical world, such as letters, glyphs, diagrams, and other symbolic techniques. Today's computer interfaces use a vast array of 2D graphical elements, such as windows, icons, menus, and pointers (WIMP), which have the ability to represent actions, functions, ideas, and concepts, as well as something from the physical world. Because most of these graphical elements are intrinsically 2D and because the majority of digital displays still present information within a single 2D plane, there is in the HCI community much more knowledge about how to design computer interfaces that make use of 2D elements than to design interfaces with 3D graphics. However, due to their abstract nature, these 2D elements do not necessarily benefit from the intrinsic connection between our human visual systems and the 3D physical world.

(a) (b)

Figure 3.2: (a) the model in figure 3.1 projected using a COP and POV above the table; and (b) a video of changing one's POV from above the table to the side. The still image of the video shows the side view. figure 3.3a shows what one might expect to see from the side of the table.

On a multitouch tabletop display, the connection between the visuals we are seeing and the physical world becomes increasingly important, because we are able to reach out and "grab" the virtual image. While, over millennia, we have become familiar with the presentation of 2D visual information, sensory-motor information has still been exclusively presented in the 3D physical world. The ability to perform 3D actions, such as twisting, flipping, piling, and so on has persisted, even with artifacts containing 2D information, such as books and paper. Some application areas also benefit greatly from the combination of 3D data presentation and the natural collaboration and face-to-face communication affordances of tabletop systems—examples include tasks such as the planning of surgical procedures with 3D body models, urban planning discussions using models of real-world structures, and collaboration over 3D visualizations. In addition, tasks in multi-display environments may require that individual displays be able to indicate other surfaces or data in the real world (e. g., an object may need to be oriented towards a partner object on another display).

Displaying 3D objects on tables, however, presents new problems for designers of tabletop applications. The representation of a 3D virtual scene on a 2D surface such as a tabletop



(a) (b)

Figure 3.3: (a) the model in figure 3.1 projected using a COP and POV at the side of the table; and (b) a video of changing one's POV from one side of the table to the other. The still image of the video shows the COP and POV at opposite ends of the table.

requires the projection of the virtual 3D objects onto the display surface—the choices made in creating this 3D image, such as where the centre of projection (COP) is for the image, or whether a perspective or parallel geometry is used, can have dramatic effects on the appearance of the resulting scene. If the projection is poorly designed, the resulting image on the tabletop appears distorted, and it becomes difficult for the viewer to determine the shape and orientation of objects in the 3D scene.

To illustrate this problem, I first present a 3D model in figure 3.1. This figure contains both a static image of the model and an embedded video image to clarify the shape of the model. Note that it contains a collection of buildings, many of which are not entirely perpendicular. I will use this model to illustrate much of the ongoing discussion. Typically, 3D graphics applications tend to use a COP directly above the table. Figure 3.2a shows a still image of what one would see if hovering above the table and viewing such a scene. In this figure, the COP and POV coincide. Figure 3.2b shows a video of what it might be like to float from above the table to being at its side. The still image shows the result of moving one's POV to the side, with the COP remaining above the table. Ideally when standing beside a digital table one would like to get a good 3D impression of this model. For instance, one might expect to

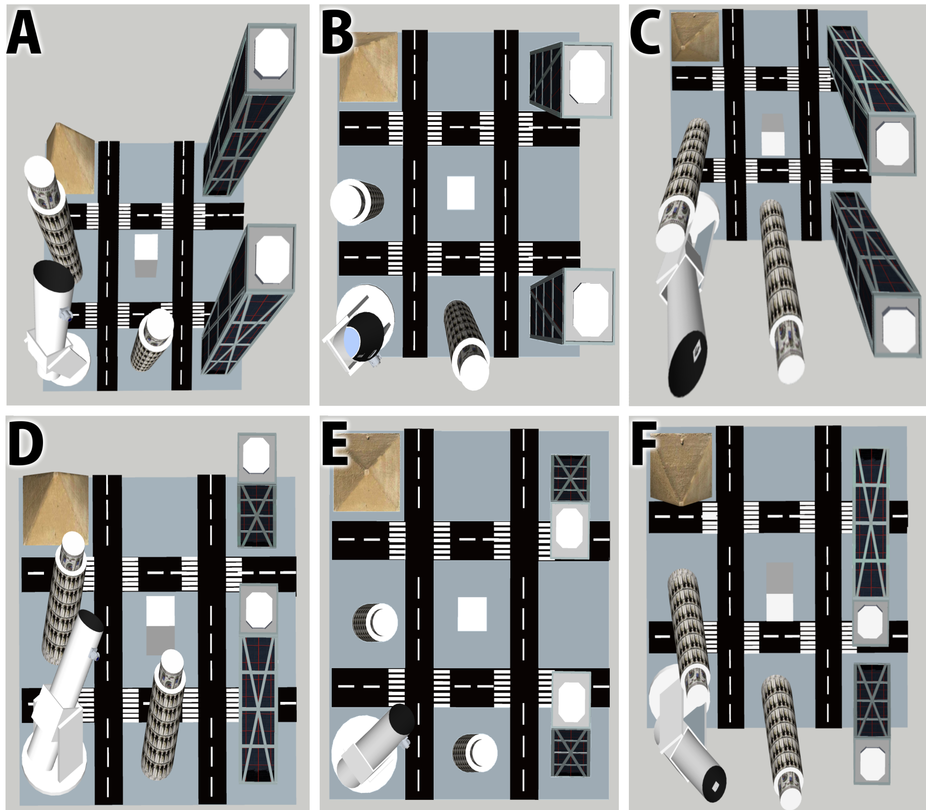


Figure 3.4: The images rendered to the tabletop display in figure 3.5.

see the model as shown in figure 3.3a. This image again has a cop and pov which coincide, but this time from the side of the table. However, if one were to walk to the opposite side of the table one would get the experience shown in the video in figure 3.3b. The still image shows the view from the opposite side. To further illustrate this problem, this same model is projected into the 2D images shown in figure 3.4 using different projection geometries and cops. The letters A–F correspond to the same labels in figure 3.5, which shows what the images would look like when observed from different povs at a tabletop display.

This problem is intensified when people work together with 3D data around a table. Common 3D projections have only a single cop centred in front of the screen at a common viewing distance. People working at a tabletop display are pretty much guaranteed to not be floating above the display and in addition many people may be looking at the display from

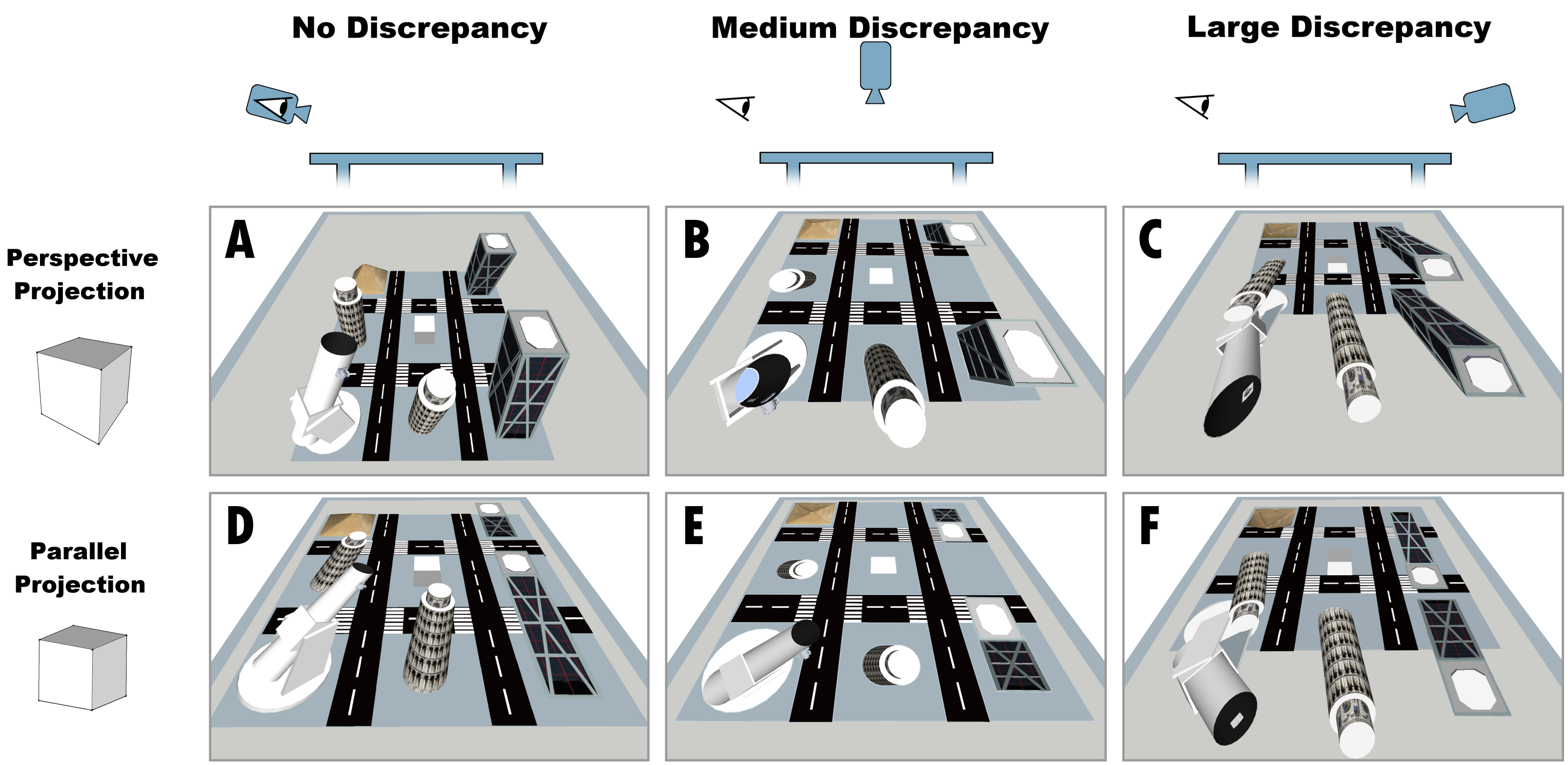


Figure 3.5: Appearance of 3D models rendered on a table with different levels of discrepancy between point of view (POV) and centre of projection (COP) using parallel and perspective projection geometries.

many different angles and heights. However, usually, 3D projections on tabletops still have only the common single virtual viewpoint. As a result some, if not most, of the collaborators around a table will see a distorted view of some of the 3D objects. In tasks where the group needs to discuss details of the model such as shape, orientation, and tilt, these distortions could cause misunderstandings and difficulty in communicating about the model.

Little is known about the problem of interpreting 3D models on 2D tables—about what types of actions are most affected, about the severity of the errors that people make, or about how to choose a projection that minimizes the negative effects. To investigate these issues, we[1] carried out a series of experiments in which people were asked to estimate the angle, pattern, or orientation of a 3D object, projected onto a tabletop display with different COPs and different projection geometries. We found that as the COP moved further away from the observer, their error in estimating orientation significantly increased, but this effect was not present for estimation of angle or pattern. However, we also found that when parallel projection geometry is used in combination with a neutral COP (i. e., between the two viewers), participants were more accurate in estimating angle, pattern, and orientation. Therefore, optimizing the perspective for one person may cause major problems for the others in the group, especially when estimating orientation—but providing a neutral COP and parallel projection geometry may help to mitigate the problem. Furthermore, participants performed better when interactivity was provided than when it was not, which suggests that this problem can also partly be mitigated by providing direct interaction.

[1]The perception study in this chapter is largely taken from the materials published in Hancock et al. [2009a] and the mitigating techniques presented in this chapter were taken from materials published in Hancock and Carpendale [2007]. The use of the first-person plural in this chapter refers to the authors of these papers: Mark Hancock, Miguel Nacenta, Carl Gutwin, and Sheelagh Carpendale.

3.2 Overall Study Design: Analyzing Issues of Discrepancy

While there have already been many studies exploring the perception of 3D objects represented in pictorial space, and these have generated some specific (if somewhat controversial) theories about the effect of COP/POV discrepancies, how these theories apply to tabletop display environments is still largely unknown. We do not know what the effects of discrepancy are on horizontal surfaces (experimental setups so far have been vertical), how discrepancy interacts with different types of projection geometries (perspective vs. parallel), or whether motion parallax can help compensate for the distortions created by the highly oblique POVs typical for tables.

In our study, we focus on the analysis of COP/POV discrepancies to inform design choices of 3D tabletop applications. We explore three levels of discrepancy (figure 3.5): when the COP coincides with a person's POV, when the COP is directly above the table, and when the COP coincides with someone else's POV. We are also interested in the combination of discrepancy with the two main types of projection geometries (perspective and parallel), and with motion parallax (section 2.2.2.2), which is reproduced through the real-time tracking of the participant's head. Our study is designed to answer three main questions:

- What are the effects of discrepancy on error?
- How do different projection geometries and motion parallax affect errors due to discrepancy?
- Are there any special cases that designers could use to alleviate errors?

In this series of experiments, the effect of the discrepancy between the COP and POV locations on three perceptual tasks is investigated:

Estimating Angle estimating the angular difference internal to a 3D object within the pictorial space. Specifically, the angle between two edges of a triangle oriented within the pictorial space.

Object Structure Matching estimating the structure of a 3D object within the pictorial space. Specifically, whether that structure matches that of another 3D object represented in the same space.

Estimating Orientation estimating the orientation of 3D objects from the pictorial space into the physical environment.

In contrast with previous research in the area, we focus on the interactive tabletop scenario, and therefore we explore factors and conditions that are relevant for interactive tabletops. For example, although a large proportion of the studies mentioned in section 2.2.3 restrict participants to monocular perception, the use of a single eye to work on tabletops is not reasonable; all our tests are, instead, binocular. Previous research shows that binocular and monocular observation of pictures from non-coincident POV and COP is different [Vishwanath et al., 2005].

We also compare perspective and parallel projections because, although parallel projections are incapable of generating a geometrically accurate retinal image, they have been shown to look more natural than their perspective counterparts [Hagen and Elliott, 1976], and are extensively used in architecture and engineering for their accuracy. Similarly, we include motion parallax in our study conditions because it provides potentially strong cues [Cutting, 1997; Dijkstra et al., 1995] and has not generally been considered in pictorial research, which is usually more concerned with static pictures.

3.3 Study Description

This section describes the experimental setup for a series of four experiments. These experiments were run in two parts: experiments 1 and 2 were performed in the same session by the first set of participants and experiments 3 and 4 were performed in another session by a second set of participants. These four experiments were designed to share some common apparatus, procedure, experimental conditions, hypotheses, and analysis. The common elements will be described in this section. Section 3.4 describes the details of experiment 1 and is followed by its results in section 3.4.5. Sections 3.5 and 3.6 describe the details of experiments 2 and 3, respectively, and are followed by the results of a combined analysis in section 3.6.5. Section 3.7 describes the details of experiment 4 and is followed by its results in section 3.7.5.

3.3.1 Participants

All participants were recruited from the local community. Although people were recruited in pairs, the experimental design was symmetric, so each participant was analysed separately. That is, the participants shared the main display and performed tasks simultaneously, but responses were entered and timed separately. Participants stood at opposite ends of the table and the conditions were mirrored such that, for each condition, the experience at one end would occur in a later condition on the other.

Part 1: Twenty-four participants (13 female, 11 male) completed the first part, which involved experiments 1 and 2. Their ages ranged from 18 to 33 (Mdn = 26, SD = 5.2) and were recruited in pairs (4 female, 3 male, and 5 mixed).

Part 2: Twenty-four more participants (11 female, 13 male) completed the second part, which involved experiments 3 and 4 and were again recruited in pairs (3 female, 4 male, and 5 mixed). Their ages ranged from 19 to 36 (Mdn = 28, SD = 4.5).

3.3.2 Common Apparatus & Procedure

Figures 3.7, 3.9, 3.10 and 3.12 show diagrams of the experimental setup. For all four tasks, participants stood at the ends of a 146 cm × 110 cm bottom-projected tabletop display with a resolution of 2800 × 2100 (19 pixels / cm). The POV of each participant was tracked using

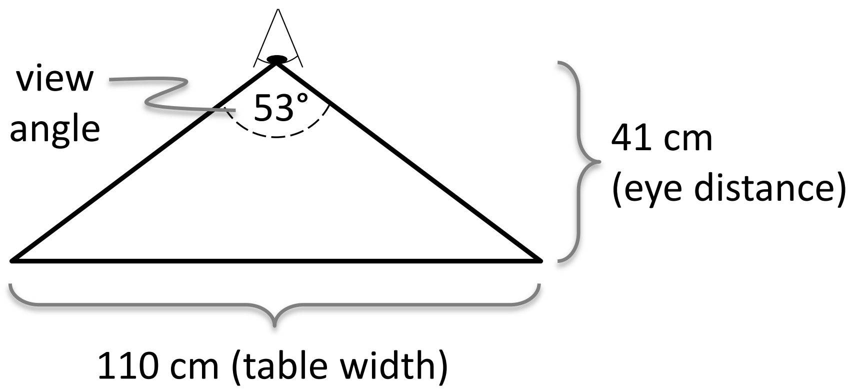


Figure 3.6: The view volume in the medium discrepancy condition with perspective projectoin geometry had a COP 41 cm from the table and a view angle of 53°. The remaining conditions used the same near plane, but the COP was matched to either the participant (in the no-discrepancy condition) or their partner (in the large-discrepancy condition).

a Vicon [2006] motion tracking system and markers placed on hats that the participants wore throughout the experiment. Figure 3.6 shows the parameters for the view volume used throughout the experiment. These parameters determine the COP for the medium discrepancy condition as well as the near plane for all the remaining conditions.

Prior to the experiment, each participant was asked to complete a background questionnaire, including six questions from the Purdue Visualization Test [Guay, 1977] and following each block of trials, was asked to respond to two to four 7-point Likert-scale questions relevant to the specific task (sections A.2.2 and A.3.2). Table 3.1 shows the list of Likert-scale questions.

	Question	**Tasks**
Q1	I found it easy to see the shape of objects in this mode	Experiments 1 to 4
Q2	I found it easy to compare objects in this mode	Experiments 1 to 3
Q3	I found it difficult to know what my partner saw	Experiments 1 and 2
Q4	I would prefer to use a tabletop display with this setting	Experiments 1 to 4

Table 3.1: A list of Likert-scale questions asked for each condition.

In experiments 1 to 3, the objects rendered to the screen were 'molecules'. For the entirety of this chapter, the term *molecule* is meant to be a set of nodes, represented as shaded spheres, and a set of edges, represented as shaded cylinders, that each connect two nodes.

3.3.3 Common Conditions & Design

The focus of our study was on how the degree of discrepancy between the centre of projection (COP) and the observer's point of view (POV) affects perception of object orientation. We thus varied the *discrepancy* between these two points as the primary factor in all four experiments. The three levels of this factor correspond to likely choices when designing a 3D application for a tabletop display: no discrepancy (when the COP and POV are the same), medium discrepancy (when the COP is directly above the table), and large discrepancy (when the COP is set to someone else's POV). A secondary factor in this series of experiments was *motion parallax*—that is, whether the COP dynamically followed the participant's POV, resulting in perspective changes as the participant moved their head. The motion parallax condition was tested both under the circumstance where the COP moved with the participant's POV and when the COP moved with their partner's POV. That is, in the latter condition, participants would be watching the effect of motion parallax for their partner.

		Discrepancy		
		none	medium	large
Motion	absent	NA	MA	LA
Parallax	present	NP		LP

Table 3.2: The discrepancy-parallax conditions. The first letter describes the level of discrepancy and the second letter describes the level of motion parallax.

With a medium discrepancy (when the COP is directly above the table), it is not sensible to introduce motion parallax, as there is no person to move with the COP. Thus, the first two factors combine into five discrepancy-parallax conditions shown in table 3.2. These combinations are coded in the remainder of this chapter so that the first parameter describes the discrepancy condition (none, medium, or large) and the second describes the motion parallax condition (present or absent) as follows: none-present (NP), none-absent (NA), medium-absent (MA), large-absent (LA), and large-present (LP).

Another secondary factor in our study was *projection geometry*, which had two levels: perspective and parallel. Strictly speaking, a parallel projection does not have a COP, but instead uses a direction of projection (DOP). However, the same COP which determines the perspective geometry can be used to generate an "equivalent" parallel geometry by using the vector from the centre of the near plane to the COP as the DOP. Thus, the discrepancy and motion parallax factors, which rely on the COP, have equivalent levels when using a parallel geometry, and thus we use the same terminology. For example, in the no discrepancy condition, the COP used to generate the perspective geometry (the participant's POV) is also used to generate an "equivalent" parallel geometry that has no discrepancy.

3.3.4 Common Hypotheses

All four experiments were designed to test the following primary hypotheses:

Hypothesis 3.1: As the discrepancy increases, perceptual error will increase.

Hypothesis 3.2: When there is no discrepancy, perspective geometry and motion parallax will reduce perceptual error; when there is discrepancy, perspective geometry and motion parallax will increase perceptual error.

Hypothesis 3.3: Medium discrepancy (COP directly above the table) will be a special case that decreases perceptual error.

All four experiments were also designed to test the following secondary hypothesis:

Hypothesis 3.4: The use of motion parallax will require more time for the participants to complete the task.

3.3.5 Common Analysis

We performed a full factorial repeated-measures analysis of variance (ANOVA) on our data and a series of planned comparisons for post-hoc analysis of an expected interaction between the discrepancy-parallax condition and projection geometry (hyp. 3.2). Our planned comparisons correspond to our primary hypotheses as follows:

- To test hyp. 3.1, we perform pairwise comparisons in the order of least to most discrepancy: NA to MA and MA to LA.
- To test hyp. 3.2, we additionally compare the two motion parallax conditions to the endpoints of discrepancy: NP to NA and LA to LP. An effect of perspective geometry would appear as a main effect of the ANOVA.
- To test hyp. 3.3, we additionally compare the MA condition with the remaining two conditions (NP and LP).

We performed these planned comparisons either by combining data from the two projection geometries (to analyze a significant main effect of the discrepancy-parallax condition) or separately for each projection geometry (to analyze a significant interaction between the discrepancy-parallax condition and projection geometry) and used a Bonferroni-corrected type I error threshold ($\alpha/6$ or $\alpha/12$).

3.4 Experiment 1: Estimating Angle

This experiment was designed to test people's perception of a virtual object's shape *within* the 3D model being presented. Specifically, our interest was in whether presenting 3D virtual shapes to multiple people at a tabletop display might result in perceptual errors for some. The participant was given the task of determining an internal angle of the simple shape of a triangular molecule.

3.4.1 Apparatus & Procedure

Participant responses were entered using styli on two 25 cm × 18 cm Tablet PCs each set to a resolution of 1024 × 768 (41 pixels / cm). Each tablet was placed next to a participant on a music stand, which was tilted at a 30° angle to match the target. Besides this restriction in angle, the participants were free to alter the height of the stand or to place it on either their left or right side, depending on their handedness or preference (some right-handed participants chose to place the stand on their left side, due to the tilt angle that was required).

3.4.1.1 Task Description

This experimental task required participants to determine the angle within a target triangular molecule represented on a 3D plane. For each trial, a 3D target molecule was displayed on the screen in one of two locations: near and to the left, or far and to the right (figure 3.7). Each molecule was a triangle with two fixed nodes (red and green) and one randomly placed third node (blue) connected by grey edges. The random placement was such that the red-green-blue angle and the green-red-blue angle were both acute (strictly less than 90°). The blue node was always placed on the same plane, which was 30° from the plane of the table, rotated along the x-axis. To provide their answer for each task, participants would tap at some location on the Tablet PC to indicate the relative position of the blue node with respect to the red and green nodes. Feedback was provided on the tablet such that, before tapping, red and green circles connected by a grey edge were drawn at the same orientation as on the table, and after tapping, a blue circle would appear at the point tapped connected by grey edges to both the red and green circles. The participant was free to tap again or drag the pen to correct their answer as much as necessary before tapping a confirm button to complete the trial. Once both participants had pressed the confirm button, the next trial would begin. The distance from the red to green nodes on the table's surface in the medium discrepancy condition with parallel geometry was 462 pixels (24.3 cm). The distance from

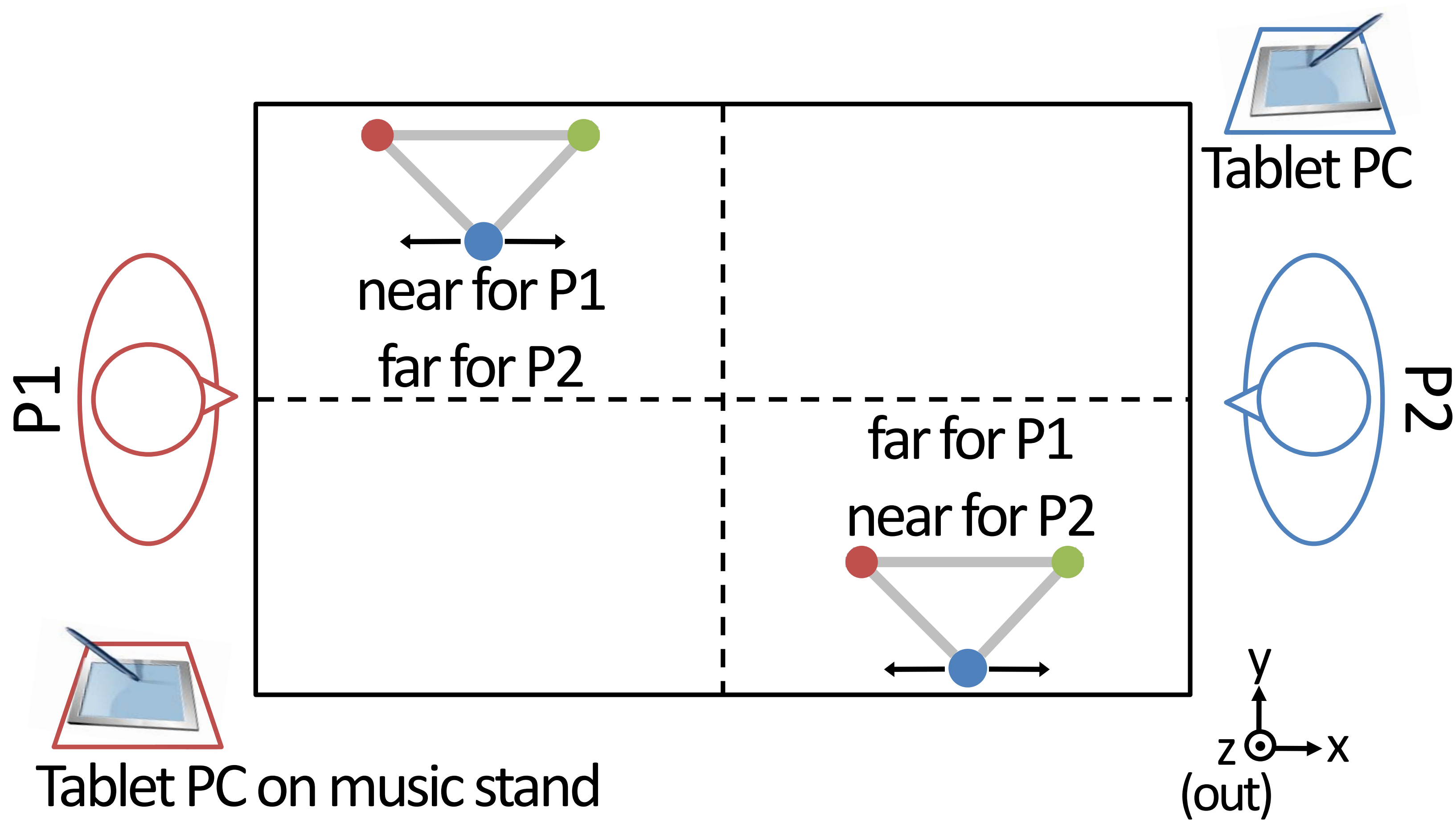


Figure 3.7: A diagram of the experimental setup for experiment 1.

the red to green circles on the tablets was 323 pixels (7.9 cm). To remind the participants about the projection geometry being used, three groups of four cube frames were shown in the empty quadrants. These cubes were rendered in the same 3D model as the target objects in all conditions.

3.4.2 Conditions and Design

In this task, targets could appear either in the 'near' or 'far' halves of the table, as shown in figure 3.7. Thus, participants carried out two practice trials (one near and one far) and eight testing trials (four repetitions of near and far) for each condition-projection combination, for a total of 100 trials (including practice). This experiment therefore used a 5 discrepancy-parallax condition × 2 projection geometry × 2 location fully-crossed within-participants design. The 10 condition-projection pairs were counterbalanced between participants using a random Latin Square.

3.4.3 Measures

Both error and task completion time (TCT) were measured. The error was broken down into error along the x-axis and error along the y-axis:

E_x: the error measured along the x-axis of the table. This error is calculated as the difference between the measured blue node and the target blue node on the tablet and is reported as a ratio of this difference to the distance between the green and red nodes. Thus, this value has no units and can be multiplied by either the distance on the table or the distance on the tablet to obtain an actual distance (see section 3.4.1 for these values).

E_y: the error measured along the y-axis of the table. This error is calculated as a ratio with the same denominator as E_x and is again reported without units.

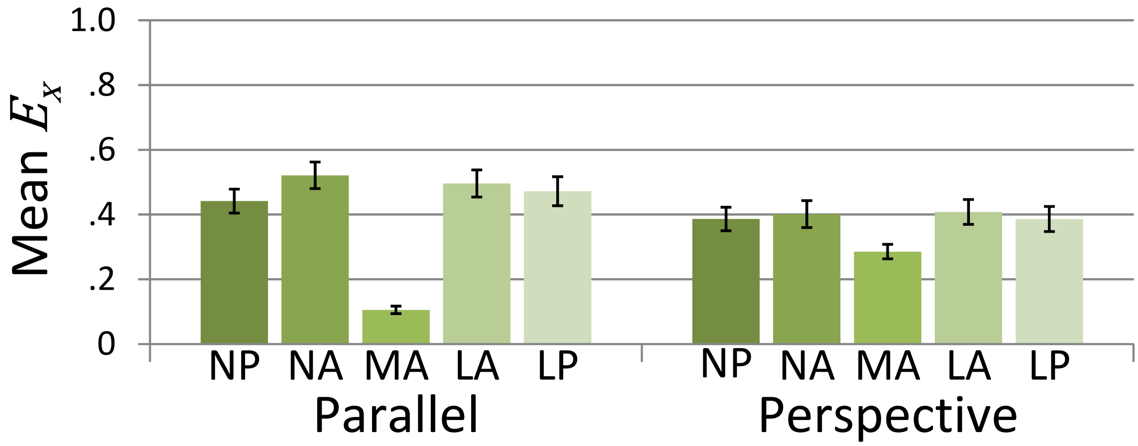


Figure 3.8: Mean error in the x-axis (E_x) for the ten condition pairs.

3.4.4 Hypotheses

The study was designed to isolate errors along the x-axis of the table (E_x) due to the change in discrepancy. Thus, hyps. 3.1 to 3.3 applies to the E_x measure, and the following additional hypothesis applies to the E_y measure:

> **Hypothesis 3.5**: The type of projection geometry will affect E_y.

3.4.5 Results

We performed a 5 discrepancy-parallax condition × 2 projection geometry × 2 location repeated measures anova, and the planned comparisons for pairwise differences.

3.4.5.1 Error in X (E_x)

What was the effect of discrepancy on error? (hyp. 3.1) There was a significant main effect of discrepancy-parallax condition on E_x ($F_{4,92} = 25.5$, $p < .001$). However, the mean differences do not follow the trend predicted by hyp. 3.1 (figure 3.8). Instead, the ma case (when the cop is directly above the table) was less error prone than all other cases.

What were the effects of projection geometry and motion parallax on error? (hyp. 3.2) There was no main effect of projection geometry on E_x ($F_{1,23} = 3.0$, $p = .10$), with mean

		Experiment 1			
		E_x ($\alpha = .0042$)		TCT ($\alpha = .0042$)	
		Parallel	Perspective	Parallel	Perspective
hyp. 3.1	NA to MA	$p <$ **.001**	$p =$ **.002**	$p = .007$	$p = .065$
	MA to LA	$p <$ **.001**	$p <$ **.001**	$p =$ **.004**	$p = .239$
hyp. 3.2	NP to NA	$p = .044$	$p = .512$	$p = .230$	$p = .896$
	LA to LP	$p = .318$	$p = .433$	$p = .004$	$p = .328$
hyp. 3.3	MA to NP	$p <$ **.001**	$p <$ **.001**	$p <$ **.001**	$p = .047$
	MA to LP	$p <$ **.001**	$p = .006$	$p <$ **.001**	$p = .033$

Table 3.3: Planned comparisons for experiment 1.

error for perspective conditions ($M = .41$, $SD = .03$) only slightly higher than for parallel conditions ($M = .37$, $SD = .03$). There was also no significant difference for the presence of motion parallax (table 3.3 and figure 3.8). Therefore, hyp. 3.2 is not confirmed.

Is it a special case to have CoP directly above the table? (hyp. 3.3) There was a significant interaction between discrepancy-parallax condition and projection geometry for E_x ($F_{4,92} = 16.7$, $p < .001$). Our planned comparisons indicate that the interaction was due to the special case of medium discrepancy when using a parallel projection. As can be seen in figure 3.8, error in MA with perspective geometry was substantially higher than error in MA with parallel geometry. Furthermore, in both the parallel and perspective case, the medium discrepancy condition (MA) was less error prone than both the no discrepancy conditions (NA and NP) and the large discrepancy conditions (LA and LP), which confirms hyp. 3.3. Furthermore, this effect is more pronounced for the parallel case than the perspective case.

3.4.5.2 Error in Y (E_y)

There was no significant main effect for condition ($F_{4,92} = 0.8$, $p = .52$) or projection geometry ($F_{1,23} = 0.1$, $p = .72$), nor was there an interaction ($F_{4,92} = 0.2$, $p = .93$). Thus, hyp. 3.5 cannot be confirmed.

3.4.5.3 Task Completion Time

We performed the same ANOVA on task completion time (TCT) using a log transform. There were significant main effects of both discrepancy-parallax condition ($F_{4,92} = 7.8$, $p < .001$) and projection geometry ($F_{1,23} = 19.6$, $p < .001$), as well as an interaction between the two ($F_{4,92} = 3.0$, $p = .02$). The planned comparisons revealed only three significant differences, all when using parallel projection geometry (table 3.3). Using this projection, the MA condition ($M = 9.8$ s, $SE = 1.1$ s) was significantly faster than the NP ($M = 13.0$ s, $SE = 1.1$ s), LA ($M = 11.8$ s, $SE = 1.1$ s), and LP ($M = 13.4$ s, $SE = 1.1$ s) conditions. This trend provides further support for hyp. 3.3. There was also a marginally significant pairwise difference between LA and LP ($p = .0042$), providing partial support for our hypothesis that motion parallax costs more time (hyp. 3.4).

3.5 Experiment 2: Object Structure Matching

This experiment was designed to test people's perception when asked to compare the shape of a 3D virtual object with another in the same tabletop display environment. We were again interested in the collaborative aspect of the perception of 3D virtual objects, and so this experiment and the next (experiment 3) differ slightly from experiments 1 and 4 in the way that the primary conditions were presented to the participants. That is, instead of all virtual objects being rendered using the same projection geometry across the entire display, in the large discrepancy conditions, participants were asked to compare objects rendered using 'their own' COP to those rendered 'for their partner'.

3.5.1 Apparatus & Procedure

Participant responses were entered using styli on the same two Tablet PCs as in experiment 1. In this experiment, the participants were free to alter both the height and angle of the stand or to place it on either their left or right side, depending on their handedness or preference (some right-handed participants chose to place the stand on their left side, as it was already placed there from experiment 1).

3.5.1.1 Task Description

In this experimental task, each participant was presented with a pattern molecule at their end of the display and four target molecules in the centre of the display which they were asked to compare and match with the pattern. To provide their answer for each task, participants were presented with four squares on the tablet that corresponded to the four target molecules and were asked to tap any and all quadrants that matched their pattern. When tapped, a square would become highlighted, and participants could change their answers by tapping again before tapping the confirm button. Once both participants had tapped the confirm button, the non-matching molecules were hidden to indicate the correct answer. Once both participants again tapped the start button, the next trial would begin. Every molecule (both matching and non-matching) had the same basic structure of a central node (grey) to which all other nodes were connected via a lighter grey edge. Three other nodes were always present: a blue node directly above the central node along the z-axis, a red node directly to the right of the central node on the x-axis, and a green node directly up from the central node on the y-axis. The remaining nodes were all grey and determined the pattern of the molecule.

In this task, we were interested in comparing performance when people had to compare objects rendered in 'their own' point of view with objects rendered 'for their partner'. Thus, in the large discrepancy condition (see section 3.3.3), the pattern was rendered with no discrepancy (i. e., using a cop that coincides with the participant's pov), while the target was rendered with large discrepancy (i. e., using a cop from their partner's pov).

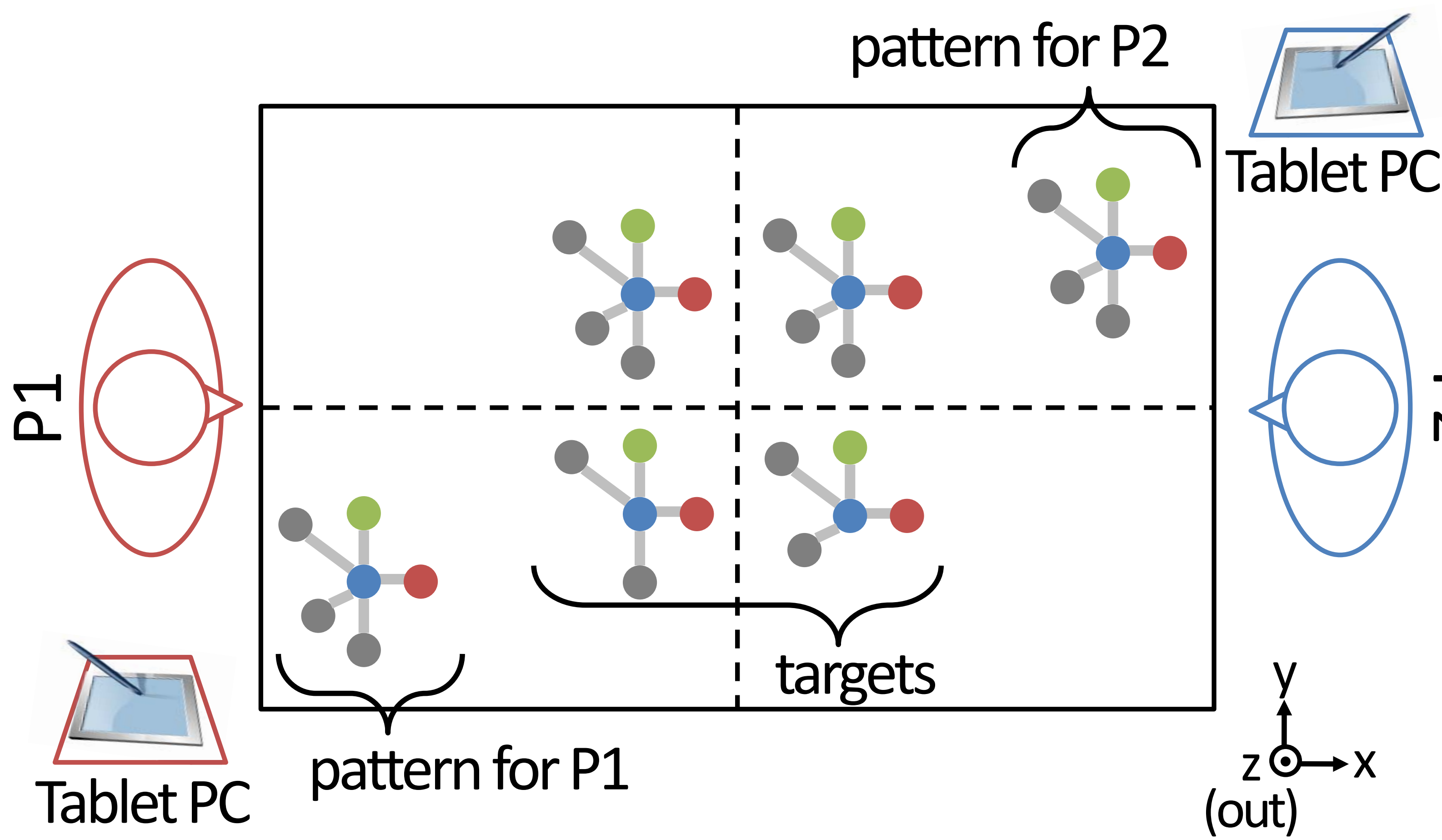


Figure 3.9: A diagram of the experimental setup for experiment 2.

3.5.2 Conditions, Design, Measures, and Hypotheses

Experiment 2, although run with a separate set of participants, shared the same conditions, design, measures and hypotheses as experiment 3. This commonality was intentional, so that a combined analysis could be performed using the between-participants factor of interactivity (present or absent). The details are discussed in the following section.

3.6 Experiment 3: Interactive Object Structure Matching

This experiment was designed to test the effect of interactivity on people's perception of 3D virtual objects in a tabletop display environment. Specifically, this experiment replicated as closely as possible the setup of experiment 2, with the exception that participants were able to interact with the pattern molecule.

3.6.1 Apparatus & Procedure

The apparatus and procedure for this experiment was the same as for experiment 2, with two exceptions: the participants could rotate their own pattern molecule by touching and dragging it with their finger and participants provided their answers via direct touch on the table instead of through the Tablet PC. The direction that the participants dragged their fingers determined the axis of rotation; the axis of rotation was always parallel to the plane of the table, orthogonal to their direction of motion, and coincident with the centre point of the central grey node. This interaction technique is the same as the pitch and roll rotation from the two-touch (section 5.2.4; second finger) and three-touch (section 5.2.5; third finger) techniques described in chapter 5. In this task, we were interested in how interacting through direct touch would interact with issues of discrepancy, motion parallax, and projection geometry.

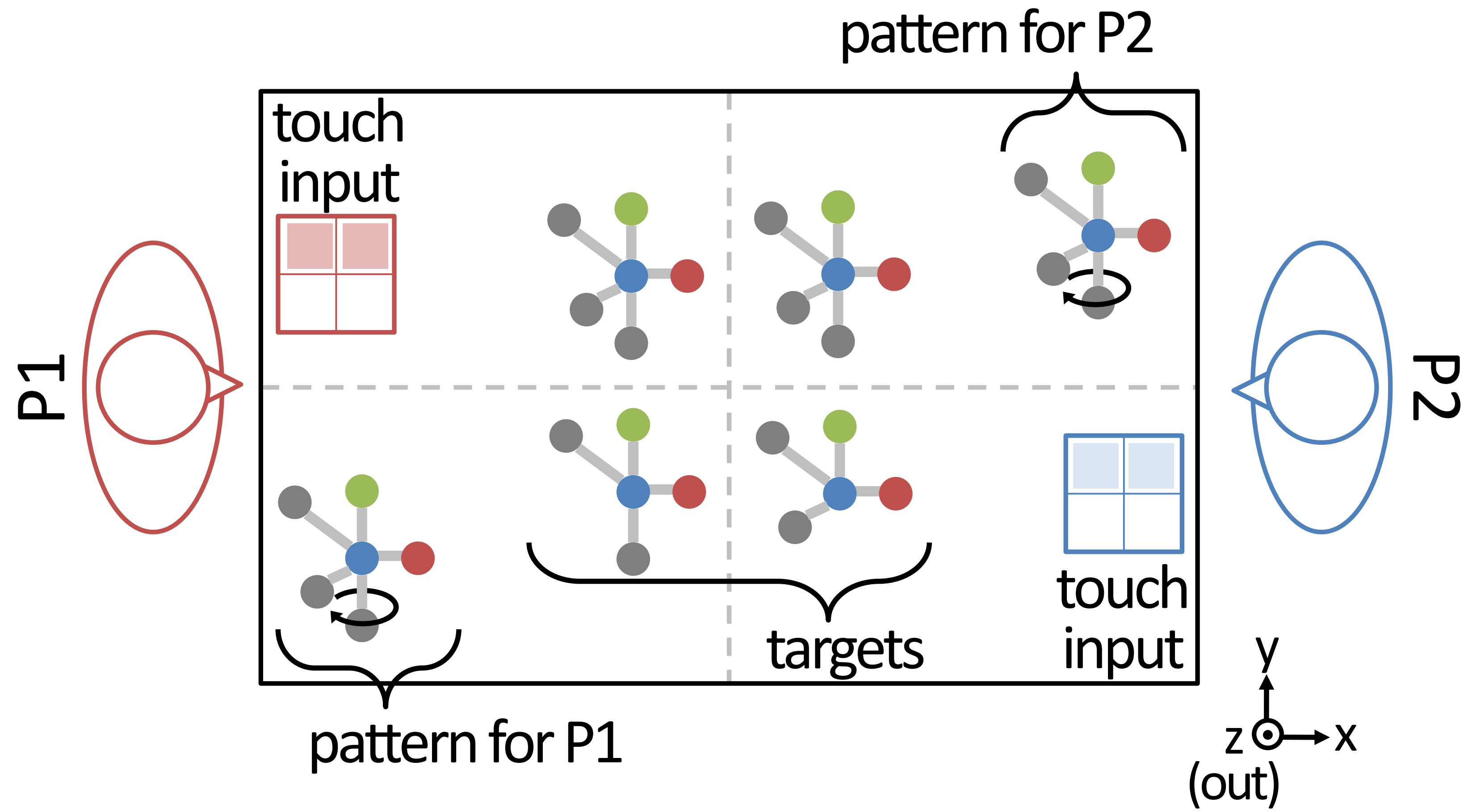


Figure 3.10: A diagram of the experimental setup for experiment 3.

3.6.2 Conditions and Design

In each of the ten condition-projection combinations, participants performed all combinations of the remaining factors in random order. Participants carried out one practice trial and two testing trials for each combination, for a total of 30 trials.

The study therefore used a 5 discrepancy-parallax condition (within) × 2 projection geometry (within) × interactivity (between) fully-crossed mixed design. The data were aggregated across both experiments, in order to obtain the between-participants interactivity factor. The 10 condition-projection pairs were counterbalanced between participants using a random Latin Square.

3.6.3 Measures

Error and TCT were both measured. The error was defined as:

> E_g: the number of incorrectly matched patterns. This measure included both false positives and false negatives and was in the range 0–4.

3.6.4 Hypotheses

The error measure in these experiments could not be separated into that along the x-axis and y-axis of the table, and so the expected perceptual error was expected to encapsulate both. That is, the number of incorrect guesses was expected to increase as perceptual error increased. Thus the primary hypotheses applied to the E_g error measure.

The inclusion of interactivity in experiment 3 allowed us to test the following additional hypothesis:

> **Hypothesis 3.6**: The effects of error will be greater in experiment 2 than in experiment 3, due to the ability to interact with the pattern molecule.

		Experiments 2 and 3					
		E_g (α = .0042)		E_g (α = .0021)			
		Both		Non-Interactive		Interactive	
		Parallel	Perspective	Parallel	Perspective	Parallel	Perspective
hyp. 3.1	NA to MA	p = **.001**	p = .393	p = **.001**	p = .242	p = .531	p = .677
	MA to LA	p = .006	p = .736	p = .041	p = .551	p = .074	p = .522
hyp. 3.2	NP to NA	p = .154	p = .724	p = .006	p = 1.000	p = .127	p = .417
	LA to LP	p = .196	p = .671	p = .091	p = .965	p = .695	p = .749
hyp. 3.3	MA to NP	p = .034	p = .392	p = .358	p = .159	p = .036	p = .445
	MA to LP	p < **.001**	p = .805	p < **.001**	p = .693	p = .081	p = .356

Table 3.4: Planned comparisons for experiments 2 and 3. These pairwise comparisons were performed using the non-parametric Wilcoxon Signed-Ranks test.

3.6.5 Results for Experiments 2 and 3

We performed a 5 discrepancy-parallax condition × 2 projection geometry × 2 interactivity (interactive, non-interactive) repeated measures ANOVA, and the planned comparisons for pairwise differences. Because each measured data point had only five possible values (0–4), we performed non-parametric pairwise tests. Specifically, we used a Wilcoxon Signed-Ranks test to compare related pairs.

3.6.5.1 Error in Matches (E_g)

What was the effect of discrepancy on error? (hyp. 3.1) There was an interaction between the discrepancy-parallax condition and projection geometry ($F_{4,184} = 3.4$, $p = .01$). We performed our planned comparisons by combining data from experiments 2 and 3. However, these comparisons did not reveal any differences that might correspond to the discrepancy between COP and POV. They instead provide support for hyp. 3.3 (see below). Thus, hyp. 3.1 cannot be confirmed.

What were the effects of projection geometry and motion parallax on error? (hyp. 3.2) There was a main effect of projection geometry ($F_{1,46} = 7.3$, $p < .01$). Participants had less errors using parallel geometry ($M = 0.8$, $SE = 0.06$) than using perspective geometry (M

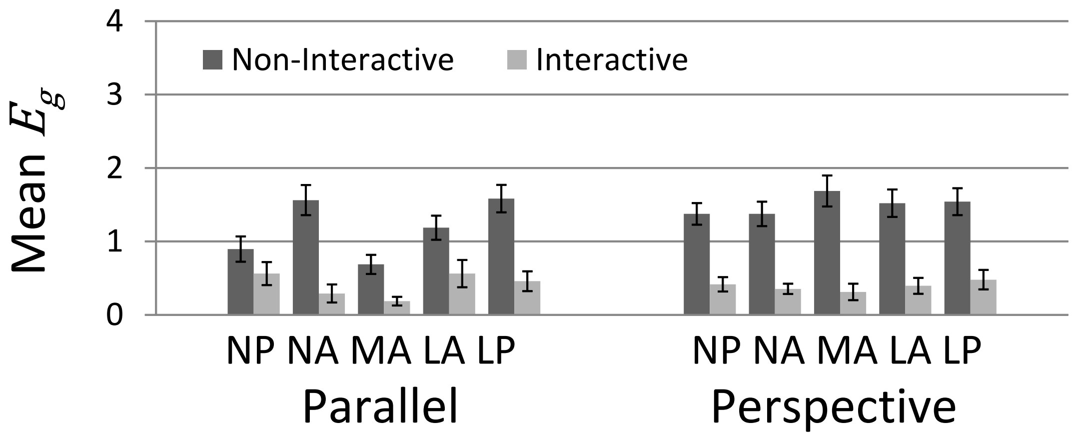


Figure 3.11: Mean number of incorrectly matched patterns (E_g) for the ten condition-projection pairs, separated by interactivity.

= 0.9, *SE* = 0.06). However, this effect depended on the interactivity condition (see below). As can be seen in table 3.4, our planned comparisons revealed no significant differences between the conditions with parallax (NP and LP) and those without (NA and LA). Therefore, hyp. 3.2 cannot be confirmed.

Is it a special case to have CoP directly above the table? (hyp. 3.3) Our planned comparisons revealed only two significant differences and only when using parallel projection geometry: the MA condition ($M = 0.4$, $SE = 0.07$) had significantly less errors than both the NA condition ($M = 0.9$, $SE = 0.12$) and the LP condition ($M = 1.0$, $SE = 0.12$). This result provides further support for the special case of having the COP directly above the table. Thus, there is partial support for hyp. 3.3.

Does the presence of interactivity with the pattern reduce error? (hyp. 3.6) There was a main effect of interactivity ($F_{1,46} = 81.5$, $p < .001$). Participants in the non-interactive condition ($M = 1.3$, $SE = 0.07$) had more incorrectly matched patterns than those in the interactive condition ($M = 0.40$, $SE = 0.07$). There was also an interaction between interactivity and projection geometry ($F_{1,46} = 9.5$, $p < .01$). The participants in the non-interactive condition had less errors when using the parallel geometry ($M = 1.2$, $SE = 0.08$) than when using perspective geometry (M = 1.5, SE = 0.08, p = .0025), but this difference was not significant for the participants in the interactive condition (p = 1.000). There was also a three-way interaction between interactivity, discrepancy-parallax condition, and projection geometry ($F_{4,184}$ = 3.2, p = .02). We performed the same planned comparisons for each projection geometry for both the participants who had interactivity and those who did not (table 3.4 and figure 3.11). These comparisons revealed that the pairwise differences for the two-way interaction between the discrepancy-parallax condition and projection geometry were only significant for participants without interactivity. There were no pairwise differences in either the parallel or perspective geometries for participants with interactivity. Thus hyp. 3.6 is confirmed.

3.6.5.2 Task Completion Time

We performed the same ANOVA on TCT using a log transform. There were main effects of both projection geometry ($F_{1,46}$ = 28.1, p < .001) and interactivity ($F_{1,46}$ = 5.6, p = .02), as well as an interaction between the discrepancy-parallax condition and projection geometry ($F_{4,184}$ = 3.1, p = .02). Participants were faster when using parallel projection geometry (M = 32.7 s, SE = 1.1 s) than when using perspective geometry (M = 45.8 s, SE = 1.0 s) and were faster without interactivity (M = 35.4 s, SE = 1.1 s) than with interactivity (M = 42.4 s, SE = 1.1 s).

The planned comparisons revealed that, using parallel projection geometry, the MA condition (M = 24.8 s, SE = 1.1 s) was faster than both the NA (M = 33.9 s, SE = 1.1 s) condition and the LP condition (M = 36.1 s, SE = 1.1 s). Thus there is partial support for hyp. 3.4, and further support for hyp. 3.3.

	Parallel	Perspective
NA to MA	p = **.001**	p = .393
MA to LA	p = .006	p = .736
NP to NA	p = .154	p = .724
LA to LP	p = .196	p = .671
MA to NP	p = .034	p = .392
MA to LP	p < **.001**	p = .805

Table 3.5: Planned comparisons for TCT for experiments 2 and 3. Pairwise tests were performed on the combined means for both tasks ($\alpha = .0042$).

3.7 Experiment 4: Estimating Orientation

The final experiment was designed to test people's perception of the orientation of 3D virtual objects with respect to their physical surroundings. That is, when an object is rendered so as to point 'out' of the display, can people perceive this intention and will the discrepancy between the COP and POV result in perceptual errors?

3.7.1 Apparatus & Procedure

For experiment 4, Vicon [2006] markers were placed at the end of a string which was attached to the tabletop corner. Participants manipulated these strings with an attached wand in order to record answers about the angle of the target. Details of participant input entry are described in detail in the next section.

3.7.1.1 Task Description

The experimental task asked participants to determine the orientation of target objects. For each trial, two 3D target objects were displayed on the screen (one per participant, always displayed on the half of the display at the participant's right). Targets were either in the 'near' half or the 'far' half of the table. Each object was a composite of a long thin cylinder inside a shorter thicker cylinder, each with the same axis. To provide their answer for each task, participants moved the wand until the string, stretched tight, was oriented at the same angle

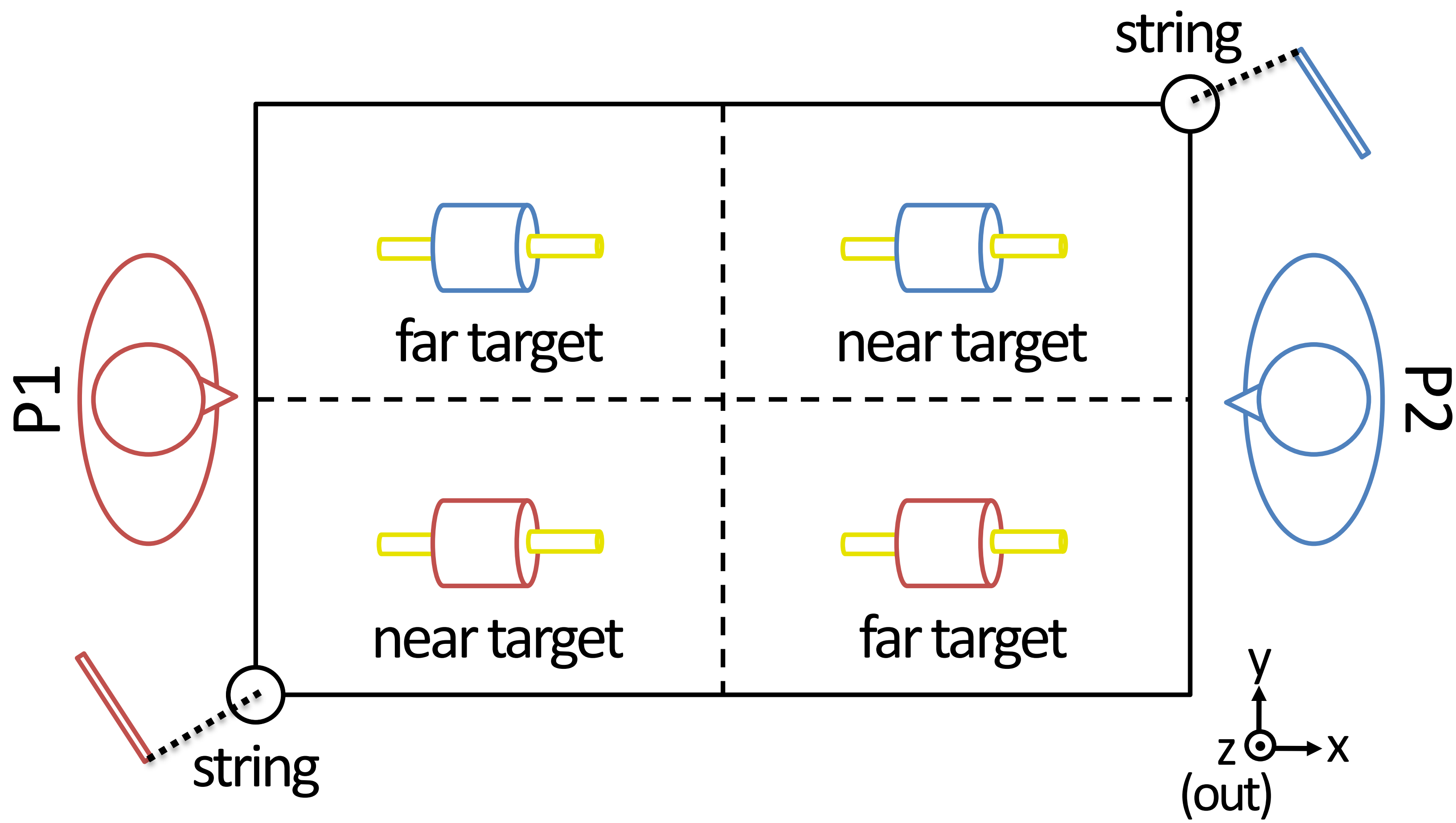


Figure 3.12: A diagram of the experimental setup for experiment 4.

as the main axis of the target. Once both participants had indicated the angle and pressed a 'done' button on the table surface, the next trial would begin. To remind the participants about the projection geometry being used, two groups of four cube frames were shown in the empty quadrants. These cubes were rendered in the same 3D model as the target objects in all conditions.

3.7.2 Conditions and Design

In this task, targets could appear either in the 'near' or 'far' halves of the table, as shown in figure 3.12. Targets were also shown in three different angular orientations, and in two locations. As shown in figure 3.14, targets could be at either 0° (laying flat on the table and pointing towards the end of the table where the participant was located), 60° (pointing upwards towards the end of the table), or 90° (pointing straight up from the table). Targets never leaned to the left or right; that is, they always stayed coplanar with the longitudinal vertical plane.

In each of the ten condition-projection combinations, participants performed all combinations of the remaining factors in random order. Participants carried out six practice trials (one for each combination of location and angle), and twelve testing trials (two repetitions of each combination) for each condition-projection combination, for a total of 180 trials.

The study therefore used a 5 discrepancy-parallax condition × 2 projection geometry × 2 location × 3 angle fully-crossed within-participants design. The 10 condition-projection pairs were counterbalanced between participants using a random Latin Square.

3.7.3 Measures

Error in angle and TCT were both measured. The error in angle had two parts (figure 3.14, right):

E_w: the error within the longitudinal vertical plane (e. g., if the target pointed to-

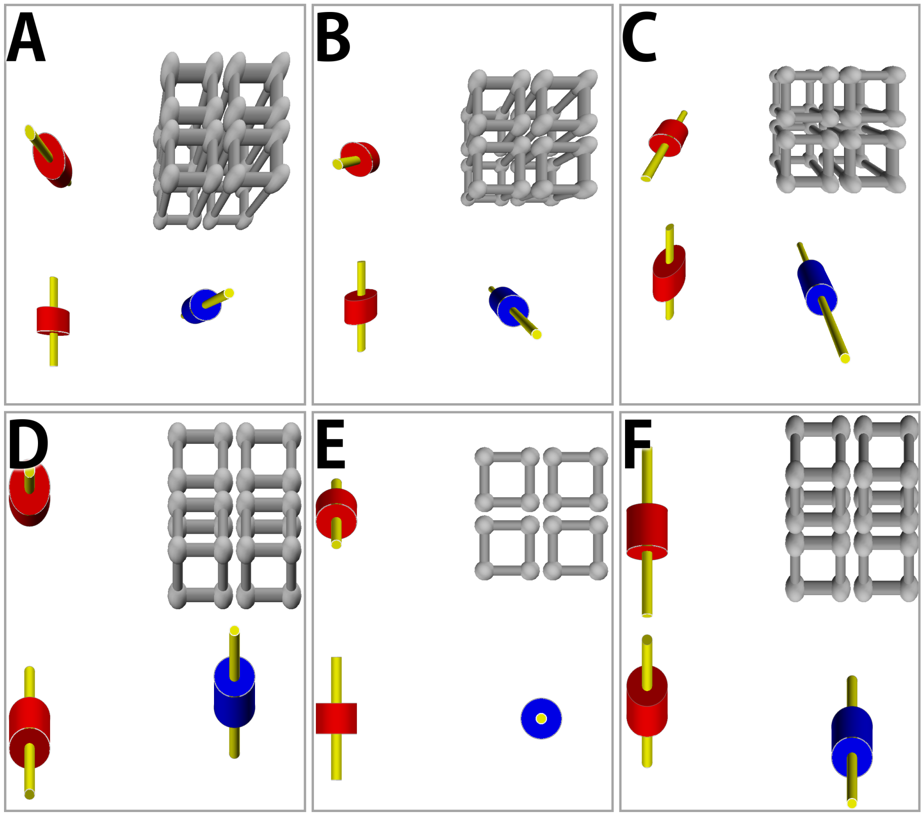


Figure 3.13: The three angles and visual hints rendered to the display in each discrepancy condition (matching the order of figure 3.5).

wards the end of the table with an angle of 60°, and the participant held the wand such that the string had an angle of 65° in the longitudinal vertical plane, the error was 5°).

E_r: the error away from the longitudinal vertical plane (no targets leaned left or right in this way, but we measured this error since some projections can make targets appear to lean).

3.7.4 Hypotheses

This experiment was designed to isolate errors along the longitudinal vertical plane due to the change in discrepancy. Thus the primary hypotheses apply to the E_w measure. We also arrived at some secondary hypotheses:

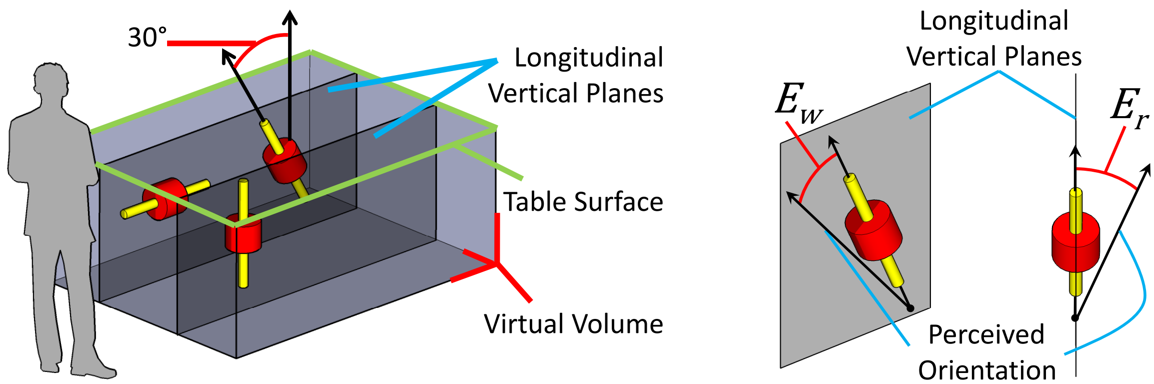


Figure 3.14: (left) The experimental setup and (right) the two types of error.

Hypothesis 3.7: The type of projection geometry will affect E_r (i. e., errors outside the plane in which the angle is varied).

Hypothesis 3.8: E_w will be least when the angle of the object is horizontal (0°) and most when the object is vertical (90°). This hypothesis will corroborate the differential rotation effect [Goldstein, 1987] (see section 2.2.3 for an explanation of this effect).

3.7.5 Results

We performed a 5 discrepancy-parallax condition × 2 projection geometry × 2 location × 3 angle repeated measures ANOVA, and the planned comparisons for pairwise differences.

3.7.5.1 Error within the Longitudinal Plane (E_w)

What was the effect of discrepancy on error? (hyp. 3.1) There was a significant main effect of discrepancy-parallax condition on E_w ($F_{4,44} = 39.5$, $p < .001$). As can be seen in figure 3.15, error increased overall as the discrepancy between COP and POV increased, confirming hyp. 3.1.

What were the effects of projection geometry and motion parallax on error? (hyp. 3.2) There was no main effect of projection geometry on E_w ($F_{1,11} = 2.6$, $p = .14$), with mean

		Experiment 4				
		E_w (α = .0042)		E_r (α = .0042)		TCT (α = .0083)
		Parallel	Perspective	Parallel	Perspective	Both
hyp. 3.1	NA to MA	p = .591	p < **.001**	p = **.002**	p = .017	p = .110
	MA to LA	p < **.001**	p < **.001**	p = 1.00	p = .050	p = .125
hyp. 3.2	NP to NA	p = .246	p = .797	p = 1.00	p = .453	p = .057
	LA to LP	p = .796	p = .617	p = .152	p = .634	p = .017
hyp. 3.3	MA to NP	p = .950	p < **.001**	p = .028	p = **.001**	p = .016
	MA to LP	p < **.001**	p < **.001**	p = 1.00	p = .037	p < **.001**

Table 3.6: Planned comparisons for experiment 4.

error for perspective conditions ($M = 37.1°$, $SD = 2.0°$) only slightly higher than for parallel conditions ($M = 33.4°$, $SD = 2.9°$). There was also no significant difference for the presence of motion parallax (table 3.6 and figure 3.15). Therefore, hyp. 3.2 is not confirmed.

Is it a special case to have CoP directly above the table? (hyp. 3.3) There was a significant interaction between discrepancy-parallax condition and projection geometry for E_w ($F_{4,44} = 6.8$, $p < .001$). Our planned comparisons indicate that the interaction was due to the special case of medium discrepancy when using a parallel projection. As can be seen in figure 3.15, error in MA with perspective geometry was substantially higher than error in MA with parallel geometry. Furthermore, in the perspective case, the medium discrepancy condition (MA) followed the expected trend of being in between the no discrepancy conditions (NP and NA) and the large discrepancy conditions (LA and LP), whereas in the parallel case, the MA condition was still less than the LA and LP conditions, but had as little error as the NP and NA conditions (table 3.6 and figure 3.15). Therefore, there is limited support for hyp. 3.3.

Did different target angles lead to different error in experiment 4? (hyp. 3.8) There was a significant main effect of target angle on error ($F_{2,22} = 7.1$, $p < .01$). Post-hoc analysis showed that errors for all three angles were significantly different ($p < .05$) in the order we predicted with hyp. 3.8 ($M_{0°} = 26.3°$, $M_{60°} = 36.7°$, $M_{90°} = 42.7°$).

Target angle was also involved in several interactions, including significant two-way in-

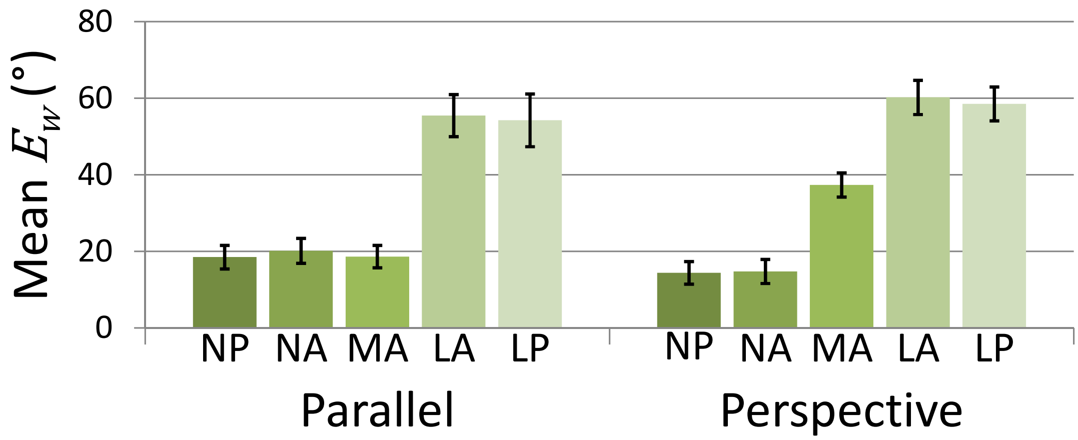


Figure 3.15: Mean error in angle (E_w) for the ten condition-projection pairs.

teractions between angle × condition ($F_{8,88} = 7.9$, $p < .001$) and angle × projection ($F_{2,22} = 9.6$, $p = .001$), a significant three-way interaction between angle × condition × projection ($F_{8,88} = 8.6$, $p < .001$) and a significant four-way interaction ($F_{8,88} = 2.4$, $p = .02$). Figure 3.16 shows that these interactions are largely explained by the special case of the medium discrepancy condition (hyp. 3.3). Specifically, for horizontal (0°) and vertical (90°) angles using a parallel projection, participants were able to determine the orientation with a high degree of accuracy, going against the trend predicted in hyp. 3.8, which held or was not significant for all other conditions.

Other analyses All other main effects and interactions were not significant. Specifically, location did not have a significant effect on E_w, nor did it interact with other factors.

3.7.5.2 Left-to-Right Error (E_r)

There were significant main effects of condition ($F_{4,44} = 8.3$, $p < .001$) and angle ($F_{2,22} = 29.4$, $p < .001$), a significant condition × projection interaction ($F_{4,44} = 24.7$, $p < .001$), a significant condition × angle interaction ($F_{8,88} = 9.4$, $p < .001$), and a significant three-way condition × projection × angle interaction ($F_{8,88} = 13.3$, $p < .001$). No other main effects or interactions were significant.

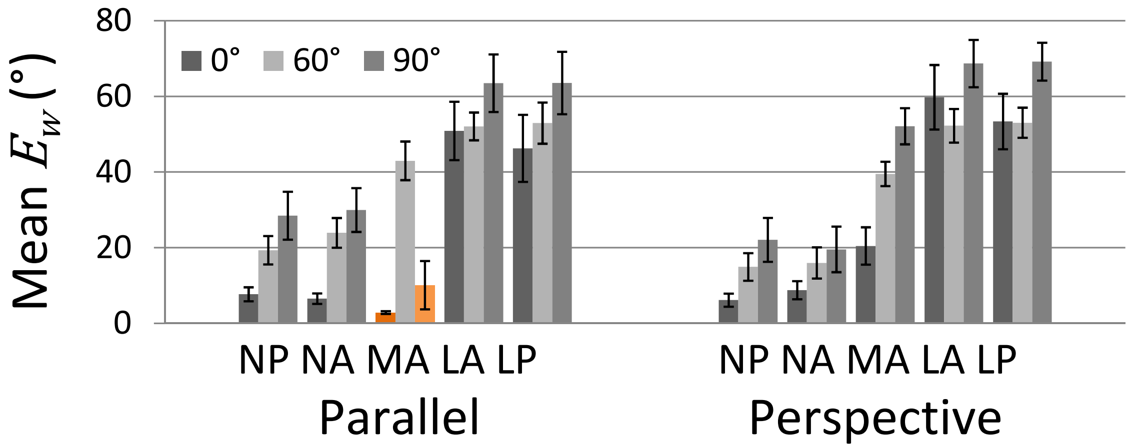


Figure 3.16: Mean $E_w$separated by angle of the object. For the MA-parallel condition, 0° and 90° were exceptions.

Our planned comparisons showed only two significant differences. One in the parallel projection showed significantly more error in the NA condition than the MA condition ($p < .01$) and one in the perspective projection showed significantly more error in the MA condition than the NP condition ($p = .001$). These two cases are best explained through the three-way interaction (figure 3.17). For the horizontal angle (0°), the left-right error was small for all conditions. For the other two angles, participants had more errors in the parallel projection when there was no discrepancy (NA and NP), as well as in the perspective projection with medium discrepancy (MA).

3.7.5.3 Task Completion Time

We performed the same factorial ANOVA on TCT using a log transform. There was a significant main effect of condition ($F_{4,40} = 4.5$, $p < .01$). Using the same planned comparisons, aggregated across both projection geometries, revealed that the MA condition was significantly faster than the LP condition ($p = .001$). Although the other pairwise differences were not significant, there is a clear trend that both conditions involving motion parallax ($M_{NP} = 6.7$ s, $M_{LP} = 7.5$ s) were slower than the three without ($M_{NA} = 5.6$ s, $M_{MA} = 5.4$ s, $M_{LA} = 6.1$ s).

There was also a significant interaction between angle and location ($F_{2,20} = 9.9$, $p < .01$),

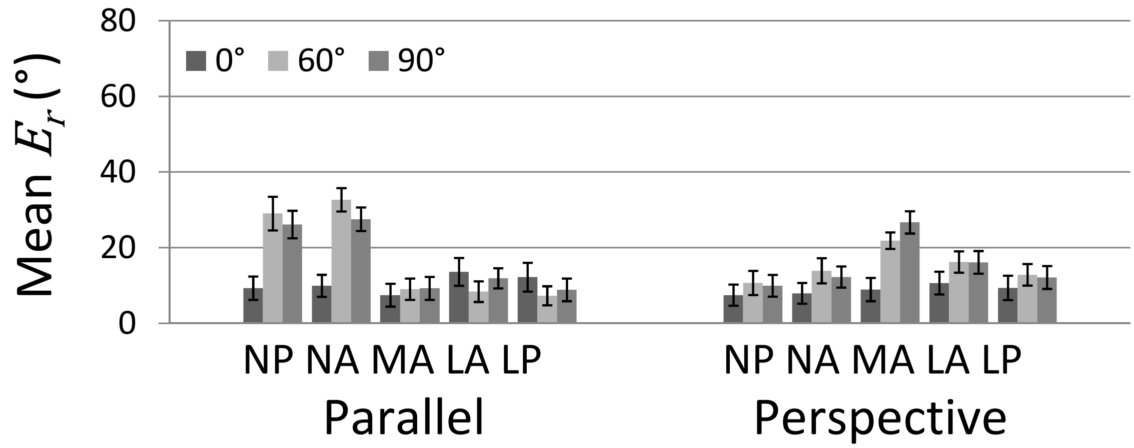


Figure 3.17: Remaining error (E_r) for each condition, projection, and angle.

but we did not investigate further, as it did not involve the discrepancy-parallax condition, nor projection geometry. No other main effects or interactions were significant for TCT.

3.8 Discussion

Table 3.7 shows a summary of the results of the analysis with respect to each hypothesis and each task. We summarize the main findings of our study as follows:

- As the discrepancy between COP and POV increases, so does the error in people's ability to judge the orientation of 3D objects (hyp. 3.1, experiment 4).
- The error in people's ability to judge patterns and angles did not increase as the discrepancy between COP and POV increased (hyp. 3.1, experiments 1 to 3).
- Motion parallax did not improve people's ability to judge angle, pattern, and orientation, nor did it make the error worse when the discrepancy was large (hyp. 3.2, experiments 1 to 4).
- For parallel projections, the case where the COP is directly above the table is special and may reduce the problem of discrepancy for judgement of orientation. For judgement

	Hyp. 3.1	Hyp. 3.2	Hyp. 3.3	Hyp. 3.4
Experiment 1	N	N	C	P
Experiments 2 and 3	N	N	P	P
Experiment 4	C	N	P	P

Table 3.7: Summary of common hypothesis confirmation (C = confirmed, N = not confirmed, P = partially confirmed).

of angle and pattern, this special case results in fewer errors than any other level of discrepancy (hyp. 3.3, experiments 1 to 4).

- People will take more time to acquire and process information when motion parallax is available (hyp. 3.4, experiments 1 to 4).
- Left to right error was not affected by the projection geometry (hyp. 3.5, experiment 1 & hyp. 3.7, experiment 4).
- The ability to interact with a 3D artifact reduces perceptual that may be caused by discrepancy between COP and POV when matching its structure to another artifact on the screen (hyp. 3.6, experiments 2 and 3).
- The differential rotation effect (DRE) effect was corroborated for all conditions, with the exception of the case where the COP is directly above the table (MA). In this case, the DRE does not predict the pattern of perceptual error (hyp. 3.8, experiment 4).

Our main result was that error increased with increasing discrepancy between COP and POV when people made judgements about the orientation of 3D objects. This relationship follows our expectations based on the idea of the differential rotation effect (DRE). This effect (as described in the "Uncle Sam" example) causes objects oriented perpendicular to the picture plane to seem as though they point toward the observer. This phenomenon can explain the increasing error that we saw for the three target angles: objects that were at 90° (fully perpendicular) were difficult for participants to judge, whereas objects at 0° were interpreted with high accuracy. These difficulties arising from DRE suggest that this effect must be considered when designing tabletop systems.

However, the effect of discrepancy was not significant for judgments of the angle internal to the 3D object, nor for the ability to match two separate 3D virtual objects. While the 3D model's relationship to the physical world appears to suffer from off-axis viewing (experiment 4), the discrepancy between COP and POV seems to have little effect on these other judgements (experiments 1 to 3). However, errors in these remaining tasks was still high (experiment 1: 38.1 %; experiments 2 and 3: 0.9 incorrectly matched patterns). These errors may be in part due to people's expectations when looking at a 3D scene on a 2D pictorial plane.

Our study showed that a neutral COP and a parallel projection geometry is a special case. One possible explanation is that the frame of the table may suggest that objects are oriented relative to that frame, and not to the POV. However, this explanation does not account for the difference between perspective and parallel projections. A perhaps more likely explanation is that objects rendered in this geometry lose their 3D appearance and become perceivable as 2D within the plane of the table. That is, a cylinder pointing out of the display becomes a circle, and when pointing horizontally becomes a rectangle. This explanation is also consistent with the differences we found for each object angle (figure 3.16).

Another special case identified by this study was that of the presence of interactivity. In experiment 3, participants had significantly less errors than those in experiment 2. This result suggests that interactivity provides a strong cue about the overall pattern or shape of a 3D object projected onto a 2D display. Thus, providing the ability to manipulate the 3D objects on the screen may be another means through which a designer can reduce errors in perception that may be caused by a discrepancy in POV and COP.

3.8.1 Study Limitations

It is well-known that any single experiment will have to balance three desirable properties: precision, generalizability, and realism [McGrath, 1984]. In designing this series of experiments, we opted for high precision and generalizability, perhaps at the cost of some realism. For instance, the tasks performed by the participants were in a lab setting, repeated many times, and were abstract in nature. These choices made it possible to precisely determine the measured error and to be able to generalize the findings to a variety of situations that involve estimating angle, pattern matching, and estimating orientation. However, in a more realistic situation, other factors may influence a person's perception. For instance, experiments 2 and 3 provide evidence of a positive effect, and further studies could tease out these other factors.

3.8.2 Discussion Summary

There are several practical recommendations that have emerged from our study. First and foremost, we have shown that decisions made about the projection geometry are important and that the designer cannot blindly use the 'defaults' from 3D graphics. In particular, attaching the COP to one person's POV can introduce errors in perception of up to 49.6 % in angle (experiment 1), 1.0 of 4 incorrectly matched patterns (experiments 2 and 3), or 60° in orientation (experiment 4) for another person at the table, and using a COP above the table together with a perspective projection can introduce errors of up to 28.6 % in angle (experiment 1), 1.0 of 4 incorrectly matched patterns (experiments 2 and 3), or over 40° in orientation (experiment 4). This series of experiments provides some evidence that a parallel projection with a COP directly above the table may alleviate some of the problems introduced by this discrepancy (down to 10.5 %, 0.4 of 4, or 20°, respectively) and that the use of interactivity may reduce errors in matching of patterns (down to 0.4 of 4). Note also that the use of perspective versus a parallel geometry is not necessarily a binary choice; a COP that is very far above the table may reap some of the benefits of a parallel projection, while maintaining some of the perspective depth cues.

3.9 Multiple Viewpoint Solutions

While this study of human perception of 3D objects at a table offers the promising solution of using a parallel projection with a COP directly above the table, there are many cases where the perspective depth cue provides useful information. For example, an aerial view of map data can provide useful additional information about the appearance of buildings and landscapes, even though satellite imagery may provide more accurate perceptual information for multiple viewers at tabletop display. In this section, I present a simplified model to compensate for off-axis viewing of 3D objects and generalize this technique for use by many people around the display. After describing both fixed and customizable alternatives for multiple viewers, we discuss the implications of each for collaboration.

3.9.1 Correcting Off-Axis Distortion

We present a general method to compensate for distortion caused by off-axis viewing of 3D objects projected onto a 2D surface. We first describe how to adjust this distortion for a single person's perspective and then describe several techniques for extending this method for multiple viewing angles. Note that our model for correcting distortion for a single user can be accomplished using existing techniques [Agrawala et al., 2000; McKenna, 1992]. We present a simpler model with minimal changes to the current 3D projection methods, which allows us to more easily extend the technique to multiple users and to explore a large variety of alternative projections.

3.9.1.1 Correcting for One User

The standard method for projecting 3D graphics onto a 2D surface assumes that the viewpoint is at the centre and directly in front of the display. A common method to achieve this projection is by transforming points in a 3D model to a canonical viewing volume [Hill, 2001]. For a perspective projection, the eye is assumed to be at the origin $(0, 0, 0)$ and the near plane is assumed to have its centre along the z-axis $(0, 0, -N)$ and to be perpendicular to that axis, and so has the equation $z = -N$, where N is the distance from the eye to the near plane. Thus, all points in the model are projected by intersecting the line from the eye to the point with this near plane. Such a line would have parametric equations:

$$x = P_x t, \quad y = P_y t, \quad z = P_z t$$

Thus, this intersection would result in a projected point (x', y') as follows:

$$x' = N\frac{P_x}{-P_z}, \quad y' = N\frac{P_y}{-P_z}, \quad z' = \frac{NP_z}{-P_z} = -N$$

The corresponding transformation matrix is:

$$\begin{bmatrix} N & 0 & 0 & 0 \\ 0 & N & 0 & 0 \\ 0 & 0 & N & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

This matrix can then be included in the standard graphics pipeline [Hill, 2001]. When viewing a display from off-axis (i. e., not along the z-axis), the assumption that the viewer is directly in front of the display is invalid. The degree to which the viewer is off the centre axis is particularly high when viewing a large display from one side. We introduce a method for rendering 3D objects that, instead of using a perpendicular near plane, uses an arbitrary near plane with the equation:

$$Ax + By + Cz = D$$

In order to preserve the property that the z-axis intersects the plane at a distance of N from the eye, we set $C = -1$ and $D = N$. Thus, A and B represent the slope of the plane in the x and y directions, respectively. Points in the model are again projected by intersecting the line from the eye to the point with this arbitrary near plane. This intersection results in the projected point (x', y', z'):

$$x' = N\frac{P_x}{AP_x + BP_y - P_z}, \quad y' = N\frac{P_y}{AP_x + BP_y - P_z} z' = \frac{aP_z + b}{AP_x + BP_y - P_z}$$

The transformation matrix is thus only slightly modified:

$$\begin{bmatrix} N & 0 & 0 & 0 \\ 0 & N & 0 & 0 \\ 0 & 0 & N & 0 \\ A & B & -1 & 0 \end{bmatrix}$$

With this simple modification to the projection matrix, 3D objects can be rendered to compensate for off-axis viewing. This method introduces no added complexity and does not interfere with the response time of interaction.

3.9.1.2 Correcting for Multiple Users

The above method allows 3D objects to be rendered correctly for a single off-axis viewpoint, but tabletop displays lend themselves to many people gathering around them, each with their own viewing angle. Thus, a single viewpoint rendering may not be sufficient. Objects can each be rendered with a different perspective transformation, depending on the posi-

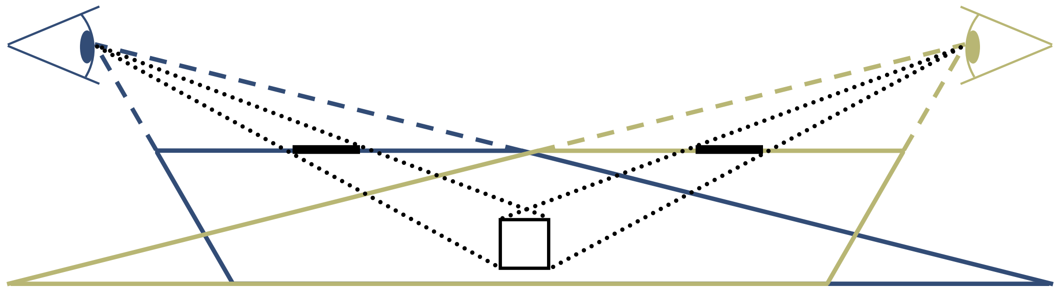


Figure 3.18: Partitions aligned with touching near planes result in intersecting view volumes.

tion of the object. By altering the perspective matrix in proportion to the object's position, this method essentially provides an arbitrarily-shaped near surface. We present alternative methods of projecting 3D objects for multiple users at a table.

Partitioning Viewpoints One method of altering the perspective for many users is to provide several dedicated areas that each optimize the viewing angle for a particular portion of the display. For example, an obvious partition for a rectangular table is to divide the table in four parts and optimize the view for the closest side to each part. Essentially, each partition provides a different "window" through which to look at the underlying 3D model. Thus, the eye position of each partition can be chosen in three different ways.

The partitions can be aligned so that the near planes of each provide the boundaries (figures 3.18 and 3.23c). This method results in view volumes that intersect one another. When an object is within the bounds of an intersection, the objects can either be displayed at all viewpoints, or some decision must be made as to which viewpoint to use.

Alternatively, the partitions can be aligned so that the separate view volumes do not intersect (figures 3.19 and 3.23d). This model has the advantage that objects cannot be within two views at the same time. However, when crossing the boundary of two views, the change in projected position can be both large and discontinuous, which may make interaction confusing. Also, this method creates a volume in the model between the partitions where

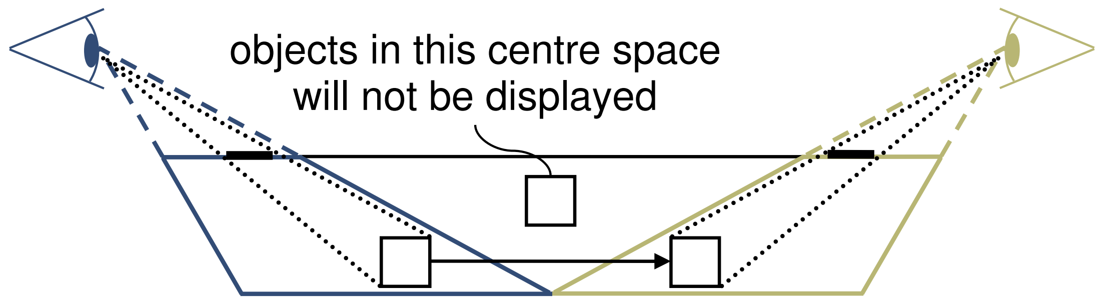


Figure 3.19: Partitions aligned with touching far planes result in disjoint view volumes. Objects moving between between view volumes may appear to jump.

an object can exist without being visible.

A third method of providing correct perspective for multiple partitions is to keep the eye position above the centre of the table, but to slope the near plane differently in each partition (figures 3.20 and 3.23e). Essentially this provides a view in each partition that is set for the correct viewing angle, while avoiding intersection of view volumes and discontinuities when objects move across boundaries. This method still compensates for off-axis distortion, but provides smoother interaction. However, because the near plane is distorted, applications requiring realism may prefer a partition with a "correct" eye position for each side.

A side effect of not moving the eye position as the slope changes is that objects become further from the near plane as they are moved closer to the table's edge. This effect can be corrected by adjusting the near plane distance so that the centre of the object is always projected to the original near plane. This new near plane distance, N', can be calculated by substituting the centre's projected point in standard perspective $(x', y', -N)$ into the equation for the plane:

$$N' = Ax' + By' + N$$

All three partitioning methods can be achieved by setting either A or B to the desired slope of the near plane.

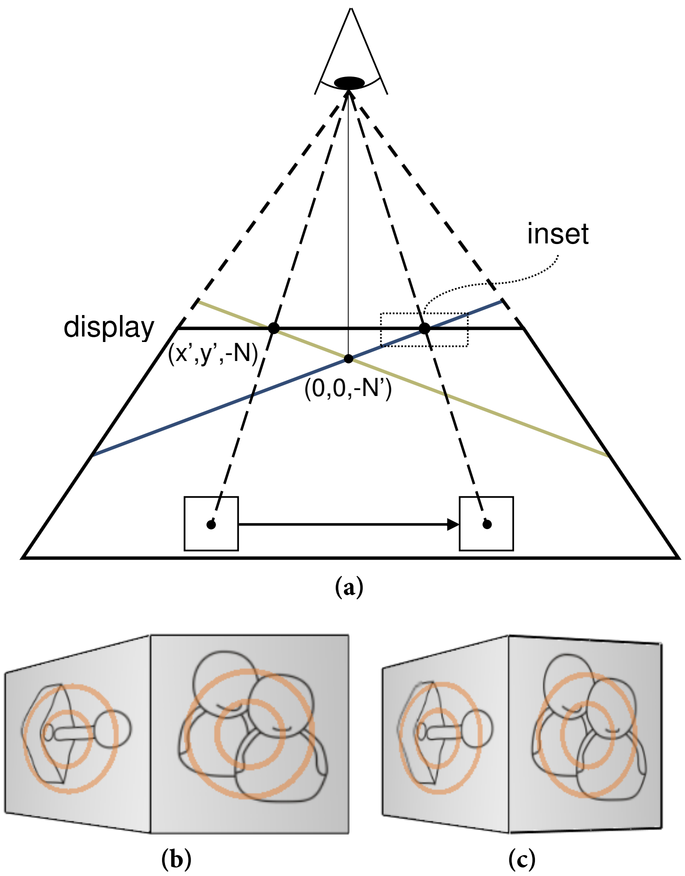


Figure 3.20: (a) The slope of the near plane can be adjusted to be appropriate for the adjacent table edge. Objects moving across the partition boundary do not jump and are not replicated. Insets show how the object would appear using (b) standard perspective and (c) when projected to the sloped plane.

Continuous Viewpoints Instead of dividing the table into discrete parts, objects can be projected so that the viewing direction changes continuously. That is, objects can be projected so that the viewpoint is determined by the rotation (θ) about the z-axis. This method prevents objects from having to cross partition boundaries and suddenly switch viewpoints; instead, the transition is smooth across the entire display. The transition can also be made smooth at the centre of the display by adjusting the plane's slope (m) according to the distance (r) of the object from the centre of the display. This method essentially provides a hemi-spherical near plane (figures 3.21 and 3.23f) and can be achieved by setting $A = mr\cos\theta$

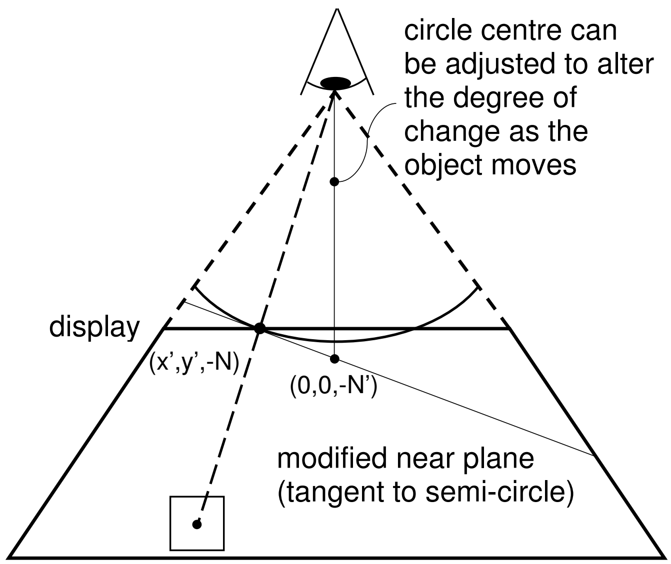


Figure 3.21: In continuous spherical perspective projection, as objects are moved toward the edges of the display, the near plane they are projected to increases in slope. As they move toward the centre, the near plane becomes a standard perspective. Objects will appear more correct the closer they are to the viewer's edge.

and $B = mr \cdot \sin\theta$.

Customizable Viewpoints We provide customizable views to allow users to manipulate the slope of the near plane and resize the area of influence of this near plane (figures 3.22 and 3.23b). This slope can be controlled with a virtual handle that can be adjusted to provide the appropriate view. Again, it is possible to implement customizable views so that the eye position moves as the slope changes. However, it can be advantageous to not move the eye location so that the objects do not change their projected positions as the handle is dragged.

3.9.1.3 Effect on Collaboration

Choices made between off-axis projections influence the environment and how people coordinate and communicate within that environment. We discuss some of the implications and limitations of the corrections we suggest for multiple viewers.

Figure 3.22: The left area is a custom view corrected for a viewpoint at the bottom-left and the right custom area is for a viewpoint at the top-right.

Viewpoint Relativity In all of the suggested multiple-viewer corrections, at any point in time, each object appears "correct" for only one viewpoint around the table. In order to make it appear correct for another, a person must either move the object along the surface or move a handle to adjust that viewpoint. Thus, our solution solves the multiple-viewer problem in a very different way than providing different views for each person [Agrawala et al., 1997; Kitamura et al., 2006]. The effect of our approach is that an object in "someone else's" space would still appear distorted. This increased distortion would be disadvantageous for perception of errors, as discussed in section 3.2. On the other hand, this distortion may provide an added hint to a collaborator not to expect the object to appear correct. For example, a person watching the actions of another across the table would see a highly distorted image, and therefore not expect to see the opposite side of a virtual cube, as they would with a real object.

Sensitivity to Frames If a 3D scene is projected inside a framed box, people tend to compensate for off-axis inconsistencies. This effect may be due to the tendency to achieve shape constancy under object rotations [Pizlo, 1994]. Since the tabletop display is itself a frame, this could lead to an expectation different from the correction that we provide. Because the frame is a fixed physical entity, viewing the frame from different viewpoints (i. e., by different people) cannot be corrected for. There is some evidence for the framing effect in our study (section 3.2), as people had less errors when using a POV above the table with a parallel projection. However, this framing did not seem to be as much of a factor for the perspective condition, which is used by all of the corrections provided in this section.

Framing can also be an issue in our customizable viewpoint correction, since we use a rectangular box (parallel to the edges of the display) to indicate the affected areas. It may be preferrable to use a frame that matches the correction to provide a better visual cue to the people at the display. Matching the frame to the correction may also improve a person's ability to parse objects in "someone else's" area, despite the increased distortion.

Fixed vs. Customizable Corrections Both partitioned and continuous views provide a fixed correction for the entire display. These fixed corrections can be chosen based on an expected scenario of use. With these solutions, users will not have to learn additional controls and may be mostly unaware that the distortion has been corrected. Thus, this correction may become invisible to the user.

It can also be beneficial to allow people to control what areas of the screen are best viewed from what side. By providing this freedom, natural communication gestures may become available. For example, with this solution, it is possible for people to set up personal areas within which artifacts look correct as they work independently, and then can share their work by adjusting the viewpoint to be correct for another person. In general, providing this ability supports many of the mechanics of collaboration [Pinelle et al., 2003], including gestural messages ("this is what mine looks like"), visual evidence ("this is how they were looking at it"), and obtaining and reserving resources ("I'll look at this area from my viewpoint"). However, providing this added freedom can also add cognitive load, in that portions of the display can be set for different viewpoints, and users are assigned the added task of adjusting to the correct viewpoint.

3.10 Chapter Summary

In this chapter, I presented an experimental evaluation of people's perception of 3D virtual objects presented on a large 2D tabletop display surface. This study provides evidence of perceptual difficulties that exist when perceiving this 3D information, as well as several phenomena that may facilitate the design of a tabletop application which minimizes these difficulties. Specifically, the designer can choose a COP directly above the table together with parallel projection geometry, or can choose to introduce the ability to manipulate the 3D virtual objects to improve perception of their structure.

Furthermore, I introduce several techniques for mitigating the perceptual error when these choices are not available to the designer. When an application may benefit from the presence of a perspective depth cue, these mitigating techniques can allow the designer to either attempt to reduce the perceptual errors for a specific usage scenario, or to provide people with control over the discrepancy between COP and POV.

Specifically, the contributions from this chapter are:

- A study providing evidence that, when projecting 3D onto a horizontal table using standard 3D graphics techniques, there is an established viewing location, and perception errors will increase as the viewer moves away from this location.
- This same study provides evidence that a parallel projection with a centre of projection (COP) directly above the table may reduce these perceptual errors.
- This same study also provides evidence that providing direct-touch interaction with the virtual artifacts being perceived will also reduce these perceptual errors.

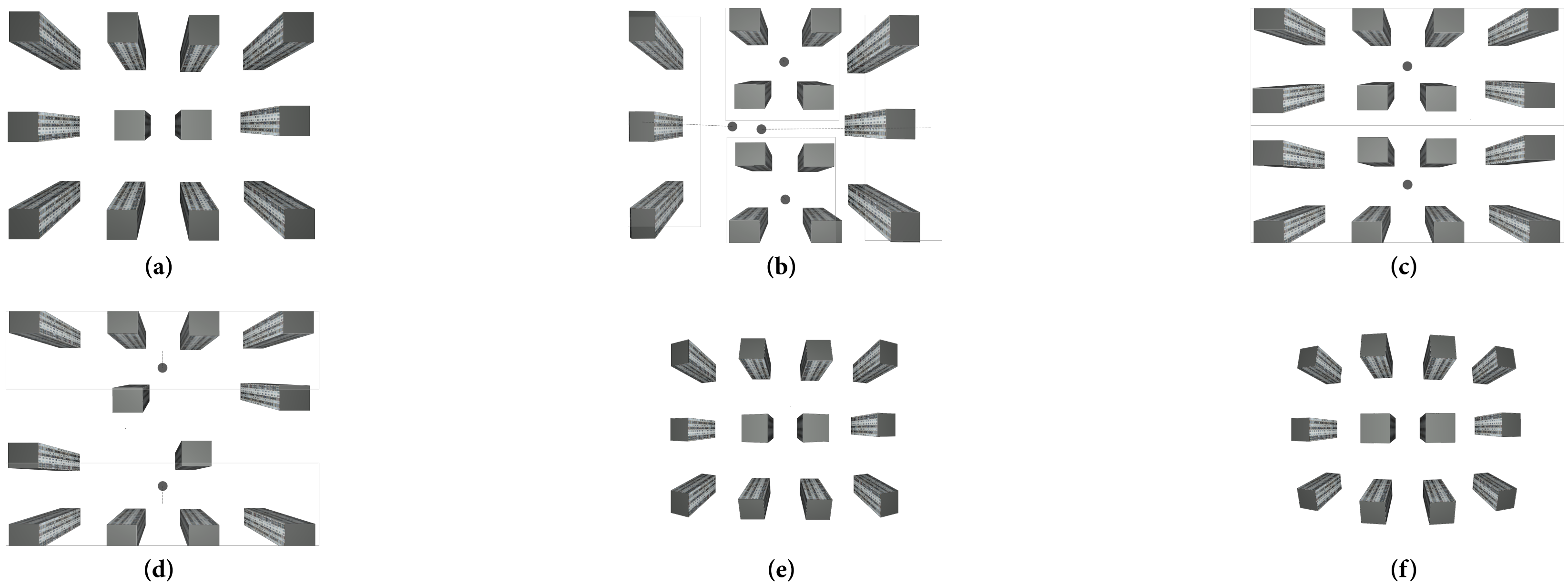


Figure 3.23: (a) shows a series of 12 buildings using a standard perspective projection, which works well for someone standing directly above the table, but no one else. (b) shows four areas, each with its own customized viewpoint. The top and bottom areas are best viewed from the top and bottom respectively. The left area is best viewed from near the centre of the display and the right area is best viewed from the left side. (c) shows a partitioning of these same building into two parts. The top partition is best viewed from the top side, and the bottom is best viewed from the bottom. Note that in this correction, the view volumes intersect and so the middle buildings are replicated in both views. (d) shows two partitions with the same optimal viewing angles, but with non-intersecting view volumes. (e) shows a partitioning into four parts (left is best viewed from left, top from top, etc.), but only the planes are sloped; the eye position is kept at the centre and above the display. (f) shows a continuously changing viewpoint as objects move across the display. Objects are projected so they are best viewed along the axis from the centre of the display to the object.

- The design and implementation of a set of non-standard 3D projections that can be used to mitigate the problem of multiple viewpoints and viewpoint discrepancies.

In the remainder of this dissertation, the designs presented make use of the results of the study presented in this chapter. The interfaces presented in the remaining chapters have the following properties:

- They are designed using a parallel projection geometry with the COP directly above the centre of the table. When the perspective depth cue is desirable, a COP is chosen that is very far above the centre of the table.
- They are designed so that the 3D virtual objects can be manipulated.

In the following chapter, I introduce several techniques for manipulating 3D virtual objects using a multitouch tabletop display and present a comparitive study to evaluate their performance.

DEFINING INPUT AND DEGREES OF FREEDOM 4

The focus of chapter 3 was on the perception of 3D information that is presented on a tabletop display. One of the results of this empirical work was that the ability to interact with a 3D virtual artifact may improve a person's perception of its structure. But, this realization leads to the question: how can a designer enable 3D interaction with these virtual artifacts? The introduction of multitouch technology to these large display environments indicates a promising direction. One of the significant factors of multitouch tabletops is that the input space—the surface one's fingers touch—is also the output space—the surface that displays the virtual artifacts. This combination of input space with output space makes it possible for people to directly touch or even feel like they can "grab" the virtual artifacts being displayed. In this chapter, I present a formalism that can help to describe some of the subtleties of multitouch interaction. This formalism builds upon the work on understanding manipulation of 3D virtual artifacts (see section 2.3.1). However, this work predates many of the recent advances in multitouch technology, and so the purpose of this revisit is to tease out some of the subtleties of these devices and to provide a formal way to describe the interaction techniques presented in chapter 5.

There has been a significant amount of HCI research that has focused on the classification of input devices based on the number and types of degrees of freedom (DOF) that they sense [e. g., Buxton, 1983; Mackinlay et al., 1990]. There has also been a lot of attention paid to the "transfer function" [Zhai, 1995] or "resolution function" [Mackinlay et al., 1990], both of which translate the sensed input into visual output. While it is possible to extend these taxonomies directly to classify multitouch input devices, there are some subtleties of multitouch classification that merit further discussion. This discussion is intended to help the reader to understand the thought process that went behind the creation of the interaction techniques described in chapter 5. It is the understanding of these subtleties that led to some of the insights in creating the one-, two-, and three-touch interaction techniques. Note that this chapter is not intended as a novel contribution to the field of HCI, as Mackinlay et al.'s [1990] taxonomy is likely sufficient to classify the variety of multitouch input devices. It is rather a demonstration of my thought process that has been useful to me as a designer when inventing the interaction techniques described in chapter 5.

The chapter is presented as follows: I first explain Mackinlay et al.'s [1990] taxonomy, along with a subtle modification, and provide two examples of how to apply this modified formalism: the mouse and keyboard. I then define the terms degrees of freedom (DOF) and magnitude of freedom (MOF) and use the same two examples to illustrate their meaning. This is followed by the application of this formalism to describe two types of multitouch devices: the DiamondTouch [Dietz and Leigh, 2001] and camera-based techniques, such as frustrated total internal reflection (FTIR) [Han, 2005], or diffuse illumination (DI). I also provide an alternate description of these devices as multipoint input using the same formalism. I end the chapter with a summary of the information this formalism provides, and describe how this exercise helps to inform the designs used in the following chapter.

4.1 Mapping Input to Output

In this section, I will describe Mackinlay et al.'s [1990] taxonomy of input devices, as well as a subtle modification. The purpose of this exercise is to later use this formalism to describe multitouch input, and then in chapter 5, to make use of this formal description to inform the design of multitouch interaction techniques. However, to understand the complexities of this taxonomy, I will illustrate its application using two much simpler examples: the mouse and keyboard. In the following sections, I refer to these examples to highlight some nuances of the term degrees of freedom and to introduce a new term, magnitude of freedom.

Mackinlay et al. [1990] describe an input device as a 6-tuple[1]: $< M, I, S, f, O, W >$, where:

- M is "a *manipulation operator*", which classifies the physical properties sensed by the input devices according to three dimensions: position vs. force, rotary vs. linear, and absolute vs. relative.
- I is "the *input domain set*", or the set of possible input values.
- S is "the current *state* of the device."
- f is "the *resolution function* that maps from the input domain set to the output domain set."
- O is "the *output domain set*", or the set of possible states of the output.
- W is "a general purpose set of device properties that describes additional aspects of how the input device *works*."

The limitation in this formalism that I will deviate from is their description of the resolution function, f. There are some subtle restrictions in this formalism that make its application awkward. That is, they restrict f to be a mapping from I to O, as follows:

$$f : I \rightarrow O$$

However, the input sent from the device is typically provided in the form of events, and so any number of frames of this input can be used to determine the next state. Similarly, the state of the model in each of the previous frames can be stored and used to determine the new state. Thus, I rename f as an *interaction technique* and describe it as a function that

[1] their description used the letters $< M, In, S, R, Out, W >$, respectively

maps any number of frames of input and any number of previous output states onto a new output state as follows:

$$f : I^n \times O^{n-1} \to O \tag{4.1}$$

where n is the number of frames used. For many interaction techniques, only a subset of the input and output history is required. For example, the current and previous input (i. e., the change in input) and the previous output may be all that is needed, so this function would more simply be stated as:

$$f : I \times I \times O \to O$$

The number and types of parameters in I and O can vary drastically, as is demonstrated in the following two examples. Mackinlay et al. [1990] also provided a way to visualize elements of the 6-tuple that define the input device, which I will demonstrate in the following examples.

Example 4.1 (Mouse)**.** Consider the interaction technique of moving a mouse, which causes the pointer to move on the screen. In this case, mouse movement is typically provided to the computer as the relative change in horizontal and vertical movement of the physical device, or a pair $(\Delta m_x, \Delta m_y)$. Thus, I is the set of all possible pairs of such values ($I = \mathbb{R}^2$). The pointer can be most simply modelled as a position or pixel on the screen, which is again a pair (p_x, p_y). However, these values can usually only take on a finite number of values. Thus, O is the cross product of the possible values for p_x and the possible values for p_y. This set will be used frequently in the remainder of this thesis, hence the following definition:

$$D = \{0, \ldots, w-1\} \times \{0, \ldots, h-1\} \tag{4.2}$$

where w and h are the width and height of the display in pixels. So, the interaction technique

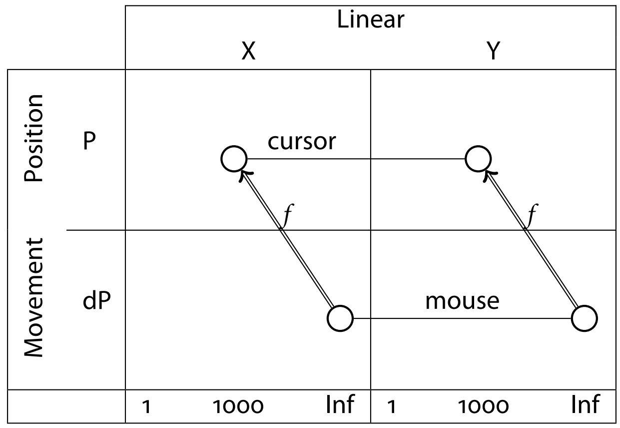


Figure 4.1: Mapping of mouse to cursor, using Mackinlay et al.'s [1990] taxonomy. The mouse senses relative position or movement (dP) of two linear dimensions (X and Y), which is then mapped via an interaction technique, f, to the cursor's location, which is measured as an absolute position (P) in the same two linear dimensions.

that translates mouse input into cursor movement would be the following mapping:

$$f : \mathbb{R}^2 \times D \to D$$

$$f(\Delta m_x, \Delta m_y, p_x, p_y) = (p_x + \Delta m_x, p_y + \Delta m_y)$$

Figure 4.1 shows how this interaction technique can be visualized using Mackinlay et al.'s [1990] taxonomy. The circles are used to indicate which of the physical properties are being sensed along the vertical axis, and in which dimension it is being sensed along the horizontal axis. For example, the mouse senses movement (i. e., relative change in position) along both the linear x and y dimensions and the virtual cursor is described by its absolute position. Note that the physical properties being sensed can also include force or change in force and the dimensions can include the linear z dimension as well as three rotary dimensions (rotation about x, y, and z).The position in the column indicates the size of the domain set used to describe these physical properties. In the case of the mouse, these sets are infinite, whereas in the case of the virtual cursor, these sets are typically in the 1000's. Solid lines connect the circles of a devices composite parts (what Mackinlay et al. [1990] described as a "merge" composition). The double-lined arrow indicates a connection between the mouse and cursor (what Mackinlay et al. [1990] described as a "connection" composition), which is achieved via the interaction technique, f.

However there are some subtleties missing from this diagram. Namely, this interaction makes use of the previous state of the output (i. e., the previous pointer location), which is not shown. Their taxonomy instead encapsulates this subtlety in the additional parameter, W. Note also that the previous state of the input (i. e., the relative change which occurred in the previous frame) is not needed, and so is left out of the mapping.

Example 4.2 (Keyboard)**.** Consider the interaction technique of using the keyboard to type letters on the screen. To simplify this example, let us assume that the keyboard only contains the 26 letters of the English alphabet plus the *Shift* key, and that the possible output is one of the 52 upper and lowercase letters of the English alphabet. Typically, the input sent from the keyboard will be which, if any, of the 27 keys are pressed ($I = \{0,1\}^{27}$). Thus, a reasonable interaction technique mapping might look something like the following equation:

$$f : \{0,1\}^{27} \rightarrow \{\text{‘a’}, \ldots, \text{‘z’}\} \cup \{\text{‘A’}, \ldots, \text{‘Z’}\} \cup \{\text{‘’}\}$$

$$f(a, \ldots, z, shift) = \begin{cases} \text{‘a’}, & \text{if } (a, \ldots, z, shift) = (1,0,0,\ldots,0,0); \\ \text{‘b’}, & \text{if } (a, \ldots, z, shift) = (0,1,0,\ldots,0,0); \\ \ldots & \\ \text{‘z’}, & \text{if } (a, \ldots, z, shift) = (0,0,0,\ldots,1,0); \\ \text{‘A’}, & \text{if } (a, \ldots, z, shift) = (1,0,0,\ldots,0,1); \\ \text{‘B’}, & \text{if } (a, \ldots, z, shift) = (0,1,0,\ldots,0,1); \\ \ldots & \\ \text{‘Z’}, & \text{if } (a, \ldots, z, shift) = (0,0,0,\ldots,1,1); \\ \text{‘’}, & \text{else.} \end{cases}$$

Note that this description of keyboard interaction neither makes use of the previous output nor the previous input. More complex (and useful) mappings would involve an entire

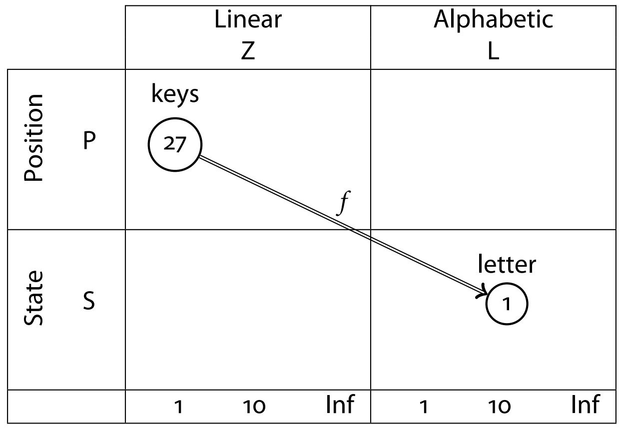


Figure 4.2: Mapping of keyboard input to lowercase text, using Mackinlay et al.'s [1990] taxonomy. Uppercase text can be depicted, but becomes difficult to parse visually.

document state (including a cursor position) as output. Note also that there are many key combinations which are mapped to '' (no character). In this mapping, nothing happens when someone presses (for example) the 'r' and 'l' keys at the same time.

It is not as clear how to visualize this interaction technique using Mackinlay et al.'s [1990], but figure 4.2 shows how it could be adapted to do so. The 27 keys of the keyboard could be depicted as 27 separate circles connected by a solid line, but this is simplified to a single circle with the number 27 inside. It is not clear, however, which of the four physical properties (position, movement, force, or change in force) applies to a letter, nor is it clear what dimension this property belongs to and whether it is linear or rotary. To accommodate this uncertainty, I added an additional "alphabetic" dimension in which the "state" property is what is needed to describe the letter. Thus the interaction technique, f, maps the state of the 27-key keyboard to a single letter (again depicted using a double arrow).

4.1.1 Summary of Mapping Input to Output

In the previous section, I demonstrated how to make use of a modified version of Mackinlay et al.'s [1990] taxonomy to describe two input devices: the mouse and keyboard. The purpose of using these two examples was twofold: (1) to illustrate the application of Mackinlay et al.'s [1990] taxonomy with something more simple than multitouch, and (2) to provide specific examples to refer to when defining the terms degrees of freedom and magnitude of freedom. In the next section, I first explain the need for some clarification in the HCI community about the term degrees of freedom, and then provide a more precise definition that will be used throughout this thesis. I then introduce the term magnitude of freedom and demonstrate why consideration of this additional information may be useful to the design of new interaction techniques.

4.2 Degrees of Freedom

The term degrees of freedom (DOF) has appeared throughout HCI literature over the last few decades. It is commonly taken to refer to the movement and rotation of a physical object or input device. For example:

> "**Degrees of freedom (DOF)** The number of independent directions a solid object or joint can move. Three-dimensional free space has six-DOF, three for translation along the x, y, and z axes, and three for rotation in the xy, yz, and xz planes. The joints of the wrist and fingers have over 20-DOF total." [Westerman, 1998, p. xxv]

This definition is consistent with the definition in kinesiology, but this term is also used in other fields, and mathematically typically refers more generally to the minimum number of parameters needed to describe the state of an object, space, or equation. The definition adopted in the HCI community also becomes problematic when attempting to refer to something that is not a rigid body. For example, how many DOF would be needed to describe moving through a set of tabs or is it even possible to determine? More specifically related to this dissertation, how many DOF are provided from a multitouch input device, and how can these DOF be mapped onto the DOF of the virtual artifacts being presented on the display?

A general definition for degrees of freedom (DOF) can be provided for both the input and output:

Definition 4.1. The *degrees of freedom of the output (DOFO)* is the minimum number of parameters required to describe the state of the computer model that represents the output.

Definition 4.2. The *degrees of freedom of the input (DOFI)* is the minimum number of parameters required to describe an input event.

Throughout the HCI literature, many claims are made that a helpful goal for an interaction designer is to attempt to match the DOFI to the DOFO [e. g., Buxton, 1983; Mackinlay et al., 1990; Zhai and Milgram, 1998; Fröhlich et al., 2006]. Indeed, this is one of the design goals used in chapter 5, and so I will state it explicitly here:

Design Guideline 4.1. An interaction technique should match the degrees of freedom of the input (DOFI) as closely as possible to the degrees of freedom of the output (DOFO).

With this goal in mind, let us reconsider our two examples. In example 4.1, it takes two parameters to describe the output (p_x and p_y), and so the DOFO is 2. It also takes two parameters to describe the input (Δm_x and Δm_y), and so the DOFI is 2, as well. Thus, this example is consistent with this design goal. In example 4.2, however, it only takes one parameter to describe the output, and so the DOFO is 1. However, it takes 27 parameters to describe the input (a,...,z, and $shift$), and so the DOFI is 27. Note that, using the visualizations depicted in figures 4.1 and 4.2, one can read these values by counting the number of circles connected by solid lines (or reading the value when they have been collapsed).

These examples illustrate a discrepancy when comparing DOFI to DOFO. Specifically, in example 4.1, the DOFI is the same as the DOFO (2), but in example 4.2, the DOFI (27) is far greater than the DOFO (1). Indeed, this discrepancy indicates some redundancy in the latter example, which partly explains the number of cases which get mapped to ‘’. However, a single key would not be sufficient to provide the same output, and so this design goal is not sufficient to explain what makes keyboard interaction successful. Because the input and output sets are finite, it is possible to also look at the size of these sets, and the consideration of these sizes will help to inform why the keyboard is a reasonable input device for text entry.

4.3 Magnitude of Freedom

In Mackinlay et al.'s [1990] taxonomy, and as depicted in figures 4.1 and 4.2, the size of the input and output domains also play a factor in how the devices can be classified, and in how an interaction technique can map the input to the output. I define the magnitude of freedom (MOF) to be the size of these sets, and this can also be defined for both the input and output:

Definition 4.3. The *magnitude of freedom of the output (MOFO)* is the number of possible states of the computer model that represents the output.

Definition 4.4. The *magnitude of freedom of the input (MOFI)* is the number of possible input events that can practically occur.

The definitions for DOF provide measures for the *number of parameters* required to describe the information being transmitted. These measures, in effect, give the minimum number of parameters that fully describe one event. The definitions I have given for MOF provide measures for the *amount* of information being transmitted. One can think of this as the scale of the range from which these DOF parameters can be chosen. Our research community has for decades been thinking about how to match the number of DOF between the input and the output. This has been useful as a guideline, and perhaps adding the matching of information about the quantity of the MOF will also lead to improvements in design. Thus, I introduce a parallel design goal:

Design Guideline 4.2. An interaction technique should match the magnitude of freedom of the input (MOFI) as closely as possible to the magnitude of freedom of the output (MOFO).

Note that both the MOFO and MOFI can be infinite. However, practical limitations can often be used to describe them in a finite way. For instance, in example 4.1, the change in mouse movement within a single frame can theoretically take on any real value. Practically, however, a person can only move the mouse a fixed distance in such a short amount of time. Similarly, the physical device can detect small changes in movement, but there is still a practical limit to this capability (e. g., for an optical mouse, this would be limited by the camera's resolution). For instance, one might assume that a mouse can only be moved about 2 cm in one frame and can accurately detect changes as small as 0.01 cm. Thus, I would be more precisely defined as:

$$I = \{-2.00, -1.99, -1.98, \ldots, 1.98, 1.99, 2.00\}^2$$

and therefore the MOFI would be $201^2 = 40,401$. For a display with 1024×768 resolution, the MOFO would be $786,432$. Note that these numbers are consistent with actual pointer movement in that movement in one frame is typically only able to traverse about 5% of the entire screen (however, the numbers chosen here are an educated guess).

For example 4.2, the MOFO would be 53 (the 52 lower and uppercase letters, plus the empty character) and the MOFI would be $2^{27} = 134,217,728$. A chorded keyboard solution [Noyes, 1983] could much more closely match these values. For example, a keyboard with only 6 keys would have a MOFI of $2^6 = 64$.

These examples illustrate that matching the MOFI to the MOFO can be a complex task, and may involve domain sets which are drastically different in size. However, as I will demonstrate in the next section and in chapter 5, when examining the application of these definitions to multitouch input, attempting to bring these numbers closer together has been beneficial in my research. Specifically, the challenge presented to the interaction technique designer is to provide a balance between guidelines 4.1 and 4.2.

4.4 3D Manipulation on a Multitouch Tabletop Display

In this section, I will make use of the formalism described in section 4.1, and demonstrated through the mouse and keyboard examples, to describe 3D manipulation on a multitouch tabletop display. The section will use the following structure: first I will formalize the output parameters involved in 3D manipulation of a virtual object, then I will describe the input parameters for two multitouch devices and compare them using guidelines 4.1 and 4.2, and lastly I will describe an alternative parametrization of multitouch devices that better satisfies these guidelines and will be used in chapter 5.

4.4.1 3D Manipulation Output (DOFO and MOFO)

In the case of 3D tabletop display interaction, the state of each 3D artifact can be described with three positional parameters (x, y, and z) and three rotational parameters (rotation about x, y, and z). Thus, interaction techniques to manipulate the position and rotation of a single 3D virtual artifact have 6DOFO ($I = \mathbb{R}^6$). A reasonable finite approximation of this set could be:

$$O = V \times R^3$$

where V is the view volume (i. e., possible positions), discretized at pixel resolution:

$$V = \{0, \ldots, w\} \times \{0, \ldots, h\} \times \{0, \ldots, d\}$$

and R is the set of possible rotations in one dimension, discretized at 1°:

$$R = \{0, \ldots, 360\}$$

It is not clear how many discrete levels of depth can be perceived using 3D projection geometry (see chapter 3), so for this dissertation I will assume that the perceived levels of depth will be no greater than either the number of horizontal or vertical pixels available ($d = \max\{w, h\}$). Using this approximation for a 1024×768 pixel display, the MOFO would be $1024 \times 768 \times 1024 \times 360^3 \approx 3.76 \times 10^{16}$.

4.4.2 Multitouch Input

There are a variety of devices that can provide multitouch tabletop display input. In this chapter, I consider two distinct types of input devices to illustrate the determination of DOFI: the DiamondTouch and camera-based technologies, such as FTIR and DI.

4.4.2.1 DiamondTouch: Comparing DOFI to DOFO

The DiamondTouch [Dietz and Leigh, 2001] input device uses an array of antennas along both the horizontal and vertical axes of a table. A person can sit or stand on a pad that capacitively couples them to the device, and thus when touching the display, they will activate some of the antennas in both the horizontal and vertical arrays. Each antenna provides a value proportional to the area being touched. Thus the input space for a DiamondTouch is:

$$I = \mathbb{R}^h \times \mathbb{R}^v = \mathbb{R}^{h+v}$$

where h and v are the number of horizontal and vertical antennas. Thus, the number of parameters required to describe I is $h + v$, which provides $h + v$ DOFI.

However, this description of the input space does not completely describe the capabilities of the device, because this input can be detected and uniquely identified for up to four people (on different pads). Thus a more complete description would be as follows:

$$I = \left(\mathbb{R}^{h+v}\right)^4 = \mathbb{R}^{4h+4v}$$

so the device has $4h + 4v$ DOFI. Assuming a 1024×768 arrangement of sensors, this device

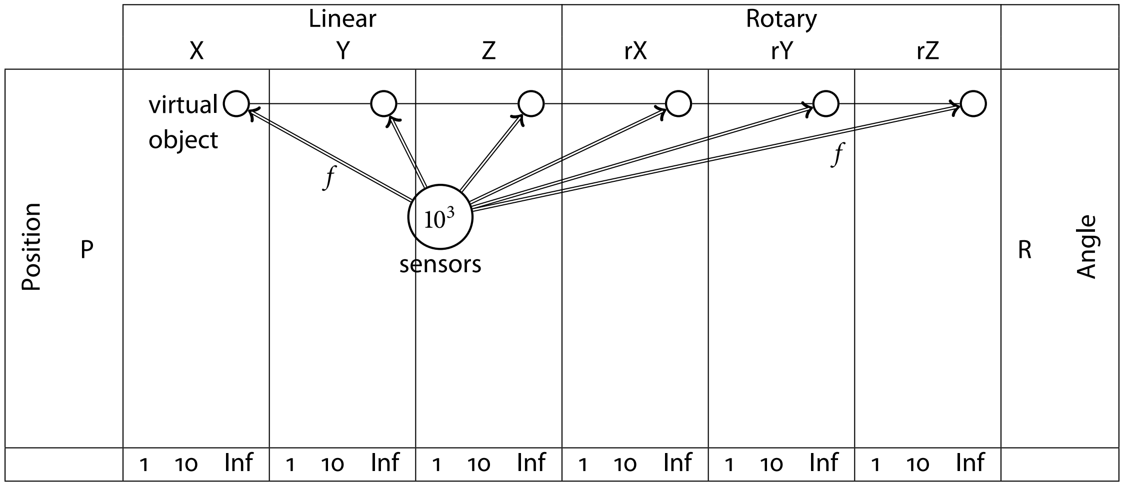


Figure 4.3: Mapping of DiamondTouch input to 3D manipulation using Mackinlay et al.'s [1990] taxonomy. There are 1024 + 768 sensors (depicted at the order of magnitude 10^3) that each provide an on/off state when a sensor is active (i. e., a contact point is at the "touched" location in the Z axis). The values of these sensors would then need to be mapped to the 6DOFO of 3D virtual object manipulation.

provides 1792DOFI, which is much greater than the 6DOFO required for 3D virtual object manipulation. Thus, based on guideline 4.1, providing a good interaction technique will be challenging.

4.4.2.2 FTIR and DI

Both FTIR [Han, 2005; SMART Technologies, 2008] and DI [Microsoft Corp., 2008] use one or more cameras to detect IR light that is reflected down from the table's surface when the table is touched. The amount of light that is reflected can vary based on both the reflective properties of the finger/hand/object being used to touch the table, as well as the pressure. This variance also depends largely on which technique is used, as well as the conditions of the table's setup, such as lighting and choice of hardware. There are also many factors about the placement of the cameras and the algorithms used to combine images from multiple cameras which affect the resolution of the input. However, a careful setup of equipment can result in at least pixel resolution. That is, at some point in the processing algorithm, an image is provided with a resolution equal to that of the display such that each pixel has a

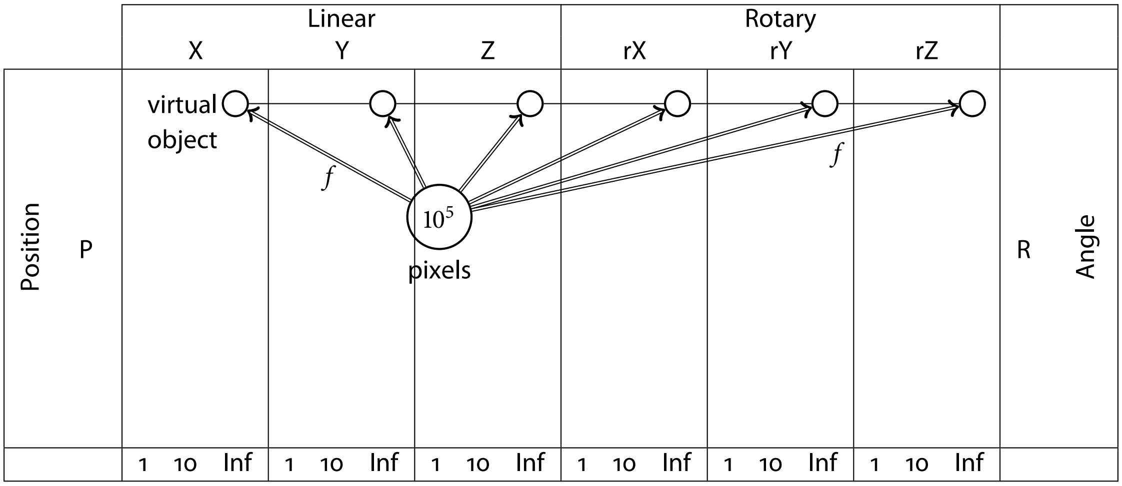


Figure 4.4: Mapping of camera-based input to 3D manipulation using Mackinlay et al.'s [1990] taxonomy. There are 1024 × 768 pixels (depicted at the order of magnitude 10^5) that each provide an on/off state when a pixel is illuminated (i. e., a contact point is at the "touched" location in the Z axis). The values of these pixels would then need to be mapped to the 6DOFO of 3D virtual object manipulation.

value which represents the intensity of IR light being reflected at that point of the tabletop display. Thus, the input space for these devices can be described as follows:

$$I = \mathbb{R}^{w \times h}$$

where w and h are the width and height of the image provided. Thus, the number of parameters required to describe I is $w \times h$, and so the device has $w \times h$ DOFI. Assuming a camera with 1024 × 768 pixels of resolution, these devices provide 236,739DOFI, which is even larger than the DOFI provided by the DiamondTouch, and is far greater than the 6DOFO required for 3D virtual object manipulation. Thus, based on guideline 4.1, providing a good interaction technique will again be challenging.

4.4.2.3 Comparing DiamondTouch to FTIR and DI

For both the DiamondTouch and the camera-based techniques, while the devices provide a range of values for each antenna or pixel, in practical use, a threshold is typically used to indicate for each one whether it is *on* or *off*. Thus, the two input spaces could be discretized as follows:

$$I_{DT} = \{0,1\}^{4h+4v} \quad \text{(DiamondTouch)}$$

$$I_{cam} = \{0,1\}^{w \times h} \quad \text{(camera-based)}$$

Thus, the MOFI for the devices would be 2^{h+v} and $2^{w \times h}$, respectively. Assuming a 1024×768 resolution display, these values would be $2^{1792} \approx 2.79 \times 10^{539}$ and $2^{786,432} \approx 4.18 \times 10^{236,739}$. The MOFI of both these devices far exceeds the MOFO of 3.76×10^{16}. Based on guideline 4.2, this large discrepancy may indicate a possible mismatch between the input device choice and a person's ability to use these devices to control the 6DOFO of 3D manipulation. However, as was shown for keyboard entry, when the DOFI and MOFI far exceed the DOFO and MOFO, respectively, a reasonable interaction technique can still be achieved. On the other hand, this discrepancy could also indicate that, once 3D manipulation is made possible, there remains a significant amount of information that can be used to control other parameters in the tabletop display application.

Perhaps more importantly, this precise definition of what is meant by DOF and MOF provides some insight about the subtle differences between these two devices and their capabilities. From these observations, it is clear that the DiamondTouch's use of two sets of arrays provides far less freedom in how this information can be interpreted. In particular, this lack of information makes it much more difficult to distinguish between more than two contacts, when these contacts come from the same sensor pad (typically associated with a particular person). This difficulty was identified in the original publication of this input technology [Dietz and Leigh, 2001], but this formulation provides a more general underlying cause.

Despite the large discrepancy between the DOFI and DOFO as well as the MOFI and MOFO of these devices, there is another method of describing these devices which can much more closely match these values. In section 4.4.3, I will describe this description and show how it

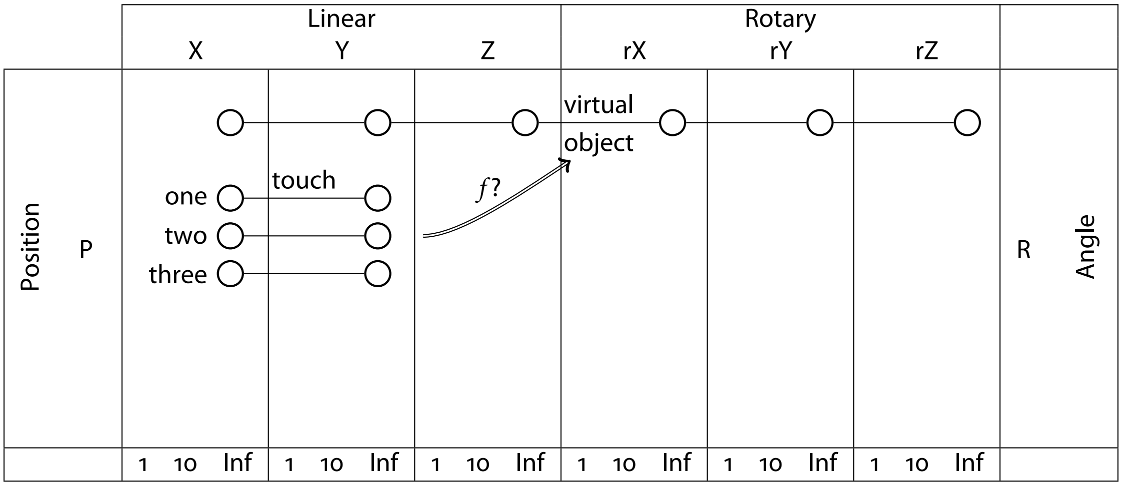


Figure 4.5: Mapping of multipoint input to 3D manipulation using Mackinlay et al.'s [1990] taxonomy. The centre of each active blob provides an absolute position in both the X and Y dimensions. A combination of these touch points can be combined and mapped to the 6DOFO of 3D virtual object manipulation.

better satisfies guidelines 4.1 and 4.2. Thus, it is this next formulation that is used in chapter 5 to inform the design of the one-, two-, and three-touch techniques described there.

4.4.3 Multipoint Input

For both the DiamondTouch and the camera-based techniques, these descriptions of the input space provide a reasonable representation of the input and show that they have vastly more DOFI and MOFI than are required to control the 6DOFO for 3D manipulation. However, these descriptions of the input space do not provide a mechanism for movement along the table's surface by a finger/hand/object. That is, the parameters used to describe the input only vary between on and off (like a button or key of a keyboard), and there is no continuity implicit in the model between antennas or pixels. This difference between touch-sensing devices and those requiring an intermediary (such as a mouse) was identified early in the HCI literature [Buxton, 1983]. For example, if a person drags their finger from pixel (109, 322) to (110, 321) on a camera-based device, this would be indistinguishable from touching (109, 322) in one frame, lifting their finger and touching (110, 321) in the next. Nevertheless, the latter

	Input					Output
	DT	Camera	Multipoint			
			$n = 1$	$n = 2$	$n = 3$	
DOF	7,168	786,432	2	4	6	6
MOF	2.79×10^{539}	$4.18 \times 10^{236,739}$	786,432	6.18×10^{11}	4.86×10^{17}	3.76×10^{16}

Table 4.1: The DOF and MOF of the two multitouch devices.

feat is difficult to accomplish physically, so the former is far more likely to be the cause of this input sequence.

Using thresholds and heuristics, it is possible to create a more useful model of the input events that can be used in an interaction technique. Namely, with the camera-based techniques, it is possible to approximate the centre-points of each contact on the display and to track them across frames using blob-detection algorithms [Lindeberg, 1994]. Similar algorithms can be used to track movement on the DiamondTouch for each capacitively-coupled person. By considering the input events that are provided after this processing of the information, the input model might look something more along the lines of the following set:

$$I = D^n$$

where D is the set of pixels on the display (eq. (4.2)) and n is the number of contacts that can be simultaneously and reliably detected using these algorithms. The centre-detection would thus be an intermediary mapping from the true input set to D^n as follows:

$$b : \{0,1\}^{w \times h} \to D^n \quad \text{(camera-based centre detection)}$$

Using this formulation, the DOFI and MOFI would be relative to the number of contacts. Specifically, the DOFI would be $2n$ and the MOFI would be $786,432^n$ for a 1024×768 display. Thus, if a person were to use 3 touches to match the 6DOFI to the 6DOFO, the MOFI would be approximately 4.86×10^{17}, which is much closer to the MOFO of 3.76×10^{16}. Table 4.1

summarizes the calculations of DOF and MOF using the formulations shown in this chapter, and clearly demonstrates a good match between the input space and output space for both the DOF and MOF when using a multipoint formulation.

4.5 Chapter Summary

In this chapter, I have formalized the definition of degrees of freedom and introduced and defined magnitude of freedom for both the input space and output space of an interaction technique. This formulation builds upon the previous literature on the taxonomies of input devices [Buxton, 1983; Foley et al., 1984; Mackinlay et al., 1990; Zhai, 1995] and highlights some of the subtleties particular to multitouch input. By precisely examining both the DOF and MOF of the input and output, one can get a better sense of these subtle differences and be better prepared for the task of designing an interaction technique. This precise formulation leads into the next chapter, where a point-based formulation of the input provided from a DiamondTouch input device is used to map its input using one, two, and three fingers to the 6DOFO of 3D manipulation of a virtual object on a tabletop display. This way of thinking about DOF and MOF has helped to clarify the limits, complexity, and potential of multitouch input. It has been a fundamental part of my design process.

Interacting with 3D Virtual Objects 5

In this chapter, I describe the design, implementation, and evaluation of a series of interaction techniques for manipulating 3D virtual objects on a multitouch digital table. In chapter 3, the result that the ability to interact reduces errors when matching 3D virtual artifacts indicates the potential benefits to comprehension of the 3D space. In addition to comprehension through exploration, the ability to manipulate 3D artifacts can enable some additional freedoms we enjoy when manipulating physical objects, such as flipping, stacking, and piling. In chapter 4, I define and describe how to identify the degrees of freedom (DOF) of a multitouch tabletop display. In this chapter, I use that formulation to describe the interaction techniques, as well as to highlight what technology these techniques are appropriate for.

This work is partly motivated by the recent surge of multitouch interactive display technology, such as the DiamondTouch [Dietz and Leigh, 2001], FTIR [Han, 2005], the SMART Table [2008], and the Microsoft Surface [2008]. It is also motivated by the common 3D nature of use of traditional tables. On traditional tables, people frequently use the third dimension to pile, sort and store objects. This common place use of 3D has long been effective and informative for organization. Digitally it has been approximated through single point plus menu interaction in Bumptop [Agarawala and Balakrishnan, 2006] and explorations are ongoing that investigate the technical possibility of sensing additional input from above [Hilliges et al., 2009] or below [Wigdor et al., 2006] such surfaces, to realise full control of 3D virtual objects. My challenge has been to stay with direct touch, maintaining interactive contact with the 2D surface which holds the displayed 3D objects.

The use of some, often shallow, 3D effects to support interaction is common in windowing environments. The layering and shadowing effects both enhance the visual appeal of the interfaces and provide natural metaphors for switching documents and workspaces into and out of focus. Some commercial interfaces further extend the 3D effects, using animations to clarify feedback effects such as distorting windows and icons to show the relationship between pre- and post-action states (e. g., Mac OS®X). Researchers are also investigating problems and solutions that arise from moving between layers on the desktop. Dragicevic [2004], for example, describes 3D visuals of dog-ears, folding and shuffling to make working with overlapping windows more intuitive. Agarawala and Balakrishnan's [2006] 'BumpTop' wholeheartedly adopts the emulation of reality on the desktop, using both rich 3D visuals and physics modelling to enrich interaction—objects can be piled on top of one another or flipped onto their backs; objects can be thrown at others, and the visual effects of collisions depends on their mass and velocity. The reality of the lustrous environment, however, is hindered by its constraint to a single point of interaction through a stylus input device. For comparison with the reality it attempts to emulate, though, consider the awkwardness of manipulating objects on your physical desk using only one index finger.

Researchers have shown that these surfaces support direct manipulation techniques that naturally emulate 2D rotation, translation, and scaling [Kruger et al., 2005; Hancock et al., 2006; Shen et al., 2004]. However, since the focus in this chapter is on manipulating 3D information, there is an apparent discrepancy between this 2D input space and the 3D control being provided. Thus, the main challenge in this work is in mapping the input available on such multitouch devices to the manipulation of 3D virtual artifacts displayed on a 2D surface.

To address this discrepancy, we[1] first consider the concept of *shallow-depth 3D*—full 3D visuals with full 3D interaction, but extremely limited depth—as a potential interaction space.

[1]The research presented in this chapter is largely taken from the materials published in Hancock et al. [2007], and so any use of the first-person plural refers to these authors: Mark Hancock, Sheelagh Carpendale, and Andy Cockburn.

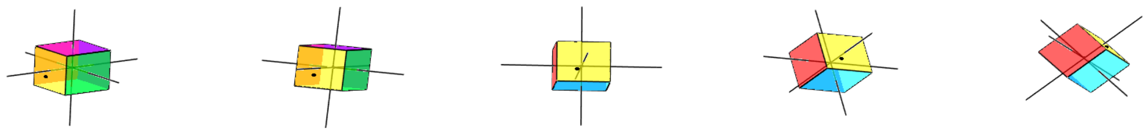


Figure 5.1: A sequence of motion using one-touch interaction in shallow-depth 3D. The black dot represents the point of contact of the person's finger.

We focus on shallow-depth interaction in the z-plane for two reasons: first, interactions on physical desktops take place within a shallow-depth field (e. g., riffling, sorting and manipulating piles, and rotating or flipping objects on the surface); second, current desktop graphical user interfaces are similarly limited to a shallow-depth field. We argue that providing shallow-depth 3D capabilities allows for a more engaging and rich experience. Providing this shallow-depth 3D interaction extends the previous 2D direct interactions of rotation, translation and scaling to include flipping—that is, planar movement (x, y translational freedom) with full 3D rotation (yaw, pitch and roll).

To empirically explore shallow-depth 3D as an interaction space, we consider the task of moving and rotating a small 3D object (e. g., a cube) across a tabletop (see figure 5.1). We first present design guidelines for direct-touch 3D interaction. Next we discuss candidate interaction techniques for supporting these manipulations using one, two and three points of contact, formally demonstrating how two-dimensional surface interactions can be used to directly manipulate shallow-depth 3D objects. We then describe a usability study that compares the speed and accuracy of the techniques as well as the participants' subjective perceptions of them. In closing we discuss the implications and suggest alternative techniques based on the results of this study. In particular, I show how the three-touch technique can be extended to provide full 3D control, by adding z-translational freedoms to the shallow-depth 3D interaction. Full 3D direct manipulation control that only uses the multitouch capabilities of the surface is one of the major contributions of this thesis.

5.1 Design Guidelines for 3D Multitouch Interaction

Our eventual goal is to enable the existing freedom of object manipulation available in the physical world within digital tables. In this endeavour, we attempt to combine the benefits of both 3D interaction and multitouch tabletop displays. In order to be successful, we need to take care in the design of new techniques to support interaction. We suggest the following design guidelines for interaction on tabletop displays in 3D.

5.1.1 Provide Separate and Simultaneous Control

Interactions including flipping of objects, storage and communication through small adjustments of objects become possible by allowing full rotation and translation in three dimensions (6DOFO, see chapter 4). On the one hand, it can be useful to control each of these DOFO separately, so as to specify a precise position and orientation, but on the other hand it may improve performance to be able to combine the movements and rotations into one movement. The design of an interaction technique must consider the balance between *simultaneity* and *independence* of rotation and translation:

- *Simultaneity of Rotation and Translation*: In the real world, people are capable of simultaneously activating a combination of muscles to perform a single action that both moves and rotates an object. Similarly, a 3D tabletop interface should allow people to simultaneously rotate and translate an object.

- *Independence of Rotation and Translation*: Along the same lines, people should be able to activate rotation and translation using distinct actions, in the same way that different muscle groups are used to perform these actions in reality. Thus, people could combine these actions at the cognitive level instead of combining them in a potentially awkward way through the interaction technique itself.

5.1.2 Provide Connected Manipulation (Sticky Fingers)

Connected manipulation is direct manipulation in which a constant visual and physical connection with the object is maintained *throughout the entire interaction*. We distinguish this from direct input—when input space and display space are superimposed—by specifically requiring that the object being manipulated, not just the display space, remain in physical contact with the input mechanism.

This guideline is also important when manipulating 2D objects on a digital table [Hancock et al., 2006; Kruger et al., 2004]. In 2D multitouch interfaces, the two-finger move / rotate / scale interaction technique has become ubiquitous. Because one's fingers stay in touch with the virtual object in the location they initially contact, this can be referred to as a *sticky-finger* interaction. This perception of touching the virtual object persists through the interaction, providing some of the feedback one might expect in the physical world. The scaling action of spreading one's fingers also maintains stickiness, still providing a person with the feeling that they are controlling two points on the virtual object. However, this scale aspect would be impossible in the physical world (at least for rigid bodies), thus it combines the partial physicality of the sticky fingers with the potential for magic that the computer offers.

Sticky fingers works well in 2D, providing move (in x and y), spin (rotate about z) and scale. In 3D the first two of these capabilities can be directly mapped giving move and spin, however in 3D two additional factors are missing: lift and flip. We emphasize the importance of maintaining connected manipulation or sticky fingers in 3D, as it may be more tempting to ignore this constraint when controlling these additional DOFO.

5.1.3 Prevent Cognitive Disconnect

By allowing people to interact directly with 3D objects, one should avoid actions that (though possible with technology) do not conform to what people expect. For example, since it is impossible to push one's finger through the display, limiting the depth of interaction maintains that expectation and prevents disconnect. Furthermore, on traditional tabletops people can interact on the surface of the table and the space between them and the top of the table; however, most interaction takes place in the first few inches. Limiting interaction to a small finite z depth places a virtual surface just below the actual display, providing a similar few inches for 3D interaction.

5.1.4 Support Many Identifiable Contacts for Each Person

Current tabletop display hardware provides a mosaic of supporting technology. Some technologies allow for a large number of points of contact, without identifying information [Han, 2005], and others provide identifiable input for a single point of contact for a small number of people [Dietz and Leigh, 2001]. In order to fully support direct-touch 3D interaction on tables for multiple people, the hardware needs to support identification of not only *where* a finger is touching, but also *which finger* of *which person* is touching.

5.2 Multipoint Direct-Touch Interaction Techniques

We have designed three new direct-touch interaction techniques for manipulating 3D objects on a tabletop display. These designs were in part informed by our suggested guidelines, but mostly have helped to generate them. Before describing the one-,two-, and three-touch interaction techniques, I first describe the input and output space in terms of the formalism presented in chapter 4. Then, in the description of the techniques, I use this formalism to clarify the touch input to 3D output mapping.

5.2.1 Input Space

As described in chapter 2, multitouch display hardware availability has been in flux during this research. Thus when first implementing these one-, two-, and three-touch techniques, I only had access to either a two-touch camera based technology [SMART Technologies, 2003] or the DiamondTouch [Dietz and Leigh, 2001], which only offered one identifiable touch per pad. I chose to extend DiamondTouch input for my needs. Thus the multi-point techniques were implemented using the DiamondTouch input device [Dietz and Leigh, 2001] by attaching distinct sensors to both the index finger and thumb of a right-handed insulating glove. The third touch-point was provided with a regular DiamondTouch pad through the left hand. Thus the input space is a subset of:

$$I_{DT} = \mathbb{R}^{3h+3v}$$

Furthermore, we mapped each of these sensors to a different point of contact so that we could track the movement of each finger across the antennae. This intermediate mapping uses the maximum antennae value in x and y for each sensor as the (x, y) coordinate of the contact point:

$$\begin{aligned} g : I_{DT} &\rightarrow D^3 \\ g(x_{11}, \ldots, x_{1h}, y_{11}, \ldots, y_{1v}, \\ x_{21}, \ldots, x_{2h}, y_{21}, \ldots, y_{2v}, \\ x_{31}, \ldots, x_{3h}, y_{31}, \ldots, y_{3v}) &= (x_1, y_1, x_2, y_2, x_3, y_3), \\ \text{where } x_i = \max\{x_{i1}, \ldots, x_{ih}\}, &\; y_i = \max\{y_{i1}, \ldots, y_{iv}\} \end{aligned}$$

This mapping provided three continously tracked points of contact. Furthermore, each contact point is identifiable, such that when a person lifts any finger and re-touches the display, the computer can identify the finger as being the same. Thus, the one-touch technique made use of only the index finger on the dominant hand (2DOFI, $I = D$), the two-touch technique made use of the index finger on the dominant hand and any finger on the non-dominant hand (4DOFI, $I = D^2$), and the three-touch technique made use of the index finger and thumb of the dominant hand, plus any finger from the non-dominant hand (6DOFI, $I = D^3$).

To use these same algorithms using FTIR or DI technology, these fingers would not be identifiable (i. e., lifting a finger and then re-touching the display would not be distinguishable from lifting with one finger and touching with another). Nonetheless, the algorithms below can be easily modified so that the order of touches determines which finger performs which action. However, this modification alters the experience in the following subtle way: a person must hold their first finger on the table to use their second and similarly hold two fingers to use the third.

5.2.2 Output Space

The shallow-depth 3D *output* we wish to provide has the following output space:

$$O = D \times R^3, \text{where } R = \{0°, \ldots, 360°\}$$

This output space leads to the following five DOFO:

- position $(x, y) \in D$—the position on the surface of the table
- yaw ($\psi \in R$)—object rotation about the z-axis (planar)
- roll ($\phi \in R$)—object rotation about the y-axis (side-to-side)
- pitch ($\theta \in R$)—object rotation about the x-axis (front-to-back)

For the one-touch technique, the rotational parameters will be specified using the axis-angle representation $(\vec{a}, \theta)$, which is equivalent to the Euler angles (ψ, ϕ, θ) based on Euler's rotation theorem [Palais et al., 2009]. Note that we also describe how z (lift) could be manipulated in some cases, but this feature was not implemented for the purpose of this study. Section 5.3.3.3 describes a technique that incorporates lift and chapters 6 and 7 describe how the use of this technique can benefit tabletop application design. Alternative designs could apply the adjustments made to z to some other DOFO (e. g., scaling).

This formulation of the output space completely describes the possible new states that can be achieved through a manipulation in shallow-depth 3D (i. e., the right-hand side of eq. (4.1)). However, the interaction techniques make use of some additional parameters of the display output in order to determine the next state. Namely, the algorithms used to implement these interaction techniques make use of the projection geometry used to render the scene and the z-value of the object in the model in order to determine the movement and rotation of the manipulation. When necessary, these techniques will include these parameters on the left-side of eq. (4.1).

5.2.3 One-Touch Input

We can achieve 5DOFO movement with a single point (2DOFI) by extending the RNT algorithm [Kruger et al., 2005] into the third dimension. The 3D object can be moved so that the point of contact remains under one's finger and the axis of rotation can be determined from the same point of contact (see figure 5.2).

5.2.3.1 Mapping

This technique uses the following mapping:

$$f : D \times D \times O \times Z \to O$$

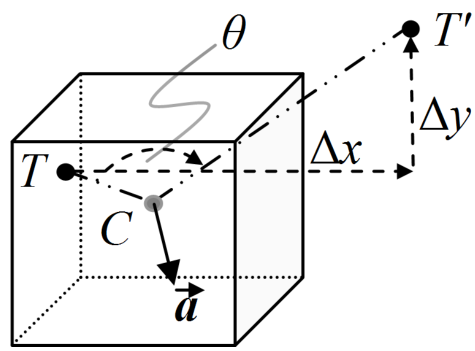


Figure 5.2: A video demonstrating the one-touch interaction in shallow-depth 3D. The still image is a diagram of the parameters that change in this interaction technique.

where $Z = \mathbb{R}^{w \times h}$ are the values in the z-buffer (i. e., the z-value of the "nearest" point rendered at each pixel).

Let $T = (x, y) \in D$ and $T' = (x', y') \in D$ be the initial and final points of contact.

Let $\overline{T} = (x, y, z)$, where z is the value in the z-buffer at (x, y)

Let $\overline{T'} = (x', y', z)$ (using the same z-value as $\overline{T}$).

Let $C = (x_c, y_c, z_c)$ be the initial centre of the object in the 3D scene.

Then the output for $f(\overline{T}, \overline{T'}, C)$ is:

$$\begin{aligned} \Delta x &= x' - x, \quad \Delta y = y' - y \\ \vec{a} &= \overrightarrow{CT} \times \overrightarrow{CT'} \\ \theta &= \angle \overline{T} C \overline{T'} \end{aligned}$$

5.2.3.2 Description

While the discrepancy in mapping between 2DOFI and 5DOFO is high, the actual action can be described in physical terms (see figure 5.1). Touching a point on the cube works like a sticky finger in that the contact point will rise toward the surface and the leading direction, causing the cube to rotate in x, y, and z until the contact point is as close to the surface and the lead direction as the shape of the cube will allow. Rotating the chosen side to the surface merely involves touching that side and dragging. This can require a re-touch for an initially occluded side.

Despite the fact that this technique provides the ability to rotate and translate a 3D object to any position and orientation, it is common to want to perform more constrained interaction, such as translation alone or planar rotation. We provide this ability through dedicated areas on the object. For polygonal objects, a circular area about the centre of each face is reserved for translations and a doughnut-shaped region around that circle is reserved for planar rotations (using the 2D RNT algorithm). For non-polygonal objects, a more abstract central location can be chosen on some surface of the object, about which the circle and doughnut shapes can be drawn.

5.2.4 Two-Touch Input

Five or six DOFO can be achieved using only two points of contact (4DOFI). The first point of contact can use the RNT algorithm [Kruger et al., 2005] to achieve both translation in x and y as well as yaw. The second point can be used to specify pitch and roll (see figure 5.3). If z motion is desirable, this can be manipulated according to the change in distance between the two points.

5.2.4.1 Mapping

This technique uses the following mapping:

$$f : D^2 \times D^2 \times D \rightarrow O$$

Let $T_i = (x_i, y_i) \in D$ and $T'_i = (x'_i, y'_i) \in D$ be the initial and final points for the ith point of contact, where $i \in \{1, 2\}$.

Let $C = (x_c, y_c) \in D$ (the projection of the centre of the object in the 2D surface).

Then the output for $f(T_1, T_2, T'_1, T'_2, C)$ is:

$$\begin{aligned}
\Delta x &= x'_1 - x_1 \quad \Delta y = y'_1 - y_1 \\
\Delta z &= |\overrightarrow{T'_1 T'_2}| - |\overrightarrow{T_1 T_2}| \quad \text{(if desired)} \\
\Delta \psi &= \angle T_1 C T'_1 \quad \text{(about } T_1\text{)} \\
\Delta \phi &= K_1(x'_2 - x_2), \exists K_1 \in \mathbb{R} \\
\Delta \theta &= K_2(y'_2 - y_2), \exists K_2 \in \mathbb{R}
\end{aligned}$$

In our user study, T_1 was provided through the index finger of the dominant hand and T_2 through the index finger of the non-dominant hand. However, the technique is not limited to this configuration; other sensible configurations include reversing these two fingers or using the index finger and thumb on the same hand.

5.2.4.2 Description

This technique provides easy causal movement coupled with rotation that maintains the vertical orientation of the object's projection. If the vertical orientation needs adjusting, for example if the right side of a cube is not at the surface, it can be adjusted with a finger on the non-dominant hand.

As with the one-touch technique, it is often desirable to perform constrained translation-

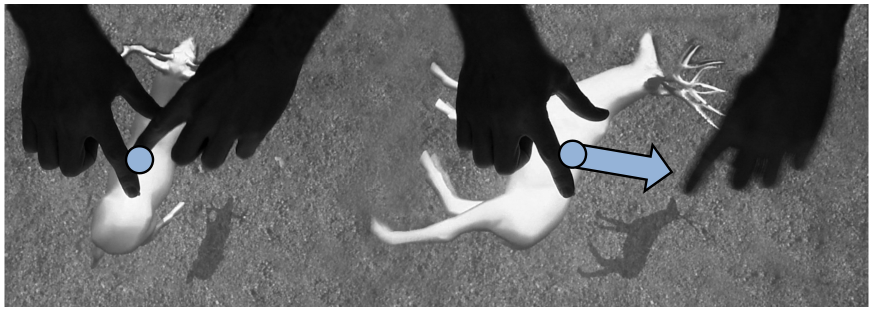


Figure 5.3: Using the two- and three-touch techniques, this video shows how a person can perform pitch and roll rotations with a finger on the non-dominant hand. The still image shows a diagram of how this technique works.

only movement. This is again provided at the centre of each face of a polygon or an abstract central location on the surface of a non-polygonal object.

5.2.5 Three-Touch Input

Our three-touch interaction technique maps 6dofi to 5 or 6dofo (see figure 5.4). In this mapping, the first point of contact is used for translation, the second point for yaw about the first point, and the third point for pitch and roll about the centre of the object. The depth can be specified by the difference in distance between the first and second touch points. The order of the points can be specified either by taking the points in order of contact with the table or in a predefined order (if the source of each point is identifiable).

5.2.5.1 Mapping

This technique uses the following mapping:

$$f : D^3 \times D^3 \rightarrow O$$

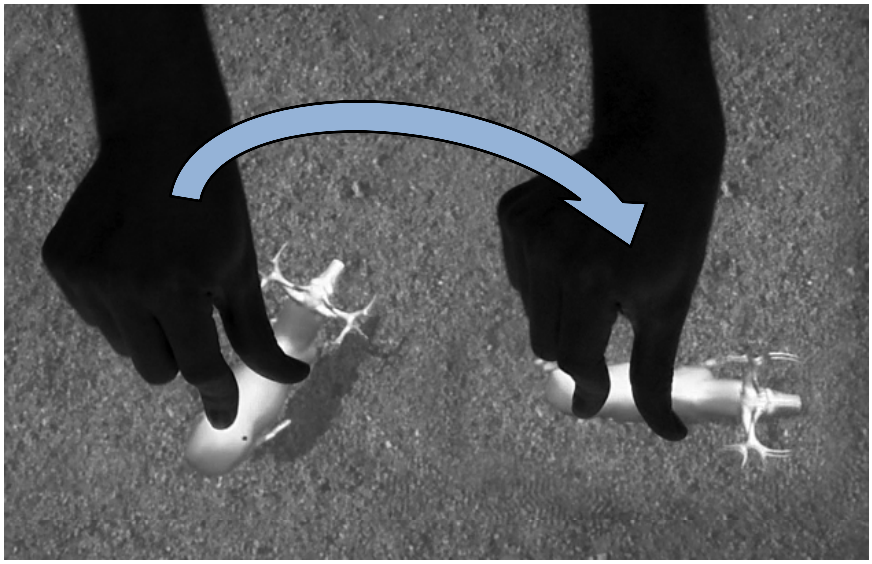


Figure 5.4: Using three-touch interaction, a person can perform a simultaneous translation and rotation on the surface of the table, as shown in this video. The still image shows a diagram of how this rotation works. They can also simultaneously rotate the object in pitch and roll with a finger on the non-dominant hand (see figure 5.3).

Let $T_i = (x_i, y_i)$ and $T'_i = (x'_i, y'_i)$ be the initial and final points for the ith point of contact, where $i \in \{1, 2, 3\}$.

Then the output for $f(T_1, T_2, T_3, T'_1, T'_2, T'_3)$ is:

$$
\begin{aligned}
\Delta x &= x'_1 - x_1, \quad \Delta y = y'_1 - y_1 \\
\Delta z &= |\overrightarrow{T'_1 T'_2}| - |\overrightarrow{T_1 T_2}| \quad \text{(if desired)} \\
\Delta \psi &= \angle T_2 T_1 T'_2 \quad \text{(about } T_1\text{)} \\
\Delta \phi &= K_1(x'_3 - x_3), \exists K_1 \in \mathbb{R} \\
\Delta \theta &= K_2(y'_3 - y_3), \exists K_2 \in \mathbb{R}
\end{aligned}
$$

In our user study, T_1 and T_2 were provided through the index finger and thumb of the dominant hand, respectively. T_3 was provided through the index finger of the non-dominant hand. Note that this technique provides the ability to perform constrained motion without the need for dedicated areas.

5.2.5.2 Description

In principle this interaction is quite simple. For example, a single touch with one's index finger supports translation only, including one's thumb adds rotation around the z-axis and the addition of a finger from one's other hand provides the other two rotations.

In theory, three-touch allows the most efficient means for control because one can concurrently and independently manipulate all 6DOFO. However, there is a risk that this freedom may be confusing. Furthermore, both the two- and three-touch techniques may disconnect the object from the initial touch location upon rotation in pitch or roll. This disconnect may add to the confusion, creating an advantage to the one-touch technique. Hence, empirical comparison of the techniques is necessary.

5.3 User Study: Interaction Technique Comparison

To better understand how people interact with these three manipulation techniques (one, two and three finger), we conducted a study that compares them in terms of their speed, accuracy and the subjective preferences of the participants. Since these techniques vary considerably in interaction styles, conducting an empirical comparison can shed light on which balance of design tradeoffs are the most effective and satisfying for people.

For example, one-touch interaction is likely to be slow, but people may appreciate its simplicity and reliability; three-touch interaction may be fast if the participants can adapt to its comparative power and complexity, but they may report a higher cognitive load if it fails to be perceived as 'natural'.

5.3.1 Method

This section describes the experimental setup, including a description of the participants, the equipment, and a description of the three tasks, including the conditions and design for each task.

5.3.1.1 Participants

Twelve students (6 male, 6 female) from a local university participated in the study. Both national and international students were selected from a variety of disciplines. Five participants reported no prior experience with 3D gaming and seven reported some. The experience of these seven varied from once a year to four times a month. Ages ranged from 21 to 33 (M = 26.3, SD = 3.8). All participants were right-handed and no participant reported any colourblindness.

5.3.1.2 Apparatus

The experiment was performed on a front-projected 1024 × 768 pixel tabletop display using DiamondTouch [Dietz and Leigh, 2001] input with an 87 cm × 66 cm display area (12 pixels / cm). Multi-finger input was provided as described in section 5.2.1. The display surface was 72 cm above the floor and participants were provided with an adjustable chair. A parallel 3D projection was used with a neutral COP (see chapter 3) to render objects on the display. Objects were all full 3D objects but to provide the shallow-depth environment there was no movement in z. That is, objects could roll, tumble, and flip but the object's centre remained at a fixed z depth. Thus interaction was limited in all conditions to 5DOFO. Software automatically logged the participants' actions and task times.

5.3.1.3 Common Procedure

For each technique (one-, two-, and three-touch), participants performed three tasks in the same order. The order of techniques was counterbalanced between participants using a Latin square. Afterwards, each participant was asked to complete a questionnaire to provide both background and feedback about their experience. Participants were then interviewed by the experimenter.

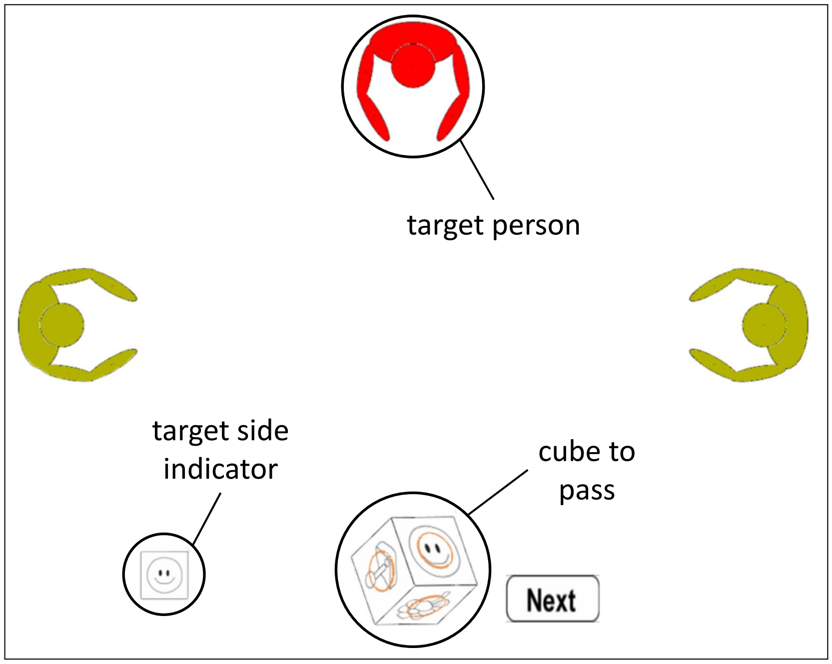


Figure 5.5: In the passing task, participants were asked to pass a cube to a target person with the target side facing up and toward the "virtual" person. The start position of the cube was close to the centre of the table.

The primary dependent measure in the two formally analysed tasks (tasks 1 and 2) was the task completion time. We additionally analysed data characterising how participants interacted with the techniques, including the time spent touching, translating and rotating the objects, and also the locations on the objects that the participants touched.

5.3.1.4 Task 1: Passing

In order to determine a person's ability to use each technique for communication with other people, our first task required participants to pass a cube to one of three "virtual" people with a specific side of the cube facing upward and toward the virtual person (see figure 5.5). This task was modelled after the study done on the 2D RNT rotation technique [Kruger et al., 2005].

Each side of the cube had a distinct black and white icon. At the start of each trial, the cube was in the same location immediately in front of the participant, with the "top" side (a happy face) uppermost. Virtual participants were located to the participant's left, right and opposite. To start each trial an icon appeared on the screen and one virtual participant was indicated in red. The participant then matched that trial icon with one on the cube and passed the cube to the indicated virtual participant with the correct icon facing upward. The task was repeated six times—once for each side of the cube—for each target destination. Participants performed six practice trials each time they started with a new interaction technique.

Data from this task were analysed using a within-participants analysis of variance for the following three factors:

- *Technique*: one-touch, two-touch, three-touch;
- *Destination*: left (40 cm), opposite (38 cm), right (40 cm);
- *Target side*: top, bottom, left, right, back, front.

5.3.1.5 Task 2: Docking

To explore performance differences in the three techniques, we asked participants to complete a docking task. This task was a variation of the task developed by Zhai and Milgram [1998] and used more recently to compare GlobeFish and GlobeMouse to other 6DOF techniques [Fröhlich et al., 2006].

In this task, participants were asked to dock a pyramid inside another of equal size (see figure 5.6). Spines around the vertices were used to indicate docking tolerance. The vertices and edges of the pyramids were coloured to aide the participants in determining object orientation and the edges were haloed to aide in depth perception. When a given vertex was moved within target range, the vertex would change colour. Once all four vertices were in place for 700 ms, the source pyramid would disappear and the participant could begin the next trial by pressing the start button. Each trial had a 40 second time limit, after which the trial was abandoned and the next trial automatically began.

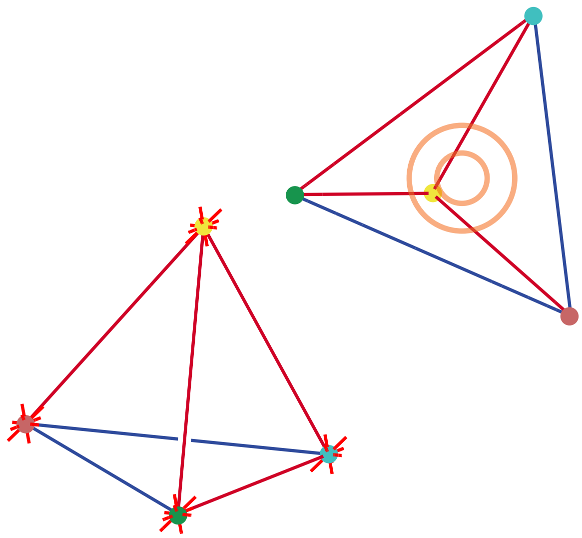


Figure 5.6: In the docking task, participants were asked to dock a pyramid object (right) in another pyramid (left).

Trials were repeated for three levels of difficulty and for two levels of orientation. The levels of difficulty varied the size of tolerance bars at each vertex on the destination pyramid—easy trials had a 54 pixel tolerance, medium trials 38 pixels, and hard trials 23 pixels. The two levels of orientation allowed us to compare the techniques' support for planar rotations with more complex rotations—planar rotations used a 135° rotation about the z-axis, and complex rotations used a 135° rotation about the x-y-z-axis.

Participants performed five repetitions of each combination of difficulty and starting orientation. Each time they began again with a new technique, participants performed six practice trials (each combination of difficulty and starting position was performed once).

Data from the docking task were analysed using a three-factor within-participants analysis of variance on the factors:

- *Technique*: one-touch, two-touch, three-touch;
- *Difficulty*: easy, medium, hard;
- *Rotation*: planar, spatial.

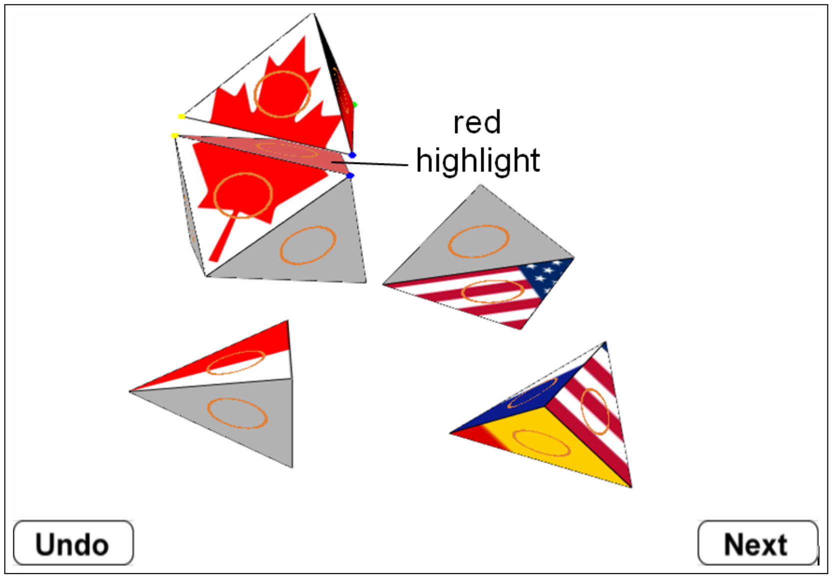


Figure 5.7: The puzzle task.

5.3.1.6 Task 3: Puzzle

This task was used to examine how the participants chose to use each of the techniques when completing a more realistic and less constrained task. Participants were asked to assemble a pyramid-shaped puzzle composed of four smaller pyramid shapes and a centre piece (see figure 5.7). Participants performed this task once for each interaction technique. Although software logged the participants' actions, data from this task was not formally analysed; our interest here was in observations of use, subjective preferences and comments about the techniques.

5.3.1.7 Order of Trials

In summary, each participant was asked to complete the following for each technique (in counterbalanced order):

- 18 passing trials in random order (6 sides × 3 destinations)
- 30 docking trials in random order (5 repetitions × 3 difficulties × 2 rotations)
- a puzzle task

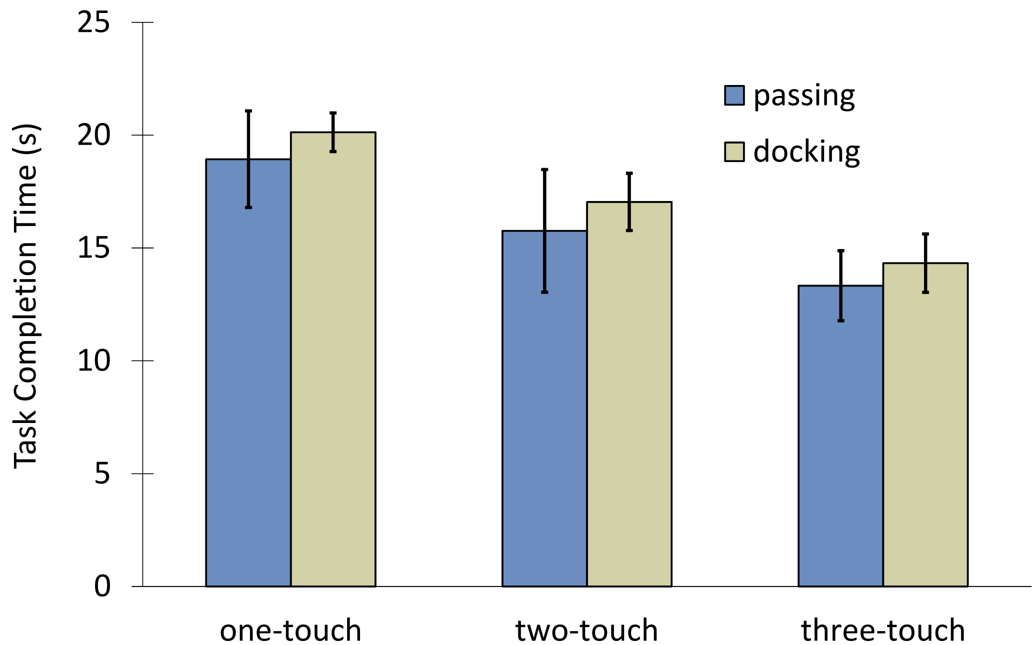


Figure 5.8: Mean TCT for both the passing task and the docking task. In both cases, three-touch interaction is fastest, followed by two-touch, and one-touch is slowest.

The random ordering of passing and docking trials was different for each participant and technique.

5.3.2 Results & Discussion

In this section, I present the results of the analysis integrated with a discussion of the implications of these results.

5.3.2.1 Task Completion Times

Data from the task completion times (TCTs) violated Mauchly's test of sphericity for the repeated measures analysis of variance (ANOVA). We therefore report results using the Greenhouse-Geisser correction (influencing df, F and p values).

TCTs in both the passing and docking tasks showed the same trend, with participants successfully exploiting the more expressive capabilities of the two- and three-touch interaction techniques. These results are summarised in figure 5.8.

Passing Task There was a marginally significant main effect of *technique* ($F_{1,16} = 3.4$, $p = .07$), with mean times reducing from 18.9 s ($SE = 2.1$ s) with one-touch, through 15.8 s ($SE = 2.7$ s) with two-touch, to 13.3 s ($SE = 1.5$ s) with three-touch; a 30% reduction in task time across the three conditions. Post-hoc pairwise comparisons only showed a significant difference between one-touch and three-touch techniques ($p < .01$). Despite the comparative efficiency of the three-touch technique, it is worth noting that even its mean task times were high—few tasks involving passing real objects would take this long, regardless of the level of precision required. We return to this issue in the discussion.

There was no significant effect of *destination* ($F_{1,18} = 0.06$, $p = .91$), nor were there significant interactions between it and the other two factors, suggesting that performance with the techniques is not substantially influenced by the direction of information sharing.

The *target side*, however, did have a significant effect on task performance ($F_{3,36} = 11.8$, $p < .001$). The time to attain the top-side target ($M = 10.6$ s, $SE = 1.7$ s) was markedly lower than all others (bottom: $M = 17.1$ s, $SE = 2.0$ s; left: $M = 17.2$ s, $SE = 2.1$ s; right: $M = 18.3$ s, $SE = 1.9$ s; back: $M = 17.3$ s, $SE = 2.3$ s; and front: $M = 15.5$ s, $SE = 1.6$ s). This effect is explained by the lack of rotation necessary with the top side as the target. Such tasks, therefore, predominantly involved translation and planar rotation rather than the more complex spatial rotations required with the other sides. Post-hoc analysis showed pair-wise differences ($p < .05$) between the top side and all other sides, and between the right and front sides. This latter difference is likely due to the combination of the facts that forward motion can more easily combine the required translation and rotation (advantaging discovery of the front side) and that our participants were right-handed, causing occlusion and disadvantaging trials involving the right side.

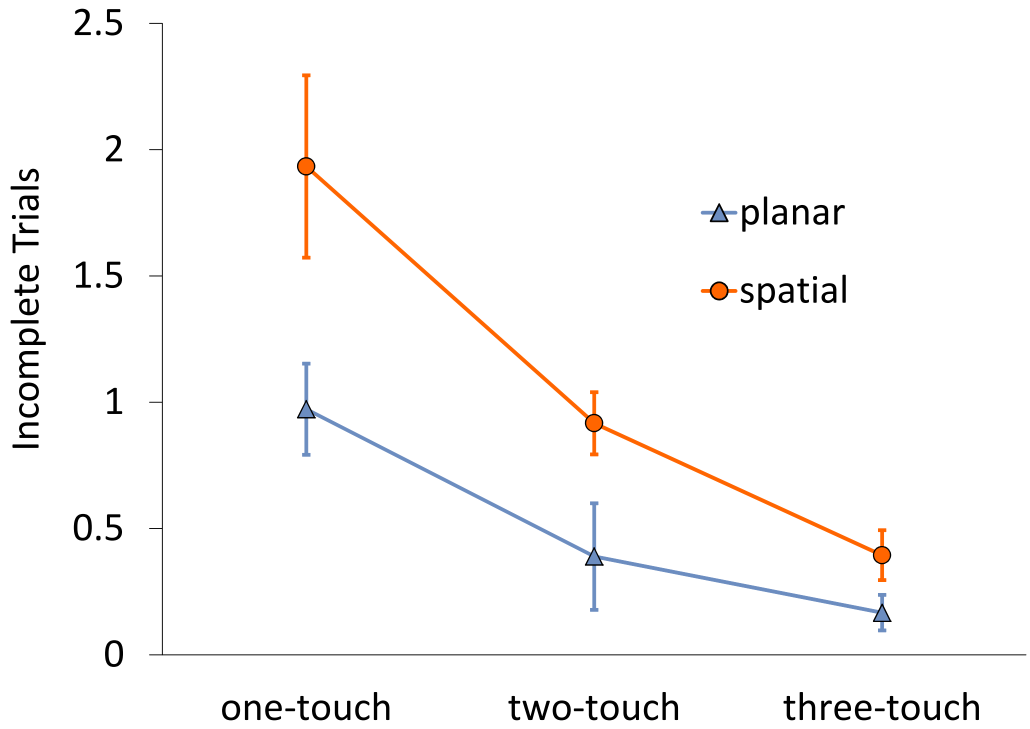


Figure 5.9: There was a significant interaction between technique and required rotation for the number of incomplete trials. For one-touch interaction, the difference between planar-only trials and trials requiring spatial rotation was larger than for two- and three-touch interaction.

Docking Task The results for the docking task showed similar trends to those for the passing task, but with stronger significance. Mean performance of the docking task with the three techniques improved significantly ($F_{2,19} = 14.2$, $p < .001$) as the number of touches increased from one to three. Means for the one-, two- and three-touch techniques were 20.1 s ($SE = 0.9$ s), 17.0 s ($SE = 1.3$ s) and 14.3 s ($SE = 1.3$ s) respectively (see figure 5.8), showing a similar overall percentage improvement between one- and three-touch (29%) to that observed in the passing task. Post-hoc pairwise comparison showed significant differences between one-touch and both others ($p < .01$), and a marginal difference between two- and three-touch ($p = .06$).

As anticipated, there was a significant effect of *difficulty* ($F_{2,20} = 39.6$, $p < .001$), with means rising from 15.0 s on easy tasks, through 16.9 s on medium ones, and 19.6 s on hard tasks (post-hoc pairwise significant for all pairs at $p < .01$). Somewhat surprisingly, though, there was no technique × difficulty interaction ($F_{3,29} = 0.5$, $p = .77$). We had anticipated that one-touch may suffer more than the other techniques on high precision tasks because it does not allow independent manipulation of each DOFO, but the data did not support this hypothesis.

Complex rotations ($M = 22.6$ s, $SE = 1.1$ s) were significantly slower than planar ones ($M = 11.7$ s, $SE = 1.3$ s): $F_{1,11} = 66.7$, $p < .001$. But again, there was no evidence that any of the techniques was particularly good or bad for complex manipulations (technique × rotation interaction, $F_{2,20} = 1.2$, $p = .33$).

Only tasks that were completed within the 40 s time limit were included in this analysis. To check that these results were not adversely influenced by different rates of incomplete trials in different conditions, we analysed the number of incomplete trials using the same 3×3×3 ANOVA. This analysis further supports the results above. Timed-out tasks were significantly more prevalent when using fewer points of contact ($F_{1,16} = 7.3$, $p = .01$), with means of 1.3 ($SE = 0.3$), 0.6 ($SE = 0.2$) and 0.2 ($SE = 0.1$) timeouts per condition with one-, two- and three-touch respectively. There were significant effects for difficulty ($F_{1,14} = 9.4$, $p < .01$) and rotation ($F_{1,11} = 14.8$, $p < .01$); but there was additionally a significant *technique* × *rotation* interaction ($F_{1,16} = 7.9$, $p < .01$), due to a much more dramatic increase in timed-out tasks between planar and complex tasks with one-touch than with two- and three-touch (figure 5.9). As before, the technique × difficulty interaction was not significant ($F_{3,31} = 0.8$, $p = .15$).

5.3.2.2 Characterising Interaction with the Techniques

The analysis above shows that the participants completed tasks more quickly when given more points of contact for interaction, and that the benefits of doing so become larger in more complex tasks. In order to better understand the strengths and weaknesses of each of

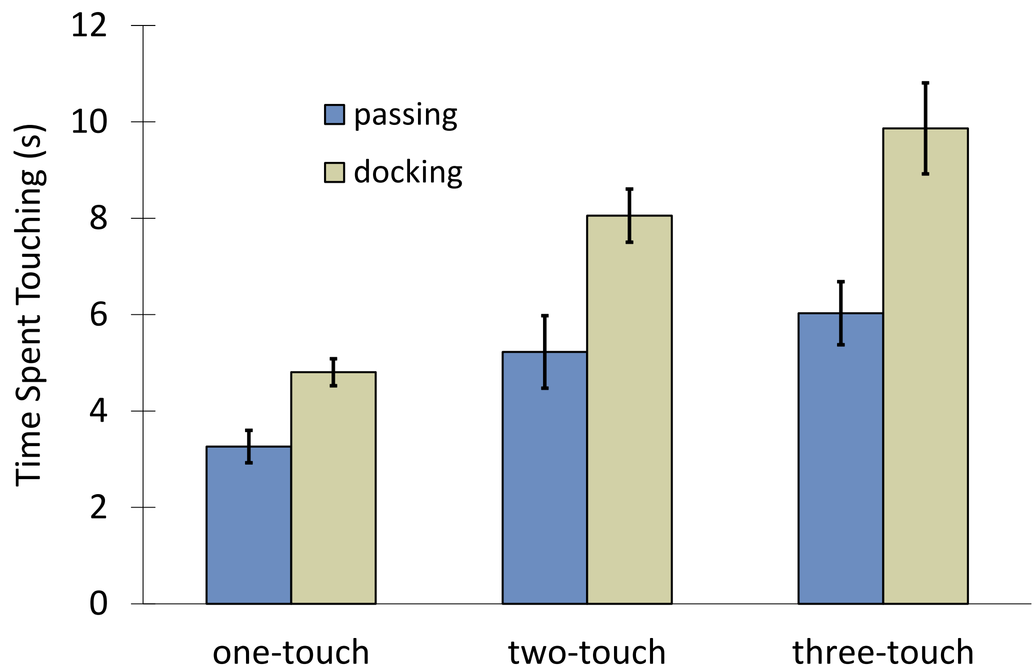


Figure 5.10: Mean time spent touching the object, separated into translations, planar rotations and spatial rotations.

the techniques for particular types of object manipulations, we now further scrutinise data on the time spent conducting particular types of manipulations, and the object regions used to do so.

To conduct this analysis we broke down the TCTs into time spent performing translation, planar rotation and spatial rotation. For the one-touch technique, this can be separated by time spent touching each dedicated area on the objects. For the two-touch technique, it is done by separating time spent inside and outside the translation-only area, and by measuring time spent using the second finger. For the three-touch technique, it is separated into time spent touching with each finger. Note that the sum of all movement types can be more than the TCT for the two- and three-touch techniques, since the participant can perform multiple movements at the same time.

We analysed the decomposed TCTs using a 3 × 3 within-participants ANOVA for factors technique (one-, two-, three-touch) and movement type (translation, planar rotation, and spatial rotation). We present the results for the passing and docking tasks together, as the main effects and interactions were sufficiently similar.

There was a main effect of technique (passing: $F_{2,18} = 8.5$, $p < .01$; docking: $F_{2,18} = 16.4$, $p < .001$). Figure 5.10 shows mean time spent for each technique. Post-hoc comparisons show that participants touched the screen significantly less with the one-touch technique than with the two-touch (passing (marginal): $p = .06$; docking: $p < .01$) and three-touch techniques (both: $p < .001$). The difference between the two-touch and three-touch techniques was not significant (passing: $p = .12$; docking: $p = .08$). This effect is in direct contrast to the main effect of technique for TCTS alone. This contrast suggests that participants spent more time performing cognitive processing than interaction with less DOFI and that this resulted in higher TCTS. Experimenter observations also confirmed that participants tended to have more difficulty with mental rotations when using the one-touch technique. Note, however that the measures fail to discriminate between manipulations that occur in parallel and in series, so this result should be cautiously appraised.

There was a significant interaction between technique and movement type (passing: $F_{2,23} = 18.7$, $p < .001$; docking: $F_{3,29} = 16.5$, $p < .001$) shown in figure 5.11. Post-hoc comparisons show that for one-touch interaction, participants spent significantly more time performing spatial rotations than either translations (both: $p < .001$) or planar rotations (both: $p < .001$) and that for three-touch interaction, participants spent significantly more time performing translations than either planar rotations (both: $p < .001$) or spatial rotations (both: $p < .01$). All other pairwise differences were not significant (both: $p > .05$). This interaction shows that participants typically spent an approximately equal amount of time performing rotations with all three techniques. Furthermore, the larger amount of translations in the three-touch condition may be because participants were able perform translations in tandem with the other types of rotation. Though it was not required by the hardware, we observed that

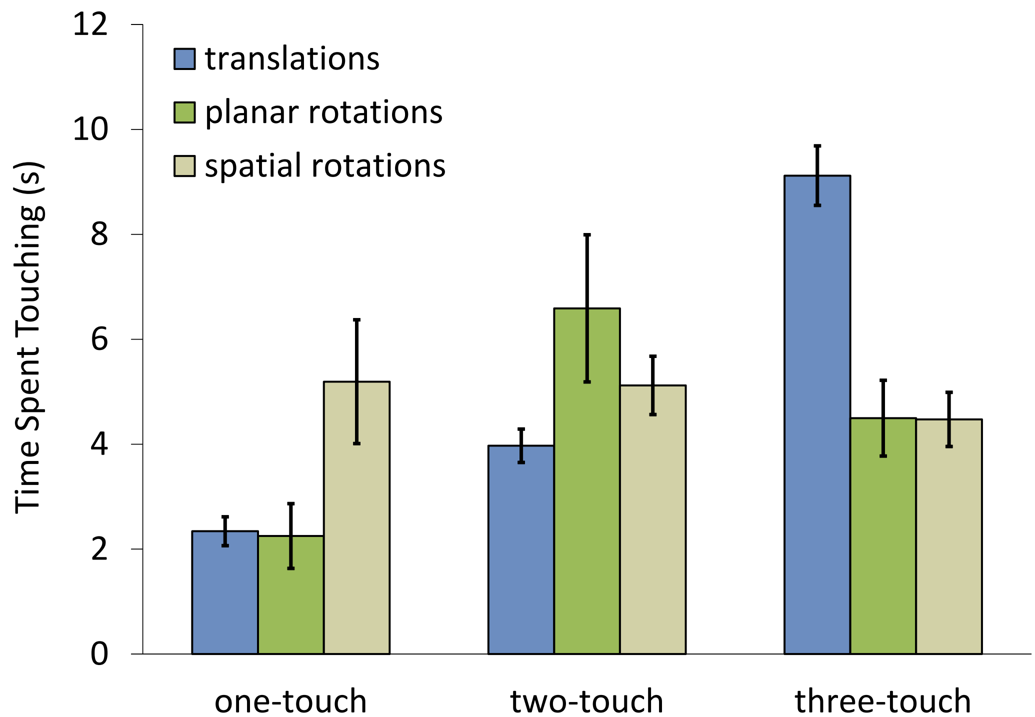


Figure 5.11: There was a significant interaction between technique and movement type for the passing task. For the one-touch technique, participants performed more spatial rotations than translations or planar rotations and for the three-touch technique, participants spent more time performing translations than planar or spatial rotations. The difference in movement type did not differ significantly for the two-touch technique.

participants tended to "hold" the object with their first finger while performing spatial rotations. This result illustrates very well that simultaneity of movements provides a strong advantage for multiple DOFI interaction.

Touches We observed during the experiment that participants tended to use object corners for spatial rotations much more with some techniques. We recorded the locations of every touch intended for spatial rotation made by each participant and rendered each point using a constant transparency. Patterns clearly show that for the one-touch technique, participants concentrated their touches at the corners and for the two- and three-touch techniques, the touch locations were more central. Figure 5.12 shows a typical face of both the

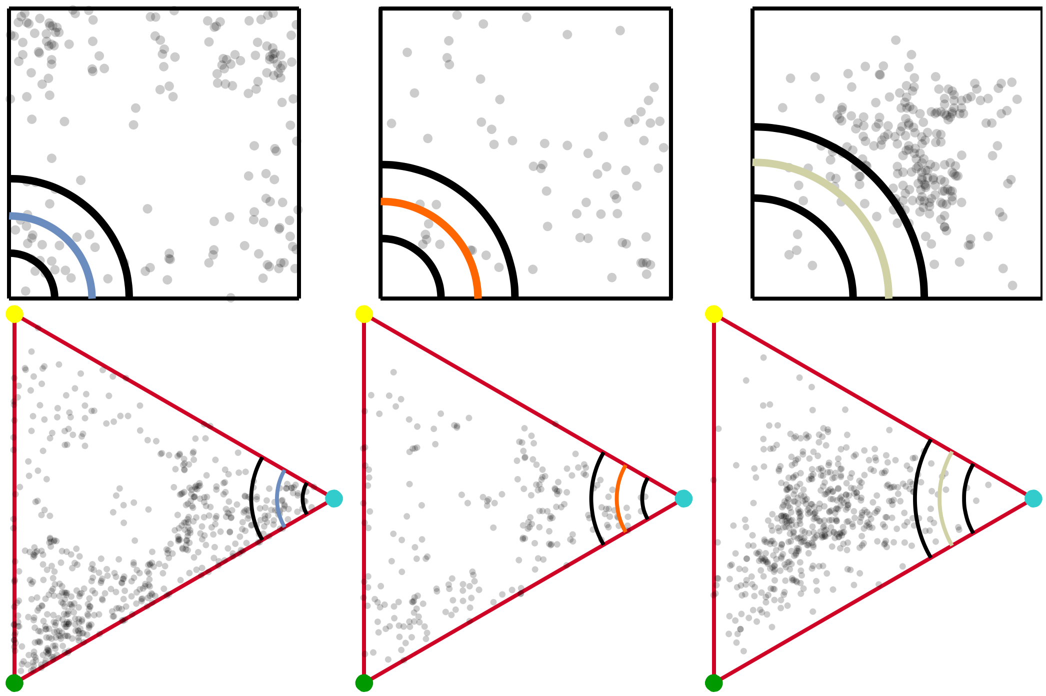


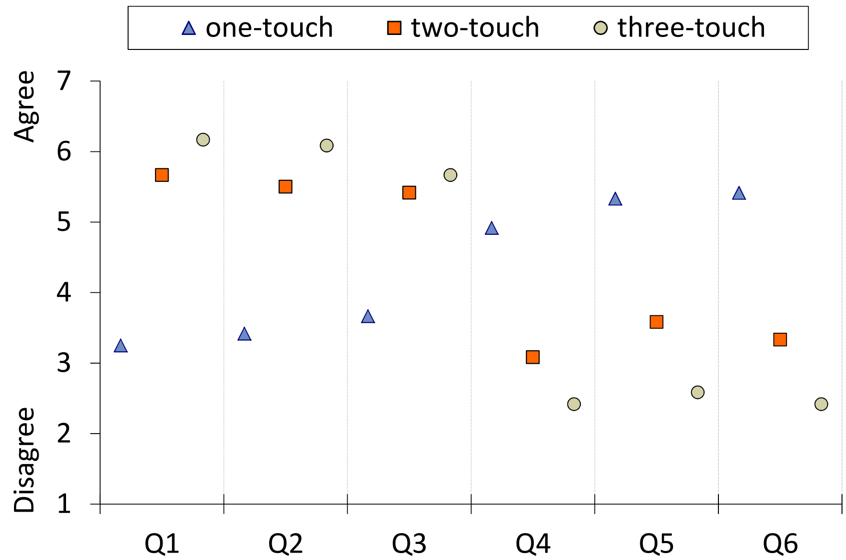
Figure 5.12: Touch locations on a typical face of the cubes in the passing task (top) and pyramids in the docking task (bottom) separated into one-touch (left), two-touch (middle), and three-touch techniques (right). The coloured arcs represent the mean distances to the nearest corner for each touch location, black arcs represent the standard deviation from these means.

cube from the passing task and the pyramid from the docking task for each technique. We also recorded the number of times the participants missed the objects completely and found that this occurred most frequently with the one-touch technique.

5.3.2.3 Subjective Ratings

Figure 5.13 shows the average scores on the follow-up questionnaire. For the docking task, 9 participants preferred the three-touch technique and 3 preferred the two-touch technique. For the passing task, 7 participants preferred the three-touch technique, four preferred the two-touch, and 1 preferred the one-touch technique. Overall, 7 participants preferred the three-touch, 3 preferred the two-touch, and 1 claimed there was no clear winner.

All subjective data shows a clear order of preference from three-touch (best), two-touch,



Q1	I found the technique easy to use
Q2	With this technique, the object reacted as I expected it to
Q3	I could easily move an object to where I wanted
Q4	I found the technique difficult to control
Q5	I found it difficult to turn objects in the plane
Q6	I found it difficult to roll objects over

Figure 5.13: Mean ratings and (standard deviations) on the follow-up questionnaire. Participants rated their level of agreement on a scale from 1 (strongly disagree) to 7 (strongly agree).

to one-touch (worst). Participants consistently rated the three-touch technique as the easiest to use (Q1) with the most appropriate reaction (Q2), as the least difficult to control (Q4) and rotate—both in the plane (Q5) and spatially (Q6). Also, the three-touch technique was most preferred for docking, passing and overall. The two-touch technique was rated second in all categories and the one-touch third, though with much higher variance.

5.3.2.4 Overall Discussion

Our study showed that the techniques that use a higher number of touches were better both in terms of performance and preference. These benefits likely appeared because the higher number of touches provided participants with the opportunity to independently control more DOFO. This type of freedom provides increased flexibility for how people decide to perform the interactions.

Our study showed that one-touch interaction was rated as difficult to use and resulted in the worst performance. This result implies that one-touch interaction (as designed here) was not efficient for interacting in shallow-depth 3D on the table. One response would be to redesign the one-touch interaction technique (one alternative is discussed below). Perhaps a more important consequence is that our results suggest that *multiple independent inputs for each person* at the table will be beneficial for both performance and satisfaction.

One concern we had when initially developing the techniques was that the complexity of multitouch interactions would prove confusing and deter people from its acceptance. In contrast, allowing separate and simultaneous control of rotation and translation provided a more preferred interaction with better performance. From watching people use these techniques, one could see that their interactions became more natural and easy as the number of touch points increased. Participants were not only capable of this more engaged, complex control, but prefer it.

The apparently contrasting effect that people spent more time actually moving and rotating the object with more fingers may in part be explained by the ability to support epistemic actions—"physical actions that make mental computation easier, faster, or more reliable" [Kirsh and Maglio, 1994]. These epistemic actions can facilitate problem solving, rather than requiring a person to only use internal thought processes. With the two- and three-touch techniques, participants may have been more able to perform these epistemic actions, and therefore more able to externalize some of the problem solving necessary for the task, which led to the combined result of more time spent moving and rotating with less total task time.

Generally, participants in our study were both intrigued and excited by all three techniques. This enthusiasm is likely due to the novel ability to use digital objects in a way that was more similar to their experiences with physical objects on tables. Participants commented that with these techniques it felt "more like I was touching it" and that "I almost want to look" under the objects. However considering the actual TCTS in comparison with what people are capable of with physical objects, there is considerable scope for future research refining these and other new techniques for manipulating shallow-depth 3D objects. Nonetheless, these techniques do provide the first steps toward enabling the more complex 3D interactions with which we are familiar.

5.3.3 Alternative Techniques

In light of the results of our study, we have explored alternative designs for our interaction techniques. Specifically, we believe that a redesign of the one-touch technique might make for a feasible method for interacting on tables incapable of multi-person, multitouch, direct-touch interaction. Furthermore, our multitouch techniques typically assign object transformations based on the movement of *every* finger. Another way of implementing bimanual, multitouch rotation would be to use the additional touches to introduce constraints that limit chosen aspects of the interaction. Such interactions have been shown to be an approach that people can easily cope with, due to the kinesthetic feedback [Sellen et al., 1992]. Lastly, we discuss how the use of the sixth DOFI could be used to provide the sixth DOFO (lift), while still maintaining our design consideration of connected manipulation (i. e., sticky fingers).

5.3.3.1 Alternative One-Touch Technique

The results of our experiment showed that, while spatial rotation interactions were accessible from both edges and corners, people typically made almost exclusive use of the corners. We also found that participants had difficulty acquiring the corners and would frequently miss the object entirely. In our new design, the 3D rotation previously available on the entire surface of the object is only allowed at the corners and the corner can be acquired by selecting anywhere inside a sphere about each vertex of the polygon. The object still has a translate-only region in the centre of each face, but the remaining parts of the object allow only planar RNT interaction. This new technique still benefits from the property that the selected point remains under one's finger.

5.3.3.2 Alternative Multitouch Technique

One of the disadvantages of both multitouch techniques used in our study is that the point of contact may not remain under one's finger once a rotation is performed with the finger on the non-dominant hand. We propose an alternative three-touch technique that constrains the effect of the primary finger based on the presence or absence of contact of the thumb and/or the finger on the non-dominant hand. When a person manipulates the object with their primary finger and no other finger is touching, the object reacts as it would in the one-touch technique. When both the thumb and the index finger are touching, planar rotation is performed as in the three-touch technique. A person can then limit the movement to translation-only by touching the table with a finger on the non-dominant hand. This technique also has the advantage that the point of contact remains under one's finger. It also corresponds to the way physical objects react, in that additional points of contact allow for more precise, constrained motion.

5.3.3.3 Sticky Fingers & Opposable Thumbs

We extend the three-touch technique to create a technique to manipulate all 6DOFO of the 3D virtual objects rendered using a perspective projection (see chapter 3). This technique controls the first 5DOFO (x, y, θ, ϕ, ψ) as described in section 5.2.5, and the change in z is controlled via the distance between the first two touches. That is, as the distance between the fingers gets larger, the virtual object moves towards the perspective viewpoint causing the object to appear larger (figure 5.14). The distinction between this method and that described in section 5.2.5 is that the first and second touches remain at the exact point on the 3D model that they first touch. Note that the virtual object's size in the 3D model will not change, only its distance to the viewpoint. The output for the z component of the technique would be a

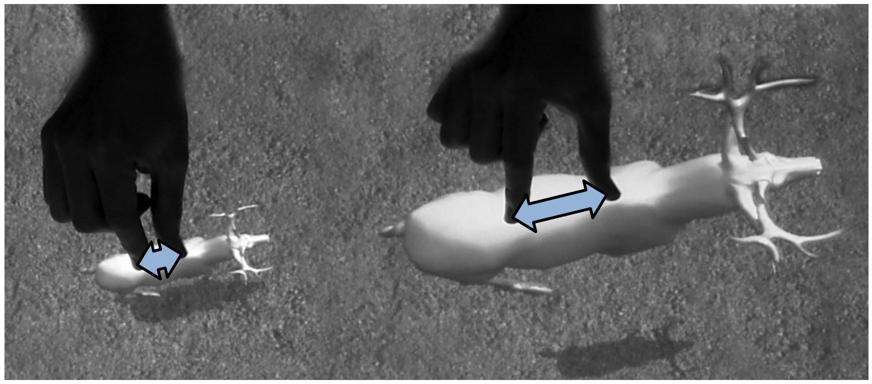


Figure 5.14: Lift in z for the sticky fingers and opposable thumbs technique is shown in this video. The still image is a diagram of this interaction.

simple change from:

$$\Delta z = |\overrightarrow{T_1' T_2'}| - |\overrightarrow{T_1 T_2}|$$

to:

$$z' = \frac{z|\overrightarrow{T_1' T_2'}|}{|\overrightarrow{T_1' T_2}|} \quad \text{(note: } T_1 \text{ is not used)}$$

With two sticky fingers alone, one cannot flip a virtual object over while maintaining the stickiness property, since the initial contact points are likely to become hidden. The third finger, which rotates about x and y as before, is the *opposable thumb*. Unlike actual thumbs, one can use any finger to provide the virtual flipping functionality that our opposable thumbs provide in the real world.

It is possible to maintain the stickiness property of the first two fingers when the third finger is active by using the axis defined by these two fingers as the axis of rotation. The disadvantage, however, is that movement along this axis with the third finger would not have any effect on the virtual object, and achieving the desired rotation may require defining a new axis of rotation (by lifting one's fingers and reselecting the axis with the first two fingers). This disadvantage led to the design decision to use relative interaction for the third touch for 3D rotations, instead of maintaining connected manipulation.

Additions When using this technique to select individual objects, it is not always clear when different contact points have different behaviours (e. g., what happens when the second touch is on another object?). However, this is easily solved by requiring that each touch starts on the object that it relates to. This allows the system to assign the order per object, making it possible for multiple people, fingers, or hands to control different objects simultaneously (or even for multiple people to control a single object together, if desired).

Because a touch has to start on the object it relates to, it can be difficult to manipulate small objects. To alleviate this problem, we implement 'crossing' (similar to the idea presented by Apitz and Guimbretière [2004]). The finger can be swept across the table surface, having no effect until it touches an object, and is assigned to an object as soon as it touches it. Because an object starts to move as soon as a finger touches it, this is another way in which the finger can be considered sticky.

5.4 Chapter Summary

In this chapter, I presented design guidelines for creating interaction techniques for manipulating 3D virtual objects. I designed and implemented one-, two-, and three-touch techniques that make use of these guidelines and performed an empirical evaluation to compare these three designs. The results showed that people preferred and were faster when using more fingers. Specifically, when the DOFI was closer to the DOFO, people were more easily able to control the 3D virtual objects. Furthermore, I described further iterations of these designs in light of the results of the study.

Specifically the contributions from this chapter are:

- A study providing evidence that using more fingers to control the six degrees of freedom of the output (DOFO) of manipulation improves performance and is preferred.

- The design, implementation, and evaluation of a set of interaction techniques that use multitouch to manipulate 3D virtual objects.

Since the original publication of this work [Hancock et al., 2007], the availability of multitouch hardware has increased dramatically, and other techniques have been introduced to enable 3D virtual object manipulation [Wilson et al., 2008; Reisman et al., 2009; Hilliges et al., 2009]. However, while these other techniques themselves have many advantages, none have been shown to be empirically better than the three-touch technique described here, nor do any provide the complete control of all 6DOFO of 3D virtual object manipulation using only the multiple touch points available with the more recent hardware (and not, for example, the space above the table), as the sticky fingers & opposable thumbs technique provides.

In the remainder of this dissertation, the designs presented make use of the results of the study on interaction described in this chapter. Specifically, chapter 6 describes a framework for the design of 3D tabletop applications that includes the use of the sticky fingers and opposable thumbs technique, which is based on the three-touch technique found to be fastest and most preferred. Chapter 7 describes how to apply this framework and makes use of this same technique in the application domain of sandtray therapy.

Force-Based Interaction / Sticky Tools 6

In this chapter, I describe *sticky tools*—a combination of the 3D visuals that were found to avoid the perceptual errors discovered in chapter 3, and the sticky fingers & opposable thumbs interaction technique presented in chapter 5. I use sticky tools to demonstrate how virtual 3D artifacts can take on a large variety of meanings and that, by providing this variety of meanings, they can be used to create rich 3D tabletop display environments—applications that can perform functions beyond movement and rotation of the 3D artifacts themselves. I then provide a general framework to illustrate how manipulating virtual objects can fit within the larger context of physical force-based interaction.

In the physical world, an object reacts to a person's actions depending on its physical properties and the forces applied to it. For example, a book can be stacked on top of another because it has two flat sides or a pencil can be rolled along a desk because it is cylindrical. People often make use of the unique properties of objects to make them affect other objects in different ways. People use pencils to write, hammers to insert nails, and utensils to cook food. In the virtual world, how objects react to human intervention depends on a particular mapping of physical movement to computer feedback (as described in chapters 4 and 5). There are benefits to both worlds; in the physical world, people become familiar with the capabilities of the tools they use regularly; in a virtual world the result of a person's actions can be made to either use or extend physical limits.

Since tabletop displays afford direct touches for interaction, the techniques have a feeling of being more physical than, for example, mouse or keyboard interaction. This directness

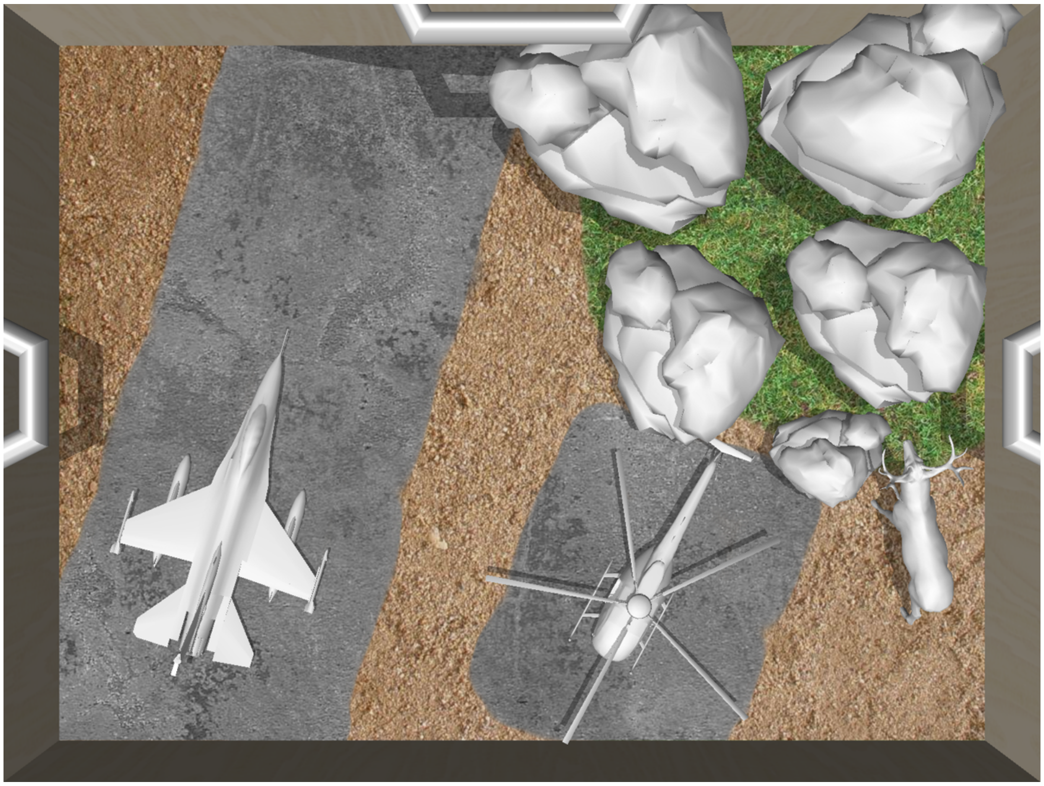


Figure 6.1: A screenshot of a 3D virtual scene.

of interaction with virtual objects opens up the potential for interactive tables to simultaneously leverage the benefits of both the physical and the virtual. A significant portion of the techniques that have been designed specifically for digital tables are based (either explicitly or implicitly) on how objects react in the physical world. However, these techniques typically resort to techniques such as symbolic gestures [Agarawala and Balakrishnan, 2006; Wu and Balakrishnan, 2003] or menus [Shen et al., 2004] to provide full functionality. Thus, the following research question arises: how does one maintain the feeling of physical interaction with the full capabilities to manipulate a 3D scene such as in figure 6.1?

I introduce sticky tools to allow force-based interaction to provide full control of a system, without the need for gestures or menus. I first introduce sticky tools, and then demonstrate how sticky tools can be used to assign richer meanings to virtual objects. The chapter ends with a discussion of how sticky tools leverage the largely unexplored research area of how virtual objects interact with other virtual objects and describe how this research direction can overcome existing limitations in tabletop interaction.

6.1 Sticky Tools

The sticky fingers and opposable thumbs technique described in chapter 5 enables the full control of a single object in a 3D virtual scene. *Sticky tools* combines this interaction technique with the concept of virtual tools to take this full control of a single object and use it to enable full functionality within that system. Thus, sticky tools are a mechanism to improve upon existing force-based interaction techniques so that they can provide full functionality to a multitouch table, without the need for symbolic gestures or menus.

6.1.1 Virtual Tools

While together sticky fingers and opposable thumbs provide a way to select and fully manipulate a single virtual object in either 2D or 3D, more complex interactions, such as using an object to push another object around or changing an object's properties (e. g., density, colour, etc.) are not possible. We introduce virtual tools to enable more complex interactions on virtual objects. A virtual tool is a virtual object that can act on other virtual objects and is able to cause changes to the receiving object. Any virtual object that is controlled with sticky fingers and opposable thumbs becomes a *sticky tool*.

While virtual tools can exist in any virtual environment, we[1] realized our virtual tools within a simulated real world by using a physics engine [NVIDIA Corporation, 2009]. Thus, when a person interacts with a virtual object, it is placed under kinematic control so that other virtual objects will react physically to its movement, but the contact with the sticky fingers gives control of the object to the fingers. Thus, the object can now be used to hit

[1] This framework was also published in Hancock et al. [2009b], and so the use of the first-person plural in this chapter refers to: Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale.

other objects, but will not be knocked from the sticky contact. When the sticky tool makes contact with another object, it can cause physically familiar behaviour but these contacts can also be detected and made to invoke abstract actions, such as re-colouring the object.

The concept of sticky tools is useful in explaining previous work. The technique introduced by Wilson et al. [2008] can be thought of as an example of a very simple virtual tool. Their interaction technique can be described as controlling the 2D position of many invisible virtual objects and these invisible objects interact with other objects in the scene through the use of a physics engine. In this framework the proxies can be considered to be a virtual tool whose behaviour is always to invoke frictional and opposing forces on other virtual objects. Similarly, the joint technique used in BumpTop [Agarawala and Balakrishnan, 2006] allows 3D icons to act as virtual tools that cause collisions that invoke behaviour on other 3D icons.

Table 6.1 shows a comparison of sticky tools (ST) to the features of the joints (J) and proxies (P) techniques provided by Wilson et al. [2008] (discussed in section 2.4.2.3) to the sticky fingers (SF) and sticky fingers with opposable thumbs (SF+OT) techniques presented in chapter 5. They are compared on many commonly provided multitouch interactions. Sticky fingers and opposable thumbs offer a more complete set of these interactions than any other; however, all have some gaps and this is not a complete list of all possible functionality. For any of these approaches the gaps can be addressed by virtual tools. That is, with virtual tools the functionality of any of the unchecked cells in table 6.1 can be enabled. For example, sticky fingers and opposable thumbs can use a virtual tool to push or surround other objects. This is also true for the joints technique or sticky fingers alone (without opposable thumbs). A virtual tool could be used in combination with either the joints technique or the proxies technique to lift objects in the third dimension. For example, a platform could be introduced that objects could be moved onto. The platform could then be used to lift the objects through use of a dial, a slider, or elevator virtual object. Similar virtual objects could also be imagined that could enable flipping and spinning of virtual objects. Virtual tools

Feature	J	P	SF	SF +OT	ST
Lift (move in z)			✓	✓	✓
Drag (move in x & y)	✓	✓	✓	✓	✓
Spin (rotate about z)	✓		✓	✓	✓
Flip (rotate about x / y)				✓	✓
Push		✓			✓
Toss	✓		✓	✓	✓
Surround		✓			✓
Additional Points		✓			✓

Table 6.1: Comparison of different techniques for interacting with 3D virtual objects on a table. The columns are abbreviated as: joints (J), proxies (P), sticky fingers (SF), sticky fingers with opposable thumbs (SF+OT), and sticky tools (ST).

also offer new potential for additional functionality not possible with any previous single technique.

6.2 Understanding Virtual Objects

In essence, the difference between the use of virtual tools and previous techniques comes down to the ability to assign richer meaning to virtual objects. This assignment of meaning is analogous to a similar discussion introduced by Underkoffler and Ishii [1999] for their luminous tangible system. They showed how tangible objects could be assigned richer meaning to expand interaction possibilities. We parallel their discussion on luminous tangible object meanings with a discussion on virtual tool object meanings. The discussion on virtual tool meanings is followed by a generalized model of how force-based interaction can be used to provide all this functionality by changing the complexity of how people control both physical and virtual objects, as well as how those physical and virtual objects can control each other.

6.2.1 Virtual Object Meanings

In this section we provide examples of how using virtual objects to control other virtual objects can enrich interaction. We demonstrate this richness by mirroring Underkoffler and Ishii's [1999] description of how tangible objects can take on different object meanings along the spectrum:

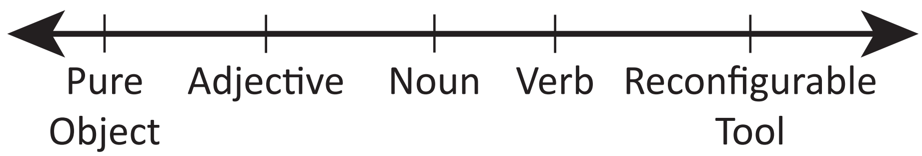


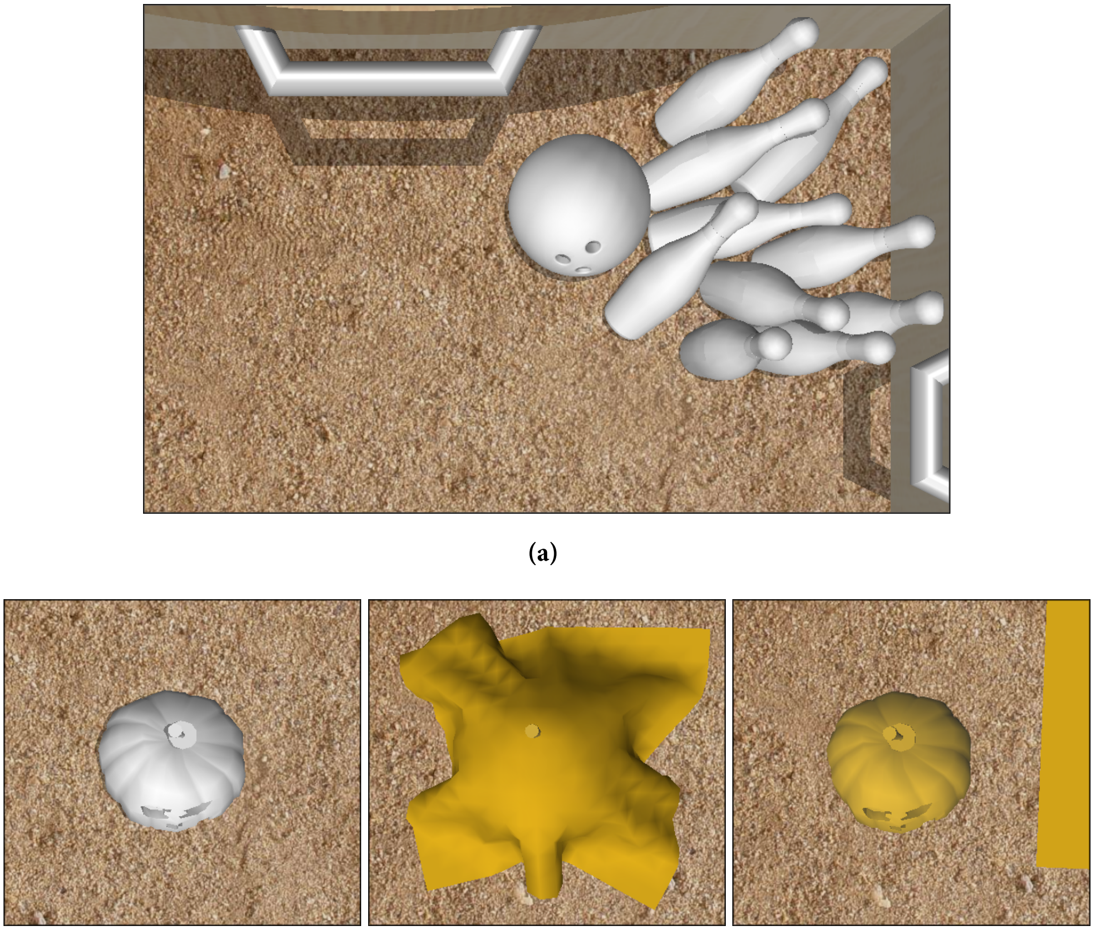
Figure 6.2: The spectrum of object meanings used in Urp [Underkoffler and Ishii, 1999] to describe tangible devices.

In each of the following subsections, we first state the definition used by Underkoffler and Ishii to describe the different object meanings (with modifications so that they describe a virtual environment, instead of a tangible system) and then describe an example of a sticky tool whose meaning can be interpreted using this definition. We thus show that sticky tools enable virtual objects to take on all of the possible meanings of tangible luminous objects.

6.2.1.1 Virtual Objects as Nouns

> These objects occupy the center of the axis and are likely the most obvious in their behavior. They are fully literal, in the sense that they work in their [virtual] context very much the way objects 'operate' in the real world—an Object As Noun exists in our applications simply as a representation of itself: an immutable thing, a stand-in for some extant or imaginable part of the real-world. [Underkoffler and Ishii, 1999, p. 392]

A virtual object as a noun stands for itself—in the virtual world, that is, for what it appears to be. Thus, if it looks like a ball it should behave like a ball. In a virtual 3D environment, we



(a)

(b)

Figure 6.3: Virtual objects as (a) nouns (shown in video and still) and (b) verbs.

can render any mesh of triangles that has been modeled. Thus, rigid bodies of virtually any shape can be added to the environment and made to interact with other rigid bodies using the physics engine. Thus, these virtual objects can operate in the virtual world in a way that is similar to how they behave in the real world. For example, a set of bowling pins in the environment can be knocked over using a virtual bowling ball (figure 6.3a).

6.2.1.2 Virtual Objects as Verbs

As we move to the right along the continuum, away from Object As Noun, inherent object meaning is progressively abstracted in favor of further—and

> more general—functionality...It is not understood as 'present' in the [virtual] world...but exists to act on other components that are, or on the environment as a whole. [Underkoffler and Ishii, 1999, p. 392]

A virtual object as verb exists as a virtual object but embodies actions. That is the appearance of the object symbolizes the possibility of an action. In our virtual environment, we include a cloth that embodies 'wrapping' (figure 6.3b). Dropping a cloth on another object wraps that object. We leave evidence of this wrapping by changing the affected object's colour, providing a way to colour objects. The act of covering another virtual object with a cloth can be further abstracted to provide a variety of different functions. We also provide a lamp sticky tool that embodies the actions of shedding light, casting shadows, and can be used as the sundial in Urp to simulate changing the time of day. This sticky tool differs from the tangible device in Urp in that the lamp can be made to disobey the law of gravity and to pass through other objects in the environment.

6.2.1.3 Virtual Objects as Reconfigurable Tools

> This variety of object-function is fully abstracted away from 'objecthood', in a way perhaps loosely analogous to a GUI's mouse-plus-pointer. [Underkoffler and Ishii, 1999, p. 392]

A virtual object as a reconfigurable tool is an object that can be manipulated to affect other objects. It does not stand for itself as a noun, or imply an action as a verb, but instead symbolizes a functionality. We create a compound sticky tool consisting of a drawer object and a dial (figure 6.4). When a figurine is placed inside this drawer, the dial can be rotated to grow or shrink the figurine. This compound sticky tool could be reconfigured to perform any action that involves changing a one-dimensional property of another virtual object along a continuous axis. For example, it could be used to change an object's density or elasticity.

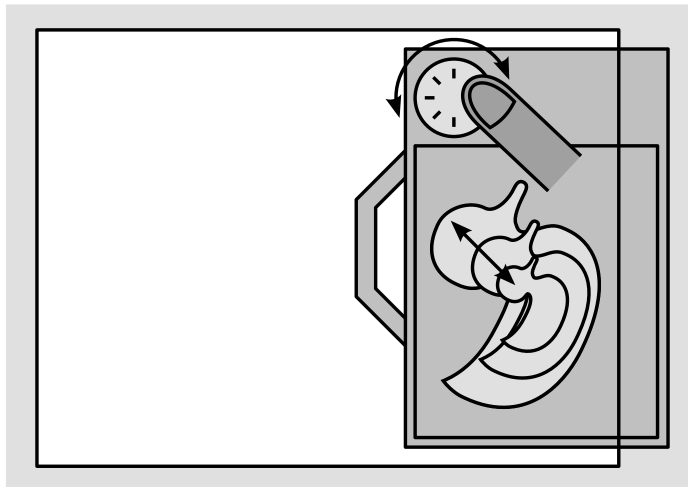


Figure 6.4: A video and still image of virtual objects as reconfigurable tools.

6.2.1.4 Virtual Objects as Attributes

> As we move to the left away from the center of the axis, an object is stripped of all but one of its properties, and it is this single remaining attribute that is alone considered by the system. [Underkoffler and Ishii, 1999, p. 392]

A virtual object as attribute represents one and only one of its attributes. For an example, we create another compound sticky tool for painting the background of the environment (figure 6.5). This sticky tool includes a group of four buckets that each contains a different texture and a hose that extends from below the buckets. In the case of the bucket, the only attribute that matters is its texture. The shape, size, density, location and all other attributes are abstracted from this virtual object. To paint the background a person selects a bucket with one finger to activate the hose and then, with the other hand, can move the hose's nozzle to indicate the area of the background to paint. Movement in the z-direction affects the area of influence of the hose (the farther from the background, the larger the radius of influence). Touching the texture bucket activates the texture that flows along the hose into

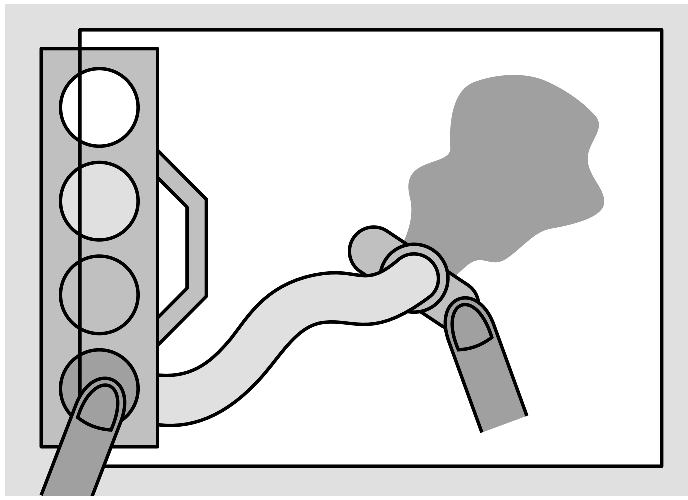


Figure 6.5: A video and still image of virtual objects as attributes.

the environment.

6.2.1.5 Virtual Objects as Pure Objects

> This last category is the most extreme, and represents the final step in the process of stripping an object of more and more of its intrinsic meanings. In this case, all that matters to a [virtual] system is that the object is knowable as an object (as distinct from nothing). [Underkoffler and Ishii, 1999, p. 392]

A virtual object as pure object is a symbol and stands for something other than itself. We create a sticky tool that allows the storage of the locations of all of the figures in a scene to be symbolized by a pure virtual object. Which virtual object will perform this symbolic function is established by placing an object in a "save" drawer. Thereafter, the scene is essentially stored in this virtual object and can be reloaded by placing that same figure in the empty environment. Thus, any virtual object can be stripped completely of its intrinsic meaning, and the locations of the remaining virtual objects can be, for example, "put in the dinosaur". That is, the dinosaur now stands for the scene.

6.2.2 Force-Based Interaction

We have introduced sticky tools and discussed them in relation to joints, proxies, particle proxies, and tangible devices. In this section, we generalize from these approaches to provide a framework that encompasses these techniques and indicates how existing functionalities in force-based interactions can be expanded.

Using physical forces to control virtual objects has the appeal of being easy to understand and learn due to our ability to transfer knowledge from our experience in the physical world. However, in order to simulate physical behaviour in the digital world, two primary components are required: a sensing technology, and a display technology. The sensing technology takes actions from the physical world and translates them into messages that can be understood by the computer, and the computer can then translate those messages into something virtual that can be understood as a physical reaction to the initial action.

In sensing and translating this information, there are several places that the complexity of the force-based action-reaction can vary. First, new sensing technologies can be invented to be able to identify more and more complex physical forces. Essentially, the computer can become better at understanding how people control physical objects (in multitouch, through a person's fingers or in tangible, through a person's use of a physical object). Second, as seen in our sticky tools, Wilson et al. [2008] and Agarawala and Balakrishnan [2006] the mapping from what is sensed to the system response can be made to include complex physics algorithms that better simulate real-world responses. Third, a largely unexplored possibility is the introduction of complexity through how the system's response propagates in the virtual environment. That is, virtual objects can control other virtual objects.

These real and virtual interaction possibilities can be summarized by: (1) *people controlling physical objects*, (2) *physical objects controlling physical objects*, (3) *people controlling virtual objects*, (4) *physical objects controlling virtual objects*, and (5) *virtual objects controlling*

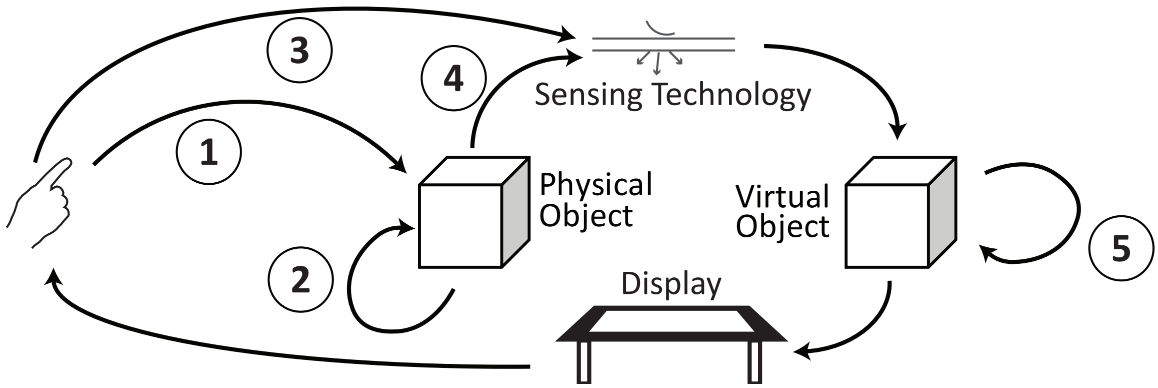


Figure 6.6: A diagram of the five major components of force-based interaction.

virtual objects. Figure 6.6 provides a diagram of this space. The first four aspects have been well-researched; the introduction of virtual objects controlling virtual objects (i. e., sticky tools) is a contribution of this thesis.

6.2.2.1 People Controlling Physical Objects

We are accustomed to interacting in the physical world and are adept at controlling physical objects with fine motor movements. It can be advantageous to leverage these natural abilities when creating computer interfaces. Although the interactions that are available in the physical world are arguably highly complex, they are familiar. From the interaction design perspective, this familiarity makes it easier to predict what a person might expect. For example, the designer might predict that people will expect large objects (i. e., objects with more mass) to require more force to push than smaller objects.

6.2.2.2 Physical Objects Controlling Physical Objects

Some tasks require more precision or more power than most people's physical abilities. For these tasks, we use tools (hammers, levers, needles). Good design can make a tool that can be both useful and usable [Norman, 1998], making it possible for people to extend their physical capabilities.

6.2.2.3 People Controlling Virtual Objects

To enable a person's actions to cause reactions in the virtual world, the physical movement of the human body must be sensed in some way. This sensed information has typically been mouse movement or a key press, though many more complex devices exist [Fröhlich et al., 2006; Zhai, 1998]. The sensed information can be then used to cause a virtual action. However, the necessary translation can interfere with familiarity. The interactions we learned from the physical world may not predict virtual interactions. However, re-introducing the familiarity of physicality is considered a positive goal, which with current massively multi-touch capabilities seems increasingly possible. A variety of techniques now approach this goal [Kruger et al., 2005; Liu et al., 2006; Shen et al., 2003, 2004; Hancock et al., 2006; Cao et al., 2008] and recently have been extended to provide more physically realistic interactions with 3D virtual objects (e. g., chapter 5; [Wilson et al., 2008]).

6.2.2.4 Physical Objects Controlling Virtual Objects

Tangible computing focuses on how physical objects can be used to control virtual objects. This line of research suggests that the richness of interactions with physical objects can be leveraged by using them directly in an interface. The use of physical objects to control virtual objects relates to the human tendency to use physical tools to control other physical objects. Tangible computing devices can be seen as a form of tool that provides a mechanism for designers to introduce complexity. One of the disadvantages of using physical objects to control virtual objects is the need to sense the behaviour of the physical object.

6.2.2.5 Virtual Objects Controlling Virtual Objects

As noted in the discussion of virtual tools, virtual objects can be used to control other virtual objects. Since these are virtual objects, they are already fully described computationally. These virtual objects can then be assigned meaning in the same way that tangible devices can; they can be used as pure objects, attributes, nouns, verbs, and reconfigurable tools [Underkoffler and Ishii, 1999].

Interface components in general are virtual tools, however, the simulation of physical forces implies that virtual objects could have the capabilities of real physical objects—where objects behave as themselves (nouns) and can be used to act (verbs). The use of virtual objects to control other virtual objects expands the methods for creating complex interactions. One direction is to leverage people's familiarity with physical objects through use of the software support of a physics engine. Another is to take advantage of the fact that, in a virtual world, physical laws do not have to be obeyed. These two ends of an interaction spectrum of course have rich possibilities of combinations. An important factor in the potential of virtual tools is that since these objects are already virtual, there is no need for a sensing interface between the action and reaction. Now that there is a simple interface between a person and a virtual object in a 3D forced-based environment the possibility of exploring the potential of virtual tools is open.

6.3 Creating Applications with Sticky Tools

The goal of this chapter is to demonstrate how the combination of 3D visuals and 3D interaction can be used to create real applications. Section 6.2.1 demonstrates that sticky tools can take on a variety of meanings and section 6.2.2 describes how sticky tools fit inside the larger context of physical interaction (in the real world and with tangible user interfaces (TUIs)). The conclusion that can be drawn from this chapter is that, by using sticky tools, the creation of tabletop display applications can more closely resemble the design of physical objects and physical interactions. That is, where the HCI challenge has perhaps traditionally been to bridge the gap between a human and the capabilities of the computer, this framework presents the challenge of creating and designing virtual artifacts whose meaning is established through their use.

For example, an interface designer might be faced with the challenge of creating an interface to allow a person to draw a straight line. A typical solution might be to provide a menu or gesture to select the "straight line" command that, once selected, allows the user to provide the start and end points of the line [e. g., Kurtenbach et al., 1997]. In contrast, a design that makes use of sticky tools might introduce a "straight edge" virtual tool and a "pencil" virtual tool that can both be manipulated using sticky fingers and opposable thumbs. To create a straight line, a person can move and rotate the straight edge as desired and then drag the pencil along that straight edge.

Thus, the HCI challenge introduced in this chapter requires both the design of virtual tools (e. g., through a computer automated design (CAD) program) together with some thought about how the virtual objects will respond to one another. Note that physically impossible tasks are still possible in the digital world through the interaction between the virtual tools, instead of between the person and the device. Note also that this framework makes functionality beyond movement and rotation possible, however this framework may not be the most appropriate approach for some applications. For example, applications that require a significant amount of text-entry may be more suited to a keyboard interface, applications that require a person's attention to be somewhere other than the interaction space (e. g., when driving) may be hindered by an interface that requires so much visual attention, and an interface which has limited screen-space (e. g., a mobile phone) may be difficult to use with three fingers. Despite these limitations, there are many situations that could benefit from force-based interaction (e. g., chapter 7).

6.4 Chapter Summary

In this chapter, I introduced sticky tools, which combine the 3D visuals from chapter 3 with the 3D interaction techniques from chapter 5 to provide full functionality to a tabletop display application. I describe how sticky tools can provide meaning along the entire spectrum introduced by Underkoffler and Ishii [1999] and then situate virtual object interaction within a larger force-based interaction framework.

Specifically the contribution from this chapter is:

- A description of how to combine 3D interaction with 3D visuals to control 3D virtual tools, thus providing the ability to do more complex actions in a virtual world. This description provides a framework for how to create 3D tabletop applications.

This chapter demonstrates in an abstract way how sticky tools can enable full functionality within a tabletop display application, without the need for gestures or menus. Chapter 7 provides an example of how to use this framework in designing a real application. Specifically, the next chapter demonstrates the use of sticky tools and force-based interaction to provide a digital sandtray to be used for art therapy—a form psychotherapy in which the client creates art as a medium for the therapist to gain insight about the client's psyche.

Case Study: Digital Sandtray Therapy 7

In this chapter, I describe a case study that demonstrates how to apply the concepts from chapters 3 to 6 in the design of a tabletop display application. This case study involves the cooperative design of a digital sandtray for child psychotherapy. The final prototype makes use of a perspective projection with a very distant COP directly above the table. This is a direct application of the results from chapter 3 in order to support the collaborative aspect of this application: the client-therapist communication. The prototype also uses the sticky fingers & opposable thumbs technique (chapter 5), and virtual tools (chapter 6) to further support this communication. More specifically, this combination provides the potential for the virtual objects to take on meaning relevant to the therapy.

First, I provide some background about physical sandtray therapy and motivation for why the digital medium is potentially beneficial. I then describe the cooperative design of a virtual sandtray as it has iteratively evolved in collaboration with three therapists who make use of sandtray in their therapeutic practices. This design is followed by a discussion of how sticky tools made therapy possible in this digital medium combined with some suggested improvements to the design. I conclude the chapter with a description of the implications from this case study to other tabletop display application designs.

7.1 Motivation for a Digital Sandtray

Social workers and therapists are developing new ways of reaching and working with children who are troubled or have experienced traumatic events and difficult life situations. One

Figure 7.1: A sandtray typically used for sandtray therapy. Sandtray therapists typically observe clients creating a scene or "playing" in the sand to gain insight into their psyche (courtesy of Walter [a]).

such method is sandtray therapy [Kalff, 1980; Bradway and McCoard, 1997; Cunningham, 2009]—a type of play or art therapy [Carpendale, 2009] in which the use of figurines in a tray of sand is a vehicle for establishing interaction and rapport between the therapist and the child (figure 7.1). Children placing and moving the figurines in the sandtray provides a venue by which therapists can observe the manner in which the child thinks about their experiences and feelings. In response to an idea from a practicing sandtray therapist, we[1] developed a virtual sandtray. In this paper we present its iterative evolution in collaboration with three therapists who make use of sandtray in their therapeutic practices.

There were many reasons for the request that triggered this research. Sandtray therapy, often considered a type of art therapy [Carpendale, 2009] because the therapy session involves the clients creating a scene out of available supplies, has particular characteristics

[1]This chapter presents material on the participatory design that was also published in Hancock et al. [2010], and so the use of the first-person plural in this chapter refers to: Mark Hancock, Thomas ten Cate, Tobias Isenberg, and Sheelagh Carpendale, as well as the sandtray therapists involved in our study, who will go unnamed for confidentiality purposes.

which make it well suited to an interactive tabletop. These include factors of age stereotyping, the characteristics of the sand itself, and the types of interaction that are therapeutically beneficial. In terms of age stereotyping, the associations of sandtrays (or sandboxes) are with activities we did as young children. While this works well in establishing rapport with young children, it can pose problems with young teenagers and pre-teens to whom activities in a sandbox may seem just too 'uncool'. This age group (10 to about 13 or 14) is a particularly difficult age group for therapists to reach and is also a particularly sensitive age group since so much is changing in their lives at these ages. One of the therapists we are working with suggested that the 'wow' factor of a digital table might prove a great bridging factor. Also, some children, perhaps due to their own response to traumas, dislike the feeling of sand and refuse to play with it and may find a digital sandbox more to their liking.

Perhaps most important is the potential for interaction; sandtrays offer special interaction advantages. For example, other forms of media used in art therapy [Carpendale, 2009], such as pencils, paints, and clay, result in the client creating a 'snapshot' (e. g., pencil drawings, paintings, sculptures) as an end result. In contrast, the temporary nature of the sand invites play and, therefore, the creation of a narrative. The process of creating a 'scene' containing several objects, in which the narrative can unfold, can be particularly informative to the therapist, who can often infer self-representation in one object and, from there, the relationships to other objects in the scene. Thus, the client's interaction with the objects is of particular importance to the therapist, and the possibilities of multitouch interaction make this application particularly suitable to adoption with tabletop display technology. The direct nature of touch input to tabletop displays affords observation of these interactions, and the fact that the display is digital makes the scene being created similarly temporary.

While this is an interesting application for tabletops, there are also particularly interesting and challenging research issues. The therapists we worked with explained how in the physical world they have become sensitive over time to understanding the possible implications of the active (manipulation of physical objects) part of the sandtray work. We were particularly interested in whether this professional skill would hold for *virtual* 3D interaction. In particular, is the virtual medium rich enough for a scene to be constructed that the therapists can understand, or are viewing issues, such as the need to project onto a 2D screen, too limiting? Can a virtual object take on a variety of meanings so as to enable the telling of a story to the therapists, or will they be interpreted as mere data or information? Are the interactions on a virtual table rich enough to convey meaning to the therapist about a client's psyche, or is the disconnect between a person's actions and the surface's reaction too great? Our more general goal was to discover whether the therapists felt that they could effectively perform therapy with this digital sandtray or some future design iteration.

We present the cooperative design of a virtual sandtray through three phases of design: initial face-to-face meetings, iterative remote collaborative design, and a face-to-face feedback session. We end with an in-depth discussion of the results of this collaboration and a description of how to make use of our findings beyond the digital sandtray.

7.2 Methodology

We used a cooperative design process [Schuler and Namioka, 1993; Greenbaum and Kyng, 1991] with three sandtray therapists, who were involved throughout this research to provide us with expert domain knowledge. There were three phases of this design process: Phase I involved initial face-to-face meetings with one sandtray therapist (T1), Phase II was an iterative distance collaboration (via phone and email) with two therapists (T1 & T2) throughout the design and implementation of the prototype, and Phase III was a face-to-face discussion with two therapists (T1 & T3) to provide feedback about the working prototype. The first phase involved two meetings, a pre-planning meeting and a follow-up meeting. At the former meeting, the idea of implementing a digital sandtray had not yet been conceived. This meeting was a demonstration of existing technology to T1, which triggered a discussion about how a digital tabletop sandtray might provide a solution to some of the comfort problems some clients feel with traditional sandtrays. The second meeting was to plan how we could collaboratively design our prototype at a distance, as the therapists were in a city approximately 630 km from our research lab. The second phase in our design process involved extensive discussion via phone and email with both T1 & T2, who described in detail what they felt were the essential elements of sandtray therapy. We include direct quotes from our email communication with T2 throughout the next section (section 7.3). We responded to this communication with descriptions of design ideas and questions about what made the elements important. These conversations were highly iterative and led to the design considerations that are described next. During design and development of the prototype, we maintained contact with both T1 & T2 to iterate and refine the design. When the prototype was finished, T1 & T3 joined us in a day-long workshop to use our prototype first-hand and provide feedback (section 7.4).

Our in-depth discussions with the therapists revealed that it is common for them to constantly be collecting artifacts to use in therapy sessions from the environment (e. g., sticks, leaves, plastic cups, etc.). For example, one could consider our first meeting with T1 to be an example of her 'collecting' our technology. More generally, their practice frequently involves the evaluation of the suitability of tools and techniques for use in therapy. This skill is learned over time and does not directly involve their clients. We thus focused our research on the therapists; our research asks whether, in the virtual world, therapists can still interpret a person's actions in a meaningful way to perform therapy. We therefore decided to only include therapists (and not clients) in our design process. This decision has the consequence that our findings should be interpreted with a therapist-focused lens. We did, however, include a mock therapy session in our day-long workshop, where one of the designers played the role of a client. Indeed, part of the training for students learning to do sandtray therapy involves participating in a session as a client themselves. The therapists in our study explained that these mock therapy sessions were necessary in order to better understand the experience for the client. Thus, it seemed particularly appropriate for our own understanding to undergo a similar experience. We make use of the interaction techniques described and validated in chapter 5 and use the force-based framework to build the application.

7.3 Phase I & II: Designing the Virtual Sandtray

There are two possible avenues to explore for features in a digital sandtray: those already offered by a physical sandtray, perhaps adapted for use on a tabletop, and new options that do not exist in the physical world, but are made possible by the virtual. Our list of design considerations (DC1–DC4) contains some features from traditional sandtrays that the therapists thought were important to maintain and some digital features they thought it would be interesting to explore. In our bottom-up approach, we started with nothing, adding features that are deemed valuable for therapy, until a sufficiently rich environment was created. Our communication with sandtray experts was the main guide in determining the most worthwhile features. A secondary concern was the cost of a feature in terms of interface and interaction complexity.

Maintain narrative potential (DC1): Without characters, there can be no story; thus, the use of figurines to represent characters and objects is essential (figure 7.2). It must be possible to add figurines when a new character or object is introduced, to move them around as the story progresses, and to remove them when their part is over. Stacking objects, such as a balanced stack of rocks or an animal on a house, also has significant psychological connotations.

> "[T]he temporary and unfixed nature of the sand-tray pieces invites play and therefore the creation of narrative. Most other media result in a 'snapshot' in which the narrative is implied but not played out." –T2

Maintain the sandtray's characteristic as an associative medium (DC2): With open media,

Figure 7.2: A shelf of figurines used for sandtray therapy (courtesy of Walter [b]).

such as paint and clay, the artist (client) has a sense that the art piece as a whole takes on meaning as an expression of themselves. With a sandtray, meaning is primarily associated with the objects, and usually only one of the objects is the self representation. Thus, to enable the development of associated meaning that allows the client to express their particular story, a broad range of objects or figurines is desirable. Some therapists group the figurines they offer by category, which makes it easier to find related figurines; others prefer a completely random presentation in which all figurines are mixed, which can trigger more spontaneous associations. Although a digital system could allow both options, we chose to use a random presentation to encourage free association. From a commercial library of 3D models, around 160 figurines of many different categories were selected for use in the virtual sandtray.

> "With a sand tray it is rare that a client will ask for a specific object, precisely because inspiration tends to start with associations to the presented repertoire of objects. This puts the client in the position of immediately symbolizing and associating. While the (relatively) fixed nature of the objects limits the expressiveness of the work, the way that they are animated and placed becomes the client's means of articulating their own meanings regarding those objects." –T2

However, the sand itself can be used as an open medium that can be shaped at will and made into a backdrop for the story. Because directly simulating the behaviour of sand is computationally intensive, we decided to provide a different type of background open media in textured 'paint'. The specific types of paint we included represent different surfaces, such as sand, grass, concrete and water.

Also the temporary nature of the sand was described by the therapists as being important. When a client sees a box of sand, they immediately recognize that whatever they create in that sandbox can be easily erased with a simple swipe of the hand. A digital display is in this way similar to a physical sandbox, because the pixels drawn on the screen are also in some sense temporary. By clearing the screen or turning off the monitor, whatever the client creates can be easily erased. Recording is still possible, much like a sandtray session can (and often is) videotaped.

Exploring simple digital extensions (DC3): Much of the design discussion with the therapists considered which aspects of the digital medium might make useful enhancements. In the real world, it is not possible to instantly duplicate objects. In the digital world, it is trivial. Being able to add multiple copies of the same figurine allows for the creation of forests, herds and families, with little or no cost in terms of interface complexity. The sandtray therapists described this feature as being particularly worthwhile.

Another real-world impossibility is resizing rigid objects (again, digitally trivial). The size of an object has significant psychological connotations: larger objects are perceived as more important, more powerful or more menacing.

> "...resizability could be a huge advantage of a virtual play-table. It's an ongoing issue that my toy collection features a range of scales. I have some dinosaurs that are smaller than my cockroaches. Children adapt and play with this, and it sometimes suggests interesting possibilities (like a giant baby who rescues a mom from a tiny car), but scalability would give you the best of all possible worlds...It's such a rich metaphor: the sense that in our psychological/creative world things do not have 'realistic' sizes, they have metaphorical sizes." –T2

Utilize multitouch to enable client / therapist collaboration (DC4): The interactions between the client and the sandtray are the focus of the therapist's observations, and thus a key aspect of this collaboration is the awareness by the therapist of these interactions. Furthermore, sandtray therapy is sometimes used for couples or families, who will cooperatively act out a story, and even a single client can ask the therapist to participate. We must also design for such multi-person scenarios.

In the next three sections, we describe in detail how we realize our design considerations in our implementation of the virtual sandtray. We adapt existing tabletop display techniques and technology to provide rich interactions so the therapist can easily observe and infer information about the client's psyche from their interaction. Design principles (DC1–DC4) are indicated for each implementation detail.

7.3.1 Implementation Details

We implemented our prototype using the SMART Table 2008, which was specifically designed for children. Its small form factor ensures that all corners of the table can be reached by a child, and its sturdy design makes it suitable for use in a practical setting. Moreover, it uses FTIR [Han, 2005] and can detect up to 40 simultaneous touches, enabling interaction through multiple fingers for multiple people at once. The direct nature of multitouch technology supports awareness by the therapist of the client's interactions (DC4). These factors make the SMART Table an ideal choice of hardware.

> "...the principle is that a sand-tray should be just big enough to fill the field of vision. This gives the sense of an immersive world without requiring that the user look around to take the whole thing in...My sandtray is about the size that you suggest (75 by 52 cm), and that feels about right." –T2

To enhance both the feeling of realism, and the narrative abilities (DC1), we employ NVIDIA's PhysX physics simulation engine 2009 in a similar way to [Wilson et al., 2008]. This allows clients to make figurines fall down, roll around, knock each other over, and to toss them around without any extra interaction techniques or development effort.

7.3.2 Figurine Manipulation

The combination of precise control over the object being moved, together with the physical reaction of the remainder of the scene, provides the client with a platform for rich expression through their narrative (DC1). With this system, the physical movements of the client have a direct correlation with the response by the system, allowing the physical movements to be interpreted by the therapist.

> "…most children will depict battles at some point. Different varieties of aggression may be coming out in this…" –T2

> "positioning objects—includes orientation and ability to push into sand." –T2

> "moving objects—sometimes includes momentum, especially when throwing objects and lifting/dropping them." –T2

> "*stacking objects* (small objects placed on top of larger ones, balanced stacks of rocks, animals in trees or on houses etc.)" –T2

> "children love containment, frequently putting things under or within other things. Another powerful metaphor" –T2

To support this rich narrative, the client should be able to freely move and rotate objects around on the surface, but also to stack them (implying vertical movement). Thus, an interaction technique is needed that provides the full 6DOFO of 3D manipulation. We use sticky fingers and opposable thumbs (chapter 5).

Many figurines, such as human figures, will often be used in a standing position. Because it is difficult or impossible to make figurines with a small base stand upright, we add invisible pedestals at the bottom of these objects. The pedestals are configured to collide only with the ground, and thus do not cause unexpected interactions with other figurines.

7.3.2.1 Expressiveness Through Physics

One of the key advantages of the sticky fingers & opposable thumbs interaction technique is that it offers precise control over an object, further enabling narrative (DC1). The use of a physics simulation, on the other hand, implies a certain imprecision and lack of control. To get the best of both worlds, during interaction the object is put into a 'kinematic' state, essentially giving it infinite mass. Thus, the object is controlled only by the fingers and does not respond to forces in the physics engine, but all other objects are controlled by the physics engine and keep responding to the manipulated object. For example, it is possible to drag a figurine around, knocking over other figurines in its path, without losing control over the dragged object. This combination allows the therapist to interpret both the intended action on the object being controlled and the physical reaction of other objects. When an object is released, it retains the linear and angular velocity that it had in the previous animation frame. This allows objects to be tossed by moving them quickly, then releasing them. With some practice it is also possible to make an object spin or fly upwards, but this requires releasing two fingers within the same animation frame. Objects can be stopped by simply touching them. The 'crossing' feature also enables actions such as sweeping across the surface with the side of the hand. Although the actual interpretation of this gesture is very different from the proxy objects introduced by Wilson et al. [2008], the net effect of objects being moved and pushing other objects ahead of them is similar.

By enabling these familiar physical interactions, we provide the client with a language for communicating to the therapist (perhaps subconsciously) through the virtual objects themselves (DC1). The therapist can then interpret what actions such as knocking over, tossing, and sweeping objects might mean about the client's psyche (DC4).

7.3.3 Drawers and Tools

To enable some of the digital extensions (DC3) and to support the accessibility of a wide range of figurines (DC2), we introduce drawers and tools. Each of these tools was designed to respond to the client's touches with the same interaction technique as the figurines, using the physics engine to impose constraints. We also designed the drawers to be able to slide in and out of view using a handle to save screen real estate. The drawers themselves can be tossed to quickly open and close them, and figurines will bounce around them in a physical way. Three such drawers are available: one for selecting figurines, one for resizing them, and one for painting on the sandtray floor.

Figure 7.3: The ring-shaped figurine drawer.

7.3.3.1 Figurine Drawer

To allow the client to choose figurines for their story (DC2), the collection of figurines needs to be shown on the screen.

> "... what really matters is that you're always looking at the toys themselves rather than some kind of menu system. So whether you go with scrolling, zooming or a 'drawer' metaphor, what is presented to the user should always be an image of some kind of toy or another. Otherwise I think you'd be inviting a really cognitive engagement with the system, which seems contrary to the unique genius of the thing to me." –T2

Each figurine is shown as it would appear in the sandtray. However, the tabletop screen is not large enough to show all figurines at the same time with sufficient detail. We used a ring-shaped drawer similar to a 'Lazy Susan', only a small portion of which is visible at any

Figure 7.4: The resizing drawer in three different stages. Upon spinning the dial, the figurines inside the drawer will smoothly grow or shrink.

given time (figure 7.3). This allows an unlimited number of figurines to be accessed. When a figurine is removed from the drawer by dragging (using the sticky fingers and opposable thumbs technique), a copy will remain behind, allowing for quick and intuitive duplication. Figurines can be removed from the scene simply by putting them back into the drawer. This wide range of figurines is beneficial for therapist's observations, as the client has more choice about which figurines to pick. Therapists can also observe the client's browsing and decision-making processes, potentially involving inclusion, exclusion, and/or copying of figurines.

7.3.3.2 Resizing Drawer

To provide the client with the (physically impossible) ability to resize rigid objects (DC3), we provided a resize drawer. An alternative way to implement this is the two-finger 'pinch' gesture, in which two points on the object are pulled apart or pushed together to grow or shrink the object, similar to zooming. However, this gesture was already mapped to vertical movement of the figurine. The use of buttons or handles on the object would be harmful to the sense of physical realism and might be easy for the therapist to miss and therefore impede interpretation.

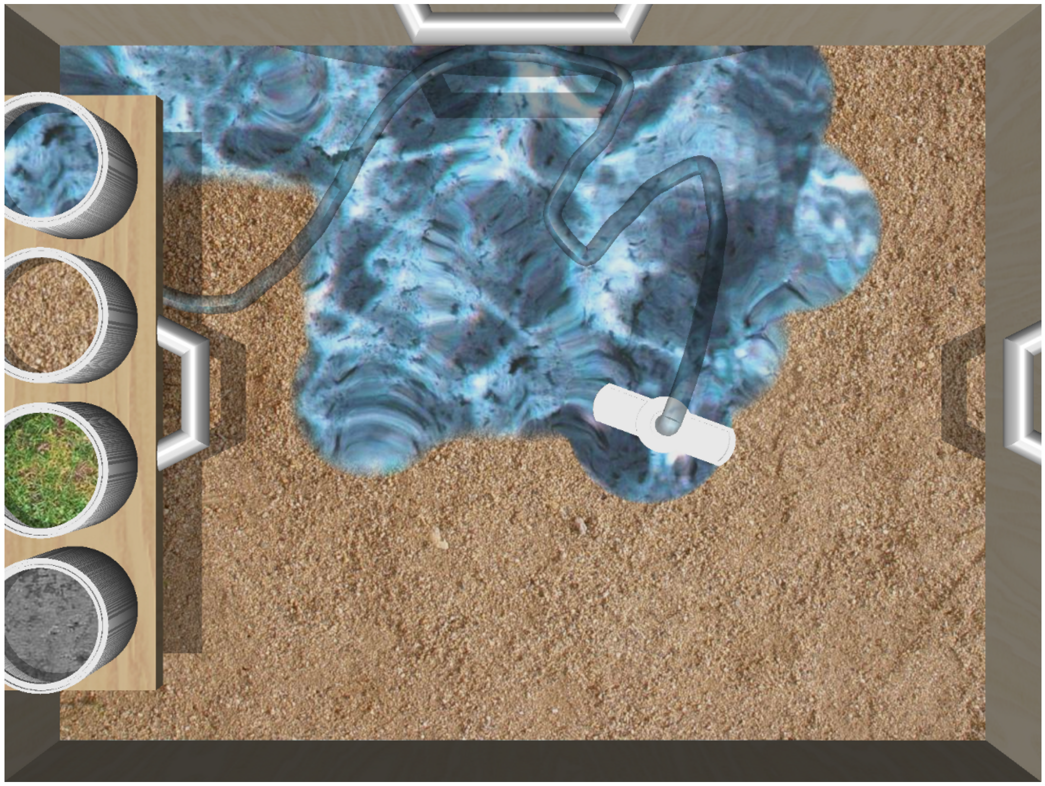


Figure 7.5: The painting system. In the drawer on the left, four buckets of paint can be seen. The hose connects the bottom of 'water paint' bucket to the nozzle, from which the water texture flows.

Instead, a drawer was added to the right side of the screen which acts as a 'resizing box'. One or more figurines can be put into the drawer. A dial is provided on the side of the drawer, with a ridged surface to suggest the ability to turn it. The dial can be turned using a single finger. When it is turned to the left, the figurines in the drawer shrink; when it is turned right, they grow (figure 7.4). A minimum and maximum size was provided to prevent figures from shrinking out of sight or from becoming bigger than the display.

7.3.3.3 Painting Drawer

To enable an open medium that the client can shape at will (DC2), we provided them with the ability to paint the background with different textures. On first thought, the interaction to 'paint' on a touch sensitive display could simply be like fingerpainting: wherever the surface is touched, the chosen paint colour appears. However, combining this technique with the other interactions in the virtual sandtray would require a mode switch. Such a mode would likely be problematic [Sellen et al., 1992], and particularly when multiple people are using the system (DC4).

The painting ability was provided through the use of a spray nozzle tool. This nozzle can paint a texture on the floor (figure 7.5). The nozzle is controlled with the same sticky fingers & opposable thumbs interaction technique. When the nozzle is picked up, it will rotate to point at the sandtray floor, which is the orientation in which it will normally be used. When the nozzle is lifted up, away from the floor, the region that is painted becomes larger. This allows for large regions to be painted quickly.

The other part of the painting system is a drawer containing buckets of paint. A hose running from the drawer to the nozzle serves as a visual cue that they are related. A bucket can be selected by touching it. As long as the bucket is touched and the hose nozzle is in use, the texture paint it contains will flow through the hose and out the nozzle. Usually, the dominant hand is used to move the nozzle, while the non-dominant hand controls the paint selection. Paint will flow from the nozzle when both the nozzle and a paint bucket are being touched simultaneously. The nozzle can be moved around without painting by releasing the bucket.

Both the resize drawer and the painting drawer involve very explicit actions on the part of the client. The system's reaction to these actions (i. e., through the physics engine) makes the consequence of those actions available for interpretation by the therapist (DC4).

7.4 Phase III: Face-to-Face Feedback Session

To validate and iterate on our design, we performed a day-long session with two sandtray therapists together with three of the designers. Neither therapist had any previous experience with digital tables. As previously stated, the focus of our research was to support the ability for therapists to recognize and interpret the actions of the client in a way that was valuable for understanding more about the child's psyche. We designed this face-to-face session to address the following questions:

- Can the therapists interpret the actions of a person using the virtual sandtray and use them to gain insight about their psyche?
- How can the design be improved to better enable the therapists to gain insight about a client's psyche?

Due to the fact that these therapists were typically distance collaborators, we also took this opportunity to share domain knowledge. While we had been iteratively discussing the prototype design remotely, this was the first face-to-face opportunity to do live demonstrations by both groups.

7.4.1 Activities

In this day-long session, the sandtray therapists participated in several activities: the therapists gave a demonstration of sandtray therapy to the designers using traditional physical figurines, the designers gave a demonstration of the virtual sandtray prototype, the therapists conducted a mock therapy session with the digital prototype, and all participated in a follow-up interview/brainstorming session.

7.4.1.1 Physical Demonstration

We began with a demonstration and instruction session from the perspective of the sandtray therapists. In this part, the sandtray therapists described how they performed sandtray therapy using physical figurines and instructed the designers about the theory, logistics, purpose, and essential components of the process of sandtray therapy. We asked questions whenever something was unclear and took written notes.

7.4.1.2 Prototype Demonstration

The designers provided an in-depth demonstration of the digital sandtray prototype. We spent approximately one hour explaining the details of how to interact with our prototype and allowed both therapists to experience using the system. We discussed our design decisions, as well as several viable design alternatives. The therapists were encouraged to share their thoughts and comments. This part was videotaped.

7.4.1.3 Mock Therapy

Following the demonstration and a lunch break, with the warning from the therapists that "some therapy might happen", a designer played the role of the client in a mock therapy session with a therapist, as the other three participants (one therapist and two designers) observed. This session lasted about an hour and was also videotaped.

7.4.1.4 Interview and Brainstorming

For the remainder of the day (around 4 hours, including breaks), the two therapists and three designers participated in an interview and brainstorming session. The interview began with several planned questions, and continued with an unstructured discussion of the benefits, limitations, and next steps for future designs of digital sandtray therapy. Designers again took written notes and this part was not videotaped.

7.5 Results & Discussion

Of central import to the sandtray therapists, throughout all three phases, was the issue of understanding the client's psyche. An essential component of the exercise of sandtray therapy is to give the therapist insight into what the client is thinking and feeling through how they interact with objects. While sandtray therapists typically use physical props to gain this insight, our design revealed that this process is also possible with virtual ones. Instead of simply being a digital representation of some underlying data or model, the virtual objects in our system can take on symbolic meaning in the same way that physical ones do, to the level of providing access into the mind of someone interacting with them.

In this section, we first describe what our research revealed to be the essential components of interaction with virtual objects. These essential components allow virtual objects to cross the boundary from a digital representation to something that can allow the therapist access to the client's psyche. We then describe specific design refinements of our prototype that could address some of the issues that arose in our iterative process. Note that in this section we are discussing the results across all three phases of our design.

7.5.1 Use of Virtual Objects

There were several repeating themes about how a client's interaction with objects can help the therapist gain some insight about what they are thinking or feeling. These themes include *construction*, *storytelling*, *actions*, and *arrangement*.

7.5.1.1 Construction

The therapists frequently identified the ability to construct as an important aspect of the client's interaction with the sandtray. Construction is important because it "stimulates imagination", and in stimulating imagination, the therapist can better access the client's psyche. Several different forms of construction became apparent in our sessions.

The first and most obvious form of construction was the construction of the scene itself. The ability to freely move and rotate the figurines made it possible to create a scene that was composed of many different parts. The ability to make multiple copies of a specific figurine played a key role in this construction. For example, the client in our mock therapy session placed four palm tree objects of different sizes around an oasis. The therapist noted that the number and size of these palm trees matched the number of members in the client's family. Furthermore, the therapist suggested that this oasis may suggest that the family made the client feel safe. This example demonstrates that this form of construction, made possible by our prototype, was a sufficiently rich interaction for the therapist to gain insight.

Another key form of construction that was described by the therapists was the ability to create barriers. This form of construction was not directly supported by our prototype and implies that another level of granularity (besides that of the figurines) might be appropriate, where the virtual objects that the client can interact with can be bent, folded, or attached to one another, like a fence or bricks.

7.5.1.2 Storytelling

Another important aspect of the sandtray interaction was the ability for both the therapist and client to "tell a story". This storytelling process might be brought about in a variety of ways. For instance, the client might be encouraged to just play in the sandtray, and then the therapist might ask the client to explain the scene or elaborate on a specific object and discuss what it means to them.

A key aspect of this storytelling is that the objects in the scene can take on a variety of meanings. On the one hand, an airplane can represent just an airplane (i. e., itself), but on the other hand, it could represent a more abstract idea in the client's mind, such as flight or a desire to escape from something. The therapist's experience with the virtual sandtray prototype led them to state that they could easily interpret these different meanings from the virtual objects.

7.5.1.3 Actions

The therapists also discussed several ways in which the actions that the client uses to interact with the objects can be key to gaining insight about their psyche. The most commonly mentioned action from physical sandtray therapy was the idea of burying an object. This burying process can vary in meaning from client to client; it can imply things that the client wishes to hide or keep secret, or it can be a sign of aggression (e. g., suffocating). The ability for the therapist to observe this burying process is fundamental to the therapy session and was described as important to the therapist-client collaboration for which we are designing. Our current prototype does not allow for this burying interaction, as our choice of physics engine does not provide this functionality.

Other actions that may be relevant to the therapy were made available through the combination of our interaction technique with the physics engine. In particular, the ability to knock over one object with another, the ability to place an object inside another, and the ability to toss an object across the screen or drop an object from high above are actions that a client may do and can help the therapist to understand what is going on in the client's mind when they create a scene.

7.5.1.4 Arrangement

The arrangement of objects in the scene was also described as being highly important to the therapy process. In using our prototype, the therapists felt confident that clients would be able to easily and freely arrange objects in a way that would be useful for a therapy session. Although we did not perform a formal evaluation of usability, we interpret this confidence together with previous formal studies [e. g., chapter 5 and Wilson et al., 2008] as a sign that the interaction technique that we included in our prototype was sufficient for the type of object arrangements that the sandtray therapy requires. Furthermore, the use of gravity through the physics engine and pedestals on the base of each figurine facilitated this arrangement process.

7.5.2 Design Refinements

In addition to the high-level results regarding the efficacy of virtual objects as meaning-carrying artifacts, we gathered feedback about our current design that will help to inform future iterations. The therapists' feedback suggests that these improvements would facilitate interpreting a client's actions.

Some words used by the therapists to describe the virtual sandtray prototype were "relaxing and pleasurable", "attractive", and "appealing". Both therapists commented that, in its current form, the prototype might perhaps already be usable for therapy. On the other hand, comments were made about the lack of sensory feedback: touch, sound, and even smell. However, the application was described as being "still quite tactile". An interesting point raised was that the virtual sandtray does not so much invite storytelling, but rather the construction of a static scene. In that light, it might be more related to art therapy, for example the making of a collage.

7.5.2.1 Figurines

A class of objects that was missing were arbitrary objects that could be brought in by the therapist or the client and play a more metaphorical role. A small cardboard box could serve as a house, a stick could be used as a sword, or a pine cone could represent a baby, covered by a handkerchief to represent a blanket. The inability to bring such objects in could be found to be limiting if many therapy sessions are performed with the same, limited collection of figurines. It was suggested that a possibility should be added to draw or otherwise create one's own figurines, but this would be difficult to implement in an intuitive fashion, but might be possible through Teddy [Igarashi et al., 1999] or Shape-Shop [Schmidt et al., 2006]. An alternative would be to use a device such as the Microsoft Surface [Microsoft Corp., 2008] capable of recognizing physical objects and have the physical props interact with the virtual ones in our prototype.

The presentation of figurines in the drawer was identified as potentially problematic; once a figurine was selected, the lack of structure made it difficult to find related figurines. Although the therapists who participated in the feedback session (T1 & T3) normally present the figurines in their therapy in an organized way (in contrast to T2), they did comment that the lack of ordering in the virtual sandtray prototype "stimulates more random aspects of the psyche". This difference in approach suggests that we should design for some choice on the part of the therapist in how the figurines are presented. We could add a way to configure whether the objects are randomly sorted or organized in some fashion. Alternatively, a hybrid approach could start with a random presentation, but allow the client to easily find related figurines once a few have been selected.

7.5.2.2 Vertical Movement

Several problems were noted in relation to vertical movement of figurines. First, with the current top-down projection, it is not clear that the object is actually moving up or down, instead of simply changing size. This confusion was strongly reflected in the terminology used while discussing this action; even though people know that the object is actually moving up and down in the scene, they often still talk about "making it bigger" and "making it smaller". This suggests a disconnect between the client's actions and the therapist's interpretation. This might partly be blamed on the fact that the shadow is cast directly below each object, which often causes the shadow to be partly or completely obscured. A second shadow, cast from the side, might improve interpretation of movement in z; a projection that is not strictly top-down could also help.

A second problem is that the 'sticky fingers' paradigm chapter 6 implicitly makes lifting an object very sensitive. Especially when the two fingers start close together, a small movement of the fingers will result in a large vertical motion. Doubling the distance between the fingers will move the object twice as close to the virtual camera, which is quite a large distance. Perhaps it is better to let go of the stickiness of one of the fingers. A more formal user study may be necessary to objectively determine which is better.

A third problem is that it is possible to move a figurine so that it becomes invisible. For example, a figurine can be pushed down right through a drawer from above, causing the figurine to become hidden underneath the drawer. While this 'feature' may be one way of achieving the burying action requested by the therapists, it may also be an unintended consequence of the client's actions, introducing ambiguity in how the therapist interprets the action. It might be better to keep a figurine always visible, by forcing it to be always above everything else or by using digital effects such as outlines, shadows or transparency.

7.6 Beyond the Digital Sandtray

On the one hand, our digital sandtray prototype is a single point in the design space of interactive tabletops. Thus, our iterative process may not yield results that generalize beyond this design. However, our method is an example of how we would recommend designing future interactive tabletop systems and we would argue that many of the design decisions that we made can be adopted on other tabletop systems. Furthermore, our feedback session provides some of the first available evidence that virtual objects can be used in a real application in a very different way than they are typically used on desktop computers. In particular, the therapists in our study pointed out the following aspects of our system as particularly compelling:

Communication through virtual objects. The therapists stated that the 3D interaction (sticky fingers and opposable thumbs) would be sufficiently rich for therapeutic purposes. Specifically, the therapists felt they would be able to gain insight about a client's psyche based on his or her interactions with a virtual object.

Repurposeable virtual objects. The therapists felt that the interaction with many individual 3D objects on an interactive table was both meaningful and usable. The therapists illustrated this during mock therapy showing how sometimes virtual objects were recognized as themselves (e. g., a rock is a rock), as well as metaphors, symbols, or archetypes (i. e., representations of things from a person's mind—for instance a rock could represent a member of one's family). This indicated that, in our system, virtual objects could be interpreted by an observer as more than just digital representations of data or information.

Deployable system. Our system is an example of an application that has been identified as usable as-is by the therapists. Our method of designing and developing this application in collaboration with sandtray therapists could be adopted for tabletop systems in other domains. Specifically, our work serves as a case study that cooperative design Greenbaum and Kyng [1991] may lead to successful interactive tabletop systems.

7.7 Chapter Summary

In this chapter, we have presented the viability of a virtual sandtray that has been identified as usable for therapy by domain experts. Beyond the specific domain of art therapy, this work also serves to inform the design process for tabletop display systems and provides some insight into how interaction with 3D objects on a table can be made useful in practice. Specifically, the use of precise interaction and a physics engine can together provide a richness that is sufficient for therapists to understand things about a client's psyche through their interactions with the virtual artifacts. These artifacts thus can take on meaning in a way that is not typically sought after in the design of traditional computer applications. Future designers of tabletop systems can use this work to inform how they can achieve similar levels of rich interaction, and therefore make a new type of interface between humans and virtual objects possible.

Specifically, the contribution from this chapter is:

- A case study of sandtray therapy to demonstrate the use of the framework described in chapter 6 to create an actual application.

In the next chapter, I summarize the contributions of the entire dissertation and describe how this work can be extended to other domains of research. I then describe some of the potential future research directions that this work may lead to.

Conclusion 8

This dissertation demonstrates the design, implementation, and evaluation of 3D tabletop display interaction. The research has been motivated by the desire to make familiar interactions such as lifting, flipping, piling, stacking, and sorting possible on an interactive table. The results of this research provide knowledge that can help other researchers to design and implement new 3D tabletop display environments. This knowledge includes insight into both the perceptual experience and the interactive experience of the people that use the tabletop display.

In this chapter, the research contributions are summarized and future directions in which this work could be taken are discussed. The progress that has been made on the issues presented in chapter 1 is described in section 8.1. The contributions to the field of HCI are then explained in section 8.2. In section 8.3, the possible ways that this work can be extended to other areas of research are discussed. The chapter concludes with a description of some of the future work that this dissertation may lead to (section 8.4) and some closing words (section 8.5).

8.1 Progress on Issues

The central goal of this dissertation has been to provide knowledge about how to design and implement a 3D tabletop display application. The research in this dissertation has addressed the following three issues:

- Issue 1: Perceiving 3D depth cues on a horizontal surface

- Issue 2: Interacting with 3D virtual objects on a horizontal multitouch surface
- Issue 3: Providing a framework to design 3D tabletop display applications

Progress on issue 1 has been made on two primary fronts. First, this dissertation provides a better understanding of the perceptual error that can result from the choices made when displaying 3D information on a 2D interactive tabletop display. Second, recommendations about how to avoid these perceptual errors and techniques to mitigate them are provided. Progress has also been made on issue 2 on two fronts. First, several techniques for manipulating 3D virtual objects were provided and, second, an empirical evaluation provides advice about which technique to use in the design of 3D tabletop display environments. Issue 3 has also been addressed by demonstrating how 3D visuals can be combined with 3D multitouch manipulation techniques to create a realistic application. This framework has also been shown to be effective in the design of a specific application for sandtray therapy.

8.2 Contributions

This research builds upon work from the fields of HCI, graphics, and perception. There are several contributions in this dissertation which can be categorized into three groups. This work contributes to our understanding of how to design 3D tabletop display applications.

8.2.1 Understanding and Mitigating 3D Perceptual Errors

When creating 3D tabletop display applications, the designer must choose how the visual feedback will be presented. These 3D visuals must in some way be projected onto the 2D display, and so a decision about projection geometry must be made. In this dissertation, I have shown that a careful choice of this projection geometry is necessary, and that poor choices may lead to perceptual errors. Specifically, the following contributions from this dissertation can help designers to understand and mitigate 3D perceptual errors:

- A study providing evidence that, when projecting 3D onto a horizontal table using standard 3D graphics techniques, there is an established viewing location, and perception errors will increase as the viewer moves away from this location.
- This same study provides evidence that a parallel projection with a centre of projection (COP) directly above the table may reduce these perceptual errors.
- This same study also provides evidence that providing direct-touch interaction with the virtual artifacts being perceived will also reduce these perceptual errors.
- The design and implementation of a set of non-standard 3D projections that can be used to mitigate the problem of multiple viewpoints and viewpoint discrepancies.

8.2.2 Creating 3D Multitouch Manipulation Techniques

When creating 3D tabletop display applications, the designer must also choose how a person can move, rotate, flip, or otherwise manipulate the virtual artifacts. In this dissertation, the focus has been on the use of multitouch input devices to provide direct input in the same space as the displayed information. The following contributions from this dissertation can help designers to create 3D multitouch manipulation techniques:

- A study providing evidence that using more fingers to control the six degrees of freedom of the output (DOFO) of manipulation improves performance and is preferred.
- The design, implementation, and evaluation of a set of interaction techniques that use multitouch to manipulate 3D virtual objects.

8.2.3 Developing a Real-World 3D Tabletop Display Application

By providing 3D visuals in the same space as the 3D multitouch manipulation techniques, people can begin to use virtual artifacts in a way that more closely resembles interaction with physical artifacts. However, designers may not yet be familiar with creating applications that make use of this physical force-based interaction. The following contributions from this dissertation can help designers to develop real-world 3D tabletop display applications:

- A description of how to combine 3D interaction with 3D visuals to control 3D virtual tools, thus providing the ability to do more complex actions in a virtual world. This description provides a framework for how to create 3D tabletop applications.
- A case study of sandtray therapy to demonstrate the use of the framework described in chapter 6 to create an actual application.

8.3 Extensions

The focus of this dissertation has been on the creation of 3D tabletop display applications, however, the results of this work can be applied to several other areas of research in the field of HCI. Specifically, the results of the series of studies that examines perceptual error can be applied to other scenarios where people will be viewing 3D information from oblique angles and from very close to the display. For example, large touch interactive wall displays require people to stand very close to the surface (so they can touch it) and to perhaps be aware of information that is both close to them and far away. Designers of 3D interactive wall displays can benefit from the knowledge that these oblique viewing angles may also result in perceptual error, and may introduce an inconsistency between people's understanding when they are at opposite ends of the wall. Moreover, when these wall displays are used in combination with interactive tabletop displays, the results presented in chapter 3 can be applied more directly. That is, when attempting to point from the tabletop display to the wall display or to some physical artifact within this environment, perceptual errors are likely to occur. While the mitigating techniques were designed specifically for 3D tabletop display environments, they are likely to be a good starting point for the design of mitigating techniques on large interactive wall displays. Furthermore, the use of a COP directly in front of the display with a parallel projection may reduce perceptual errors in this situation as well.

The 3D interaction techniques described in chapter 5 can also be applied to other multi-touch input devices. In particular, they are likely to be a good starting point for the design of new interaction techniques for manipulating 3D virtual objects on other multitouch devices, provided that the display and input are in the same space and that this space is large enough that a person can comfortably use one, two, or three fingers. Because the interaction techniques are described generally enough to apply to any multitouch surface, they can also be used in environments where multiple interactive surfaces are available (multiple display environments), for example, several wall displays and a tabletop display. By using the same interaction technique on all of these displays, the designer can achieve consistency in how the people interact with the technology.

Furthermore, these same multiple display environments may benefit from the use of force-based interaction. That is, applications on walls and other surfaces can make use of virtual tools to enable more complete functionality, without the designer having to resort to the use of gestures or menus within the interface. In this same vein, just as research in TUIs has helped to inform the force-based interaction framework described in chapter 6, the use of virtual tools in combination with tangible devices may also improve the design of TUIs. Specifically, tangible devices can introduce a physical intermediary between a person and the virtual tool being used. Tools can be created that are composed in part of physical devices and in part of virtual tools. This hybrid tool can be used to interact with other virtual artifacts in the environment.

Lastly, this work can be beneficial in the field of HCI as an exemplar of how to explore a specific design space. That is, the process of examining how people perceive a new technology together with a detailed exploration of the technological capabilities of the new technology can lead to a combined environment where people's perceptions and actions can coincide with their expectations from the physical world.

8.4 Future Work

While this dissertation provides a significant contribution to HCI and can be a good starting point for a designer choosing to create a 3D tabletop display application, there are many promising research directions which would further add to the larger body of work. As discussed in section 8.3, the results of this dissertation can already be applied to a variety of multitouch display technologies and environments, however further study is needed in order to better understand how transferable these results are. For example, people are already investigating above-surface gestural interaction; it remains to be seen how this would compare to, or perhaps integrate with, force-based touch interaction. Furthermore, several questions still remain: do the interaction techniques in this dissertation scale to both smaller devices where many fingers may not fit, or to large devices where reach may be more problematic? Similarly, can the techniques for mitigating perceptual error scale to massive walls or tables, for example in a museum installation where 50–100 people can gather simultaneously?

The series of perceptual studies from chapter 3, while extensive, were also not exhaustive. It is still not well understood how people perceive shape in 3D artifacts and, while the studies reported in this dissertation had the advantage of precisely investigating this perceptual error, further studies could examine these perceptual phenomena in a more realistic setting, either in a field study or in an observational study with fewer constraints (e. g., allowing or encouraging people to adopt a variety of collaborative strategies). The research in this dissertation also suggests a variety of other studies, including an evaluation which explores how engaged a person can be with a virtual tool when compared to their use of physical tools, as well as a study which examines the suitability of the interaction techniques described herein on other multitouch display configurations, such as walls or mobile devices. Lastly, a deployment and observation of the sandtray application developed in chapter 7 may provide further insight into the creation of 3D tabletop display applications using force-based interaction.

The framework in chapter 6 also opens a new area of research within the field of tabletop display literature. Namely, the design of 3D tabletop display applications can benefit from research into the design of 3D virtual tools. Whereas a significant amount of research in tabletop display environments has focused on either the physical hardware or on the interaction techniques to control 2D and 3D virtual artifacts, a new area of research could explore the design of virtual tools themselves. For example, the use of cloth was briefly explored in chapter 6, but cloth may provide many interactive benefits when used on an interactive table as a virtual tool beyond its ability to "cover" other objects. Further research could explore the design of cloth-based interfaces.

8.5 Closing Remarks

This dissertation has demonstrated how knowledge about how people both perceive 3D visual information and interact physically with that information can lead to an improved design for tabletop display applications. This research has provided a careful investigation of both these perceptual phenomena and the benefits to interaction uniquely provided by multitouch input devices. Furthermore, from these two solid bases, a framework has been developed that can inform the design of tabletop display applications and bring force-based interaction to these environments. This research has provided many contributions to the process of designing 3D interactive tabletop applications, but as discussed in this chapter, there are still many areas of research which could bring physical interactions beyond lifting, flipping, stacking, and sorting to interactive multitouch devices.

Bibliography

Agarawala, Anand and Balakrishnan, Ravin (2006). Keepin' it real: Pushing the desktop metaphor with physics, piles and the pen. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1283–1292. ACM Press. doi:10.1145/1124772.1124965 → p. 34, 40, 116, 117, 152, 154, 161

Agrawala, Maneesh, Beers, Andrew C., McDowall, Ian, Fröhlich, Bernd, Bolas, Mark, and Hanrahan, Pat (1997). The two-user responsive workbench: Support for collaboration through individual views of a shared space. In *SIGGRAPH '97: Proceedings of the international conference on Computer graphics and interactive techniques*, 327–332. New York, NY, USA: ACM Press. doi:10.1145/258734.258875 → p. 28, 92

Agrawala, Maneesh, Zorin, Denis, and Munzner, Tamara (2000). Artistic multiprojection rendering. In *Proceedings of the Eurographics Workshop on Rendering Techniques*, 125–136. London, UK: Springer-Verlag → p. 27, 85

Apitz, Georg and Guimbretière, François (2004). Crossy: a crossing-based drawing application. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, 3–12. New York, NY, USA: ACM. doi:10.1145/1029632.1029635 → p. 149

Apple Inc. (2007a). iPhone. http://www.apple.com/iphone. Initial release: September 5, 2007 → p. 17

Apple Inc. (2007b). iPod Touch. http://www.apple.com/ipodtouch. Initial release: September 13, 2007 → p. 17, 19

Apple Inc. (2010). iPad. http://www.apple.com/ipad. Initial release: April 3, 2010 → p. 17

Balakrishnan, Ravin and Kurtenbach, Gordon (1999). Exploring bimanual camera control and object manipulation in 3d graphics interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 56–62. New York, NY, USA: ACM. doi:10.1145/302979.302991 → p. 32

Bers, Marina Umaschi, Ackermann, Edith, Cassell, Justine, Donegan, Beth, Gonzalez-Heydrich, Joseph, DeMaso, David Ray, Strohecker, Carol, Lualdi, Sarah, Bromley, Dennis, and Karlin, Judith (1998). Interactive storytelling environments: coping with cardiac illness at Boston's Children's Hospital. In *Proceedings of the SIGCHI* *conference on Human factors in computing systems*, 603–610. New York: ACM Press. doi:10.1145/274644.274725 → p. 33, 41

Bi, Xiaojun, Bae, Seok-Hyung, and Balakrishnan, Ravin (2010). Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 65–74. New York, NY, USA: ACM. doi:10.1145/1753326.1753337 → p. 15

Bowman, Doug A., Kruijff, Ernst, Joseph J. LaViola, Jr., and Poupyrev, Ivan (2005). *3D User Interfaces: Theory and Practice*. Addison-Wesley → p. 18, 29

Bradway, Kay and McCoard, Barbara (1997). *Sandplay: Silent Workshop of the Psyche*. Routledge → p. 168

Brosz, John, Samavati, Faramarz F., Sheelagh, M. T. Carpendale, and Sousa, Mario Costa (2007). Single camera flexible projection. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, 33–42. New York, NY, USA: ACM. doi:10.1145/1274871.1274876 → p. 27

Buxton, William (1983). Lexical and pragmatic considerations of input structures. *SIGGRAPH '83: Proceedings of the 10th annual conference on Computer graphics and interactive techniques* 17(1):31–37. doi:10.1145/988584.988586 → p. 30, 35, 97, 105, 113, 115

Buxton, William, Fitzmaurice, George, Balakrishnan, Ravin, and Kurtenbach, Gordon (2000). Large displays in automotive design. *IEEE Computer Graphics and Applications* 20(4):68–75. doi:10.1109/38.851753 → p. 16

Buxton, William and Sniderman, R. (1980). Iteration in the design of the human-computer interface. In *Proc 13th. Meeting, Human Factors Assoc. of Canada*, 72–81 → p. 6

Cao, Xiang, Wilson, Andy. D., Balakrishnan, Ravin, Hinckley, Ken, and Hudson, Scott E. (2008). ShapeTouch: leveraging contact shape on interactive surfaces. In *Third Annual IEEE International Workshop on Tabletops and Interactive Surfaces (TABLETOP'08*, 129–136. doi:{10.1109/TABLETOP.2008.4660195} → p. 33, 40, 163

Carpendale, Monica (2009). *Essence and Praxis in the Art Therapy Studio*. Trafford Publishing → p. 168, 169

Cassell, Justine and Ryokai, Kimiko (2001). Making space for voice: technologies to support children's fantasy and storytelling. *Personal Ubiquitous Computing* 5(3):169–190. doi:10.1007/PL00000018 → p. 41

Coleman, Patrick and Singh, Karan (2004). Ryan: Rendering your animation nonlinearly projected. In *NPAR '04: Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering*, 129–156. New York, NY, USA: ACM Press. doi:10.1145/987657.987678 → p. 27

Collett, Thomas S. and Harkness, L. I. K. (1982). Depth vision in animals. In Ingle, Goodale MA Mansfield RJW, DJ (Ed.), *Analysis of Visual Behavior*, 111–176. MIT → p. 20

Cunningham, Joyce (Ed.) (2009). *Journal of Sandplay Therapy*, vol. 18. Sandplay Therapists of America → p. 168

Cutting, James E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance* 13(3):323–334 → p. 24, 25

Cutting, James E. (1988). Affine distortions of pictorial space: Some predictions for goldstein (1987) that la gournerie (1859) might have made. *Journal of Experimental Psychology* 14(2):305–311 → p. 24, 25, 26

Cutting, James E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments & Computers* 29:27–36 → p. 25, 53

de la Gournerie, Jules (1859). *Traité de Perspective Linéaire Contenant les Tracés pour les Tableaux, Plans & Courbes, les Bas Reliefs & les Décorations Théatrales, avec une Théorie des Effets de Perspective*. Dalmont et Dunod → p. 25, 26

Deering, Michael (1992). High resolution virtual reality. In *SIGGRAPH '92: Proceedings of the international conference on Computer graphics and interactive techniques*, 195–202. New York, NY, USA: ACM Press. doi:10.1145/133994.134039 → p. 27

Dietz, Paul and Leigh, Darren (2001). DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, 219–226. New York, NY, USA: ACM Press. doi:10.1145/502348.502389 → p. 16, 19, 32, 36, 37, 38, 98, 109, 112, 116, 121, 122, 131

Dijkstra, T.M.H., Cornilleau-Pérès, V., Gielen, C.C.A.M., and Droulez, J. (1995). Perception of three-dimensional shape from ego- and object-motion: Comparison between small- and large-field stimuli. *Vision Research* 35(4):453 – 462. doi:10.1016/0042-6989(94)00147-E → p. 53

Dourish, Paul (2001). *Where the action is: the foundations of embodied interaction*. Cambridge, MA, USA: MIT Press → p. 2

Dragicevic, Pierre (2004). Combining crossing-based and paper-based interaction paradigms for dragging and dropping between overlapping windows. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, 193–196. ACM Press. doi:10.1145/1029632.1029667 → p. 117

Ebert, David, Bedwell, Edward, Maher, Stephen, Smoliar, Laura, and Downing, Elizabeth (1999). Realizing 3d visualization using crossed-beam volumetric displays. *Commununications of the ACM* 42(8):100–107. doi:10.1145/310930.310979 → p. 24

Ellis, Stephen R., Smith, Stephen, Grunwald, Arthur, and McGreevy, Michael W. (1991). Direction judgement error in computer generated displays and actual scenes. *Pictorial communication in virtual and real environments* 504–526 → p. 24, 26

Fitzmaurice, George W., Ishii, Hiroshi, and Buxton, William A. S. (1995). Bricks: laying the foundations for graspable user interfaces. In *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems*, 442–449. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. doi:10.1145/223904.223964 → p. 15

Foley, James D. and Wallace, Victor L. (1974). The art of natural graphic man-machine conversation. *Proceedings of the IEEE* 62(4):462–471. doi:10.1145/988026.988037 → p. 30

Foley, James D., Wallace, Victor L., and Chan, Peggy (1984). The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications* 4(11):13–48 → p. 30, 115

Foxlin, Eric (2002). Motion tracking requirements and technologies. In Stanney, Kay (Ed.), *Handbook of Virtual Environment Technology*, chap. 8, 163–210. Hillsdale, N.J.: Lawrence Erlbaum Publishers → p. 18

Freeman, Dustin, Benko, Hrvoje, Morris, Meredith Ringel, and Wigdor, Daniel (2009). Shadowguides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 165–172. New York, NY, USA: ACM. doi:10.1145/1731903.1731935 → p. 37

Fröhlich, Bernd, Hochstrate, Jan, Skuk, Verena, and Huckauf, Anke (2006). The GlobeFish and the GlobeMouse: Two new six degree of freedom input devices for graphics applications. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 191–199. ACM Press. doi:10.1145/1124772.1124802 → p. 31, 105, 133, 163

Goldstein, E. Bruce (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance* 13(2):256–266 → p. 25, 26, 77

Goldstein, E. Bruce (1991). Perceived orientation, spatial layout and the geometry ofpictures. *Pictorial communication in virtual and real environments* 480 → p. 25

Google (2010). Google Nexus One. http://www.google.com/phone. Initial release: January 5, 2010 → p. 17

Gouraud, Henri (1971). Continuous shading of curved surfaces. *IEEE Transactions on Computers* C-20(6):623–629 → p. 23

Greenbaum, Joan M. and Kyng, Morten (Eds.) (1991). *Design at Work: Cooperative Design of Computer Systems*. L. Erlbaum Associates Inc. → p. 170, 191

Grossman, Tovi and Wigdor, Daniel (2007). Going deeper: a taxonomy of 3D on the tabletop. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*, 137–144. doi:{10.1109/TABLETOP.2007.18} → p. 26

Grubert, Jens, Carpendale, Sheelagh, and Isenberg, Tobias (2008). Interactive stroke-based NPR using hand postures on large displays. In *Eurographics Short Papers*, 279–282. Aire-la-Ville, Switzerland: Eurographics Association → p. 33

Guay, Roland (1977). *Purdue Spatial Visualization Test: Rotations*. W. Lafayette, IN, USA: Purdue Research Foundation → p. 55, 218, 224

Hagen, Margaret A. and Elliott, Harry B. (1976). An investigation of the relationship between viewing condition and preference for true and modified linear perspective and adults. *Journal of Experimental Psychology: Human Perception and Performance* 2(4):479–490 → p. 53

Han, Jefferson Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 115–118. New York, NY, USA: ACM Press. doi:10.1145/1095034.1095054 → p. 17, 98, 110, 116, 121, 176

Hancock, Mark and Carpendale, Sheelagh (2007). Supporting multiple off-axis viewpoints at a tabletop display. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*, 171–178. Newport, RI, USA. doi:{10.1109/TABLETOP.2007.9} → p. 51

Hancock, Mark, Carpendale, Sheelagh, and Cockburn, Andy (2007). Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1147–1156. New York, NY, USA: ACM Press. doi:10.1145/1240624.1240798 → p. 33, 34, 117, 150

Hancock, Mark, Nacenta, Miguel, Gutwin, Carl, and Carpendale, Sheelagh (2009a). The effects of changing projection geometry on the interpretation of 3d orientation on tabletops. In *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 157–164. New York, NY, USA: ACM. doi:10.1145/1731903.1731934 → p. 51

Hancock, Mark, ten Cate, Thomas, and Carpendale, Sheelagh (2009b). Sticky tools: full 6dof force-based interaction for multi-touch tables. In *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 133–140. New York, NY, USA: ACM. doi:10.1145/1731903.1731930 → p. 153

Hancock, Mark, ten Cate, Thomas, Carpendale, Sheelagh, and Isenberg, Tobias (2010). Supporting sandtray therapy on an interactive tabletop. In *Proceedings of the SIGCHI* *conference on Human factors in computing systems*, 2133–2142. New York, NY, USA: ACM. doi:10.1145/1753326.1753651 → p. 168

Hancock, Mark S., Vernier, Frédéric D., Wigdor, Daniel, Carpendale, Sheelagh, and Shen, Chia (2006). Rotation and translation mechanisms for tabletop interaction. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 79–88. Los Alamitos, CA, USA: IEEE Computer Society. doi:10.1109/TABLETOP.2006.26 → p. 32, 33, 117, 120, 163

Harrison, Chris, Tan, Desney, and Morris, Dan (2010). Skinput: appropriating the body as an input surface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 453–462. New York, NY, USA: ACM. doi:10.1145/1753326.1753394 → p. 18

Hill, Francis S., jr (2001). *Computer Graphics Using Open GL*. Prentice Hall, 2nd ed. → p. 86

Hilliges, Otmar, Izadi, Shahram, Wilson, Andrew D., Hodges, Steve, Garcia-Mendoza, Armando, and Butz, Andreas (2009). Interactions in the air: adding further depth to interactive tabletops. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, 139–148. New York, NY, USA: ACM. doi:10.1145/1622176.1622203 → p. 116, 150

Hinrichs, Uta, Carpendale, Sheelagh, and Scott, Stacey D. (2006). Evaluating the effects of fluid interface components on tabletop collaboration. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, 27–34. New York, NY, USA: ACM. doi:10.1145/1133265.1133272 → p. 39

Igarashi, Takeo, Matsuoka, Satoshi, and Tanaka, Hidehiko (1999). Teddy: a sketching interface for 3D freeform design. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 409–416. ACM Press/Addison-Wesley Publishing Co. doi:10.1145/311535.311602 → p. 188

Isenberg, Petra, Tang, Anthony, and Carpendale, Sheelagh (2008). An exploratory study of visual information analysis. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1217–1226. New York, NY, USA: ACM. doi:10.1145/1357054.1357245 → p. 37

Isenberg, Tobias, Miede, André, and Carpendale, Sheelagh (2006). A buffer framework for supporting responsive interaction in information visualization interfaces. In *Proc. C5*, 262–269. IEEE Computer Society. doi:10.1109/C5.2006.4 → p. 39

Jacob, Robert J. K. and Sibert, Linda E. (1992). The perceptual structure of multidimensional input device selection. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 211–218. New York, NY, USA: ACM. doi:10.1145/142750.142792 → p. 30

Jacob, Robert J. K., Sibert, Linda E., McFarlane, Daniel C., and M. Preston Mullen, Jr. (1994). Integrality and separability of input devices. *ACM Transactions on Computer-Human Interaction* 1(1):3–26. doi:10.1145/174630.174631 → p. 32

Jones, Rebecca K. and Hagen, Margaret A. (1978). The perceptual constraints on choosing a pictorial station point. *Leonardo* 11(3):191–196. doi:10.2307/1574141 → p. 21

Kalff, Dora M. (1980). *Sandplay: A Psychotherapeutic Approach to the Psyche*. Sigo Press → p. 168

Keijser, Jeroen, Carpendale, Sheelagh, Hancock, Mark, and Isenberg, Tobias (2007). Exploring 3D interaction in alternate control-display space mappings. In *IEEE Symposium on 3D User Interfaces* → p. 41

Kirsh, David and Maglio, Paul (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science* 18(4):513–549. doi:10.1016/0364-0213(94)90007-8 → p. 145

Kitamura, Yoshifumi, Nakayama, Tomokazu, Nakashima, Takashi, and Yamamoto, Sumihiko (2006). The IllusionHole with polarization filters. In *VRST '06: Proceedings of the ACM symposium on Virtual reality software and technology*, 244–251. Limassol, Cyprus: ACM. doi:10.1145/1180495.1180546 → p. 92

Koenderink, Jan J., van Doorn, Andrea J., Kappers, Astrid M. L., and Todd, James T. (2004). Pointing out of the picture. *Perception* 33:513–530 → p. 25

Kruger, Russell, Carpendale, M. Sheelagh T., Scott, Stacey D., and Tang, Anthony (2005). Fluid integration of rotation and translation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 601–610. ACM Press. doi:10.1145/1054972.1055055 → p. 32, 39, 117, 124, 126, 132, 163

Kruger, Russell, Carpendale, Sheelagh, Scott, Stacey D., and Greenberg, Saul (2003). How people use orientation on tables: Comprehension, coordination and communication. In *GROUP '03: Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, 369–378. New York, NY, USA: ACM Press. doi:10.1145/958160.958219 → p. 39

Kruger, Russell, Carpendale, Sheelagh, Scott, Stacey D., and Greenberg, Saul (2004). Roles of orientation in tabletop collaboration: Comprehension, coordination and communication. *Computer Supported Cooperative Work* 13(5-6):501–537. doi:10.1007/s10606-004-5062-8 → p. 32, 120

Krüger, Wolfgang and Fröhlich, Bernd (1994). The Responsive Workbench. *IEEE Computer Graphics and Applications* 14(3):12–15. doi:10.1109/38.279036 → p. 34

Kurtenbach, Gordon and Buxton, William (1993). The limits of expert performance using hierarchic marking menus. In *CHI '93: Proceedings of the SIGCHI conference on Human* *factors in computing systems*, 482–487. New York, NY, USA: ACM Press. doi:10.1145/169059.169426 → p. 33

Kurtenbach, Gordon, Fitzmaurice, George, Baudel, Thomas, and Buxton, Bill (1997). The design of a gui paradigm based on tablets, two-hands, and transparency. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 35–42. New York, NY, USA: ACM. doi:10.1145/258549.258574 → p. 32, 165

Lepinski, G. Julian, Grossman, Tovi, and Fitzmaurice, George (2010). The design and evaluation of multitouch marking menus. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2233–2242. New York, NY, USA: ACM. doi:10.1145/1753326.1753663 → p. 33

Li, Ying, Fontijn, Willem, and Markopoulos, Panos (2008). A tangible tabletop game supporting therapy of children with cerebral palsy. In *Proc. Fun and Games*, 182–193. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-540-88322-7_18 → p. 41

Lindeberg, Tony (1994). *Scale-Space Theory in Computer Vision*. Norwell, MA, USA: Kluwer Academic Publishers → p. 17, 114

Liu, Jun, Pinelle, David, Sallam, Samer, Subramanian, Sriram, and Gutwin, Carl (2006). TNT: Improved rotation and translation on digital tables. In *Proceedings of the conference on Graphics Interface*, 25–32. Toronto, Ont., Canada, Canada: Canadian Information Processing Society → p. 33, 39, 163

Lucente, Mark (1997). Interactive three-dimensional holographic displays: seeing the future in depth. *SIGGRAPH '97: Proceedings of the international conference on Computer graphics and interactive techniques* 31(2):63–67. doi:10.1145/271283.271312 → p. 24

Mackinlay, Jock, Card, Stuart K., and Robertson, George G. (1990). A semantic analysis of the design space of input devices. *Human-Computer Interaction* 5:145–190 → p. xvii, xviii, 30, 31, 35, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 110, 111, 113, 115

Malik, Shahzad, Ranjan, Abhishek, and Balakrishnan, Ravin (2005). Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 43–52. Seattle, WA, USA: ACM. doi:10.1145/1095034.1095042 → p. 41

Masliah, Maurice R. and Milgram, Paul (2000). Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 25–32. New York, NY, USA: ACM. doi:10.1145/332040.332403 → p. 31

Matsushita, Mitsunori, Iida, Makoto, Ohguro, Takeshi, Shirai, Yoshinari, Kakehi, Yasuaki, and Naemura, Takeshi (2004). Lumisight table: a face-to-face collaboration support system that optimizes direction of projected information to each stakeholder. In

Proceedings of the ACM conference on Computer supported cooperative work, 274–283. ACM Press. doi:10.1145/1031607.1031653 → p. 34

McGrath, Joseph E. (1984). *Groups: Interaction and Performance*. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc. → p. 84

McKenna, Michael (1992). Interactive viewpoint control and three-dimensional operations. In *Proc. SI3D*, 53–56. New York, NY, USA: ACM Press. doi:10.1145/147156.147163 → p. 27, 85

Meyer, Kenneth, Applewhite, Hugh L., and Biocca, Frank A. (1992). A survey of position trackers. *Presence: Teleoperators and Virtual Environments* 1(2):173–200 → p. 18

Microsoft Corp. (2008). Microsoft Surface. http://www.microsoft.com/surface. Initial release: April 17, 2008 → p. 17, 38, 110, 116, 188

Microsoft Corp. (2009). Microsoft Zune HD. http://www.zune.net/. Initial release: September 15, 2009 → p. 17

Morris, Meredith Ringel, Cassanego, Anthony, Paepcke, Andreas, Winograd, Terry, Piper, Ann Marie, and Huang, Anqi (2006a). Mediating group dynamics through tabletop interface design. *IEEE Computer Graphics and Applications* 26(5):65–73. doi:10.1109/MCG.2006.114 → p. 36, 41

Morris, Meredith Ringel, Huang, Anqi, Paepcke, Andreas, and Winograd, Terry (2006b). Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1201–1210. New York, NY, USA: ACM. doi:10.1145/1124772.1124952 → p. 37

Morris, Meredith Ringel, Paepcke, Andreas, Winograd, Terry, and Stamberger, Jeannie (2006c). Teamtag: exploring centralized versus replicated controls for co-located tabletop groupware. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1273–1282. New York, NY, USA: ACM. doi:10.1145/1124772.1124964 → p. 37

Morris, Meredith Ringel, Ryall, Kathy, Shen, Chia, Forlines, Clifton, and Vernier, Frederic (2004). Beyond "social protocols": multi-user coordination policies for co-located groupware. In *Proceedings of the ACM conference on Computer supported cooperative work*, 262–265. New York, NY, USA: ACM. doi:10.1145/1031607.1031648 → p. 36

Nacenta, Miguel A., Sakurai, Satoshi, Yamaguchi, Tokuo, Miki, Yohei, Itoh, Yuichi, Kitamura, Yoshifumi, Subramanian, Sriram, and Gutwin, Carl (2007). E-conic: a perspective-aware interface for multi-display environments. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, 279–288. Newport, Rhode Island, USA: ACM. doi:10.1145/1294211.1294260 → p. 26

Nakashima, Kousuke, Machida, Takashi, Kiyokawa, Kiyoshi, and Takemura, Haruo (2005). A 2D–3D integrated environment for cooperative work. In *VRST '05: Proceedings of the ACM symposium on Virtual reality software and technology*, 16–22. New York, NY, USA: ACM Press. doi:10.1145/1101616.1101621 → p. 28, 34

Nielsen, Jacob (1993). Iterative user-interface design. *Computer* 26(11):32–41. doi:10.1109/2.241424 → p. 6

Norman, Donald A (1998). *Psychology of everyday things*. New York, NY: Basic Books → p. 162

Noyes, J. (1983). Chord keyboards. *Applied Ergonomics* 14(1):55–59. doi:10.1016/0003-6870(83)90221-1 → p. 107

NVIDIA Corporation (2009). NVIDIA Corporation. NVIDIA PhysX. http://www.nvidia.com/object/nvidia_physx.html → p. 153, 176

Palais, Bob, Palais, Richard, and Rodi, Stephen (2009). A disorienting look at euler's theorem on the axis of a rotation. *American Mathematical Monthly* 116(10):892–909 → p. 124

Perceptive Pixel (2010). Perceptive Pixel. http://www.perceptivepixel.com → p. 17

Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception and Psychophysics* 14:13–18 → p. 25

Phong, Bui Tuong (1975). Illumination for computer generated pictures. *Communications of ACM* 18(6):311–317 → p. 23

Pinelle, David, Gutwin, Carl, and Greenberg, Saul (2003). Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction* 10(4):281–311. doi:10.1145/966930.966932 → p. 93

Piper, Anne Marie and Hollan, James D. (2008). Supporting medical conversations between deaf and hearing individuals with tabletop displays. In *Proceedings of the ACM conference on Computer supported cooperative work*, 147–156. San Diego, CA, USA: ACM Press. doi:10.1145/1460563.1460587 → p. 42

Pizlo, Zygmunt (1994). A theory of shape constancy based on perspective invariants. *Vision Research* 34(12):1637–1658 → p. 93

Polhemus (1994). Polhemus Fastrak. http://www.polhemus.com → p. 18

Reisman, Jason L., Davidson, Philip L., and Han, Jefferson Y. (2009). A screen-space formulation for 2d and 3d direct manipulation. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, 69–78. New York, NY, USA: ACM. doi:10.1145/1622176.1622190 → p. 150

Richards, James G. (1999). The measurement of human motion: A comparison of commercially available systems. *Human Movement Science* 18(5):589 – 602. doi:10.1016/S0167-9457(99)00023-8 → p. 18

Rosenberg, Ilya and Perlin, Ken (2009). The unmousepad: an interpolating multi-touch force-sensing input pad. In *SIGGRAPH '09: Proceedings of the international conference on Computer graphics and interactive techniques*, 1–9. New York, NY, USA: ACM. doi:10.1145/1576246.1531371 → p. 18

Ryokai, Kimiko and Cassell, Justine (1999). StoryMat: a play space for collaborative storytelling. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 272–273. New York: ACM. doi:10.1145/632716.632883 → p. 41

Saunders, Jeffrey A. and Backus, Benjamin T. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision* 6(9):933–954. doi:10.1167/6.9.7 → p. 24, 25

Schmidt, Ryan, Wyvill, Brian, Sousa, Mario C., and Jorge, Joaquim A. (2006). ShapeShop: sketch-based solid modeling with BlobTrees. In *ACM SIGGRAPH 2006 Courses*, 14. Boston, Massachusetts: ACM. doi:10.1145/1185657.1185775 → p. 188

Schuler, Douglas and Namioka, Aki (Eds.) (1993). *Participatory Design: Principles and Practices*. L. Erlbaum Associates Inc. → p. 170

Scott, Stacey D. (2005). *Territoriality in Collaborative Tabletop Workspaces*. Ph.D. thesis, University of Calgary → p. 36

Scott, Stacey D., Carpendale, M. Sheelagh T., and Inkpen, Kori M. (2004). Territoriality in collaborative tabletop workspaces. In *Proceedings of the ACM conference on Computer supported cooperative work*, 294–303. New York, NY, USA: ACM Press. doi:10.1145/1031607.1031655 → p. 36

Scott, Stacey D., Grant, Karen D., and Mandryk, Regan L. (2003). System guidelines for co-located, collaborative work on a tabletop display. In *Proceedings of the European conference on Computer-supported cooperative work*, 159–178. Helsinki, Finland → p. 36

Sedgwick, H. A. (1993). *The effects of viewpoint on the virtual space of pictures*, 460–479. Taylor & Francis, Inc. → p. 24, 25

Sellen, Abigail J., Kurtenbach, Gordon P., and Buxton, William A. S. (1992). The prevention of mode errors through sensory feedback. *Human-Computer Interaction* 7(2):141–164. doi:10.1207/s15327051hci0702_1 → p. 146, 182

Shen, Chia, Lesh, Neal, and Vernier, Frédéric (2003). Personal digital historian: story sharing around the table. *Interactions* 10(2):15–22. doi:10.1145/637848.637856 → p. 38, 163

Shen, Chia, Vernier, Frédéric D., Forlines, Clifton, and Ringel, Meredith (2004). DiamondSpin: an extensible toolkit for around-the-table interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 167–174. New York, NY, USA: ACM. doi:10.1145/985692.985714 → p. 32, 37, 117, 152, 163

Shoemake, Ken (1992). Arcball: a user interface for specifying three-dimensional orientation using a mouse. In *Proceedings of the conference on Graphics Interface*, 151–156. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. → p. 33

SMART Technologies (2003). SMART Digital Vision Technology. White paper → p. 33, 36, 38, 122

SMART Technologies (2008). SMART Table interactive learning center. http://www.smarttech.com/. Initial release date: October 23, 2008 → p. 17, 19, 38, 110, 116, 176

Ståhl, Olov, Wallberg, Anders, Söderberg, Jonas, Humble, Jan, Fahlén, Lennart E., Bullock, Adrian, and Lundberg, Jenny (2002). Information exploration using the pond. In *Proceedings of the 4th international conference on Collaborative virtual environments*, 72–79. ACM Press. doi:10.1145/571878.571890 → p. 34, 38

Sturman, David J. (1992). *Whole-Hand Input*. Ph.D. thesis, Massachussetts Institute of Technology → p. 33

Tang, Anthony, Tory, Melanie, Po, Barry, Neumann, Petra, and Carpendale, Sheelagh (2006). Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1181–1190. New York, NY, USA: ACM. doi:10.1145/1124772.1124950 → p. 37

Tse, Edward and Greenberg, Saul (2004). Rapidly prototyping single display groupware through the sdgtoolkit. In *Proc Fifth Australasian User Interface Conference, Volume 28 in the CRPIT Conferences in Research and Practice in Information Technology Series*, 101–110. Dunedin, NZ: Australian Computer Society Inc. → p. 38

Tse, Edward, Shen, Chia, Greenberg, Saul, and Forlines, Clifton (2007). How pairs interact over a multimodal digital table. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 215–218. New York: ACM. doi:10.1145/1240624.1240659 → p. 33

Underkoffler, John and Ishii, Hiroshi (1999). Urp: a luminous-tangible workbench for urban planning and design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 386–393. New York, NY, USA: ACM Press. doi:10.1145/302979.303114 → p. 2, 155, 156, 158, 159, 160, 164, 166

Vicon (2006). Vicon MX. http://www.vicon.com/. Initial release: 2006 → p. 18, 55, 73

Vishwanath, Dhanraj, Girshick, Ahna R, and Banks, Martin S (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience* 8(10):1401–1410. doi:10.1038/nn1553 → p. 25, 26, 53

Walter, Kristina (a). http://en.wikipedia.org/wiki/File:Sandspiel1.jpg → p. 168

Walter, Kristina (b). http://en.wikipedia.org/wiki/File:Sandspiel_Figuren2.jpg → p. 173

Wang, Yanqing, MacKenzie, Christine L., Summers, Valerie A., and Booth, Kellogg S. (1998). The structure of object transportation and orientation in human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 312–319. ACM Press. doi:10.1145/274644.274688 → p. 32

Ware, Colin (2004). *Information Visualization: Perception for Design*. Morgan Kaufmann, 2nd ed. → p. 20, 21, 24

Ware, Colin, Arthur, Kevin, and Booth, Kellogg S. (1993). Fish tank virtual reality. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 37–42. Amsterdam, The Netherlands: ACM. doi:10.1145/169059.169066 → p. 23, 24, 27

Welch, Greg and Foxlin, Eric (2002). Motion tracking: No silver bullet, but a respectable arsenal. *IEEE Computer Graphics and Applications* 22:24–38. doi:10.1109/MCG.2002.1046626 → p. 18

Wellner, Pierre (1993). Interacting with paper on the digitaldesk. *Communications of the ACM* 36(7):87–96. doi:10.1145/159544.159630 → p. 14, 15

Westerman, Wayne (1998). *Hand Tracking,Finger Identification, and Chordic Manipulation on a Multi-Touch Surface*. Ph.D. thesis, University of Delaware → p. 32, 104

Wigdor, Daniel, Leigh, Darren, Forlines, Clifton, Shipman, Samuel, Barnwell, John, Balakrishnan, Ravin, and Shen, Chia (2006). Under the table interaction. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, 259–268. New York, NY, USA: ACM. doi:10.1145/1166253.1166294 → p. 116

Wigdor, Daniel, Shen, Chia, Forlines, Clifton, and Balakrishnan, Ravin (2007). Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 473–482. San Jose, California, USA: ACM. doi:10.1145/1240624.1240701 → p. 26

Wilson, Andrew D., Izadi, Shahram, Hilliges, Otmar, Garcia-Mendoza, Armando, and Kirk, David (2008). Bringing physics to the surface. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, 67–76. New York, NY, USA: ACM. doi:10.1145/1449715.1449728 → p. 34, 39, 40, 41, 150, 154, 161, 163, 176, 178, 187

Wobbrock, Jacob O., Morris, Meredith Ringel, and Wilson, Andrew D. (2009). User-defined gestures for surface computing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1083–1092. New York, NY, USA: ACM. doi:10.1145/1518701.1518866 → p. 37

Wu, Mike and Balakrishnan, Ravin (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, 193–202. New York, NY, USA: ACM. doi:10.1145/964696.964718 → p. 33, 37, 152

Zeil, Jochen (2000). Depth cues, behavioural context, and natural illumination: some potential limitations of video playback techniques. *Acta ethologica* 3:39–48 → p. 20

Zhai, Shumin (1995). *Human Performance in Six Degree of Freedom Input Control*. Ph.D. thesis, University of Toronto → p. 29, 31, 35, 97, 115

Zhai, Shumin (1998). User performance in relation to 3D input device design. *SIGGRAPH '98: Proceedings of the international conference on Computer graphics and interactive techniques* 32(4):50–54. doi:10.1145/307710.307728 → p. 163

Zhai, Shumin and Milgram, Paul (1998). Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 320–327. ACM Press. doi:10.1145/274644.274689 → p. 105, 133

Zorin, Denis and Barr, Alan H. (1995). Correction of geometric perceptual distortions in pictures. In *SIGGRAPH '95: Proceedings of the international conference on Computer graphics and interactive techniques*, 257–264. New York, NY, USA: ACM Press. doi:10.1145/218380.218449 → p. 27

Materials for Perception Study A

This appendix includes the material used in the study described in chapter 3.

A.1 Informed Consent Form

The following pages show the informed consent form that was presented to the participants, and the participants were asked to sign, before the study began.

Name of Researcher, Faculty, Department, Telephone & Email:
Mark Hancock, Ph.D. Student, Dept. of Computer Science, (403) 210-9499, msh@cs.ucalgary.ca

Supervisor:
Sheelagh Carpendale, Associate Professor, Dept. of Computer Science, (403) 220-6055, sheelagh@cs.ucalgary.ca

Title of Project:
Computer Feedback in Co-Located Collaborative Environments

Sponsor:
National Sciences and Engineering Research Council, Canadian Foundation for Innovation, and Alberta Ingenuity

---

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Purpose of the Study:

We are currently investigating how technology can be used to help support collaboration. We are investigating both the effects of technology on how you collaborate and on evaluating how well the software and hardware supports collaborative interactions.

What Will I Be Asked To Do?

You will be asked to use different interaction techniques and devices that offer a variety of support for organizing and sharing information such as documents and images. This may involve such tasks as moving digital objects to different locations, passing digital objects between collaborators, and organizing digital objects for shared access.

We will be observing and programmatically capturing your actions, as well as videotaping you during the course of the session. This videotaping is optional and you can still participate if you choose not to be videotaped. You will also be asked to complete a pre- and post-session questionnaire to further our investigation. It is estimated that your involvement will take approximately one hour, and you will be remunerated for your time.

Participation in this study is completely voluntary and you may refuse to participate altogether or at any time during the study without any repercussions.

What Type of Personal Information Will Be Collected?

Should you agree to participate, you will be asked to provide both your gender and age. No other personal identifying information will be collected in this study, and all participants shall remain anonymous.

There are several options for you to consider if you decide to take part in this research. You can choose all, some or none of them. Please put a check mark on the corresponding line(s) that grants me your permission to:

I grant permission to be audio taped: Yes: ___ No: ___

I grant permission to be videotaped: Yes: ___ No: ___

I grant permission to have video or still images of me used in a publication: Yes: ___ No: ___

Are there Risks or Benefits if I Participate?

There are no known risks involved in participating in this experiment. You will be given $10 as compensation for your time. If you decide for any reason to withdraw from the experiment, you will still receive this compensation.

What Happens to the Information I Provide?

Participation is completely voluntary and confidential. Note that because you will be working with others to complete the experiment, the participants in the group besides yourself may be able to identify you. Besides this limitation, you will remain completely anonymous. You are free to discontinue participation at any time during the study. If you decide to discontinue participation, data collected up to that point will be destroyed. The research data, including your answers to the questionnaire and the video recording, will only be viewed by the researcher, his supervisor, and a team of graduate students. There are no names on the questionnaire. Only group information will be summarized for any presentation or publication of results. The questionnaires are kept in a locked cabinet only accessible by the researcher, his supervisor, and a team of graduate students. The anonymous data will be stored for three years on a computer disk, at which time, it will be permanently erased.

Signatures (written consent)

Your signature on this form indicates that you 1) understand to your satisfaction the information provided to you about your participation in this research project, and 2) agree to participate as a research subject.

In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Participant's Name: (please print) ____

Participant's Signature ____Date: _______________

Researcher's Name: (please print) ____

Researcher's Signature: ____Date: ________________

Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Mr. Mark Hancock and Dr. Sheelagh Carpendale
Department of Computer Science
(403) 210-9499, {msh,sheelagh}@cs.ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact Patricia Evans, Associate Director, Research Services Office, University of Calgary at (403) 220-3782; email plevans@ucalgary.ca.

A copy of this consent form has been given to you to keep for your records and reference. The investigator has kept a copy of the consent form.

A.2 Part 1 Materials

This section contains the study materials that were used for Part 2 of the study in chapter 3, which consisted of experiments 1 and 2.

A.2.1 Instructions

The following page contains the instructions which were read to the participants.

Experiment Description

Part I Description

In the first part of the study, you will be identifying the shape of a triangular molecule by indicating the position of the third (blue) atom. In each trial, you indicate the position of that atom by tapping on the tablet PC at your end of the table. You may tap as many times as you like to correct your answer and when you are satisfied with your answer, please tap the "Confirm" button. At each stage, please make sure the trial number on the tablet PC matches the trial number on the large table. We will be measuring only the accuracy of your answers and not how fast you finish each trial.

Part II Description

In part two, you will be matching a molecule with several molecules in the centre of the table. A copy of the molecule to match will be both at your end of the table and at your partner's end of the table. You will be given a sheet of paper to write down how many objects in the centre of the table match this molecule. The two copies at each end do not count. During this phase, please do not talk with your partner about how many matches you see and please refrain from counting out loud.

Condition Instructions

Moving for left/right, Parallel

In this condition, objects will be adjusted to face (name of person on left/right) as he/she moves his/her head.

Stationary for left/right, Parallel

In this condition, objects will be adjusted to face (name of person on left/right) at the beginning of the trial, but will remain that way. When he/she moves his/her head, the objects will not be adjusted.

Neutral, Parallel

In this condition, objects will be adjusted to face directly above the table. They will not move or take into account either of your positions.

Moving for left/right, Perspective

In this condition, objects will be adjusted to appear correct for (name of person on left/right) as he/she moves his/her head.

Stationary for left/right, Perspective

In this condition, objects will be adjusted to appear correct for (name of person on left/right) at the beginning of the trial, but will remain that way. When he/she moves his/her head, the objects will not be adjusted.

Neutral, Perspective

In this condition, objects will be adjusted to appear correct directly above the table. They will not move or take into account either of your positions.

A.2.2 Study Questionnaire

The following pages show the questionnaire that each participant completed during Part 1 of the study in chapter 3, which consisted of experiments 1 and 2. Pages 2–4 of the 6-page questionnaire are removed, due to the fact that they contain copyright material. This material can be found in Guay [1977].

Group «Group»: «Participant»-«Side»

Background Questionnaire

1. Age: ____
2. Are you (circle one) Male or Female ?
3. Are you colorblind? Yes No
   a. If yes, please describe the type of colorblindness:

      ____

      ____

      ____

4. Please describe your experience with 3D games.
   a. How often do you play 3D games?

      _______ times a (circle one) year / month / week / day

   b. List some of the games you have played the most often:

      ____

      ____

      ____

      ____

5. Have you ever used a tabletop display? Yes No
   a. If yes, please explain when and where:

      ____

      ____

      ____

6. Are you a student at the University of Calgary? Yes No
   a. If yes, in what faculty/department? ________________________

Group «Group»: «Participant»-«Side»

Part 1

1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = agree, 7 = strongly agree

	Neutral, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Neutral, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

Group «Group»: «Participant»-«Side»

Part 2

1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = agree, 7 = strongly agree

	Neutral, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Neutral, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7
I found it difficult to know what my partner saw	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

A.3 Part 2 Materials

This section contains the study materials that were used for Part 2 of the study in chapter 3, which consisted of experiments 3 and 4.

A.3.1 Instructions

The following page contains the instructions which were read to the participants.

Experiment Description

Part I Description

In the first part of the study, you will be identifying the shape of a triangular molecule by indicating the position of the third (blue) atom. In each trial, you indicate the position of that atom by tapping on the tablet PC at your end of the table. You may tap as many times as you like to correct your answer and when you are satisfied with your answer, please tap the "Confirm" button. At each stage, please make sure the trial number on the tablet PC matches the trial number on the large table. We will be measuring only the accuracy of your answers and not how fast you finish each trial.

Part II Description

In part two, you will be matching a molecule with several molecules in the centre of the table. A copy of the molecule to match will be both at your end of the table and at your partner's end of the table. You will be given a sheet of paper to write down how many objects in the centre of the table match this molecule. The two copies at each end do not count. During this phase, please do not talk with your partner about how many matches you see and please refrain from counting out loud.

Condition Instructions

Moving for left/right, Parallel

In this condition, objects will be adjusted to face (name of person on left/right) as he/she moves his/her head.

Stationary for left/right, Parallel

In this condition, objects will be adjusted to face (name of person on left/right) at the beginning of the trial, but will remain that way. When he/she moves his/her head, the objects will not be adjusted.

Neutral, Parallel

In this condition, objects will be adjusted to face directly above the table. They will not move or take into account either of your positions.

Moving for left/right, Perspective

In this condition, objects will be adjusted to appear correct for (name of person on left/right) as he/she moves his/her head.

Stationary for left/right, Perspective

In this condition, objects will be adjusted to appear correct for (name of person on left/right) at the beginning of the trial, but will remain that way. When he/she moves his/her head, the objects will not be adjusted.

Neutral, Perspective

In this condition, objects will be adjusted to appear correct directly above the table. They will not move or take into account either of your positions.

A.3.2 Study Questionnaire

The following pages show the questionnaire that each participant completed during Part 2 of the study in chapter 3, which consisted of experiments 3 and 4. Pages 2–4 of the 6-page questionnaire are removed, due to the fact that they contain copyright material. This material can be found in Guay [1977].

Group B«Group»: «Participant»-«Side»

Background Questionnaire

1. Age: ____
2. Are you (circle one) Male or Female ?
3. Are you colorblind? Yes No
   a. If yes, please describe the type of colorblindness:

   ____

   ____

   ____

4. Please describe your experience with 3D games.
   a. How often do you play 3D games?

      _______ times a (circle one) year / month / week / day

   b. List some of the games you have played the most often:

   ____

   ____

   ____

   ____

5. Have you ever used a tabletop display? Yes No
   a. If yes, please explain when and where:

   ____

   ____

   ____

6. Are you a student at the University of Calgary? Yes No
   a. If yes, in what faculty/department? ________________________

Group B«Group»: «Participant»-«Side»

Part 1

1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = agree, 7 = strongly agree

	Neutral, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Parallel						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Neutral, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Moving for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for Me, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

	Stationary for My Partner, Perspective						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7

Group B«Group»: «Participant»-«Side»

Part 2

	1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = agree, 7 = strongly agree																																		
	Neutral, Parallel							**Moving for Me, Parallel**							**Moving for My Partner, Parallel**							**Stationary for Me, Parallel**							**Stationary for My Partner, Parallel**						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

	Neutral, Perspective							**Moving for Me, Perspective**							**Moving for My Partner, Perspective**							**Stationary for Me, Perspective**							**Stationary for My Partner, Perspective**						
I found it easy to see the shape of objects in this mode	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I found it easy to compare objects in this mode	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I would prefer to use a tabletop display with this setting	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Materials for Shallow-Depth Study B

This appendix includes the material used in the study described in chapter 5.

B.1 Informed Consent Form

The following pages show the informed consent form that was presented to the participants, and the participants were asked to sign, before the study began.

Name of Researcher, Faculty, Department, Telephone & Email:

Mark Hancock, Ph.D. Student, Department of Computer Science, (403) 210-9499, msh@cs.ucalgary.ca

Supervisor:

Sheelagh Carpendale, Associate Professor, Department of Computer Science, (403) 220-6055, sheelagh@cs.ucalgary.ca

Title of Project:

Computer Feedback in Co-Located Collaborative Environments

***Sponsor*:**

National Sciences and Engineering Research Council, Canadian Foundation for Innovation, and Alberta Ingenuity

---

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Purpose of the Study:

We are currently investigating how technology can be used to help support collaboration. We are investigating both the effects of technology on how you collaborate and on evaluating how well the software and hardware supports collaborative interactions.

What Will I Be Asked To Do?

You will be asked to use different interaction techniques and devices that offer a variety of support for organizing and sharing information such as documents and images. This may involve such tasks as moving digital objects to different locations, passing digital objects between collaborators, and organizing digital objects for shared access.

We will be observing and programmatically capturing your actions, as well as videotaping you during the course of the session. This videotaping is optional and you can still participate if you choose not to be videotaped. You will also be asked to complete a pre- and post-session questionnaire to further our investigation. It is estimated that your involvement will take approximately one hour, and you will be remunerated for your time.

Participation in this study is completely voluntary and you may refuse to participate altogether or at any time during the study without any repercussions.

What Type of Personal Information Will Be Collected?

Should you agree to participate, you will be asked to provide both your gender and age. No other personal identifying information will be collected in this study, and all participants shall remain anonymous.

There are several options for you to consider if you decide to take part in this research. You can choose all, some or none of them. Please put a check mark on the corresponding line(s) that grants me your permission to:

I grant permission to be audio taped: Yes: ___ No: ___

I grant permission to be videotaped: Yes: ___ No: ___

I grant permission to have video or still images of me used in a publication: Yes: ___ No: ___

Are there Risks or Benefits if I Participate?

There are no known risks involved in participating in this experiment. You will be given $10 as compensation for your time. If you decide for any reason to withdraw from the experiment, you will still receive this compensation.

What Happens to the Information I Provide?

Participation is completely voluntary and confidential. Note that because you will be working with others to complete the experiment, the participants in the group besides yourself may be able to identify you. Besides this limitation, you will remain completely anonymous. You are free to discontinue participation at any time during the study. If you decide to discontinue participation, data collected up to that point will be destroyed. No one except the researcher and his supervisor will be allowed to see or hear any of the answers to the questionnaire or the interview tape. There are no names on the questionnaire. Only group information will be summarized for any presentation or publication of results. The questionnaires are kept in a locked cabinet only accessible by the researcher and his supervisor. The anonymous data will be stored for three years on a computer disk, at which time, it will be permanently erased.

Signatures (written consent)

Your signature on this form indicates that you 1) understand to your satisfaction the information provided to you about your participation in this research project, and 2) agree to participate as a research subject.
In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Participant's Name: (please print) ____

Participant's Signature ____Date: ________________

Researcher's Name: (please print) ____

Researcher's Signature: ____Date: _________________

Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Mr. Mark Hancock and Dr. Sheelagh Carpendale
Department of Computer Science
403-210-9499, {msh,sheelagh}@cs.ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact Patricia Evans, Associate Director, Research Services Office, University of Calgary at (403) 220-3782; email plevans@ucalgary.ca

A copy of this consent form has been given to you to keep for your records and reference. The investigator has kept a copy of the consent form.

B.2 Experiment Instructions

The following pages were given to the participant to read prior to the beginning of each task that they performed.

Experiment Instructions

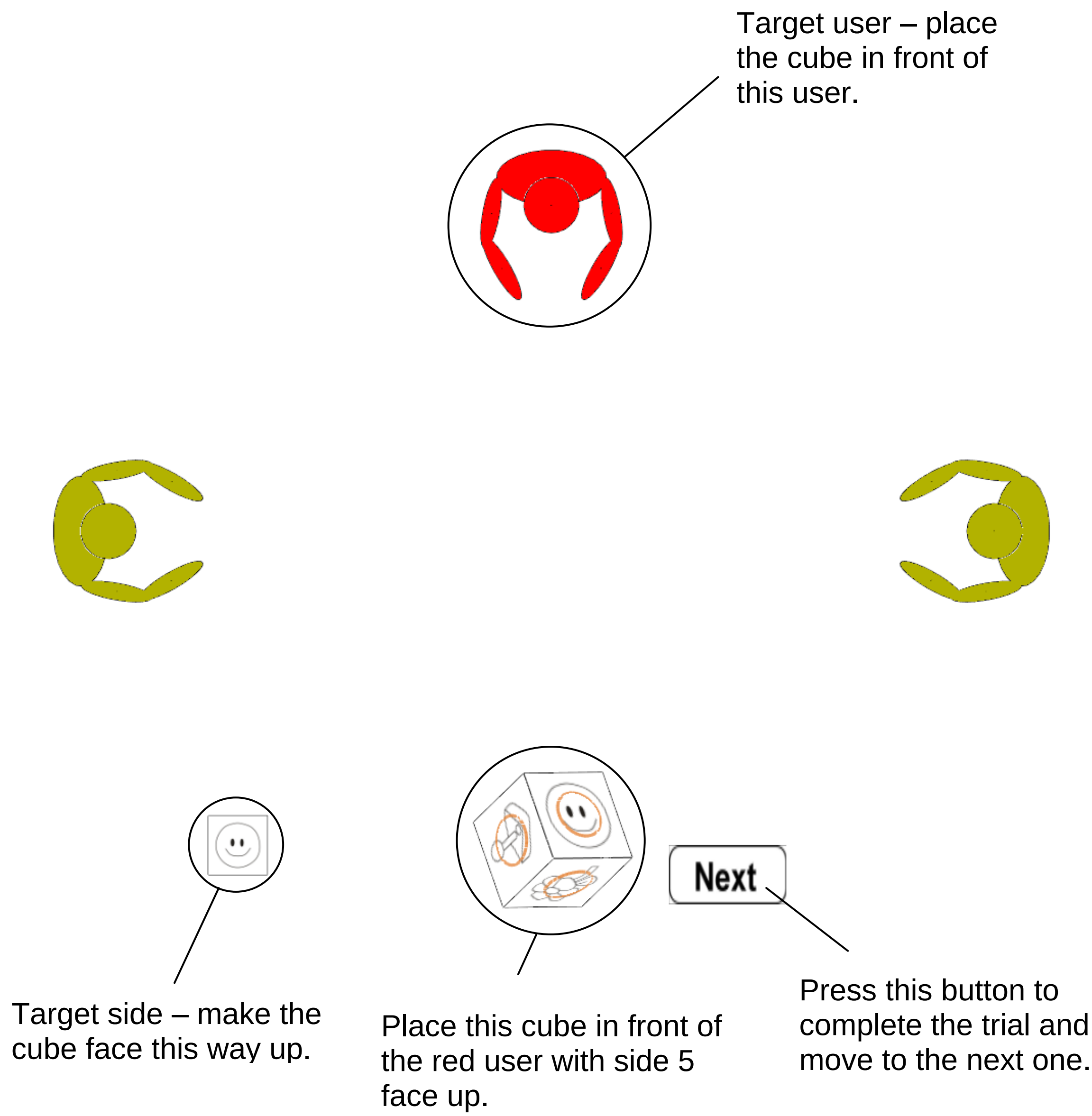
You will be asked to complete three tasks: *docking tetrahedrons*, *passing a cube* to (virtual) people, and assembling a *puzzle*. You will complete these tasks on a touch-sensitive table. You will be asked to wear a glove on your right hand to drag objects on the screen and **you may at any time take a break and remove the glove** before continuing.

You will be able to control the objects on the table in three different ways, using both the glove in your right hand and by directly touching the display with your left. Before each group of trials, the experimenter will describe how the technique you will be using works. You will get a chance to practice each technique before beginning.

In each part of the experiment, we will be measuring your speed, so please try to complete the task as *quickly* as you can.

Passing Cube

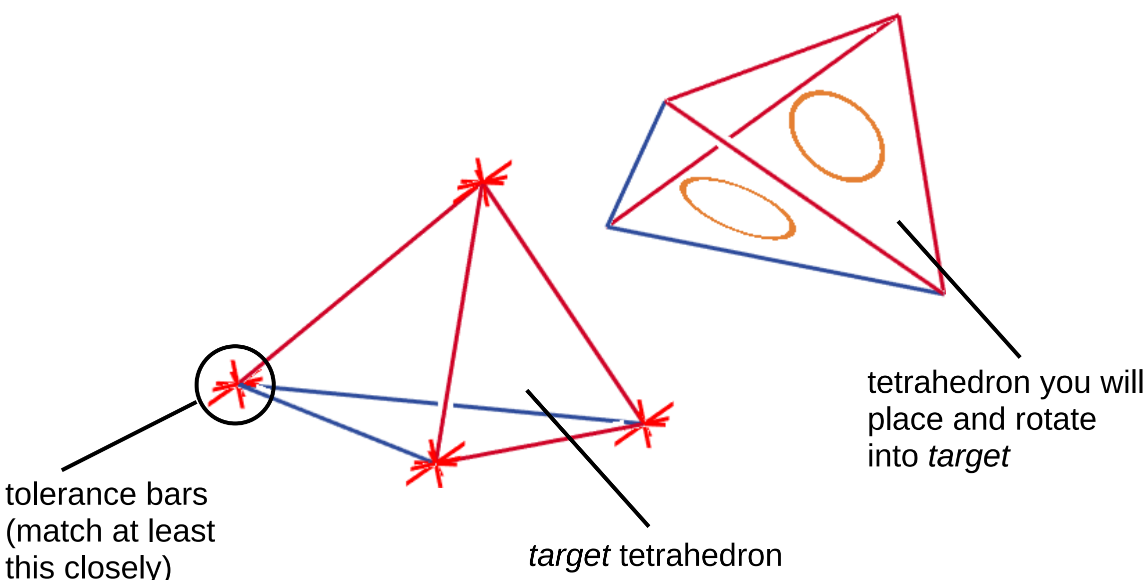
In this task, you will place the *large* cube in front of the (virtual) person who is highlighted in red (see Figure below) with the side indicated by the *small* cube in the bottom-left of the screen facing up. You will do this many times in a row. The next trial will begin as soon as you press on the “Next” button. Please try to pass the cube as quickly as you can, but make sure that it is the right side up.



Docking Tetrahedrons

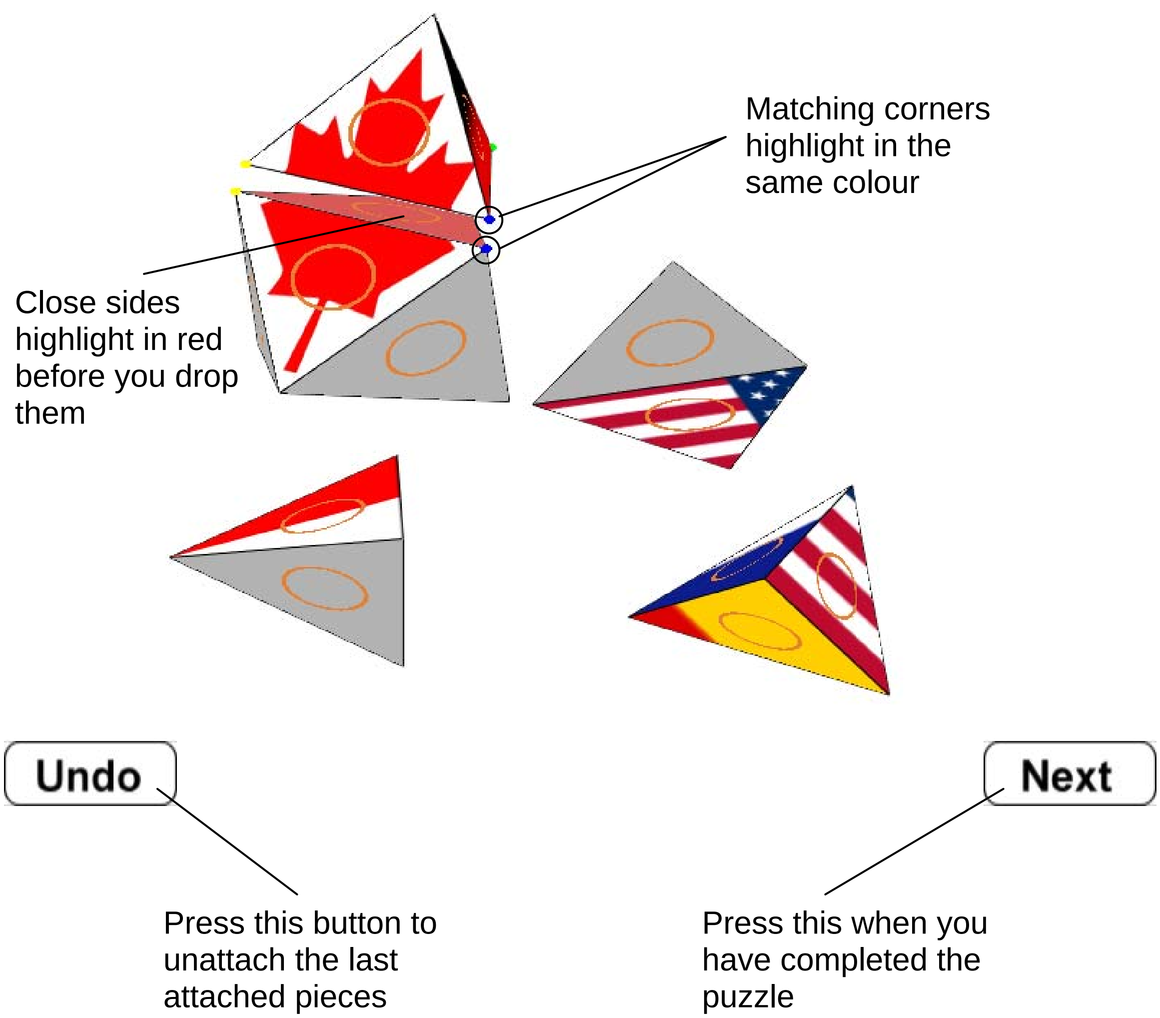
In this task, you will place the starting tetrahedron in the target tetrahedron (see Figure below). You will be required to place the tetrahedron in the *same* position and orientation as the target. The two blue edges should match up when you have the correct orientation. The corners of the object need only be as close as is indicated by the tolerance bars on the target. Note that these will *change* throughout the experiment. When you place a corner close enough to the target, it will highlight with a blue dot. Please try to dock each tetrahedron as quickly as you can.

You will be asked to do this many times in a row. To start each trial, simply press on the "Start" button located on the edge nearest you on the table. The trial will automatically complete when you have properly docked the tetrahedron.



Puzzle

In this task, you will be asked to assemble a 3D puzzle. The puzzle will always have the same shape, but the pictures on each side will be different. You can attach pieces of the puzzle by dragging them close to each other and then releasing them by lifting your finger. When two sides are close enough that they will attach, those sides will both change to a red colour and the matching corners will highlight with the same colour. If you would like to unattach two pieces that you have just attached, press the “Undo” button at the bottom-left of the screen. When you have finished the puzzle, simply press the “Next” button at the bottom-right.



B.3 Post-Experiment Questionnaire

The following pages contain the questionnaire that was given to the participants at the end of the experiment.

«Participant»

Background Questionnaire

1. Age: ____
2. Sex (circle one): M F
3. Are you colorblind? Yes No
   a. If yes, please describe the type of colorblindness:

   ____

   ____

   ____

4. Please describe your experience with 3D games.
   a. How often do you play 3D games?

   _______ times a (circle one) year / month / week / day

   b. Please check *all* types of 3D games you have played (if known):

☐ platformer	☐ racing	☐ turn-based games
☐ adventure	☐ shooting	☐ Real Time Strategy (RTS)
☐ role-playing game (RPG)	☐ fighting	☐ Stealth-based game
☐ first person shooter (FPS)	☐ action	☐ Tactical shooter
☐ third person shooter	☐ puzzle	☐ Action adventure
☐ sports	☐ simulation	☐ Other: _______________

   c. List some of the games you have played the most often:

   ____

   ____

   ____

   ____

«Participant»

5. Have you ever used a tabletop display? Yes No
    a. If yes, please explain when and where:

       ____

       ____

       ____

«Participant»

	One-Finger							Two-Finger							Three-Finger						
	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
I found the technique easy to use	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I found the technique difficult to control	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
With this technique, the object reacted as I expected it to	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I found it difficult to turn objects in the plane	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I found it difficult to roll objects over	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
I could easily move an object to where I wanted	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Please state which technique you preferred when *docking tetrahedrons* (circle one):

One-Finger Two-Finger Three-Finger No Winner No Preference

Please state which technique you preferred when *passing the cube* (circle one):

One-Finger Two-Finger Three-Finger No Winner No Preference

Please state which technique you preferred when *assembling the puzzle* (circle one):

One-Finger Two-Finger Three-Finger No Winner No Preference

Please state which technique you preferred overall (circle one):

One-Finger Two-Finger Three-Finger No Winner No Preference

B.4 SPSS Output of Analyses

B.4.1 Results of Passing Task ANOVA

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
technique	Sphericity Assumed	3415.153	2	1707.577	3.387	.052	.235
	Greenhouse-Geisser	3415.153	1.422	2401.468	3.387	.073	.235
	Huynh-Feldt	3415.153	1.573	2171.208	3.387	.067	.235
	Lower-bound	3415.153	1.000	3415.153	3.387	.093	.235
Error(technique)	Sphericity Assumed	11091.101	22	504.141			
	Greenhouse-Geisser	11091.101	15.643	709.004			
	Huynh-Feldt	11091.101	17.302	641.022			
	Lower-bound	11091.101	11.000	1008.282			
person	Sphericity Assumed	15.371	2	7.685	.063	.939	.006
	Greenhouse-Geisser	15.371	1.624	9.463	.063	.907	.006
	Huynh-Feldt	15.371	1.866	8.239	.063	.929	.006
	Lower-bound	15.371	1.000	15.371	.063	.806	.006
Error(person)	Sphericity Assumed	2665.705	22	121.168			
	Greenhouse-Geisser	2665.705	17.868	149.191			
	Huynh-Feldt	2665.705	20.523	129.891			
	Lower-bound	2665.705	11.000	242.337			
side	Sphericity Assumed	4246.510	5	849.302	11.807	.000	.518
	Greenhouse-Geisser	4246.510	3.295	1288.904	11.807	.000	.518
	Huynh-Feldt	4246.510	4.871	871.717	11.807	.000	.518
	Lower-bound	4246.510	1.000	4246.510	11.807	.006	.518
Error(side)	Sphericity Assumed	3956.122	55	71.929			
	Greenhouse-Geisser	3956.122	36.241	109.160			
	Huynh-Feldt	3956.122	53.586	73.828			
	Lower-bound	3956.122	11.000	359.647			
technique * person	Sphericity Assumed	322.510	4	80.628	.798	.533	.068
	Greenhouse-Geisser	322.510	1.991	161.959	.798	.462	.068
	Huynh-Feldt	322.510	2.431	132.693	.798	.483	.068
	Lower-bound	322.510	1.000	322.510	.798	.391	.068

Error(technique*person)	Sphericity Assumed	4445.485	44	101.034			
	Greenhouse-Geisser	4445.485	21.904	202.950			
	Huynh-Feldt	4445.485	26.736	166.276			
	Lower-bound	4445.485	11.000	404.135			
technique * side	Sphericity Assumed	686.381	10	68.638	.949	.492	.079
	Greenhouse-Geisser	686.381	4.417	155.396	.949	.450	.079
	Huynh-Feldt	686.381	7.748	88.591	.949	.479	.079
	Lower-bound	686.381	1.000	686.381	.949	.351	.079
Error(technique*side)	Sphericity Assumed	7955.215	110	72.320			
	Greenhouse-Geisser	7955.215	48.587	163.733			
	Huynh-Feldt	7955.215	85.225	93.344			
	Lower-bound	7955.215	11.000	723.201			
person * side	Sphericity Assumed	234.413	10	23.441	.323	.974	.028
	Greenhouse-Geisser	234.413	4.001	58.585	.323	.861	.028
	Huynh-Feldt	234.413	6.575	35.654	.323	.934	.028
	Lower-bound	234.413	1.000	234.413	.323	.581	.028
Error(person*side)	Sphericity Assumed	7994.371	110	72.676			
	Greenhouse-Geisser	7994.371	44.013	181.635			
	Huynh-Feldt	7994.371	72.321	110.539			
	Lower-bound	7994.371	11.000	726.761			
technique * person * side	Sphericity Assumed	779.253	20	38.963	.560	.936	.048
	Greenhouse-Geisser	779.253	5.120	152.212	.560	.734	.048
	Huynh-Feldt	779.253	10.107	77.100	.560	.845	.048
	Lower-bound	779.253	1.000	779.253	.560	.470	.048
Error(technique*person*side)	Sphericity Assumed	15299.052	220	69.541			
	Greenhouse-Geisser	15299.052	56.315	271.670			
	Huynh-Feldt	15299.052	111.178	137.609			
	Lower-bound	15299.052	11.000	1390.823			

1. technique

Estimates

Measure:MEASURE_1

technique	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	18.934	2.138	14.228	23.639
2	15.757	2.719	9.771	21.742
3	13.327	1.549	9.918	16.735

Pairwise Comparisons

Measure:MEASURE_1

(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	3.177	2.718	.267	-2.806	9.160
	3	5.607*	1.505	.003	2.294	8.920
2	1	-3.177	2.718	.267	-9.160	2.806
	3	2.430	2.085	.269	-2.160	7.020
3	1	-5.607*	1.505	.003	-8.920	-2.294
	2	-2.430	2.085	.269	-7.020	2.160

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.601	7.527[a]	2.000	10.000	.010	.601
Wilks' lambda	.399	7.527[a]	2.000	10.000	.010	.601
Hotelling's trace	1.505	7.527[a]	2.000	10.000	.010	.601
Roy's largest root	1.505	7.527[a]	2.000	10.000	.010	.601

Each F tests the multivariate effect of technique. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

2. side

Estimates

Measure:MEASURE_1

side	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	10.577	1.660	6.924	14.230
2	17.057	2.006	12.641	21.474
3	17.238	2.139	12.530	21.946
4	18.286	1.852	14.210	22.361
5	17.348	2.3499	12.175	22.520
6	15.529	1.554	12.108	18.949

Pairwise Comparisons

Measure:MEASURE_1

(I) side	(J) side	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-6.480*	.868	.000	-8.391	-4.569
	3	-6.661*	1.180	.000	-9.258	-4.064
	4	-7.708*	1.000	.000	-9.910	-5.507
	5	-6.770*	1.407	.001	-9.867	-3.674
	6	-4.952*	1.084	.001	-7.338	-2.565
2	1	6.480*	.868	.000	4.569	8.391
	3	-.181	1.424	.901	-3.314	2.953
	4	-1.228	.618	.072	-2.589	.132
	5	-.290	1.051	.788	-2.604	2.024
	6	1.529	.974	.145	-.615	3.672
3	1	6.661*	1.180	.000	4.064	9.258
	2	.181	1.424	.901	-2.953	3.314
	4	-1.048	1.169	.389	-3.622	1.526
	5	-.110	1.490	.943	-3.390	3.170
	6	1.709	1.285	.210	-1.119	4.537

4	1	7.708[*]	1.000	.000	5.507	9.910
	2	1.228	.618	.072	-.132	2.589
	3	1.048	1.169	.389	-1.526	3.622
	5	.938	1.130	.424	-1.549	3.425
	6	2.757[*]	.696	.002	1.224	4.290
5	1	6.770[*]	1.407	.001	3.674	9.867
	2	.290	1.051	.788	-2.024	2.604
	3	.110	1.490	.943	-3.170	3.390
	4	-.938	1.130	.424	-3.425	1.549
	6	1.819	1.488	.247	-1.456	5.094
6	1	4.952[*]	1.084	.001	2.565	7.338
	2	-1.529	.974	.145	-3.672	.615
	3	-1.709	1.285	.210	-4.537	1.119
	4	-2.757[*]	.696	.002	-4.290	-1.224
	5	-1.819	1.488	.247	-5.094	1.456

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.896	12.021[a]	5.000	7.000	.003	.896
Wilks' lambda	.104	12.021[a]	5.000	7.000	.003	.896
Hotelling's trace	8.586	12.021[a]	5.000	7.000	.003	.896
Roy's largest root	8.586	12.021[a]	5.000	7.000	.003	.896

Each F tests the multivariate effect of side. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

B.4.2 Results of Passing Task ANOVA, Separated by Movement Type

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
technique	Sphericity Assumed	2510.126	2	1255.063	8.542	.002	.437
	Greenhouse-Geisser	2510.126	1.645	1525.746	8.542	.004	.437
	Huynh-Feldt	2510.126	1.897	1323.502	8.542	.002	.437
	Lower-bound	2510.126	1.000	2510.126	8.542	.014	.437
Error(technique)	Sphericity Assumed	3232.318	22	146.924			
	Greenhouse-Geisser	3232.318	18.097	178.611			
	Huynh-Feldt	3232.318	20.862	154.935			
	Lower-bound	3232.318	11.000	293.847			
movement	Sphericity Assumed	149.128	2	74.564	1.170	.329	.096
	Greenhouse-Geisser	149.128	1.319	113.089	1.170	.315	.096
	Huynh-Feldt	149.128	1.428	104.437	1.170	.318	.096
	Lower-bound	149.128	1.000	149.128	1.170	.303	.096
Error(movement)	Sphericity Assumed	1401.889	22	63.722			
	Greenhouse-Geisser	1401.889	14.505	96.646			
	Huynh-Feldt	1401.889	15.707	89.252			
	Lower-bound	1401.889	11.000	127.444			
technique * movement	Sphericity Assumed	4909.335	4	1227.334	18.691	.000	.630
	Greenhouse-Geisser	4909.335	2.056	2387.725	18.691	.000	.630
	Huynh-Feldt	4909.335	2.535	1936.620	18.691	.000	.630
	Lower-bound	4909.335	1.000	4909.335	18.691	.001	.630
Error(technique*movement)	Sphericity Assumed	2889.188	44	65.663			
	Greenhouse-Geisser	2889.188	22.617	127.745			
	Huynh-Feldt	2889.188	27.885	103.611			
	Lower-bound	2889.188	11.000	262.653			

1. technique

Estimates

Measure:MEASURE_1

technique	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	3.280	.318	2.580	3.980
2	5.003	.813	3.214	6.791
3	6.034	.616	4.679	7.390

Pairwise Comparisons

Measure:MEASURE_1

(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-1.723	.813	.058	-3.512	.066
	3	-2.755*	.563	.000	-3.994	-1.515
2	1	1.723	.813	.058	-.066	3.512
	3	-1.032	.618	.123	-2.393	.329
3	1	2.755*	.563	.000	1.515	3.994
	2	1.032	.618	.123	-.329	2.393

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

2. movement

Estimates

Measure:MEASURE_1

movement	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	5.048	.592	3.744	6.351
2	4.393	.581	3.115	5.672
3	4.876	.443	3.901	5.850

Pairwise Comparisons

Measure:MEASURE_1

(I) movement	(J) movement	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	.654	.499	.217	-.445	1.754
	3	.172	.533	.753	-1.001	1.345
2	1	-.654	.499	.217	-1.754	.445
	3	-.482	.238	.068	-1.007	.042
3	1	-.172	.533	.753	-1.345	1.001
	2	.482	.238	.068	-.042	1.007

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

3. movement * technique

Estimates

Measure:MEASURE_1

movement	technique	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	2.303	.267	1.715	2.892
	2	3.746	.706	2.192	5.300
	3	9.094	1.110	6.651	11.536
2	1	2.268	.301	1.605	2.932
	2	6.366	1.377	3.336	9.397
	3	4.545	.529	3.380	5.710
3	1	5.267	.550	4.056	6.478
	2	4.896	.789	3.160	6.631
	3	4.464	.485	3.396	5.532

Pairwise Comparisons

Measure:MEASURE_1

technique	(I) movement	(J) movement	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	.035	.221	.877	-.450	.520
		3	-2.964*	.514	.000	-4.095	-1.833
	2	1	-.035	.221	.877	-.520	.450
		3	-2.999*	.415	.000	-3.913	-2.085
	3	1	2.964*	.514	.000	1.833	4.095
		2	2.999*	.415	.000	2.085	3.913
2	1	2	-2.620	1.400	.088	-5.703	.462
		3	-1.150	.777	.167	-2.861	.561
	2	1	2.620	1.400	.088	-.462	5.703
		3	1.471	.735	.071	-.147	3.088
	3	1	1.150	.777	.167	-.561	2.861
		2	-1.471	.735	.071	-3.088	.147
3	1	2	4.548*	.651	.000	3.114	5.982
		3	4.629*	1.027	.001	2.370	6.889
	2	1	-4.548*	.651	.000	-5.982	-3.114
		3	.081	.590	.893	-1.218	1.381
	3	1	-4.629*	1.027	.001	-6.889	-2.370
		2	-.081	.590	.893	-1.381	1.218

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

B.4.3 Results of Docking Task ANOVA

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
technique	Sphericity Assumed	1213.868	2	606.934	14.212	.000	.564
	Greenhouse-Geisser	1213.868	1.718	706.744	14.212	.000	.564
	Huynh-Feldt	1213.868	2.000	606.934	14.212	.000	.564
	Lower-bound	1213.868	1.000	1213.868	14.212	.003	.564
Error(technique)	Sphericity Assumed	939.511	22	42.705			
	Greenhouse-Geisser	939.511	18.893	49.728			
	Huynh-Feldt	939.511	22.000	42.705			
	Lower-bound	939.511	11.000	85.410			
rotation	Sphericity Assumed	6344.385	1	6344.385	66.719	.000	.858
	Greenhouse-Geisser	6344.385	1.000	6344.385	66.719	.000	.858
	Huynh-Feldt	6344.385	1.000	6344.385	66.719	.000	.858
	Lower-bound	6344.385	1.000	6344.385	66.719	.000	.858
Error(rotation)	Sphericity Assumed	1046.003	11	95.091			
	Greenhouse-Geisser	1046.003	11.000	95.091			
	Huynh-Feldt	1046.003	11.000	95.091			
	Lower-bound	1046.003	11.000	95.091			
difficulty	Sphericity Assumed	786.559	2	393.280	39.613	.000	.783
	Greenhouse-Geisser	786.559	1.784	440.825	39.613	.000	.783
	Huynh-Feldt	786.559	2.000	393.280	39.613	.000	.783
	Lower-bound	786.559	1.000	786.559	39.613	.000	.783
Error(difficulty)	Sphericity Assumed	218.417	22	9.928			
	Greenhouse-Geisser	218.417	19.627	11.128			
	Huynh-Feldt	218.417	22.000	9.928			
	Lower-bound	218.417	11.000	19.856			
technique * rotation	Sphericity Assumed	93.448	2	46.724	1.173	.328	.096
	Greenhouse-Geisser	93.448	1.830	51.053	1.173	.326	.096
	Huynh-Feldt	93.448	2.000	46.724	1.173	.328	.096
	Lower-bound	93.448	1.000	93.448	1.173	.302	.096

Error(technique*rotation)	Sphericity Assumed	876.659	22	39.848			
	Greenhouse-Geisser	876.659	20.134	43.540			
	Huynh-Feldt	876.659	22.000	39.848			
	Lower-bound	876.659	11.000	79.696			
technique * difficulty	Sphericity Assumed	26.419	4	6.605	.451	.771	.039
	Greenhouse-Geisser	26.419	2.616	10.101	.451	.693	.039
	Huynh-Feldt	26.419	3.505	7.538	.451	.747	.039
	Lower-bound	26.419	1.000	26.419	.451	.516	.039
Error(technique*difficulty)	Sphericity Assumed	644.867	44	14.656			
	Greenhouse-Geisser	644.867	28.771	22.414			
	Huynh-Feldt	644.867	38.554	16.726			
	Lower-bound	644.867	11.000	58.624			
rotation * difficulty	Sphericity Assumed	92.379	2	46.189	5.079	.015	.316
	Greenhouse-Geisser	92.379	1.470	62.842	5.079	.027	.316
	Huynh-Feldt	92.379	1.641	56.288	5.079	.023	.316
	Lower-bound	92.379	1.000	92.379	5.079	.046	.316
Error(rotation*difficulty)	Sphericity Assumed	200.063	22	9.094			
	Greenhouse-Geisser	200.063	16.170	12.372			
	Huynh-Feldt	200.063	18.053	11.082			
	Lower-bound	200.063	11.000	18.188			
technique * rotation * difficulty	Sphericity Assumed	17.150	4	4.287	.492	.742	.043
	Greenhouse-Geisser	17.150	2.892	5.930	.492	.684	.043
	Huynh-Feldt	17.150	4.000	4.287	.492	.742	.043
	Lower-bound	17.150	1.000	17.150	.492	.498	.043
Error(technique*rotation*difficulty)	Sphericity Assumed	383.649	44	8.719			
	Greenhouse-Geisser	383.649	31.811	12.060			
	Huynh-Feldt	383.649	44.000	8.719			
	Lower-bound	383.649	11.000	34.877			

1. technique

Estimates

Measure:MEASURE_1

technique	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	20.129	.855	18.246	22.012
2	17.040	1.266	14.253	19.827
3	14.326	1.295	11.476	17.177

Pairwise Comparisons

Measure:MEASURE_1

(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
1	2	3.089*	.931	.007	1.040	5.138
	3	5.803*	1.017	.000	3.564	8.042
2	1	-3.089*	.931	.007	-5.138	-1.040
	3	2.714	1.287	.059	-.120	5.547
3	1	-5.803*	1.017	.000	-8.042	-3.564
	2	-2.714	1.287	.059	-5.547	.120

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

2. rotation

Estimates

Measure:MEASURE_1

rotation	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	11.745	1.258	8.977	14.514
2	22.585	1.088	20.191	24.978

3. technique * rotation

Estimates

Measure:MEASURE_1

technique	rotation	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	14.997	1.579	11.522	18.471
	2	25.261	1.032	22.989	27.533
2	1	10.710	1.512	7.383	14.037
	2	23.369	1.375	20.344	26.395
3	1	9.529	1.769	5.634	13.424
	2	19.124	1.421	15.997	22.250

Pairwise Comparisons

Measure:MEASURE_1

technique	(I) rotation	(J) rotation	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	-10.264*	2.046	.000	-14.768	-5.760
	2	1	10.264*	2.046	.000	5.760	14.768
2	1	2	-12.659*	1.391	.000	-15.720	-9.598
	2	1	12.659*	1.391	.000	9.598	15.720
3	1	2	-9.594*	1.894	.000	-13.764	-5.425
	2	1	9.594*	1.894	.000	5.425	13.764

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

4. technique * rotation

Estimates

Measure:MEASURE_1

technique	rotation	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	14.997	1.579	11.522	18.471
	2	25.261	1.032	22.989	27.533
2	1	10.710	1.512	7.383	14.037
	2	23.369	1.375	20.344	26.395
3	1	9.529	1.769	5.634	13.424
	2	19.124	1.421	15.997	22.250

Pairwise Comparisons

Measure:MEASURE_1

rotation	(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	4.287*	1.605	.022	.753	7.820
		3	5.468*	1.858	.013	1.378	9.557
	2	1	-4.287*	1.605	.022	-7.820	-.753
		3	1.181	1.859	.538	-2.910	5.273
	3	1	-5.468*	1.858	.013	-9.557	-1.378
		2	-1.181	1.859	.538	-5.273	2.910
2	1	2	1.892	1.262	.162	-.887	4.670
		3	6.138*	1.127	.000	3.658	8.618
	2	1	-1.892	1.262	.162	-4.670	.887
		3	4.246*	1.188	.004	1.632	6.860
	3	1	-6.138*	1.127	.000	-8.618	-3.658
		2	-4.246*	1.188	.004	-6.860	-1.632

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

B.4.4 Results of Docking Task ANOVA, Separated by Movement Type

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
technique	Sphericity Assumed	68588.180	2	34294.090	16.352	.000	.598
	Greenhouse-Geisser	68588.180	1.598	42928.335	16.352	.000	.598
	Huynh-Feldt	68588.180	1.826	37552.557	16.352	.000	.598
	Lower-bound	68588.180	1.000	68588.180	16.352	.002	.598
Error(technique)	Sphericity Assumed	46138.979	22	2097.226			
	Greenhouse-Geisser	46138.979	17.575	2625.246			
	Huynh-Feldt	46138.979	20.091	2296.495			
	Lower-bound	46138.979	11.000	4194.453			
movement	Sphericity Assumed	16720.255	2	8360.128	9.711	.001	.469
	Greenhouse-Geisser	16720.255	1.952	8564.350	9.711	.001	.469
	Huynh-Feldt	16720.255	2.000	8360.128	9.711	.001	.469
	Lower-bound	16720.255	1.000	16720.255	9.711	.010	.469
Error(movement)	Sphericity Assumed	18939.097	22	860.868			
	Greenhouse-Geisser	18939.097	21.475	881.897			
	Huynh-Feldt	18939.097	22.000	860.868			
	Lower-bound	18939.097	11.000	1721.736			
technique * movement	Sphericity Assumed	47287.616	4	11821.904	16.542	.000	.601
	Greenhouse-Geisser	47287.616	2.660	17779.260	16.542	.000	.601
	Huynh-Feldt	47287.616	3.587	13183.121	16.542	.000	.601
	Lower-bound	47287.616	1.000	47287.616	16.542	.002	.601
Error(technique*movement)	Sphericity Assumed	31445.816	44	714.678			
	Greenhouse-Geisser	31445.816	29.257	1074.822			
	Huynh-Feldt	31445.816	39.457	796.968			
	Lower-bound	31445.816	11.000	2858.711			

1. technique

Estimates

Measure:MEASURE_1

technique	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	23.860	1.351	20.888	26.833
2	38.488	4.111	29.440	47.536
3	48.946	4.494	39.056	58.836

Pairwise Comparisons

Measure:MEASURE_1

(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-14.628*	3.596	.002	-22.543	-6.713
	3	-25.085*	4.046	.000	-33.990	-16.181
2	1	14.628*	3.596	.002	6.713	22.543
	3	-10.458	5.381	.078	-22.302	1.387
3	1	25.085*	4.046	.000	16.181	33.990
	2	10.458	5.381	.078	-1.387	22.302

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

2. movement

Estimates

Measure:MEASURE_1

movement	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	39.366	3.469	31.730	47.001
2	30.061	1.939	25.793	34.329
3	41.868	3.417	34.347	49.388

Pairwise Comparisons

Measure:MEASURE_1

(I) movement	(J) movement	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	9.304*	2.888	.008	2.948	15.660
	3	-2.502	2.969	.417	-9.038	4.033
2	1	-9.304*	2.888	.008	-15.660	-2.948
	3	-11.807*	2.599	.001	-17.528	-6.086
3	1	2.502	2.969	.417	-4.033	9.038
	2	11.807*	2.599	.001	6.086	17.528

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

3. technique * movement

Estimates

Measure:MEASURE_1

technique	movement	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	17.444	1.229	14.738	20.150
	2	14.345	1.869	10.232	18.458
	3	39.792	4.476	29.940	49.644
2	1	34.401	5.071	23.241	45.561
	2	39.080	4.117	30.018	48.142
	3	41.984	5.708	29.420	54.548
3	1	66.251	6.355	52.265	80.238
	2	36.759	4.828	26.133	47.385
	3	43.827	4.048	34.917	52.738

Pairwise Comparisons

Measure:MEASURE_1

technique	(I) movement	(J) movement	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	3.099	1.895	.130	-1.072	7.270
		3	-22.348*	4.856	.001	-33.036	-11.661
	2	1	-3.099	1.895	.130	-7.270	1.072
		3	-25.447*	5.615	.001	-37.806	-13.088
	3	1	22.348*	4.856	.001	11.661	33.036
		2	25.447*	5.615	.001	13.088	37.806
2	1	2	-4.678	6.226	.468	-18.382	9.025
		3	-7.583	4.255	.102	-16.948	1.783
	2	1	4.678	6.226	.468	-9.025	18.382
		3	-2.904	4.096	.493	-11.919	6.111
	3	1	7.583	4.255	.102	-1.783	16.948
		2	2.904	4.096	.493	-6.111	11.919
3	1	2	29.492*	3.496	.000	21.798	37.187
		3	22.424*	5.232	.001	10.909	33.938
	2	1	-29.492*	3.496	.000	-37.187	-21.798
		3	-7.069	4.349	.132	-16.642	2.504
	3	1	-22.424*	5.232	.001	-33.938	-10.909
		2	7.069	4.349	.132	-2.504	16.642

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

B.4.5 Results of ANOVA for Incomplete Trials

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
technique	Sphericity Assumed	43.815	2	21.907	7.307	.004	.399
	Greenhouse-Geisser	43.815	1.440	30.424	7.307	.010	.399
	Huynh-Feldt	43.815	1.599	27.409	7.307	.007	.399
	Lower-bound	43.815	1.000	43.815	7.307	.021	.399
Error(technique)	Sphericity Assumed	65.963	22	2.998			
	Greenhouse-Geisser	65.963	15.842	4.164			
	Huynh-Feldt	65.963	17.584	3.751			
	Lower-bound	65.963	11.000	5.997			
rotation	Sphericity Assumed	14.519	1	14.519	14.767	.003	.573
	Greenhouse-Geisser	14.519	1.000	14.519	14.767	.003	.573
	Huynh-Feldt	14.519	1.000	14.519	14.767	.003	.573
	Lower-bound	14.519	1.000	14.519	14.767	.003	.573
Error(rotation)	Sphericity Assumed	10.815	11	.983			
	Greenhouse-Geisser	10.815	11.000	.983			
	Huynh-Feldt	10.815	11.000	.983			
	Lower-bound	10.815	11.000	.983			
difficulty	Sphericity Assumed	10.954	2	5.477	9.396	.001	.461
	Greenhouse-Geisser	10.954	1.303	8.404	9.396	.005	.461
	Huynh-Feldt	10.954	1.407	7.786	9.396	.004	.461
	Lower-bound	10.954	1.000	10.954	9.396	.011	.461
Error(difficulty)	Sphericity Assumed	12.824	22	.583			
	Greenhouse-Geisser	12.824	14.338	.894			
	Huynh-Feldt	12.824	15.475	.829			
	Lower-bound	12.824	11.000	1.166			
technique * rotation	Sphericity Assumed	6.704	2	3.352	7.932	.003	.419
	Greenhouse-Geisser	6.704	1.483	4.520	7.932	.007	.419
	Huynh-Feldt	6.704	1.660	4.039	7.932	.005	.419
	Lower-bound	6.704	1.000	6.704	7.932	.017	.419

Error(technique*rotation)	Sphericity Assumed	9.296	22	.423			
	Greenhouse-Geisser	9.296	16.313	.570			
	Huynh-Feldt	9.296	18.258	.509			
	Lower-bound	9.296	11.000	.845			
technique * difficulty	Sphericity Assumed	3.241	4	.810	1.912	.125	.148
	Greenhouse-Geisser	3.241	2.841	1.141	1.912	.151	.148
	Huynh-Feldt	3.241	3.933	.824	1.912	.127	.148
	Lower-bound	3.241	1.000	3.241	1.912	.194	.148
Error(technique*difficulty)	Sphericity Assumed	18.648	44	.424			
	Greenhouse-Geisser	18.648	31.252	.597			
	Huynh-Feldt	18.648	43.268	.431			
	Lower-bound	18.648	11.000	1.695			
rotation * difficulty	Sphericity Assumed	.454	2	.227	.459	.638	.040
	Greenhouse-Geisser	.454	1.334	.340	.459	.564	.040
	Huynh-Feldt	.454	1.450	.313	.459	.579	.040
	Lower-bound	.454	1.000	.454	.459	.512	.040
Error(rotation*difficulty)	Sphericity Assumed	10.880	22	.495			
	Greenhouse-Geisser	10.880	14.679	.741			
	Huynh-Feldt	10.880	15.948	.682			
	Lower-bound	10.880	11.000	.989			
technique * rotation * difficulty	Sphericity Assumed	4.574	4	1.144	2.833	.036	.205
	Greenhouse-Geisser	4.574	2.226	2.055	2.833	.073	.205
	Huynh-Feldt	4.574	2.817	1.624	2.833	.057	.205
	Lower-bound	4.574	1.000	4.574	2.833	.120	.205
Error(technique*rotation*difficulty)	Sphericity Assumed	17.759	44	.404			
	Greenhouse-Geisser	17.759	24.489	.725			
	Huynh-Feldt	17.759	30.986	.573			
	Lower-bound	17.759	11.000	1.614			

1. technique

Estimates

Measure:MEASURE_1

technique	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	1.333	.338	.589	2.078
2	.611	.198	.175	1.047
3	.250	.073	.090	.410

Pairwise Comparisons

Measure:MEASURE_1

(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	.722	.360	.070	-.069	1.514
	3	1.083*	.285	.003	.457	1.710
2	1	-.722	.360	.070	-1.514	.069
	3	.361	.199	.097	-.076	.799
3	1	-1.083*	.285	.003	-1.710	-.457
	2	-.361	.199	.097	-.799	.076

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

2. rotation

Estimates

Measure:MEASURE_1

rotation	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.472	.160	.120	.825
2	.991	.184	.585	1.396

Pairwise Comparisons

Measure:MEASURE_1

(I) rotation	(J) rotation	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.519*	.135	.003	-.816	-.222
2	1	.519*	.135	.003	.222	.816

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

3. technique * rotation

Estimates

Measure:MEASURE_1

technique	rotation	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.833	.303	.167	1.500
	2	1.833	.413	.924	2.743
2	1	.417	.206	-.036	.869
	2	.806	.219	.324	1.287
3	1	.167	.087	-.025	.358
	2	.333	.092	.131	.535

Pairwise Comparisons

Measure:MEASURE_1

technique	(I) rotation	(J) rotation	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	-1.000*	.259	.003	-1.571	-.429
	2	1	1.000*	.259	.003	.429	1.571
2	1	2	-.389*	.153	.027	-.725	-.053
	2	1	.389*	.153	.027	.053	.725
3	1	2	-.167	.105	.139	-.397	.064
	2	1	.167	.105	.139	-.064	.397

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

4. technique * rotation

Estimates

Measure:MEASURE_1

technique	rotation	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.833	.303	.167	1.500
	2	1.833	.413	.924	2.743
2	1	.417	.206	-.036	.869
	2	.806	.219	.324	1.287
3	1	.167	.087	-.025	.358
	2	.333	.092	.131	.535

Pairwise Comparisons

Measure:MEASURE_1

rotation	(I) technique	(J) technique	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	1	2	.417	.305	.199	-.254	1.087
		3	.667*	.243	.019	.132	1.201
	2	1	-.417	.305	.199	-1.087	.254
		3	.250	.206	.250	-.203	.703
	3	1	-.667*	.243	.019	-1.201	-.132
		2	-.250	.206	.250	-.703	.203
2	1	2	1.028*	.448	.043	.041	2.015
		3	1.500*	.357	.001	.715	2.285
	2	1	-1.028*	.448	.043	-2.015	-.041
		3	.472	.219	.054	-.010	.954
	3	1	-1.500*	.357	.001	-2.285	-.715
		2	-.472	.219	.054	-.954	.010

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

*. The mean difference is significant at the .05 level.

Materials for Participatory Design C

This appendix includes the material used in the participatory design process described in chapter 7.

C.1 Informed Consent Form

The following pages show the informed consent form that was presented to the participants, and the participants were asked to sign, before the day-long session began.

UNIVERSITY OF CALGARY

Name of Researcher, Faculty, Department, Telephone & Email:
Mark Hancock, Ph.D. Student, Dept. of Computer Science, (403) 210-9499, msh@cs.ucalgary.ca

Supervisor:
Sheelagh Carpendale, Associate Professor, Dept. of Computer Science, (403) 220-6055, sheelagh@cs.ucalgary.ca

Title of Project:
Computer Feedback in Co-Located Collaborative Environments

Sponsor:
National Sciences and Engineering Research Council, Canadian Foundation for Innovation, and Alberta Ingenuity

---

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Purpose of the Study:

We are currently investigating how technology can be used to help support collaboration. We are investigating both the effects of technology on how you collaborate and on evaluating how well the software and hardware supports collaborative interactions.

What Will I Be Asked To Do?

You will be asked to use different interaction techniques and devices that offer a variety of support for organizing and sharing information such as documents and images. This may involve such tasks as moving digital objects to different locations, passing digital objects between collaborators, and organizing digital objects for shared access.

We will be observing and programmatically capturing your actions, as well as videotaping you during the course of the session. This videotaping is optional and you can still participate if you choose not to be videotaped. You will also be asked to complete a pre- and post-session questionnaire to further our investigation. It is estimated that your involvement will take approximately one hour, and you will be remunerated for your time.

Participation in this study is completely voluntary and you may refuse to participate altogether or at any time during the study without any repercussions.

What Type of Personal Information Will Be Collected?

Should you agree to participate, you will be asked to provide both your gender and age. No other personal identifying information will be collected in this study, and all participants shall remain anonymous.

There are several options for you to consider if you decide to take part in this research. You can choose all, some or none of them. Please put a check mark on the corresponding line(s) that grants me your permission to:

I grant permission to be audio taped: Yes: ___ No: ___

I grant permission to be videotaped: Yes: ___ No: ___

I grant permission to have video or still images of me used in a publication: Yes: ___ No: ___

Are there Risks or Benefits if I Participate?

There are no known risks involved in participating in this experiment. You will be given $10 as compensation for your time. If you decide for any reason to withdraw from the experiment, you will still receive this compensation.

What Happens to the Information I Provide?

Participation is completely voluntary and confidential. Note that because you will be working with others to complete the experiment, the participants in the group besides yourself may be able to identify you. Besides this limitation, you will remain completely anonymous. You are free to discontinue participation at any time during the study. If you decide to discontinue participation, data collected up to that point will be destroyed. The research data, including your answers to the questionnaire and the video recording, will only be viewed by the researcher, his supervisor, and a team of graduate students. There are no names on the questionnaire. Only group information will be summarized for any presentation or publication of results. The questionnaires are kept in a locked cabinet only accessible by the researcher, his supervisor, and a team of graduate students. The anonymous data will be stored for three years on a computer disk, at which time, it will be permanently erased.

Signatures (written consent)

Your signature on this form indicates that you 1) understand to your satisfaction the information provided to you about your participation in this research project, and 2) agree to participate as a research subject.

In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Participant's Name: (please print) ____

Participant's Signature ____Date: _______________

Researcher's Name: (please print) ____

Researcher's Signature: ____Date: ________________

Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Mr. Mark Hancock and Dr. Sheelagh Carpendale
Department of Computer Science
(403) 210-9499, {msh,sheelagh}@cs.ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact Patricia Evans, Associate Director, Research Services Office, University of Calgary at (403) 220-3782; email plevans@ucalgary.ca.

A copy of this consent form has been given to you to keep for your records and reference. The investigator has kept a copy of the consent form.

C.2 Interview Questions

The following is a list of questions that were used to start the interview during the day-long session. Note that these questions were intended to spark conversation, and are not an exhaustive list of all that was discussed.

1. How does a person approach the sandtray?
2. What triggers storytelling?
3. How does someone choose objects?
4. What type of patients is this good for?
5. How does sandtray therapy help these patients?
6. What is the role of the therapist?
7. How do the patient's mental processes translate themselves into stories in the sandtray?
8. What limitations does a physical sandtray pose to this expressivity?

C.3 Study Notes: Comparison of Physical to Digital

The following page shows a chart that was created based on the discussion with the sandtray therapists on the relative advantages of digital and physical media.

Perspectives on the digital sandtray

Physical vs. digital

Comments	Physical	Concept	Digital	Comments
very natural	pick up, put down	moving	combination of techniques (dragging with one finger; rotating with two fingers; move up/down by changing distance between fingers; rotate in 3D with third finger)	not as natural as its physical counterpart, but still quite good
approximation: multiple objects of different sizes	not possible	resizing	scale, stretch etc.	very good for storytelling
very natural, reliable and familiar	happens automatically and cannot be controlled	collisions	collisions between objects can be turned on or off at will	objects can 'hide' inside other objects
very natural, reliable and familiar	happens automatically and cannot be controlled	gravity	gravity can be changed globally or per object	ability to make objects 'float' or be very light/ heavy
very natural, easy to sculpt, possible to bury things	sand in a tray	surface	flat surface or simulated sand, can be (locally) changed into different surface, e.g. concrete, wood, grass...	sculpting can be hard, but possibility to paint
very natural	happens automatically and cannot be controlled	visibility	objects can be made transparent or invisible	transparency to create 'ghosts', invisibility could represent hiding

Implementation Details D

The full Javadoc listing of the codebase used throughout this dissertation can be found at http://code.markhancock.ca/DissertationDoc/. In this appendix, I describe a few examples of how this codebase can be used.

D.1 Creating a Simple 3D Environment

This code sample demonstrates how to use this codebase to create a 3D virtual object that can be used on a SMART Table, Microsoft Surface, or other multitouch device, using the one-touch listener described in chapter 5. There are several important points to note about this code:

- The `DraggableObject` class represents the 3D virtual object, and can be one of many choices, including `Cube` (a simple cube) and `MeshObject` (an object that can be loaded from an .obj file).
- The `SpaceDeformation` class represents the deformation, and can be one of the many described in chapter 3, including a parallel projection with a COP directly above the table (as shown here using the `StandardDeformation` class).
- The `TouchListener` class represents the interaction technique for the draggable object, and can be one of the many described in chapter 5, including the `OneTouchListener`, as shown here.
- The `ObjectTouchManager` and `InputEventListener` handle the multitouch device and

multipoint events being used, respectively.

- The `ScreenRenderer` and `Picker` classes handle low-level rendering and picking code.
- Each of these classes can be overridden to provide customized code. For example, the `TouchListener` class can be overridden if the programmer wishes to create their own interaction technique.
- The `SimplestDemo` class provides the basic main method necessary to setup the OpenGL context and to run this code in a Java application.

```
package examples.simplestdemo;

import examples.ApplicationUtilities;
import glprimitives.core.ViewVolume;
import glprimitives.objects.*;
import glprimitives.picking.*;
import glprimitives.rendering.ScreenRenderer;

import javax.media.opengl.*;
import javax.vecmath.*;

import sandbox.ShallowDepthInteractor;
import sandbox.interaction.*;
import sandbox.listeners.InputEventListener;
import sandbox.projection.*;

public class SimplestDemoGL extends GLCanvas implements
   GLEventListener, ShallowDepthInteractor {
    private static ViewVolume VIEW_VOLUME = new ViewVolume(-4, 4,
        -4, 4, 3, 19);

    private DraggableObject object;
    private SpaceDeformation deformation;
    private TouchListener<DraggableObject> touchListener;

    private ObjectTouchManager touchManager;
    private InputEventListener inputEventListener;

    private Picker<ManagedTouchable> picker;
    private ScreenRenderer renderer;

    public SimplestDemoGL() {
        addGLEventListener(this);

        touchListener = new OneTouchListener();

        renderer = new ScreenRenderer(this);
        picker = new Picker<ManagedTouchable>(this);

        inputEventListener = ApplicationUtilities.setupListener(this
            , this);
        touchManager = new ObjectTouchManager(inputEventListener,
            picker);
    }
```

Listing D.1: Import statements, class declaration, member variables, and constructor.

```
    @Override
    public void display(GLAutoDrawable drawable) {
        touchManager.handleTouches(drawable);

        renderer.render(drawable);
        inputEventListener.getCurrentTouchState().
           transformAndDrawForScreen(drawable);
        deformation.transformAndDrawForScreen(drawable);

        repaint();
    }

    @Override
    public void displayChanged(GLAutoDrawable arg0, boolean arg1,
       boolean arg2) { }

    @Override
    public void init(GLAutoDrawable drawable) {
        drawable.setGL(new DebugGL(drawable.getGL()));

        final float b = 0.8f;
        object = new Cube(this, new Point3f(0, 0, -4));
        renderer.addRenderable(object);
        object.setPicker(picker);
        object.init(drawable);
        object.setTouchListener(touchListener);
        renderer.init(drawable);
        inputEventListener.init();
    }

    @Override
    public void reshape(GLAutoDrawable drawable, int x, int y, int w
       , int h) {
        deformation = new StandardDeformation(w, h, VIEW_VOLUME,
           ProjectionType.PARALLEL);
    }

    @Override
    public SpaceDeformation getDeformation() {
        return deformation;
    }

    @Override
    public void setDeformation(SpaceDeformation deformation) {
        this.deformation = deformation;
        repaint();
    }
} // End SimplestDemoGL class
```

Listing D.2: Overridden Methods from GLEventListener and ShallowDepthInteractor

```
package examples.simplestdemo;

import java.awt.Dimension;
import java.awt.HeadlessException;

import javax.media.opengl.GLCanvas;
import javax.media.opengl.GLCapabilities;
import javax.swing.JFrame;
import javax.swing.SwingUtilities;

import examples.ScreenType;

public class SimplestDemo extends JFrame
{
    public SimplestDemo() throws HeadlessException
    {
        super("Demo 3D");
        setDefaultCloseOperation(JFrame.EXIT_ON_CLOSE);

        GLCanvas glPanel = new SimplestDemoGL();
        glPanel.setPreferredSize(new Dimension(1024, 768));
        getContentPane().add(glPanel);

        ScreenType.WINDOW.setWindowBounds(this);
    }

    public static void main(String[] args)
    {
        SwingUtilities.invokeLater(new Runnable()
        {
            public void run()
            {
                final SimplestDemo frame = new SimplestDemo();
                frame.setVisible(true);
            }
        });
    }
}
```

Listing D.3: The main program for the Simplest Demo application

D.2 Multipoint Interaction Techniques

One potential use that a future coder might have with this codebase is to create their own interaction technique that makes use of miltipoint interaction. In order to achieve this, the coder simply has to override the `ClassicTouchListener` class and then override the `getTransform` method, which determines the transform based on the current state of the touch points. The following listings show how the `OneToucListener`, `TwoTouchListener`, and `ThreeTouchListener` are implemented to demonstrate how this can be achieved.

```
@Override
protected Matrix4f getTransform(GLAutoDrawable drawable,
        Point3f worldTrueCentre, Point3f worldCentre,
        TouchState touchState, TouchState lastTouchState)
{
    final GL gl = drawable.getGL();

    Point2f t1 = lastTouchState.getTouchPoint(0);
    Point2f t1Prime = touchState.getTouchPoint(0);
    if (t1.distance(t1Prime) > DISTANCE_OF_DISBELIEF)
    {
        return GLUtilities.newIdentity();
    }
    Point3f p1 = computeP1();
    Point3f p1Prime = computeP1Prime(gl, t1Prime, p1);

    Point3f offsetCentre = new Point3f(worldTrueCentre);
    offsetCentre.z += offset;

    Matrix4f transMat = GLUtilities.newIdentity();

    if (isDedicatedRegionActive("translate"))
    {
        translateXY(transMat, p1, p1Prime);
    }
    else if (isDedicatedRegionActive("rotate"))
    {
        twoDimensionalRNT(gl, transMat, offsetCentre, p1, p1Prime);
    }
    else
    {
        threeDimensionalRNT(transMat, offsetCentre, p1, p1Prime);
    }

    restoreZ(gl, transMat, worldCentre);

    return transMat;
}
```

Listing D.4: The transformation which happens during the one-touch interaction

```
@Override
protected Matrix4f getTransform(GLAutoDrawable drawable,
        Point3f worldTrueCentre, Point3f worldCentre,
        TouchState touchState, TouchState lastTouchState)
{
    final GL gl = drawable.getGL();

    Point2f t1 = lastTouchState.getTouchPoint(0), t1Prime =
       touchState
            .getTouchPoint(0);
    if (t1.distance(t1Prime) > DISTANCE_OF_DISBELIEF)
    {
        return GLUtilities.newIdentity();
    }
    Point3f p1 = computeP1(), p1Prime = computeP1Prime(gl, t1Prime,
       p1);

    Matrix4f transMat = GLUtilities.newIdentity();

    // Rotate roll and pitch
    rotateRollPitch(transMat, lastTouchState.getTouchPoint(1),
       touchState
            .getTouchPoint(1), worldCentre);


    // Rotate yaw and translate
    if (isDedicatedRegionActive("translate"))
    {
        translateXY(transMat, p1, p1Prime);
    }
    else
    {
        twoDimensionalRNT(gl, transMat, worldTrueCentre, p1, p1Prime
           );
    }

    // force the object back into its plane
    restoreZ(gl, transMat, worldCentre);

    return transMat;
}
```

Listing D.5: The transformation which happens during the two-touch interaction

```
@Override
protected Matrix4f getTransform(GLAutoDrawable drawable,
        Point3f worldTrueCentre, Point3f worldCentre,
        TouchState touchState, TouchState lastTouchState)
{
    final GL gl = drawable.getGL();

    Point2f t1 = lastTouchState.getTouchPoint(0), t1Prime =
       touchState
            .getTouchPoint(0);
    if (t1.distance(t1Prime) > DISTANCE_OF_DISBELIEF)
    {
        return GLUtilities.newIdentity();
    }
    Point3f p1 = computeP1(), p1Prime = computeP1Prime(gl, t1Prime,
       p1);

    Matrix4f transMat = GLUtilities.newIdentity();

    // Rotate roll and pitch
    rotateRollPitch(transMat, lastTouchState.getTouchPoint(2),
       touchState
            .getTouchPoint(2), p1);

    // Rotate yaw
    rotateYaw(transMat, t1, t1Prime, lastTouchState.getTouchPoint(1)
       ,
            touchState.getTouchPoint(1), p1);

    // Translate
    translateXY(transMat, p1, p1Prime);

    // force the object back into its plane
    restoreZ(gl, transMat, worldCentre);

    return transMat;
}
```

Listing D.6: The transformation which happens during the three-touch interaction

```
/**
 * Modifies transMat to include
 * a translation over the vector from p to pPrime.
 */
protected void translateXY(Matrix4f transMat, Point3f p, Point3f
   pPrime)
{
    GLUtilities.translate(transMat, pPrime);
    GLUtilities.translateNeg(transMat, p);
}

/**
 * Modifies transMat to perform a three-dimensional
 * rotate and translate operation that brings p to pPrime.
 */
protected void threeDimensionalRNT(Matrix4f transMat, Point3f
   worldCentre,
        Point3f p, Point3f pPrime)
{
    GLUtilities.translate(transMat, pPrime);

    Vector3f from = GLUtilities.getVector3f(worldCentre, p);
    Vector3f to = GLUtilities.getVector3f(worldCentre, pPrime);

    Vector3f axis = new Vector3f();
    axis.cross(from, to);
    float angle = from.angle(to);

    GLUtilities.rotate(transMat, axis, angle);

    GLUtilities.translateNeg(transMat, p);
}
```

Listing D.7: Helper functions for the one-, two-, and three-touch listener

```
/**
 * Modifies transMat to perform
 * a two-dimensional rotate and translate operation that brings p to
     pPrime.
 * If the projected object centre is too close to the touch location
    , instability would occur.
 * So in that case, we do translation only.
 */
protected void twoDimensionalRNT(GL gl, Matrix4f transMat,
        Point3f worldCentre, Point3f p, Point3f pPrime)
{
    GLUtilities.translate(transMat, pPrime);

    // pretend that the centre is at the same depth as the touch;
    // this feels more natural when a perspective projection skews
       the projected centre far away from the touch
    Point3f correctedCentre = new Point3f(worldCentre);
    correctedCentre.z = p.z;

    // project centre and touch onto the window to find out if the
       touch is in the dead zone
    Point3f projectedCentre = GLUtilities.project(gl,
       correctedCentre);
    Point3f projectedTouch = GLUtilities.project(gl, p);
    projectedCentre.z = 0;
    projectedTouch.z = 0;
    float angleMultiplier;
    if (projectedCentre.distance(projectedTouch) >
       TWO_D_RNT_DEAD_ZONE_SIZE)
    {
        angleMultiplier = 1.0f;
    }
    else
    {
        angleMultiplier = 0.0f;
    }
```

Listing D.8: Helper functions for the one-, two-, and three-touch listener (cont'd)

```
    // find centre to touch vectors
    Vector3f from = GLUtilities.getVector3f(worldCentre, p);
    Vector3f to = GLUtilities.getVector3f(worldCentre, pPrime);

    // flatten them to 2D
    from.z = 0;
    to.z = 0;

    // perform rotation from first to second vector
    float angle = from.angle(to);
    Vector3f axis = new Vector3f();
    axis.cross(from, to);
    GLUtilities.rotate(transMat, axis, angle * angleMultiplier);

    GLUtilities.translateNeg(transMat, p);
}


/**
 * Modifies transMat such that it performs a pitch and roll rotation
     about rotPoint.
 * The rotation axis is perpendicular to the vector from t to tPrime
     .
 * The rotation angle is determined by the length of this vector.
 */
protected void rotateRollPitch(Matrix4f transMat, Point2f t,
        Point2f tPrime, Point3f rotPoint)
{
    if (t != null && tPrime != null)
    {
        Point3f q2 = new Point3f(t.x, -t.y, 0);
        Point3f q2Prime = new Point3f(tPrime.x, -tPrime.y, 0);

        // rotation axis is in the z plane, perpendicular to the
            movement of the touch
        Vector3f delta = GLUtilities.getVector3f(q2, q2Prime);
        Vector3f axis = new Vector3f();
        axis.cross(delta, new Vector3f(0, 0, -1));

        // translate rotation centre to origin, rotate, then
            translate back
        GLUtilities.translate(transMat, rotPoint);
        GLUtilities.rotate(transMat, axis, rollPitchSensitivity
                * delta.length());
        GLUtilities.translateNeg(transMat, rotPoint);
    }
}
```

Listing D.9: Helper functions for the one-, two-, and three-touch listener (cont'd)