The Effectiveness of Three Dimensional Interaction

by

James Boritz

B. Math. Honours Co-op Computer Science, University of Waterloo, 1988

M. Math. Computer Science, University of Waterloo, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Department of Computer Science)

we accept this thesis as conforming
to the required standard

The University of British Columbia

October 1998

Abstract

Most interaction with computers today takes place in a two dimensional environment. Even when using three dimensional graphics applications, input is often still restricted to two dimensions. Many believe that the use of three dimensional input devices will alleviate this restriction and allow for a much more natural human-machine dialog.

This thesis seeks to establish how factors dealing with visual feedback and task structure affect the ability to perform interactive tasks in a three dimensional virtual environment. The factors investigated were *stereoscopic vision*, *motion parallax*, *stimulus arrangement* and *stimulus complexity*. Four tasks were studied. These tasks were: point location, docking, line tracing and curve tracing. All the tasks used a six degree of freedom input device to control a pointer in a three dimensional virtual environment.

Four experiments corresponding to the four tasks were conducted to investigate these factors. Among other things the results showed the following. Stereoscopic vision provided a strong benefit to positioning-based tasks, but this benefit was weakened in the case of tracing tasks. Motion parallax via head-tracking often had no effect upon task performance and where an effect was found it was often detrimental. The position of stimuli influenced performance across all of the tasks. The orientation of stimuli influenced performance in the task in which it was varied.

Contents

List of Tables

List of Figures

Acknowledgements

A thesis is a significant achievement in many ways. It is an academic achievement that indicates that the author is now ready to enter the academic research community as an independent researcher. However, it is also a tremendous personal achievement showing the determination and endurance of the author in pursuing a single goal over many years.

Many people provided help and support over the years during which this thesis has been a work in progress.

Kelly Booth, my supervisor, has provided me with both intellectual guidance and financial support throughout my graduate studies. He has always been available when needed to review work and provide advice. My supervisory committee consisted of Christine MacKenzie, Dave Forsey, Maria Klawe and Peter Lawrence. They provided moral support and scholarly insight at times of confused indecision. Izak Benbasat and Jack Snoeyink were my university examiners and pointed out where those less familiar with the work would have problems. Robert Jacob my external examiner provided thoughtful and positive comments that eased my worries.

Thanks to Bill Buxton for allowing me to make use of the photos of the Etch-A-Sketch and Skeedoodle toys.

Thanks to the Tuesday night crowd that has gotten me out of the house at least one a week and managed to get my mind off my thesis for a short little while.

Final thanks go to Gayla my partner, who has supported all my endeavours, particularly at those times when I thought I had no endurance left.

JAMES BORITZ

The University of British Columbia

October 1998

To my Dad.

Chapter 1

Introduction

Interacting in a three dimensional environment with a two dimensional input device imposes many limitations on users. Researchers believed that the advent of three dimensional input devices would alleviate many of these problems and allow for a much more natural human-machine dialog. This thesis seeks to establish how a few factors dealing with visual feedback and task structure affect the ability to perform interactive tasks in a three dimensional virtual environment. The factors investigated were *stereoscopic vision*,*motion parallax*, *stimulus arrangement* and *stimulus complexity*. The tasks used a 6DOF input device to control a pointer in a three dimensional virtual environment.

Stereoscopic vision was found to provide a strong benefit to positioning based tasks, but this benefit was weakened in the case of tracing tasks. Motion parallax via head-tracking often had no effect upon task performance and where an effect was found it was often detrimental. The position and orientation of stimuli were found to influence performance across a range of tasks. In the case of stereoscopic vision these results are generally consistent with many earlier studies. The results for motion parallax contradict the beliefs of most practitioners and the findings of several other research studies.

Before launching into the details of the thesis the historical thread that lead to this point is briefly reviewed.

1.1 Historical Perspective

Most interaction with computers today takes place in a two dimensional (2D) environment: the desktop mouse moves on a flat 2D surface; the cursor that it controls moves on a 2D computer display. Aside from changes of the input device (e.g. trackball, tablet, light-pen, etc.), the nature of the human-computer interface is characterized as being primarily two dimensional. For much of the software that is in widespread use today – word-processors, spreadsheets and databases – this 2D environment is sufficient. But three dimensional (3D) computer graphics has brought a new class of software where users try to build models or interact in virtual environments (VEs) that possess three dimensions like the world that surrounds us.

The "modern era" in interactive computer graphics can be said to have begun with Ivan Sutherland's "Sketchpad" system in the 1960s [91]. Sketchpad only operated in two dimensions, but it allowed interactive specification of points and constraints. Following quickly on Sutherland's work, Johnson [59] extended Sketchpad to function in three dimensions. Unfortunately, viewing and interaction were still constrained to the 2D surface of a computer display.

Sutherland sought to eliminate one of these constraints a few years later when he developed a head mounted display [92] to allow stereoscopic viewing of simulated objects. Vickers [97] developed a system for interacting in 3D using the head mounted display and a wand interface that would allow interactive specification of points in 3D space. Clark [30] used a similar system to develop a package for designing surfaces in three dimensions. Clark, and many others who followed, felt that the design of three dimensional surfaces is best done in three dimensions.

Still, until fairly recently, most interaction in computerized 3D virtual environments has required users to work through simpler 2D input devices. Input devices capable of supporting direct 3D interaction such as Vicker's Wand [97] and magnetic trackers from Polhemus [78] were first developed in the late 1970's. These devices were too costly or too fragile to be used outside of a controlled research environment. These restrictions limited the dispersion of 3D input devices into the research and user communities.

With the proliferation of faster computers came the possibility of real-time interaction within a reasonably sophisticated 3D visual environment. This spawned research in many areas, most notably the fields of *virtual reality* (VR) and *virtual environments* (VEs). This in turn has sparked renewed interest in three dimensional input.

Some of the devices that were developed to sense position in 3D also sensed orientation. This allowed researchers to jump from devices with two degrees of freedom to devices with a total of six degrees of freedom (6DOF). Most of the early studies of 3D input sought to determine what sorts of interactive techniques could be developed with devices containing higher degrees of freedom. Zimmerman et al. [113] attached a 6DOF sensor to a Cyberglove™ to allow more naturalistic input. Ware and Osborne [103] developed techniques for manipulating the eye point using a 6DOF device. Relatively little effort was spent evaluating human factors aspects of these devices. However, there were a few studies [93] [81] that demonstrated the potential benefit of 3D devices over 2D devices.

Recently researchers have begun to focus more attention upon the human factors of 3D input, 3D interaction and the 3D environment. However, very little is known about human input capabilities in simulated three dimensional environments. Most researchers now feel that while three dimensional (six degree of freedom) input devices may remove some of the restrictions of two dimensional input, some problems will remain and new problems will be introduced. At a panel on "Three Dimensional Interaction" held at SIGGRAPH '94, Dan Venolia stated that three dimensional interaction is hard because two dimensional interaction is hard. At the same panel Andries Van Dam told the audience that real interaction in three dimensions is hard, so interacting with a computerized three dimensional environment will also be hard.

Six degree of freedom input devices capable of sensing three position and three orientation degrees of freedom are becoming more easily available. Nonetheless, it is still rare to find 6 DOF input devices in use today with the exception of research labs and high-end industrial facilities. The research described here looks forward to the day when these devices are a standard component in interactive 3D computer graphic applications. To aid in the development of new interactive 3D applications it is important to understand the quality of the interaction these devices afford and what effects task structure and visual feedback have upon that interaction quality.

This thesis seeks to determine the effect of *stereoscopic vision*, *motion parallax*, *stimulus arrangement* and *stimulus complexity* upon hand-based interaction that uses a 6DOF input device to control a pointer in a 3D virtual environment. Results of the research showed that:

- stereoscopic vision was found to provide a strong benefit for positioning-based tasks, but this benefit was weakened in the case of tracing tasks

- motion parallax via head-tracking often had no effect upon task performance and where an effect was found it was often detrimental

- the position and orientation of stimuli were found to influence performance across a range of tasks

In the case of stereoscopic vision these results are generally consistent with many earlier studies. The results for motion parallax contradict the beliefs of most practitioners and the findings of several other research studies. The methodical investigation of stimulus arrangement and complexity has not been previously reported in the literature.

1.2 Three Dimensional Interaction Terminology

As already indicated, the study of interaction in three dimensions commonly occurs in the context of six degree of freedom interaction. It is important to distinguish between the dimensionality in which the interaction takes place and the number of degrees of freedom that are available for manipulation. It is also important to distinguish between different degrees of freedom and the techniques through which degrees of freedom are controlled. Thus, before going any further, it is beneficial to establish some basic terminology and common ground for the discussion that follows.

1.2.1 Location

Input that takes place in three real-world dimensions has six degrees of freedom(6DOF). I will use the term *six degree of freedom input* to describe the input that takes place in a three dimensional interactive environment.

When describing the components of a six degree of freedom input space I will use the terms position, orientation, translation and rotation. *Position* refers to the current location of an object in space, typically expressed as coordinates in a Cartesian system. *Orientation* refers to the current revolution of the object, often expressed as Euler rotations around each of the three Cartesian axes. *Translation* is the process of changing the position of an object. *Rotation* is the process of changing the orientation of an object. On occasion the term position is used to refer to both the position and orientation of an object. When the sense of the word position is not clear from the context, the term *general position* is used to indicate both the position and orientation of an object.

1.2.2 Coordinates

In order to specify the position and orientation of an object a coordinate frame is essential. Throughout this work a Cartesian coordinate system will be used. The X-axis is horizontal to the left and right with positive values to the right of the origin. The Y-axis is vertical and positive values are up from the origin. The Z-axis is backward and forward with positive values going toward the user from the origin. Rotation about each of the Cartesian axes is specified in in this right-handed[1] coordinate system. Figure 1.1 illustrates the coordinate system. The terms pitch, yaw and roll or elevation, azimuth and roll are sometimes used to specify rotations about the X, Y and Z axes respectively.

Specification of position in 3D is straightforward because there is no potential for crossover between the axes of the object and those of the world coordinate system. Orientation is more troublesome because it is possible to rotate axes in the object's reference frame onto different axes in the world reference frame. In addition, the order of rotations is important because applying rotations about the three axes in different orders will produce different results. The general convention used throughout the thesis is to apply rotations in X, Y, Z order.

Quaternions are an alternate means of specifying and manipulating rotations, but they are less intuitive. Quaternions were discovered by Hamilton in 1844 [51] and applied to the task of describing rotations by Cayley in 1845 [25]. Quaternions were first brought to the attention of the computer graphics community by Shoemake in 1985 [85], but they were used in the aerospace community as early as 1970 [47]. A detailed description of quater-

[1] To determine the direction of positive rotation in a right- or left-handed system, the thumb of the corresponding hand is pointed in the positive direction along the axis in question and the curl of the fingers indicates the direction of positive rotation. Alternately, when viewed from the positive end of an axis looking toward the origin, a positive rotation proceeds in a counter-clockwise direction.

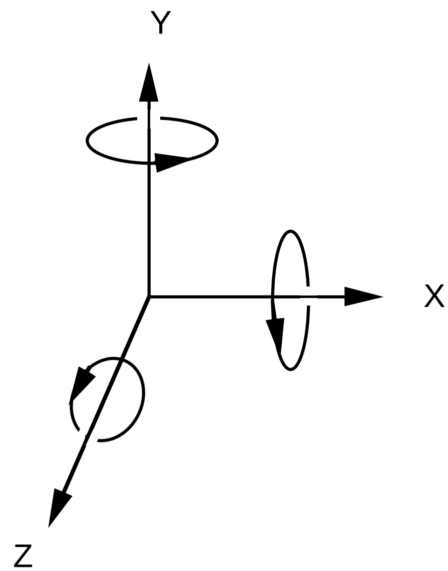


Figure 1.1: The arrangement of the right-handed Cartesian coordinate system that is used throughout the thesis. Arrows indicate the positive direction of each axis or positive rotations as appropriate. As the circles illustrating positive rotations may be ambiguous, the arrowhead should always be considered to be at the front most (closest to the reader) part of the circle.

nions from a mathematical perspective can be found in Altmann [4]. Rotations are specified by unit quaternions because these have a straightforward geometric interpretation. Briefly, a unit quaternion is four-element vector in which the first element represents the cosine of half the angle of rotation and the second, third and fourth elements are the components of a vector in three dimensions about which the rotation takes place.

1.2.3 Input and Output Devices

The terms control and display are commonly used in the study of human motor control and I will use them in their generally accepted form. A *control* is the physical object, such as a mouse, joystick, or knob, that is used to specify input. The term *display* is used to describe the virtual object that moves in response to the control, providing feedback to the user. In the field of computing, the term display is more commonly associated with the computer screen rather than the visual content. For this reason terms such as *tracker* and *cursor* are often used to avoid this ambiguity. In this chapter I will use the term tracker to refer to the object on the display (screen) whose position and orientation is acted upon by the control. In later chapters, when discussing the studies that were carried out, I will use the terms *physical pointer* to refer to the control and *virtual pointer* to refer to the display. These terms are more common in the virtual environment literature and have clear meanings for subjects.

Input devices (controls) come in different forms, each presenting its own set of control dynamics. Zhai and Milgram put forward a rough sensing-based taxonomy for classifying input devices in three dimensional interaction tasks [108]. In their taxonomy input controls can be categorized along three abstract axes. The three axes are:

- Integration: separated vs. integrated
- Mapping: position vs. rate
- Sensing: isotonic vs. isometric

An *integrated control* combines the manipulation of multiple degrees of freedom into a single physical control object. In the case of a *separable control* there may be a distinct physical control object for each degree of freedom.

A *control-display mapping* is the process by which the position[2] of the tracker on the display is altered by operations performed with the control. Poulton [77] describes several different types of control-display mappings that have been developed. A *position control* is one in which changes to the position of the control map directly to changes in the position of the tracker. The precise amount of change may be modulated by a gain factor that

[2] Unless an explicit distinction is made, position in the remainder of this section is general position and thus refers to both position and orientation.

can be linear or non-linear. Poulton also refers to a position control as a *zero order control*. Mathematically, the relationship between the control and the display can be expressed by the equation,

$$D(t) = F(C(t))$$

where $D(t)$ represents the position of the tracker on the display at time t, F is an arbitrary gain function (usually monotonic and often continuous), and $C(t)$ is the position of the control at time t.

With a *rate control*, movement of the control is used to adjust the velocity of the tracker. A linear or non-linear gain factor can also be applied in this setting. The position of the tracker changes according to its current velocity setting. Rate control is also referred to as *first order control*, and is expressed as

$$D'(t) = F(C(t))$$

where F and $C(t)$ are the same as before and $D'(t)$ is the rate of change in the position of the tracker. In this case $D(t)$, the position of the tracker on the display, is given by integrating over time from an initial tracker position

$$D(t) = \int_0^t F(C(\tau))d\tau + D(0)$$

Higher order controls have been developed, but in practice it has been found that their operation becomes increasingly difficult. In an *acceleration control* or *second order control*, the control alters the acceleration of the tracker. The position of the tracker changes according to the velocity, and the velocity changes according to the acceleration, so that a double integral is required. Position control is the most direct control-display mapping. At any given time t, $D(t)$ can be determined from $C(t)$ and F. Velocity control and acceleration control are less direct, requiring some knowledge of the history of the control to determine the display.

In the preceding discussion, $D(t)$ and $C(t)$ are both six-valued functions, one for each degree of freedom, and F is a vector of gain functions, one for each degree of freedom. Although, it is conceivable that the gain function used for position is different from that of orientation, it would be unusual for the three position degrees of freedom or the three orientation degrees of freedom to have different gain functions.

An *isotonic control*[3] is one with little or no resistance to movement. Isotonic controllers with a broad range of movement are commonly used as both rate and position controllers. The Ascension Bird™ [7] and the Polhemus Fastrak™ [76] are examples of commercially available isotonic controllers.

An *isometric control*[4] is one that provides resistance to movement. It is possible to use an isometric controller for both position and rate control, but rate control is much more common. An example of a commercially available isometric controller is the SpaceBall™ .

Isometric position controllers map force to position. They are sometimes difficult to use because they typically have a restricted range of motion and position control requires a constant force to be applied for the tracker to remain stationary anywhere other than the "home" position. Isometric rate controllers map force to velocity. Manipulations are thus often brief and benefit from the availability of a self-centering capability in the device. Zhai and Milgram [108] tested the four combinations of sensing (isometric or isotonic) and mapping (position or rate) and found that isometric position controllers provided the poorest performance. In the same study they found that isometric rate controllers and isotonic position controllers provided the best performance.

This thesis deals with isotonic position controllers because of their superior performance and more natural correspondence with free hand movement in an open environment.

[3]**isotonic** *adj.* …**2** *Physiol.* (of muscle action) taking place with normal contraction. [3]

[4]**isometric** *adj.* …**2** *Physiol.* (of muscle action) developing tension while the muscle is prevented from contracting. [3]

The control-display transformation is simple, thus placing the smallest possible cognitive burden upon the user to determine how the the display will react in response to manipulation of the control. The goal is to leave the user with as much cognitive processing capacity as possible to deal with the other factors when performing the assigned task.

1.2.4 Viewing

The perception of objects in the real world takes advantage of many pieces of visual information to enhance our perception of depth. Wickens, Todd and Seidler [106] provide a broad overview of a large number of depth cues that have been identified in the human visual system. Designers of 3D graphics applications generally make use of some set of these depth cues to enhance depth perception in synthesized displays. Depth cues can be subdivided into two principal categories, monocular and binocular. *Monocular*[5] cues do not rely upon the availabity of two distinct images in the visual system. *Binocular*[6] cues are cues that rely upon the presentation of two disparate images to each eye.

Monocular Depth Cues

There are two principal projection techniques used for viewing three dimensional images, orthographic and perspective. Both of these techniques are commonly used and each has its advantages and disadvantages.

An *orthographic projection* is a "projection of a single view of an object on a drawing surface that is perpendicular to both the view and the lines of projection" [105]. The implication of using an orthographic projection technique is that the size of an object on the display does not change as a function of its distance from the viewer. This allows accurate comparisons of distances and sizes regardless of the position of the object in the environ-

[5] **monocular** *adj.* with or for one eye. [3]
[6] **binocular** *adj.* adapted for or using both eyes. [3]

ment. An orthographic projection does not convey any depth information to the viewer. Orthographic projection is sometimes referred to as a parallel projection.

A *perspective projection* is “the technique or process of representing on a plane or curved surface the spatial relation of objects as they might appear to the eye” [105]. The implication of using a perspective projection is that the size of an object changes based upon its distance from the viewer. As objects get further away from the viewer they become smaller and as objects get closer to the viewer they become larger. It is important to note that in order to judge depth the viewer must have some knowledge about the true size of the object. Perspective projection is sometimes referred to as polar projection, a technique that makes use of linear perspective.

A variety of depth information can be generated from movement of either the viewer or of an object. *Motion perspective* is a depth cue that arises from the change in the appearance of objects as the observer moves. *Motion parallax* is a special case of motion perspective based upon side to side movement of the viewer’s head. Objects at different distances from the observer’s fixation point are perceived to move at different rates. Another motion based depth cue is the kinetic depth effect. The *kinetic depth effect* is perception of 3D structure from the movement (typically rotation) of the object itself. The kinetic depth effect is also referred to as “structure from motion.”

I will use the terms fixed viewpoint and head-tracked viewpoint to distinguish between the two primary motion-based depth cues. A *fixed viewpoint* display is one in which a default viewpoint is selected and there is no change in the rendering of the scene based upon changes in the position of the head of the viewer. A *head-tracked* display is one in which the rendering of the scene is adjusted based upon the estimated position of the viewer’s eyes making motion parallax available. In the case of a monoscopic display the midpoint between the eyes is used to generate the scene in a head-tracked display.

Occlusion is a powerful monocular depth cue. *Occlusion* or *interposition* are the names given to the effect of objects nearer to the viewer blocking the view of objects that are farther away along the line of sight . Early computer graphics systems displayed objects by showing only the wireframes of their polygonal edges. No occlusion was present and this allowed surfaces (shown as outlines) behind other surfaces to be seen, sparking many efforts in the development of hidden-line removal algorithms. The problem of correctly computing occluding objects still exists with the use of solid shaded images. The most common solution employed is the *Z-buffer algorithm*. In a sense the Z-buffer algorithm mimics the physical occlusion process, only allowing the pieces of an object that are closer to the viewpoint to be displayed.

Binocular Depth Cues

The mechanics of the visual system provide a variety of additional depth cues that includes stereopsis, vergence and accommodation. "*Stereopsis*[7] is the perception of depth based on retinal disparity from the presence of horizontally separated eyes" [74]. *Disparity* is the difference between both eyes in the projection of the image of an object onto the retina relative to the projection of the fixation point. *Vergence* is a proprioceptive muscular cue that is derived from the rotation of the eyes that is necessary to fuse the two disparate images [106]. *Accommodation* is another proprioceptive muscular cue that is based upon the adjustment of the lens in order to bring an object into focus upon the retina [106]. *Binocular fusion* is the process within the visual system that combines two disparate retinal images into a single image with depth.

Several elements combine in the perception of images containing binocular disparity. The following brief description of the elements involved in the perception of disparity

[7] **stereopsis** *n.* the perception of depth produced by combining the visual images from both eyes; binocular vision.

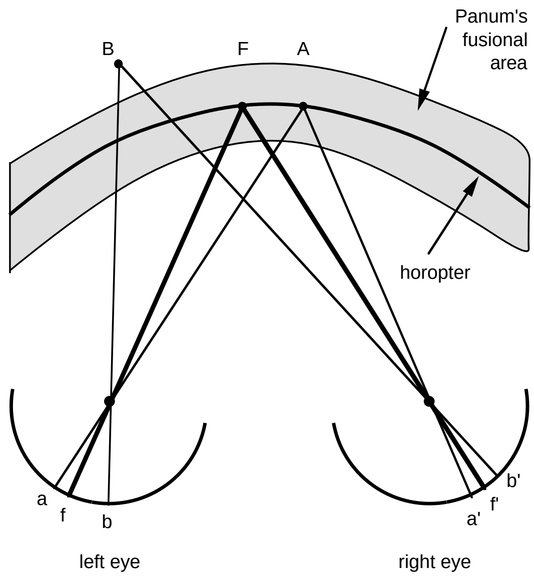


Figure 1.2: Elements of the binocular vision system. The eyes are fixating on the point F with corresponding points on the left and right retina of f and f', respectively. All points lying on the horopter (A) produce the same amount of disparity relative to F. Points in front or behind (B) the horopter produce different amounts of disparity relative to F. Panum's fusional area represents the zone for which the visual system can fuse the disparate retinal images. The point B outside Panum's area would not be successfully fused. (Adapted from Patterson and Martin [74])

information is derived from Patterson and Martin's survey paper [74]. The *fixation point* is the location in space to which the eyes have adjusted so that there is no relative retinal disparity between the two images for objects at that point. The *longitudinal horopter* is a curved line through the fixation point for which objects present equal amounts of disparity in the images for both eyes. The region in front and behind the horopter for which images can be successfully fused is known as *Panum's fusional area*. Figure 1.2 illustrates these elements. Items in front of the horopter have *crossed disparity* and items behind the horopter have *uncrossed disparity*.

When viewing images on a computer display it is assumed that vergence and accommodation act so as to place the fixation point at the surface of the display. Disparity information is generated based on this assumption. Items behind the surface of the display have uncrossed disparity and those in front of the display have crossed disparity.

I will use the terms monoscopic (mono) and stereoscopic (stereo) to distinguish between two types of computer display. A *monoscopic* display is one where no disparity information is present. A *stereoscopic* display is one that contains disparity information. In both cases, other depth cues may or may not be present.

1.3 Research Questions

The goal of a great deal of work in the field of human-computer interaction (HCI) is to understand what obstacles people are faced with when using computers. A variety of approaches are used to study interface techniques in HCI. In one approach, two or more interface techniques are compared to see which one allows better performance. A second approach is to study a particular input technique and determine how a particular set of factors affect performance. By understanding these factors, it is often possible to reduce or eliminate their effect. In this research I have followed the second approach, where a specific input technique is systematically perturbed by a set of factors in order to develop an understanding of the influence of those factors.

This research addresses questions such as:

- How accurately can we perform three or six degree of freedom input tasks in a simulated three dimensional environment?
- How does the type of visual feedback provided by the user interface affect the user's response time or accuracy in performing a task?
- Does operating in certain regions of the input space (e.g. near vs. far) affect performance?

- Are the dimensions such as top-to-bottom, left-to-right and front-to-back equivalent?
- How does incorporating an orientation component in the task influence performance?
- How does the total number of degrees of freedom under control affect performance?

1.4 Effectiveness

Most studies of interaction tasks focus on speed or accuracy as the primary measures of "goodness" for the interface. *Speed* is the catch-all term used to describe any of the potential measurements that involve time. In some cases researchers do measure and analyze the speed of movement, but more commonly they study *task completion time*, the elapsed time from the beginning to the end of the task. Similarly, *accuracy* is the catch-all term for any measure of error in the performed task. The measure of error can be the number of mistakes made, or the distance between a desired location in space and the location actually achieved.

An interaction technique that allows an operation to be performed in less time or with fewer errors than some other technique is considered to be the better of the two. Unfortunately, this can sometimes leave us with an incomplete picture of the difficulties associated with a task, especially if there are speed-accuracy tradeoffs.

I use the term *effectiveness* to refer to a collection of measures that are used to evaluate and compare input techniques. As an aid in understanding effectiveness, I have adapted the following maxims of effective input based upon Grice's [46] maxims of conversation[8].

1. speed – the input action can be performed quickly
2. accuracy – the input action matches what is required
3. conciseness – the input action contains only the information needed

[8]Maxims of conversation have been used because of my belief that the development of the user interface aims to improve the conversation between the human and the computer.

4. felicity – the input action does not put undue physical, mental or emotional strain upon the user

Any given input operation will involve a balance amongst these measures. In some cases, improving one measure will degrade another. When comparing two activities, the more effective one is the one for which more measures are improved than are degraded. A common situation is the tradeoff between speed and accuracy. Without any change to the task a user making rapid responses is likely to have more errors than a user making slow responses. Examining either of these measures alone would indicate a difference. One user is faster while the other is more accurate. In such a situation the input operations could be considered equivalent, with neither being more effective than the other.

Speed and accuracy are obviously important measures for an input operation to be effective. I have chosen to add conciseness and felicity to the set of measures by which an input is judged. These are necessarily less objective than the first two, but potentially just as important in developing an understanding of task difficulty.

1.4.1 Tradeoffs Among Measures

To determine effectiveness a large set of dependent measures are used to develop a broad idea of how the independent factors influence the task under examination. Effectiveness is not judged by some sort of single composite ranking. Rather, it relies on the interface designer to use the gathered measures to develop a deep understanding of the interactive task and how that task meshes with the requirements of the interface. The goal of the interface designer should be to balance these maxims. Sometimes one maxim may be violated in order to improve one of the others.

Conciseness – Pointing

Bolt's [15] "Put-That-There" system is a good example of an input technique that makes use of many more degrees of freedom than are needed. In "Put-That-There" a six degree of freedom input device is used to specify input on a wall, which is just a two dimensional surface. This would violate the maxim of conciseness because the user is providing more information than is needed. A more concise alternative might be to have the user physically touch points on the wall. But this would probably take longer (speed) and it might be difficult for the user to touch certain points (felicity). In this situation a more effective input mechanism was created by making use of more degrees of freedom than are actually needed.

Felicity – Head operated beam targeting

Chung's [28] study of beam targeting is an excellent example of a situation where time to completion for a task is somewhat less than satisfactory. In Chung's study participants were asked to find a beam direction that intersected a tumor while intersecting as little other brain matter as possible. He found that some beam directions were rarely explored because of the human factors involved. The particular beam directions that were avoided involved looking up or down, placing a great deal of strain on the user's neck. This finding is extremely significant because it tells us that people will avoid interaction tasks that are physically stressful or uncomfortable. An interface designer with knowledge of this constraint can adapt the interface to avoid the need for uncomfortable operations. Chung's work indicates that factors other than just speed and accuracy will affect the manner in which an input device is used. An uncomfortable action that is avoided will not appear in measures of speed and accuracy.

1.5 Motivation

I have subdivided the factors that influence effectiveness into four areas: cognitive, perceptual, biomechanical and technological. *Cognitive* elements are factors relating to the ability of a human operator to conceive of and carry out a certain motion or action. *Perceptual* elements describe the visual feedback environment that allows an operator to see his or her actions and make the necessary corrections. *Biomechanical* factors are those that relate to physical ability to make certain motions and the kinesthetic[9] feedback accompanying those motions. *Technological* elements are the capabilities or limitations of the input and output devices available (e.g. mechanical, electromagnetic, computational etc.).

In this section, I will elaborate on the issues that drive the questions of Section 1.3 in relation to the four factors influencing effectiveness stated above. While I have divided the discussion that follows into these four areas, it should be noted that there is a great deal of overlap between the cognitive and perceptual areas.

1.5.1 Cognitive

When developing interactive tasks in a simulated three dimensional setting it would be beneficial to have prior knowledge of the accuracy that can be expected and what factors might affect accuracy. Accuracy, in terms of how close the actual input is to the required input, is of central importance to this work. In most of this work I am more concerned with time as an indicator of task difficulty, not as a factor to be improved. Nonetheless, accuracy is often traded off against performance time and it may not always be possible to separate these two elements.

Orientation is often not an issue in two dimensional systems where there is only one

[9] **kinesthetic** *adj.* a sense of awareness of the position and movement of the voluntary muscles of the body. [3]

rotational degree of freedom. Few tasks in two dimensional systems require orientation to be specified. In contrast, many tasks required within three dimensional systems require orientation to be specified. The problem becomes further encumbered by the presence of three rotational degrees of freedom, rather than just a single rotational degree of freedom when there are only two spatial dimensions. When input of rotational and translational degrees of freedom is combined, the resulting six degree of freedom interaction task takes on a completely different character.

1.5.2 Perceptual

Computer systems that require the specification or manipulation of objects in three dimensions are often extremely challenging to use. Specific applications that make heavy use of three dimensional information include computer aided design and manufacturing (CAD/CAM), architectural design and virtual reality[10]. There are many aspects to this problem including the large amount of data, the time required to perform complex computations on the data and the restriction inherent in current displays for viewing output in two dimensions. Of importance here is the restriction that requires most input and output from three dimensional systems to take place in two dimensions.

Engineers in non-computerized environments have recognized the difficulty of specifying three dimensional objects within the constraints of two dimensional representations. One of the most common techniques for specifying three dimensional input is taken from the world of drafting. In drafting, three dimensional objects are presented via three orthographic projections from the top, front and side. To aid in visualization, a perspective projection of the object is almost always included as a redundant fourth view.

[10] Users of real-time computer graphics workstations are able to manipulate a virtual world in the form of images; …The step into virtual reality is made by tracking the user's head movements and using this to control the perspective view seen by the user [98].

In computer systems using this technique, users are provided with three orthographic views and a perspective view. Most interaction takes place in the orthographic views. Input in the orthographic views is limited to two degrees of freedom along the axes parallel to the plane of the display.

A perspective view may often be easier to understand but it can also be ambiguous or deceptive. Because of this deceptive quality it is essential that users be provided with some mechanism that will allow them to easily change their view to one that is more appropriate for the task at hand. Some computer applications only provide users with a single perspective view. Perspective projection allows us to view three dimensional objects in a somewhat more natural way. Nonetheless, we are still operating with a projection of three dimensional space to two dimensions (unless stereopsis used) and the possibility for confusion still exists.

For example, one common source of confusion that can arise in a perspective display is the determination of relative position between objects. The perspective projection causes a greater reduction in the display size of a distant object than a near object. Thus, what might appear as two equidistant objects of similar size may in fact be two objects of vastly different size and depth. Figure 1.3 shows an example of such a situation.

Input has posed an even bigger problem in three dimensional systems than output. Even though a perspective view may be misleading, it does a reasonably good job of assisting in the visualization of three dimensional objects. The process of computing a viewing projection for a three dimensional object takes three dimensional information and reduces it to two dimensions. The transformation goes from more information to less information. In the case of input with a two degree of freedom device we are faced with a more difficult problem, that of trying to generate at least three input degrees of freedom with only two degrees of freedom available.

Some systems [12][11][1][2] have sought to deal with this problem by allowing a

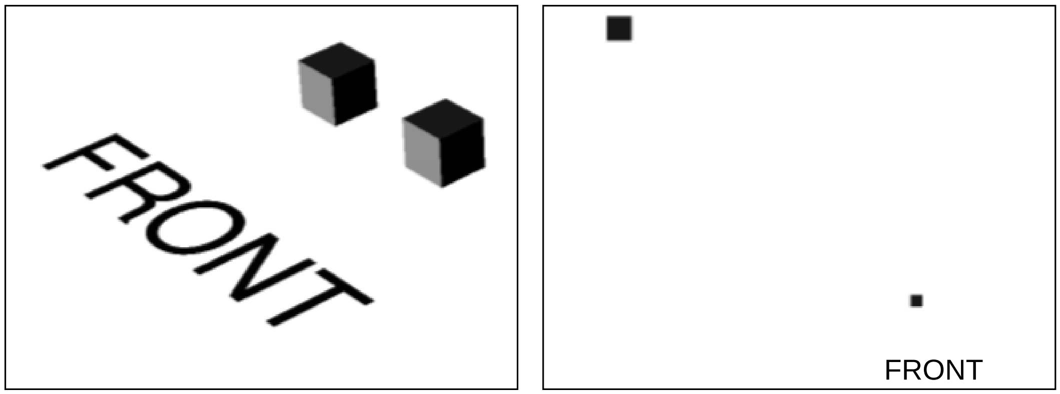


Figure 1.3: In the perspective projection (left) both cubes appear to be of equal size and at an equal distance from the front. In the orthographic view (right) taken from above the objects it becomes apparent that one of the cubes is larger than the other and is further from the front than the other. An orthographic side view (not shown) would show that the larger cube is also lower than the other cube. In the perspective view the cubes appear to be the same size because the larger cube is farther away from the viewer along the viewing axis.

user to use objects or elements of objects to constrain the interaction space. Thus, one might point at the surface of a cube in a perspective view to indicate a two-dimensional sub-space in which interaction should take place. By restricting one degree of freedom in the object space, mapping the two degree of freedom input to the two-dimensional sub-space is straightforward.

1.5.3 Biomechanical

Work by Soechting, Lacquanti and Terzuolo [88] has shown regular distortions of hand movement when subjects are asked to draw figures in three dimensional space. These distortions varied based upon the plane in which the figures were drawn. These distortions are shown to be a result of biomechanical factors and subjects are not aware of them. This suggests that accuracy in three dimensional tasks might be affected by the position and orientation in which the task is carried out.

1.5.4 Technological

Advances in technology have made it possible to provide binocular disparity in computer displays. The simulation of depth using this technology can enhance visualization by providing a more realistic image. The addition of depth to the display can eliminate many of the ambiguities present in simple perspective projections.

Another advance in technology has made it possible to specify the position and orientation of a point in three dimensions using a device that is relatively natural and unencumbering. It has always been possible to control the degrees of freedom involved using multiple controllers. New sensing technology has made it possible to unify control into a single device that can sense all six degrees of freedom simultaneously. The common 6DOF isotonic position controllers such as the Polhemus Fastrak [76] and the Ascension Bird [7], integrate all the degrees of freedom into a convenient easy-to-use control device.

The increase in the general level of computing power has made three dimensional applications more common and accessible. It has also increased the richness of the graphics that can now be displayed. Intriguing new applications such as virtual reality are enabled through a fusion of computing, input and output technology.

1.6 Research Goals

The focus of this research is to better understand how people perform interactive tasks in a computer simulated three dimensional computer environment using an isotonic 6 DOF position control input device. The research explores how different target positions in conjunction with different display modes (monoscopic vs. stereoscopic) and different head-tracking modes (fixed viewpoint and head-tracked viewpoint) affect a user's ability to perform a variety of simple tasks. The goal of this work is to establish a framework within which designers of interactive 3D applications can determine what kind of user performance is possible for a given interactive task. Alternately, knowledge of what factors affect performance might allow designers to determine how to assemble a task in order to achieve a desired performance level. By focusing on relatively simple tasks that are components of more sophisticated interactive tasks, I hope to apply these findings to a broad range of future applications.

Early studies have been performed in this area, but the results are somewhat inconsistent. Additional studies are required to enable the development of a comprehensive model of task performance in a 3D environment.

The simplest task imaginable is the specification or location of points in 3D. In a *point location* task a user must match the tip of a pointer to a fixed point in space as accurately as possible using the 6 DOF device. *Volume location* is a closely related task in which a user must move the pointer into a volume in space. *Docking* is a more sophisticated task that combines point or volume matching with orientation matching. It is probably the simplest composite 6 DOF input task. *Tracking* is a slightly more taxing form of this task where subjects must continually try to match the position and orientation of the pointer to the position and orientation of a moving object.

Path tracing is a potentially more complex task than point location that requires a user to move a pointer along a specified path in space. The path can be a straight line, varying along only one dimension, a planar curve, varying in two dimensions, or a space filling curve, varying in all three dimensions. Path tracing may be more difficult because it requires a user to constrain his or her movement during the entire task rather than just at the endpoints. On the other hand, a user may hope that movement naturally mimics the path (e.g. straight line) and that no additional effort is required.

I have conducted a series of experiments to investigate the nature of interaction in 3D. The first experiment is a point location study. The structure of this experiment and its results are described in Chapter 4. Chapter 5 describes a study of a docking task. Chapter 6 and Chapter 7 describe two path tracing experiments. The first of these path tracing experiments used straight line paths and the second used planar curves, both in three dimensional space.

When conducting many computerized tasks it is possible for the system to provide enhanced feedback to the user indicating when the user is operating within some desired performance bounds. For example, consider a 2D task where the user is to move the mouse controlled tracker (arrow cursor) into a square region. There is undoubtedly visual information that shows when the cursor is within the region. Nonetheless it is possible to provide enhanced feedback in the form of a colour change of the region, or an auditory beep. As these additional forms of feedback take advantage of alternate human perceptual channels, they were generally not used in the studies that were conducted.

Before getting into the details of the experiments, Chapter 2 outlines the literature that motivated and informed the work I have carried out. Chapter 3 describes the computer system and the common environment in which all of the experiments were conducted, as well as other elements common to all of the experiments.

After describing the outcomes of each of the experiments, Chapter 8 ties together all of the results and lays the groundwork for future studies which are discussed in Chapter 9.

Chapter 2

Prior Work

In this chapter I expand upon the earlier work in the three areas of Human-Computer Interaction and Graphics that form the basis of this thesis. These areas are

- input device taxonomies,
- three dimensional operations using two dimensional input and
- true three dimensional interaction.

The study of input devices and the taxonomies that have been developed to describe them are important because research in this area has taken into account the *pragmatics* of matching input devices with input techniques and human ability. Early device taxonomies tended to focus upon the number of simultaneous degrees of freedom under control (dimensional integrability), and whether an object was user defined or application defined. The interchangeability of devices at the program level was of more concern than the suitability of a device for a given task. Pragmatics seeks to go beyond programatic interchangeability to understanding how the different physical quantities being sensed, and the ergonomics of the sensing device affect the use of the device.

The study of three dimensional input without three dimensional input devices is relevant for three reasons. One is to identify the elements that are necessary for three dimensional input. Because of the restricted input environment, research was highly focused upon understanding the essential elements of three dimensional interaction. Second, several different interaction techniques were developed to allow users to interact with a three dimensional environment. Lastly, this body of work identifies many of the basic three dimensional interaction tasks that are commonly required.

Any description of input and interaction in three dimensions is confounded by the fact that more than three degrees of freedom are present. In fact even in two dimensions, more than two degrees of freedom are available. In many two dimensional input tasks we are only concerned with position along the two Cartesian axes. Even though orientation as a third degree of freedom is present in two dimensions, it is often overlooked or dealt with as a special case. Orientation is hardly ever specified in conjunction with position in two dimensional tasks.

The transition to three dimensions results in one more translational degree of freedom, and two more rotational degrees of freedom. So, in three dimensions there are a total of six degrees of freedom and fully half of them are devoted to rotation. The nature of most tasks performed in three dimensions makes the rotational degrees of freedom more salient than in two dimensions. As such, much of the research that studies interaction in three dimensions concerns itself with devices that sense six degrees of freedom simultaneously.

The study of six degree of freedom (three dimensional) input is the central concern of this work. Prior work in this area is sparse. Some of the research in this area has attempted to determine the types of tasks that must be performed and the different conceptual models for performing these tasks. Other work has concerned itself with performance variations between different types of six degree of freedom input devices. More recent work has begun to investigate how many degrees of freedom can be operated upon simultaneously. These studies also seek to understand how the task or its structure affect the number of degrees of freedom that can be controlled simultaneously.

2.1 Input Device Taxonomies

One of the most important efforts in the study of input devices and interaction techniques has been the development of input device taxonomies. These taxonomies are aimed at understanding the similarities and differences between input devices. Through a use of these taxonomies it is possible to determine how one device may be substituted for another, or how software can be written to make the best use of the devices available.

The development of taxonomies can be separated into two approaches to the problem that I have named the *software engineering* approach and the *user engineering* approach. The software engineering approach was developed first and, as its name suggests, it seeks to foster device interchangeability via a software encapsulation of the input provided by a device. Advocates of this approach classify devices based on logical properties such as the number of degrees of freedom a device can sense or the tasks that can be performed through its use. The user engineering approach followed later, and argued that the interchangeability promoted by the software engineering approach ignored the physical characteristics of the device and of humans. Taxonomies developed by this group distinguish devices based on the tasks[1] that a device is best suited for or the physical properties that the device senses.

The software engineering group hoped to simplify the development of interactive

[1] Because both groups use tasks to develop their taxonomies, the difference between them might appear to be minor or even non-existent. The distinction is that one group focuses on how the use of logical tasks can be used to facilitate a very high degree of interchangeability amongst devices. The other group focuses upon how suitable or unsuitable a device is for an actual task, often reducing the level of interchangeability because very low-level characteristics of the task and the device are taken into account.

software by developing a set of virtual devices. Virtual devices are logical devices to which physical devices can be mapped. A small comprehensive set of virtual devices would allow application software developers to support many input devices while needing to code for only a few. The virtual device approach would also make software more portable because developers would code for generic rather than specific devices.

The efforts of the software engineering group were well described by Foley and Wallace in 1974 [40], who reduced the large number of then-current input devices to a set of only four virtual devices: button, pick, locator, and valuator. A *pick* is used to designate user defined objects within the software system. A *button* is used to designate system defined objects. A *locator* is used to determine position and/or orientation. A *valuator* is used to determine a single value within some number space. In addition to establishing the notion of virtual devices, this extremely important paper identified the importance of decomposing interactive tasks into lexical, syntactic and semantic levels and pointed out the pitfalls of incorrectly processing input at one level that was intended for a different level.

The virtual device approach was modified for use in the ACM SIGGRAPH Graphics Standards Planning Committee system (Core)[48], [49] and later used as the basis for graphical input in the Graphical Kernel System (GKS) [50]. The GKS standard proposes six virtual input devices. *Locator*, *valuator*, *pick* and *choice* (button) are the same as in Foley and Wallace [40]. Two new virtual devices, stroke and string were added. A *stroke* is a sequence of points; a *string* is a sequence of characters. These were needed only for pragmatic reasons and provide an early sign that a pure software engineering based approach was not sufficient.

Ten years later Foley, Wallace and Chan [41] revisited the interaction standards problem and decided to take a user task centered approach instead of their earlier device centered approach. They proposed the following six interaction tasks:

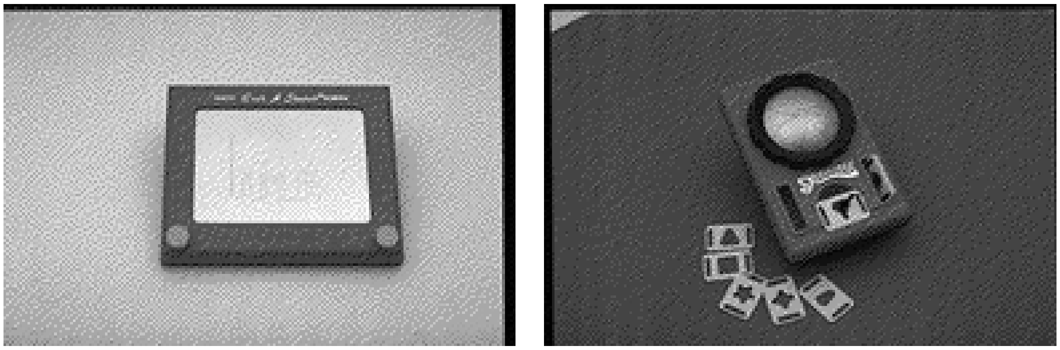
- select
- position
- orient
- path
- quantify
- text

Generally, each of these interaction tasks is closely related to the virtual devices of earlier schemes, with the exception of orient: select unifies choice (button) and pick from the earlier standards; position is roughly equivalent to locator; path to stroke; quantify to valuator and text to string. This approach recognizes that a mapping can be established between almost any input device and these logical interaction tasks. For example, an input string can be used to specify an orientation, or a series of button clicks on a virtual keyboard can be used to supply an input string. The key difference between this work and Foley and Wallace's earlier work is an expansion of the interchangeability between devices.

The ability to handle exotic input devices is an extremely interesting result of not dealing directly with devices. As long as the function of the input device can be mapped into one of the above tasks it will fit into the taxonomy. Thus, speech input as a means of selection can be classified by this scheme [41], whereas none of the approaches developed by the user engineering group can do so.

The user engineering group sought to match input devices to tasks based on the *affordances*[2] and *restrictions* offered by the input device. Buxton [23] provides an excellent

[2] According to Norman[70], the term affordance refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used.



(a) (b)

Figure 2.1: Photographs of the Etch-A-Sketch (a) and Skeedoodle (b) children's toys. The templates used to constrain the movement of the Skeedoodle joystick are visible in (b).

example of input device affordances when he compares two children's drawing toys, Etch-A-Sketch™ and Skeedoodle.™ Etch-A-Sketch uses two knobs to move the drawing tip, one for the X-axis and one for the Y-axis. Skeedoodle uses a joystick to move the drawing tip. Etch-A-Sketch affords the drawing of straight lines parallel to the X and Y axes, but lines of other slopes and curves require a careful and difficult coordination of the input controls. Skeedoodle, on the other hand, affords much greater drawing freedom for curves and lines of any slope, but perfectly straight lines are difficult to produce. In fact the Skeedoodle toy included snap in templates for the control joystick to constrain movement for drawing certain shapes. Figure 2.1 shows the Etch-A-Sketch and Skeedoodle toys.

The goal of the user engineering group is to simplify the process of matching tasks to devices. Thus, the user engineering group tends to classify devices based upon the physical property being sensed and how the device is used. Their belief is that by understanding the nature of the desired input for the software system, one can select the input device that most closely affords the sensing of this input. The intent is to indicate that even though a virtual keyboard can be used to enter a string, a real keyboard is better suited to the task.

The earliest taxonomy based on sensing phenomena was developed by Buxton [22]. He argues that it is not sufficient to classify devices based upon their *lexical* properties as proposed be the software engineering group. He states that *pragmatics* must also be considered. His categorization was based upon the number of dimensions of control and the physical property being sensed (position, motion or pressure). A sub-classification is made between devices with the same dimension of control based upon *type* of control motion. In this classification, a tablet and mouse are more closely related than a tablet and a joystick even though they all provide two dimensional locator input according to their lexical classification.

Tablets, sensing translational motion, operate over a large physical area, thereby requiring large hand and arm movements. Joysticks, sensing rotary motion, operate in a much smaller physical area with correspondingly smaller hand and arm movements. Buxton chooses to differentiate with respect to sensing properties, rather than trying to make difficult and possibly arbitrary distinctions over the amount of space used. Nonetheless, he indicates that the amount of space that a device uses is an important factor. This gives rise to questions about how the available input space of a device is used. Are there certain methods of using the input space that are easier and thus consistently favoured by users? Is movement in some directions preferred over other directions?

Mackinlay, Card and Robertson [64] assembled a more comprehensive "physically based" sensing taxonomy. They extended Buxton's notion of property being sensed, in two ways. First, they distinguish between absolute and relative quantities and, second, they separate translational and rotational movement. Instead of distinguishing based upon number of dimensions, Mackinlay et al., distinguish between the axes in which the device operates. The dimensionality of a device is indicated by composing the axes in which it operates. A

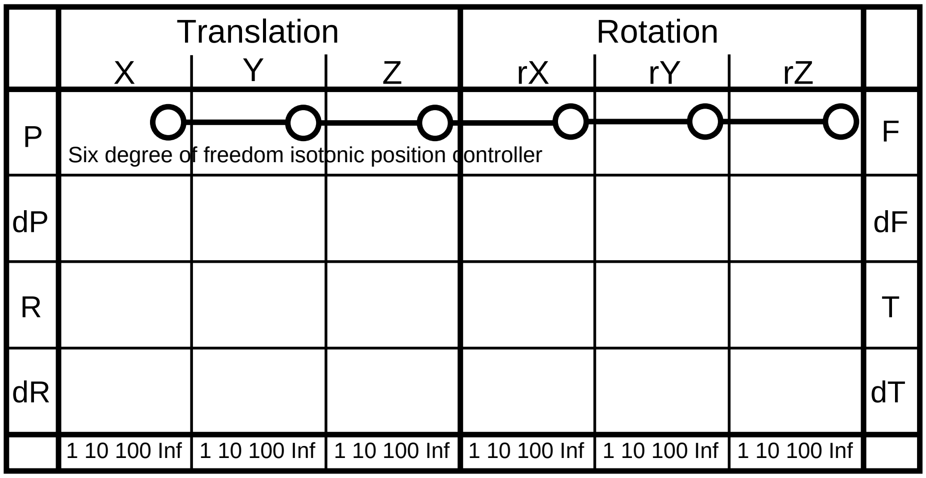


Figure 2.2: Table adapted from Mackinlay et al. [64] for classifying devices in their taxonomy, illustrating how a six degree of freedom such as a Polhemus Isotrak is described. In the table, P indicates position, F indicates force, R indicates rotation and T indicates torque. The labels across the top indicate either translation or rotation with respect to the indicated Cartesian axis. The labels along the bottom indicate the number of values that can be sensed by the device in each degree of freedom.

final distinction is made between the number of values that a device can sense in a particular axis. A copy of the table they use for making classifications can be found in Figure 2.2.

A more encompassing set of distinctions, matching task semantics to devices is still possible. According to Jacob et al. [57],

> Bleser [14], developed a device taxonomy and input model that explicitly incorporated the physical attributes of input devices, including the notion of the physical separability of input degrees of freedom, and knowledge about task requirements.

A recent study by Jacob et al. [57] has brought to the forefront the importance of matching device characteristics to the task. The study also attempts to utilize knowledge of human perception to further our understanding of input tasks. In their study, subjects were required to perform two tasks using either a mouse or a Polhemus tracker. Each task required subjects to move a cursor to a target position on the screen. In one task subjects needed to match the size of the cursor to the size of the target. In the other task subjects needed to match the gray level of the cursor to the gray level of the target. Based on prior studies in cognitive psychology by Garner [43] and Garner and Felfoldy [44], the position and size adjustment task was considered to be *integral*[3], whereas the position and gray level adjustment task was considered to be *separable*. When using a mouse for the two tasks, subjects were allowed to control position or the task element (size or gray level) but not both at the same time. When using the Polhemus, subjects were allowed to control position and the task element simultaneously. The results of their study showed that subjects performed the separable task faster with the separable device (mouse) and integral tasks faster with the integral device (Polhemus). It is especially interesting that the mouse yielded faster performance than the Polhemus on the separable task, since a naive analysis might conclude that more degrees of freedom are always better.

In their analysis of movement trajectories, Jacob et al. [57] describe a technique for determining whether the trajectory is integral or separable[4]. In an integral trajectory, movement cuts across all dimensions simultaneously. In a separable trajectory, movement is broken up into sub-movements each of which takes place in fewer than all of the dimensions.

[3]Garner and Felfoldy [44] categorized stimulus dimensions as integral and nonintegral (separable). They technically defined integral dimensions as those which produced a Euclidean metric and nonintegral dimensions as those which produced a city-block metric. A somewhat less formal definition of these terms is that integral dimensions are those which can be considered together, while separable dimensions are those which cannot. When separable dimensions are combined into a task, selective attention must be paid to each of the dimensions and this produces the city-block metric.

[4]Jacob et al. used the terms Euclidean pattern and city-block pattern, respectively, to describe the two different trajectory types

2.2 Three Dimensions from Two Dimensions

Johnson [59] was one of the first people to have to deal with the difficulties of interacting in a three dimensional environment. Johnson developed his "Sketchpad III" system on the heels of Sutherland's [91] original two dimensional Sketchpad system. In Sketchpad III users could interactively specify three dimensional objects while viewing them on the display. Three dimensional input for Sketchpad III was specified using multiple controllers, because integrated controls with three or more degrees of freedom did not exist.

Three dimensional computer graphics has been constrained for many years by limitations in both the input and output media available. Many of the most common input devices available (e.g. light pen, tablet and mouse) are able to sense position in only two dimensions or with only two degrees of freedom[5]. Almost completely overlooked is the fact that even in two dimensions, an input control can possess three degrees of freedom if rotation is taken into account. Computer displays have also been limited to output in two dimensions. Most of the input and output techniques developed have been designed to make the most effective use of what is available.

The various techniques for 3D output are fairly well known and will not be discussed further. The techniques developed to overcome the limitations of a two dimensional input space can be grouped into four main categories (adapted from Banks [8]):

- cross product
- partitioning the input space
- discarding range degrees of freedom
- overloading the input mapping

[5]Note that some mice advertised 2D + rotation sensing and some tablets (GTCO) offered two degrees of tilt and one of pressure for a total of five degrees of freedom

2.2.1 Cross Product of the Input Space

The simplest way to control more degrees of freedom is to use more locators. The additional locators can be either physical or logical. For example, two mice can be used to control four degrees of freedom. Alternately, a modifier key on the mouse or keyboard can be used to select between one of two logical mice. In general, k locators, each with n degrees of freedom, will allow $k \times n$ degrees of freedom to be controlled.

Chen [26] demonstrates an example of this approach in one condition where three virtual locators, each with one degree of freedom (sliders), are used to control the three orientation degrees of freedom.

2.2.2 Partitioning the Input Space

Partitioning the input space involves subdividing the input space either statically or dynamically. Based upon where the input is located it can be mapped onto different dimensions or degrees of freedom. Nielson [69] describes a technique for dynamically partitioning the input space based upon the location and orientation of a three dimensional cursor called a *jack*. Selecting an axis of the jack allows translational movement of the object or cursor to be constrained to the selected axis.

Chen [27][26] describes an Overlapping Sliders controller (see Figure 2.3). In this controller a 3x3 grid is superimposed upon an object to be rotated. Vertical movement along the three horizontally centered squares causes rotation about the X-axis while horizontal movement along the three vertically centered squares results in rotation about the Y-axis. Circular movement around the eight outer squares results in rotation about the Z-axis.

Chen's overlapping sliders provide good compatibility between the input and the operation to be performed. The implementation scheme bears a striking resemblance to Ledeen's [68] (Appendix VIII) character recognizer for the upper case English alphabet. In

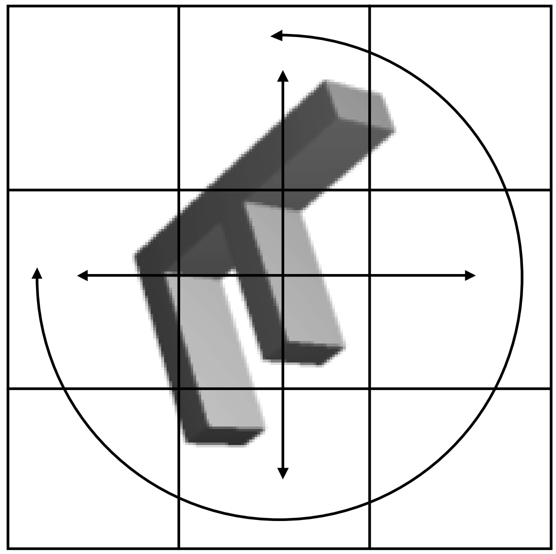


Figure 2.3: An Overlapping Sliders controller showing the movements that would cause a rotation to take place.

Ledeen's recognizer a 3x3 grid is overlayed on a character stroke and the sequence of cells visited by the stroke determines the character. This suggests that an even richer set of control gestures could be built into a similar controller allowing control of more than three degrees of freedom.

Conner et al. [31] and Zeleznik et al. [107] describe three dimensional widgets. The three dimensional widget technique makes use of several different control handles that are attached to the item being manipulated. Operation of each control handle causes one or two dimensions of input to be mapped to control of one or two degrees of freedom of an object. Because handles appear explicitly on or around the object being manipulated, the presence of many handles at different spatial locations can be used to manipulate many degrees of freedom.

It is easy to confuse the partitioning of the input space with the cross product of the input space. Both of these techniques are aimed at altering the association between control degrees of freedom and display degrees of freedom. By partitioning the input space, association changes can be adapted to the contents of the display and appear more dynamic. Using a simple cross product of the input space does not provide for the same level of dynamic coupling between control and display.

2.2.3 Discarding Range Degrees of Freedom

Discarding range degrees of freedom typically reduces the number of degrees of freedom in the range of a control to match the number of degrees of freedom available in the domain of the control. Essentially the result is that control movement in an n-dimensional manifold gets mapped to some other n-dimensional manifold.

Neilson [69] and Bier[10] make use of elements within the computer model to define constraints such as axis of rotation or plane of motion that then only allow one and two degrees of freedom, respectively, to be controlled.

Chen [26] developed the *virtual sphere* and Shoemake [86] developed the *arcball*. Both of these techniques project the mouse coordinates onto a unit sphere and use the displacement vector to determine an amount of rotation. A mapping between the positions on the sphere and the required rotation in three dimensions can be established because only two degrees of freedom are needed to describe the position of a point on the unit sphere. The discarded degree of freedom in this case is the rotation about the vector from the origin to the surface of the unit sphere.

Hanrahan and Haeberli [52] dynamically map two degrees of freedom to arbitrary surfaces. They use a modified z-buffer scheme to determine the visible object under the cursor. It is then possible to map the input degrees of freedom to a position on the two dimensional surface in question.

2.2.4 Overloading the Input Mapping

Overloading the input mapping involves carefully examining the two dimensional input and using it to selectively generate three dimensions of control. Evans et al. [35] developed a technique where circular, or stirring motions could be used to specify action in a third dimension independently of the other two dimensions of the control. This differs from Chen's approach in that the location of the circular motion is not restricted to any particular location on the display or size of motion.

2.3 Three Dimensional Input

The development of more sophisticated sensing technologies has made it possible to determine the position of a control in three dimensions. The movement into three dimensions also made the orientation of the control, that was often overlooked in two dimensions, more important. Most three dimensional input devices are also referred to as six degree of freedom devices, three positional degrees of freedom and three rotational degrees of freedom. Research into the use of six degree of freedom input has taken several different approaches as scientists and application developers have tried to understand this relatively new technology.

The researchers exploring six degree of freedom input fall into two somewhat arbitrary groups. One group is investigating how tasks that are reasonably well understood in two dimensions can be extended into three dimensions. The other group is looking at how six degree of freedom input can be used in new and exotic ways. For a lack of any other better names, I have chosen to call these "traditional input tasks" and "virtual reality tasks." These names are intended to reflect the emphasis that I place on the work, they may differ from the emphases of others.

2.3.1 Traditional Input Tasks

Traditional three dimensional input tasks are almost all directly derived from two dimensional input tasks. Zhai [111] provides the following list of two dimensional tasks that are in the process of being extended to three dimensions:

- inking
- target acquisition
- pursuit tracking
- sweeping out regions
- orientation
- navigation
- docking

Inking is the process of laying down some sort of track in space and possibly in time. It can also be used to describe freehand gestures in space. *Target acquisition* is the process of moving a cursor from some starting point to a defined end point. Two dimensional versions of this task tend not to include an orientation component, whereas three dimensional versions often include orientation. *Pursuit tracking* is the process of attempting to keep a cursor in contact with an object as it moves along a path. Once again, in two dimensions, orientation is often disregarded. In three dimensions, orientation is frequently included. *Sweeping out regions* is the process of defining an area in two dimensions. In three dimensions it is the process of defining a volume. *Orientation* is the process of determining rotation in two dimensions and has only one degree of freedom. In three dimensions orientation is more general and has three degrees of freedom. *Navigation* is the process of controlling the movement of an object or the viewpoint. *Docking* is the process of fitting one object into another, often by matching position and orientation.

It is interesting to note that rotational control is disregarded in most two dimensional tasks. In fact there are specific two dimensional tasks dealing with rotational information, namely orientation and docking. When extending the above tasks to three dimensions, orientation becomes much more critical. It is important to distinguish between a target acquisition task involving only position and one involving both position and orientation. When target acquisition includes an orientation component it becomes very similar to the docking task.

Another interesting element of this task list is that it differentiates between apparently identical tasks based upon their semantics. Inking and sweeping out regions are two tasks that are essentially composed of a sequence of points (the GKS stroke). The primary difference is that inking implies a stoke used to convey content within an application, while sweeping a region is coupled to selection.

2.3.2 Virtual Reality Tasks

A system where the software application can sense the position and orientation of a user's head in order to determine the current viewpoint is a primary requirement for virtual reality (VR) system[6]. The use of *head-tracking* makes it possible to determine the position and orientation of the user's head and, by extrapolation, the user's eyes. The displayed scene can then be recomputed based on changes in the viewpoint. The addition of six degree of freedom sensing for hand input frees the user from the restrictions of a two dimensional input

[6] The issue of what makes a system VR is a somewhat contentious one. Some argue that immersion is the critical factor and thus VR requires the user to wear a HMD or be surrounded by the 3D virtual environment as in Cruz-Neira's CAVE [32]. Others argue that the term VR can be applied to any system where the user interacts with a simulated environment rather than a real one. At the present time it seems that any graphical system can use the term VR if its designer so chooses.

device that may be unnatural in the virtual environment.

Virtual reality approaches can be grouped into three broad categories based upon the type of immersion into a real or simulated environment. These three categories, and the level of immersion they typically involve are, immersive virtual reality (high), fish tank virtual reality (medium) and augmented reality (low).

Immersive virtual reality [21][38][13] typically refers to the situation which combines head-tracking with a head mounted display (HMD) containing a separate display dedicated to each eye. The CAVE system [32] is another type of immersive virtual reality system that does not require a HMD because the images of the virtual environment are displayed on the walls surrounding the user. Immersive VR systems are generally quite costly because of the specialized hardware needed to operate the HMD and the tracking devices.

Fish Tank virtual reality [6][33] [34] or a *desktop virtual environment* refers to the setting in which head-tracking is used in conjunction with a single standard display. Fish tank VR is somewhat less costly to implement, needing only a standard display, however, there is a potential for interference between the elements of the real world and the elements of the virtual world.

Augmented reality refers to a situation in which computer generated imagery is combined with an image of the real world. There are several mechanisms through which this combination can be performed. A fairly simple approach is to project computer images on real world objects in order to allow flexible and dynamic reconfiguration of those objects. A more sophisticated approach used by Feiner et al. [37] employs a see-through display similar to a pair of eyeglasses. The computer generated images are projected onto the inner surface of the glasses and the user sees the computer generated images overlayed on the real world. The approach used by Schmandt [82] and Wang et al. [99] is to combine the views of real and computer generated objects by having users look though a half-silvered mirror; objects can be real, virtual, or a combination of the two. A final approach presents users with a display where real and computer generated images are combined using video mixing prior to their presentation. This approach is often used to overcome the luminance differences between real and computer generated images.

Subdividing the aspects of a virtual reality system into input and output is relatively straightforward. Input involves sensing the position of the user's head and hands. Output involves providing visual feedback to the user based on the current viewpoint. Input tasks are the primary concern in this work. Nonetheless, it is important to have some understanding of how human vision works because the perception of the three dimensional environment is based upon the visual simulation. There are certainly applications that make use of more inputs or where haptic and auditory output are provided in addition to the visual display, but those applications are beyond the scope of this work.

Input

In many virtual reality applications, the user operates in a fairly large environment. These applications require the user to move through the environment to arrive at objects of interest and then interact with those objects. Chung [28] makes the distinction between two modes of operation in three dimensional interaction tasks related to one's position, steering and navigation. *Steering* is used to describe the process of changing position or orientation. *Navigation* is the process of understanding one's current position and orientation relative to other objects in the scene.

Ware [103] describes three operational metaphors that can be used to understand steering tasks when using the hand to direct an isotonic six degree of freedom controller. These metaphors are eye-in-hand, world-in-hand and flight.

The *eye-in-hand metaphor* describes an interaction style where an imaginary eyeball is held in the hand of the user. In this environment the scene remains fixed and the viewpoint is under the user's control. The view of the scene changes as the virtual eyeball is translated and oriented in space relative to the objects being viewed.

With the *world-in-hand* metaphor the user holds the object or scene in his hand. In this setting the viewpoint is fixed and the scene is under the user's control. The displayed view of the scene changes as the virtual scene in the user's hand is manipulated.

The *flight metaphor* decouples location and orientation control in the interface. The eye-in-hand and world-in-hand metaphors provide reasonable control over orientation, but location control is weak when dealing with a large scene. In the flight metaphor location is operated as a rate control and orientation continues to operate as a position control. The position of the viewpoint moves according to its current velocity. Changes in the position of the control are mapped to velocity changes of the viewpoint. Orientation of the viewpoint is mapped directly from the control.

Virtual reality systems make a fundamental distinction between local and at-a-distance interaction. *Local* interaction takes place anywhere within the volume of space that a user can reach with their hand. *At-a-distance* interaction encompasses any interaction that requires a user to indicate or operate upon an object outside their immediate reach.

Selection, as a fundamental operation in almost every user interface, is an excellent example of the difference between these two forms of interaction. Local selection is considered a matter of simply moving the tracker into the volume of space occupied by the object to be selected. Selection at-a-distance has been a much more difficult task.

Bolt [15] used the location of the user's finger coupled with its orientation to cast a ray from the finger to a wall in order to select objects and specify positions. While not even close to virtual reality, the selection technique used in Bolt's system is a direct ancestor of what has been dubbed *laser selection* in virtual reality applications such as Liang and

Green's JDCAD [61]. While Bolt had the advantage of an opaque wall just a few metres from the user, many virtual reality applications attempt to allow interaction over spaces that are conceptually much larger. Small angular changes of the control lead to large changes in distance at the end of the pointer. This lead Liang and Green to develop *cone selection* as an adjunct to the laser selection technique. The need for even more control over selection lead Forsberg et al. [42] to develop what they call *aperture based* selection where a combination of gaze direction and hand position are used to control the direction and size respectively, of a selection cone.

Mine [67] provides a broad discussion of many different interaction techniques that have been developed for use within immersive environments. Mine identifies five fundamental forms of interaction as movement, selection, manipulation, scaling and menu interaction. While the specification of scaling and menu interaction as fundamental techniques appears somewhat odd, it highlights the increasingly dominant role that semantics has taken in classifying activities in virtual reality systems. Earlier approaches might have viewed scaling and menu selection as simply a sequence of movement and selection operations coupled to specific objects in the environment or locations on the display.

Output: Stereopsis and Depth

The presentation of a stereo image for computer graphics applications has taken many different approaches. Johnson [58] provides a brief overview of the techniques that had been developed by the late 1980s. While research into new display technologies is ongoing, Johnson's article is still representative of the most widely used approaches. The primary subdivision of output techniques is into time-multiplexed and time-parallel systems.

In the *time-multiplexed* or *field sequential* [63] approach, the presentation system alternates rapidly between images for each eye and a selection mechanism of some kind is used to insure that each eye sees only the image intended for it. Thus, when the image for the right eye is displayed the selection mechanism allows the image to pass through to the right eye while blocking the image from the left eye. When the next field is displayed, the selection mechanism allows the new image to pass to the previously blocked left eye, while blocking the right eye that received the previous image. The human visual system then fuses these two images into a single image that appears to have depth.

One of the most common field-sequential systems employs a high-frequency computer monitor capable of displaying fields at 120 Hz (60 Hz to each eye) and a pair of liquid crystal (LCD) shuttered glasses worn by the viewer to select which eye sees the currently displayed image. An infrared signal is used to synchronize the glasses and the monitor.

The *time-parallel* approach harkens back to the stereoscopes of the 19th century where two slightly offset images were presented to each eye by a special viewing device. A contemporary analog to the stereoscope is the ViewMaster™ children's toy. In more modern settings, two computer displays are used to present a slightly different image to each eye. In some systems full size monitors are used forcing the user to sit in a specific position, while in more recent systems small light-weight displays are mounted in a headset that is worn by the user. Sutherland [92] was one of the first to develop a stereoscopic computerized display system in the late 1960s.

Field-sequential displays tend to be more common than time-parallel display systems because they are relatively inexpensive, only requiring the addition of a pair of LCD glasses. Time-parallel displays are more expensive because two display systems are needed in addition to an apparatus to properly present the images to the eyes. Field-sequential displays are common in desktop virtual environment applications because they make use of standard display technology. Time-parallel systems are used in immersive settings where it is impractical to use a single display, or in settings where potential side effects of field-sequential displays are unacceptable.

Wickens, Todd and Seedier [106] conducted a review of the literature on depth perception and found that motion, binocular disparity (stereopsis) and occlusion are the most powerful depth cues in displays. Braunstein et al. [20] found that occlusion dominates motion in the perception of depth.

A study by Tittle and Braunstein [95] indicates that a cooperative relationship exists between binocular disparity and stereo from motion. They also showed that the perception of depth was greater for rotational motion than for translational motion.

2.4 Difficulty of Three Dimensional Interaction

When using an isotonic six degree of freedom input device in three dimensions, several issues that could reasonably be disregarded in two dimensional interaction become much more important. Conner [31] nicely summarizes this point with the following statement.

> The interface can easily obscure itself, and 3D interaction tasks can require great agility and manual dexterity. Indeed, physical human factors are a central part of 3D interface design, whereas 2D interface designers can assume that hardware designers have handled the ergonomics of device interaction.

2.4.1 Physical Constraints

Just because it is physically possible to perform some action, does not necessarily mean that the action will be performed in the course of routine work. Chung [28] describes a study that compared four head-tracked and three non-head-tracked modes for steering a beam in a simulated radiation therapy task. In addition to a number of objective measures of performance, Chung replayed the experimental trials to obtain a qualitative impression of subject's performance. Chung reports that subjects seldom followed a systematic search strategy for determining the best beam direction. While unable to describe the reasons why many beam paths were unexplored, Chung states that beam paths that would require looking straight up or straight down were often ignored. He states that the head mounted display apparatus would have exerted a large torque on a subject's neck when they looked in these directions, presumably making it uncomfortable.

The work of Soechting et al. [88] indicates that non-uniformities are present in hand movement that are a result of the kinematics of the arm. When subjects were asked to draw free-hand circles in different planes relative to the body, the actual path deviated from the intended path, yet subjects were unaware of this deviation. This suggests that performance of a movement is likely to vary if the plane, or the spatial region in which movement takes place, varies.

2.4.2 Dimensional Integrability

Multiple degrees of freedom can be controlled simultaneously, or separately. Ware and Slipp [104] performed a study in which they compared a SpaceBall, a six degree of freedom mouse and control panel input using a regular two dimensional mouse. One of the tasks required subjects to navigate into a narrowing tube with square cross sections. The six degree of freedom mouse easily outperformed the SpaceBall, but was roughly equivalent to the Control Panel. What makes this surprising is that the control panel interface required constant movement of the two dimensional mouse to operate 12 virtual buttons. In this application the ability to manipulate six degrees of control simultaneously with the six degree of freedom mouse is balanced by the more precise input control, constrained to operate in only one dimension at a time, offered by the control panel.

The work on the integrability and separability of input tasks by Jacob[57] et al. that was discussed in Section 2.1 clearly demonstrates that in certain situations a task or device may provide unused, unneeded or unwanted simultaneous control over available degrees of freedom.

2.4.3 Visual and Geometric Constraints

Occlusion is a common problem encountered in three dimensional tasks that is infrequent in two dimensions. When viewing a shaded three dimensional scene, objects close to the viewpoint obscure objects that lie further from the viewpoint along the viewing axis. This same situation exists in the physical environment around us, but changing our viewpoint to eliminate the occlusion is much easier.

One way of dealing with the occlusion problem is to render the objects in the scene as outlines only (wireframe) rather than fully shaded objects. The wireframe approach mostly eliminates the occlusion problem, but at the same time it eliminates a powerful depth cue. In addition, a new problem relating to scene density is introduced. The wireframe outlines of all the objects along the viewing axis are now visible. If there are many objects in the scene the view becomes crowded and confusing to the user. McWhorter et al. [66] conducted a study of user's subjective rankings of CAD drawings and found that users ranked wireframe stereo displays high in terms of geometric information, but lower than occlusion based techniques in their level of realism.

Translation parallel to the viewing axis and rotations about axes perpendicular to the viewing axis (parallel to the projection plane) are also problematic. In the case of translation under the control of a two dimensional input device these operations are difficult because small movements in the input map to large movements of the object being controlled. In the case of rotations, the mapping of input control to the object is non-uniform and only allows a restricted range of rotations.

Phillips et al. [75] propose a novel way of dealing with these problems by automatically adjusting the viewpoint. In the case of occlusion they search for a viewpoint that will provide an unoccluded view of the object being manipulated. In the case of geometric problems they rotate the viewpoint about the object under control to eliminate the parallel or perpendicular constraints.

2.4.4 Dimensional Bias

Gestures are a frequently used form of non-verbal communication between people. Gestures can be used in situations where visual contact can be made but voice contact can not. Gestures can also be used to provide a more natural form of interaction. In these situations gestures that are already known to users are used as input rather than requiring users to learn some new input technique.

Sturman [90] describes a variety of tasks where hand gestures in three dimensions can be used. One example comes from the gestures used by a ground based construction worker to direct the actions of a crane operator. A system that can recognize the gestures of the construction worker could be used to operate the crane. Interestingly, most of the gestures are essentially planar, performed with the hand moving freely in space.

Boritz [17] investigated recognition of gestures in three dimensions and found that a majority of gestures could be described in one or two dimensions. In his scheme a principal plane is determined for each gesture and the 3D gesture is then projected onto this principal plane. Recognition of the gesture then proceeds in the same manner as for 2D gestures.

2.4.5 Directional Bias

Zhai and Milgram [109] conducted a study to investigate accuracy in a six degree of freedom pursuit-tracking task. They used an isometric rate controller and a self-developed elastic[7] rate controller to manipulate the tracker. They found significant differences in tracking error along all three axes. In particular, tracking error in the Z direction was 40% greater than in the X direction. Tracking error in the Y direction fell in between the error rates for the X and Z directions.

Fitts' Law [39][24] does not consider direction of movement, but assumes that movement time should be uniform in all directions. Boritz, Booth & Cowan [19] conducted a study to determine if the speed and accuracy of mouse movement in a plane parallel to the desktop was isotropic or if it varied with the direction of movement. Subjects were required to perform a simple target selection task using both their dominant and non-dominant hands. Two target sizes, two target distances and eight target directions were tested in a fully within subjects factorial design. They found that movement time does vary with direction. While only a small subset of the conditions displayed a statistically significant difference, they found that for movements made by right-hand-dominant subjects using their right hands, movement horizontally to the right was always fastest and movement straight down was always slowest. When right-hand-dominant subjects used their left hand, movement straight down was almost always slowest, but the fastest condition was movement horizontally to the left.

They further showed that even in two dimensions performance is not directionally uniform. Possible causes for this non-uniformity include:

- Perceptual scanning bias would allow the targets on the horizontal axis to be detected

[7] An elastic control falls somewhere between an isometric and an isotonic control on the sensing axis, allowing more movement than an isometric control but less than an isotonic control. An elastic control is likely to possess a self-centering feature.

prior to other targets.

- Bias introduced by the use of a scanned raster display[8] would allow targets at the top of the display to be detected before those at the bottom of the display.

- Biomechanical limitations on movement that require larger limb movements to move the mouse straight down.

- More practiced and finer control for left and right movements that can be performed using only the wrist.

2.4.6 Synopsis

This chapter has reviewed the development of input techniques for three dimensional environments. It began by examining the development of input device taxonomies intended to aid in the matching of devices to input requirements. The review of interaction techniques used to manipulate objects in 3D environments highlighted a variety of problems related to interaction in 3D using devices with fewer degrees of freedom than the task demands. The discussion of three dimensional input outlined how 3D interactive tasks are being developed based upon the 2D interactive tasks that preceded them.

Viewing and interacting in a 3D environment is difficult, but the additional limitation of using a conventional 2D display makes interaction even more problematic. A number of techniques have been developed to reduce the limiting effects of the display. While these techniques have been shown to be effective in enhancing the ability to perform tasks in 3D environments, there is little known about how these techniques might interact with each

[8] Raster video displays are typically filled by an electron beam that scans from top to bottom. Presentation of a stimulus on the screen appears instantaneous but actually requires approximately 16ms (assuming a 60Hz display). The result is that items at the top of the display are shown before items at the bottom of the display. While the amount of time involved is small, it is within human perceptual limits.

other, with the content of the display and with the structure of the interactive task itself. By combining some of these factors into controlled laboratory experiments, this thesis seeks to further our understanding of interaction in desktop virtual environments. The goal is to direct others in the development of interactive techniques that are derived from our knowledge of human capabilities in these environments.

Chapter 3

Experimental Overview

As stated earlier, the central focus of this thesis is a series of four experiments. These experiments were each conducted using the same general structure and experimental procedure. In this chapter I will describe common elements shared by all of the experiments: the computer system – both hardware and software – and the preliminary screening steps in each experiment. The format will follow the structure used for the descriptions of the individual experiments in the chapters that follow.

3.1 Hypotheses

The following independent variables were manipulated although each experiment only used a subset:

- *Display mode* – monoscopic versus stereoscopic. Earlier studies by Zhai [110], Arthur et al. [6] and Sollenberger and Milgram [89], suggest that stereoscopic viewing should result in improved performance.

- *Head-tracking mode* – fixed viewpoint or head-tracked viewpoint. Earlier work by Arthur et al. [6] and Sollenberger and Milgram [89], suggest that the presence of head-tracking should result in improved performance.

- *Target position* – six positions $\pm$10cm from the starting position at the origin along one of the X, Y or Z axes. Very few studies have considered different target positions in space. Takemura et al. [93] found no differences across a wide variety of different target positions. Ware and Balakrishnan [101] found a difference between the X-axis target position and the Z-axis target position. These differing results provide little insight, however, most prior work suggests that Z-axis targets should have poorer performance than target positions along the X- and Y-axes.

- *Target orientation* – seven orientations: a default orientation and $\pm 45°$from the default orientation about one of the X, Y or Z axes. No studies have been found that distinguish between different target orientations and thus there are no expected differences between target orientations.

The following dependent measures were gathered for each task, but once again, not every dependent measure was used in each experiment. The dependent measures are all linked to the notion of effectiveness whose characterization is the overall goal of this thesis. A *click* is defined as the transition of a button from the up position, to the down position and back to the up position. In all experiments, subjects used the middle mouse button – with the mouse held in the left hand – to provide the clicks that moved them through the trials. The *ideal path* is the shortest distance between two points when no explicit path is specified. When an explicit path is specified, the ideal path is the explicit path.

Speed

- *trial completion time* is the time from the click by the subject to start the trial until the click to end the trial.

Accuracy

- *final position error* is the distance between the tip of the pointer and the tip of the target when the trial end click is received.

- *final position axis error* is the absolute value of the position difference between the pointer and the target along each of the X, Y and Z axes when the end trial click is received.

- *final orientation error* is the difference in orientation between the pointer and the target when the trial end click is received.

- *final orientation axis error* is the projection of the unit vector representing the axis of rotation onto the X, Y and Z axes.

- *RMS error* is root mean square of the distance from the tip of the pointer to the closest point on the ideal path from the start position to the target position.

- *RMS axis error* is the root mean square of the of the distance from the tip of the pointer to the closest point in the ideal path along the X, Y, and Z axes for each pointer position along the movement from the start position to the end position.

Conciseness

- *path length* is the length of the movement path of the tip of the pointer between the trial start and trial end clicks. Path length is computed as the piecewise sum of the distance between successive pointer positions.

- *orientation length* is the cumulative total amount of rotation of the pointer between the trial start and trial end clicks. Orientation length is computed as the sum of the successive absolute differences between pointer orientations.

Felicity

Felicity was measured by a set of questionnaires, one that was administered at the end of each viewing condition session and a final questionnaire that was administered at the end of the experiment after all viewing condition sessions had been completed.

The post-session questionnaire asked subjects to score the task based on the difficulty of determining some property of the stimulus and the difficulty of interactively matching or tracing that stimulus using the pointer. These questions were answered on a 5 point scale where 1 was “easy” and 5 was “hard.” The specific questions varied somewhat as the wording was adjusted for each of the experiments. Copies of the post-session questionnaires can be found in Appendices A.3 through A.6.

At the end of the experiment subjects were asked to score all viewing conditions they encountered according to task difficulty and ease of use. Subjects were also asked to rank the viewing conditions they encountered in order of their preference for the mode, the ease of use of the mode and their performance in the mode.

Dependent Measure Details

Many of these dependent measures are fairly easy to understand, but the RMS error and final orientation axis error are somewhat unusual. RMS error is defined as:

$$RMSerror = \sqrt{\frac{\sum (d_i)^2}{n}}$$

where d_i is the distance (error) from the tip of the pointer to the ideal movement path at step i and n is the number of samples in the path between the start and end clicks.

Final orientation axis error is given as the projection onto the X, Y and Z axes of the unit vector representing the axis of rotation for the rotation that takes the final orientation of the pointer to the orientation of the target. This vector is determined from the 2nd, 3rd and 4th elements of the quaternion that rotates the final orientation of the pointer to the orientation of the target. This vector must be normalized first, because in its quaternion form the magnitude of the vector is reduced by the magnitude of the rotation.

The projection of the rotation axis vector is used instead of the Euler rotations about X, Y and Z because the computation to extract the Euler axis rotations is problematic. Conversion of the single axis rotation to Euler rotations results in an implicit ordering of rotations about X, Y and Z. The conversion also limits the maximum rotation about one of the axes (usually the Y-axis) to 90°, which may unfairly bias the results. A more extensive exploration of this issue appears in the Discussion Section of the Docking Experiment.

Significance Level

For statistical testing, an α level of 0.05 was used to determine whether or not the independent variables had a significant effect.

Two-tailed tests[1] were used in all circumstances to allow outcomes contrary to initial expectations to be considered significant.

[1]A more detailed argument in favour of two-tailed tests can be found in Howell [55]. If a one-tailed test is conducted and a potentially significant result is detected in the other tail (at the original α level) then accepting this result as significant implies an α level that is actually 50% higher than initialy planned. Conducting two-tailed tests even when there is some expectation of the outcome allows for a contrary result to be found significant while maintaining the original α level.

	E1	E2	E3	E4	Total
Potential Subjects	10	10	9	11	40
Failed Screening					
Handedness	0	0	0	1	1
Colour Vision	1	0	0	0	1
Stereo Vision	0	2	1	0	3
Other	1	0	0	2	3
Accepted	8	8	8	8	32

Table 3.1: Table showing the number of potential subjects for each experiment, the number disqualified for failing one of the screening tests, the number lost for other reasons (e.g. system failure) and the final number of subjects accepted into each experiment.

3.2 Participants

Paid volunteer subjects were recruited via postings to local network newsgroups. All subjects were graduate or undergraduate computer science or engineering students. In order to reduce the variance not accounted for by the experimental factors, subjects were required to be male, right-handed, and with normal colour vision and acceptable stereo vision. Subjects were informed of these requirements in the recruiting posting and in the consent form. To ensure that subjects met these criteria they were screened prior to the experimental session. Each experiment made use of eight subjects none of whom participated in any of the earlier experiments. Table 3.1 shows the number of subjects eliminated as a result of the screening tests for each experiment.

Handedness

Handedness screening was conducted via the Edinburgh Inventory developed by Oldfield [71], modified to include a question on the hand used to control a computer mouse. A copy of the handedness screening form appears in Appendix A.2. Handedness (sinistrality) is not a binary condition where an individual is strictly right- or left-handed, but rather a continuum where an individual lies somewhere between being exclusively right- or left-handed. Oldfield [71] and Annett [5] have conducted extensive research on determining where someone lies on this continuum. Oldfield's Edinburgh Inventory was selected because it was felt to be easier to administer and score than Annett's procedure. Using the modified Edinburgh Inventory a subject was accepted to be right-handed if he indicated that he used his left hand exclusively for no more than 2 of the 21 tasks listed.

Colour Vision

Colour vision is a complex process with a variety of possible deficiencies. The stimuli used in the experiments relied heavily upon the ability to distinguish between colours. Screening for normal colour vision was performed using InSight-2 InColor [9], a software package that implements the Farnsworth Dichotomous Test for Color Blindness Panel D-15 [36]. In this test subjects are shown a series of 15 colour caps (coloured circles) and simply told to order them by closest match starting from a specified colour. Individuals with normal colour vision will arrive at a specific ordering, perhaps making one or two transpositions in the sequence. Those with a colour deficiency will form orderings in some other sequence (varying according to the specific deficiency). Under controlled conditions the D-15 Panel is a tool used to diagnose colour vision deficiencies. For screening purposes subjects were considered to have normal colour vision if they showed two or fewer transpositions of the normal colour sequence.

Binocular Vision

In most real world viewing, the depth information based upon the binocular disparity of an object in stereo image pairs matches the depth information obtained by vergence and accommodation [74]. Many of the techniques to simulate depth via binocular disparity – including the field-sequential stereo technique used here – are unable to simulate depth via focal distance. In these systems, focal depth is centered upon the surface of the display while disparity generally indicates a depth somewhere in front or behind the display. In order for these systems to work, an individual's visual system must be able to deal with the conflicting depth cues and give priority to the disparity cue. Fortunately, these depth cues can be decoupled in most individuals. However, some individuals with otherwise normal binocular vision may not be able to fuse disparity information in display systems such as the one used in this study.

Acceptable stereo vision is defined for the purposes of this study as the ability to correctly make stereoscopic depth judgments on stimuli similar to those displayed in this experiment. To test for this ability a special screening test was developed in which a subject is shown a set of three objects and asked to state which of the three – identified by number – is different from the other two in depth. The three objects are identical to the pointer and target used in the first experiment. The presentation of the stimuli is carefully controlled so that disparity is the only available depth cue. The stereo images are drawn in an orthographic projection. A screening session consists of a three trial segment used to instruct the subject about the task and verify that the equipment is functioning, followed by a randomized fifteen trial segment. A subject is accepted as long as no more than one error is made in the fifteen trial segment.

On occasion, subjects do not immediately adapt to the display and are unable to make correct depth judgments for the first three to five trials. After a series of incorrect responses (subjects are not told whether their responses are correct or incorrect), subjects make correct depth judgments in all the remaining trials. In this situation, subjects are shown an additional fifteen trial sequence and required to make no errors in these trials in order to be accepted.

According to Patterson and Martin [74] approximately 30% of the population has difficulty perceiving depth information in static stereoscopic displays of brief duration, but this deficiency is greatly reduced when display duration is increased. A study by Tittle, Rouse and Braunstein [94] showed that most subjects that had difficulty judging depth in static disparity displays were able to make use of disparity information in displays with dynamic disparity.

In the screening test, subjects were allowed to view the static stimuli until they made a response. During experiment trials, the moving pointer provided dynamic disparity information, presumably making the disparity depth judgment easier than in the screening task.

3.3 Equipment

A system for conducting experiments of 6DOF interaction tasks in a desktop virtual environment was assembled for this study. The system made use of a Silicon Graphics Indigo 2 workstation, a stereo capable monitor, mechanical and magnetic trackers, LCD glasses and custom software.

3.3.1 Hardware

A computer system for conducting experiments in an interactive virtual three dimensional environment supporting head-tracking and stereopsis (using the field sequential presentation technique) was developed to conduct these and future experiments. The system consists of a Silicon Graphics Indigo 2™ three dimensional graphics workstation with a 133MHz MIPS R4600 CPU, 64 M of memory, a 1 GB internal hard drive and a GR3-Elan graphics board. The workstation was connected to 10 Mbps Ethernet local area network with most file systems accessed via NFS. The display was a 19” colour monitor with a resolution of 1280 pixels horizontally and 1024 pixels vertically (512 pixels vertically in stereo), a Shooting Star Technologies ADL-1™ [87] six degree of freedom head tracker, a Polhemus Fastrak™ [76] six degree of freedom input device, and a pair of CrystalEyes™ LCD shuttered glasses.

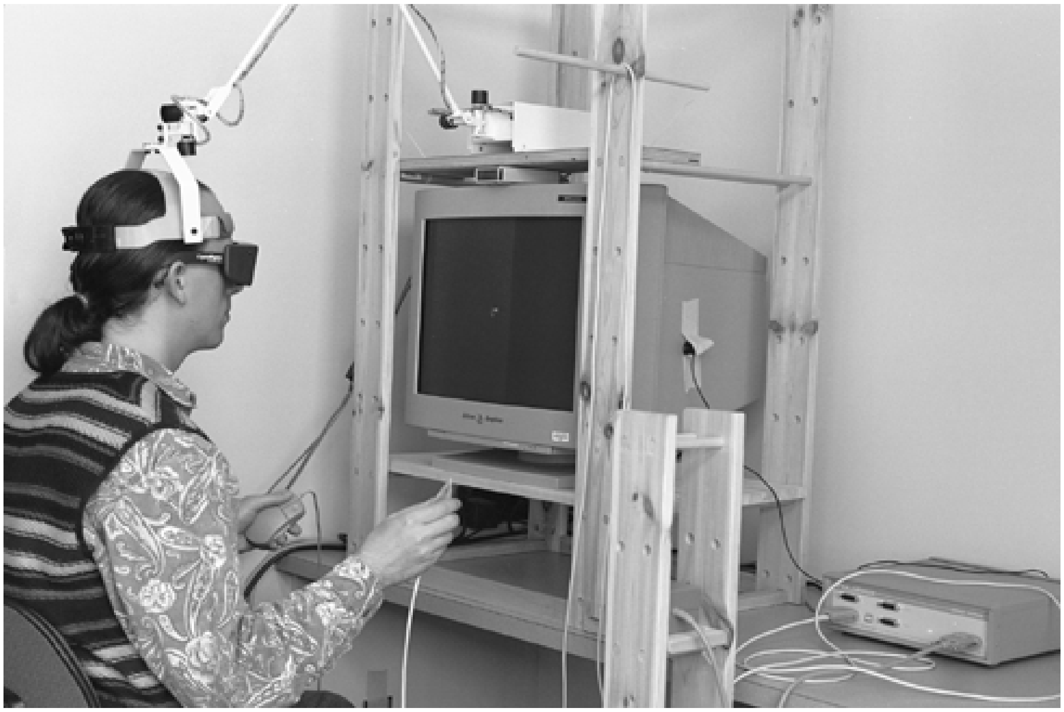


Figure 3.1: Photograph of experiment system components used in three of the four experiments. The subject seated in front of the monitor is wearing the LCD glasses and the ADL-1 head-tracker. The pointer can be seen in the subject's right hand and the mouse is visible in the subject's left hand. The Fastrak controller is visible in the lower right background. The Fastrak transmitter is mounted in the wooden structure in the foreground at the lower right.

Figure 3.1 shows the experiment system hardware used in three of the four experiments.

The ADL-1 and Fastrak input devices were selected because of their high sampling rates and low sensing lags. According to its specifications, the ADL-1 has a sensing lag of less than 2ms and accuracy of 0.5 cm. In our experiments eye position is only estimated to within a few centimetres. Trip [96] reported a somewhat lower accuracy (2.0 cm) than the manufacturer's claim, however, even this was within the accuracy we required for head position. The specifications of the Polhemus Fastrak indicate a sensing lag of 4ms and an accuracy of 0.08 cm. The sensing lag of the Fastrak is a huge improvement over the earlier Isotrak II,™ which has a sensing lag of 20ms. A formal analysis of the system lag was not

Figure 3.2: The physical pointer used in this study. The Fastrak receiver is enclosed within Crayola Model Magic™ modeling material, shaped by hand and coloured using felt markers to resemble the virtual pointer.

performed. Informally, the data timestamps indicate that the system obtained data from the two input devices at 30 Hz. Assuming an average one frame delay until the data is rendered to the display would imply the average sensing lag of the system is 50 ms.

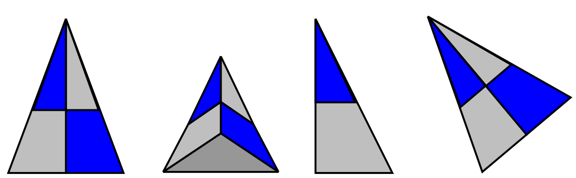
The receiver of the Polhemus Fastrak was encased in Crayola Model Magic™ modeling material that was shaped and coloured to resemble the the virtual pointer displayed in the virtual environment. When dry the modeling material has weight and consistency similar to that of styrofoam used to make cups. The physical pointer was coloured using felt tipped Crayola Markers™ as required for each experiment. Figure 3.2 shows the physical pointer as it appeared in the last three experiments.

There was no attempt to match the size of the physical and virtual pointers. The virtual pointer was intended to be small in comparison to the size of the display. The physical pointer was intended to be comfortable to hold. The virtual pointer was a tetrahedron-like object, in which the height (1.73cm) was twice the width (0.87cm). The difference in size along one dimension allowed one vertex to be easily identified as the tip and the opposite face to be identified as the base.

The physical pointer, the virtual pointer and the target had identical appearances in each experiment, but the colouration changed after the first experiment. In the first experiment the two front faces were drawn with a single colour. For the other experiments, a checkerboard scheme was used on the two front faces. Figure 3.3 shows the pointer as it appeared in the last three experiments.

The field-sequential system used here employs a high-frequency computer monitor to display the images and a pair of liquid crystal shuttered lenses to select which eye sees the currently displayed image. An infrared signal is used to synchronize the glasses and the monitor.

When operating in stereoscopic mode the decay rate of the phosphor may not be fast enough to achieve a good separation between the stereo pairs [62]. A phosphor with a slow decay rate may persist on the display into the next field when the image for the opposite eye is being drawn. The resulting field-to-field persistence or "bleeding" between the images for each eye interferes with the fusion process and produces a double image. In practice, the



(a) (b) (c) (d)

Figure 3.3: Appearance of the pointer and target in the default (a), $-45°$X (b), $+45°$Y (c) and $+45°$Z (d) orientations.

decay rates of the red and blue phosphors have been found to be fast enough, but the decay rate of the green phosphor is slower and can bleed between fields in certain situations.

Two methods were used to deal with this problem. In the first two experiments no colours making use of the green phosphor were used. In the third and fourth experiments, where some shaded objects were used, a gray background was used to mask potential bleeding into the field for the opposite eye. The intent was for the intensity of the gray background to be stronger than the intensity of any phosphor persistence from the prior field.

Subjects used the mouse held in their left hand to advance through the experiment trials. The mouse provided a hand held button that could be easily operated without any need to be located visually. A button was deliberately not attached to the Polhemus receiver. Attaching a button to the receiver would have added an additional degree of freedom to the control in the right hand. The need to operate a button on the receiver would have limited the potential hand posture used by subjects to grip and manipulate the pointer. In these experiments subjects were encouraged to use the hand posture that they felt best suited their use of the control. The keyboard was removed to provide a physically unobstructed environment for operation of the control.

3.3.2 Software

A software system was developed by the author using C++ and the Silicon Graphics GL graphics library to conduct the experiments. The software system was developed in order to allow extensive control over all the components used within the experiments. At the outset of the work described here, the following broad requirements were established for the software.

- display of visual information during the experiment
- capture and coordination of input device operations

- time-stamped logging of input
- complete software controlled scripting of experimental trials
- facility for real-time playback of experimental trials

The software system allows an experimenter to assemble an experiment from a variety of components. The reference coordinate frame for the virtual environment uses a right-handed coordinate system with the origin (0,0,0) located at the centre of the monitor's phosphor surface (pixel location (639.5, 511.5)), the positive X-axis pointing to the right, the positive Y-axis pointing upwards and the positive Z-axis pointing out toward the viewer. The pointer moves in a similar right-handed coordinate system whose axes are aligned parallel with the axes of the virtual display, but whose origin is offset to result in a comfortable hand and arm position for a subject seated in front of the display. There is a one-to-one mapping between the coordinate system used for the control and the coordinate system used for the display.

I would like to briefly discuss the importance of each of the requirements stated above and relate the benefits that have accrued from developing the software to support them.

Accurate display of visual information was crucial to the experiments that were conducted. The system supports the presentation of three dimensional objects in the virtual environment along with two dimensional information such as buttons or text. In three dimensions the display simulated a viewing portal into a small virtual environment located behind the glass of the monitor. Rendering of objects on the the display mimicked what a viewer would see if the virtual objects were replaced with real objects of the same sizes and at the same positions in space. The general approach described by Deering [33] for head-tracked stereo was followed with the exception of the corrections for the refraction and curvature of the screen.

The different viewing conditions used in this study would change the detailed nature of the display. In the perspective-only condition (no binocular disparity and no head-tracking) the display represented a perspective projection of the environment from a fixed view position. The head-tracked perspective condition allowed the display to be adjusted to mimic what the viewer would see from their current view point. It should be noted that in this study eye-point was estimated based on the sensed head position and not measured directly. The same eye-point estimation was used for all viewers.

Binocular disparity could be added to the display with or without head-tracking. The binocular (stereoscopic) display was generated by producing an estimated position for each eye. In the non head-tracked condition these positions are fixed as in the monoscopic case. In order to make the displayed disparity more accurate, the inter-pupilary distance (IPD) can be specified. In this study the IPD of each subject was measured using a ruler[2] and used in all the stereoscopic displays they were shown. An unfortunate side effect of the hardware system used to generated stereoscopic imagery is that the vertical resolution of the display is reduced by half from its monoscopic counterpart. One potential means of keeping the vertical resolution constant between both display modes would be to render the monoscopic display at half the resolution as well.

During the course of the experiments the system had to capture and record information from multiple sources. Two sensing devices were used, the Polhemus Fastrak for pointer position and orientation and the Shooting Star ADL-1 for head position and orientation. These devices sense data in their own coordinate frames and do not provide any timing information. As data comes in from each device it must be timestamped and adjusted

[2] A device known as a pupilometer can be used to measure inter-pupilary distance. In a conversation with some optometrists they indicated that in the absence of such a device, a ruler can be used to give good measurements of IPD by measuring from bridge of the nose to the outside edge of the pupil.

to the reference coordinate frame. In this system, the data from both devices are collected via a polling loop and are given the same timestamp by the software system after they have been retrieved. During the experiments the middle mouse button is used by the subjects as a hand held button to indicate the start and end of a trial. The button click event is timestamped when it is extracted from the system event queue. To avoid unpredictable system response and dependencies upon the file system, the experiment system stores all data in memory buffers that are flushed to disk only at the end of a block of trials. The buffers are not flushed after each trial as this could result in unpredictable delays between trials.

All of the significant input that occurs during the course of an experiment was captured to allow for future playback. For the work described here, there were two event streams. The first stream consists of all the timestamped data from the two sensing devices. The second stream contains the trial start and stop events. These streams are saved to separate files during the experiments and recombined during playback. The playback facility allows the experimenter to replay any trial and extract additional information. In a system where a high frame rate and low lag are important, the computation performed while the subject is interacting with the system can be kept to a minimum. Upon playback, when interactive response is not as critical, instrumentation can be added to the system to gather additional statistics.

In the tracing experiments (Experiments three and four) where a path of some sort had to be displayed, cylindrical sections (only one in the case of the straight line path) were used to connect the start and end points rather than a single pixel line. A single pixel line is difficult to see and fuse. The cylindrical sections were composed of polygonal facets and a GL lighting mode was used to shade the polygons and smooth the edges between facets. In order to make the lighting somewhat natural a white light source was used. The implication was that highlights on the surface of the rendered cylinders may appear white rather than the colour assigned to the curve.

In experiments one and two, green was avoided because of the lengthier decay time of the green phosphor that might cause bleeding between frames in the stereoscopic display mode. Because of the lighting a small amount of green may appear on the display. To mask potential bleeding a dark gray background was used rather than the black background used in earlier studies.

3.4 Procedure

An experiment is composed of the following five components. The first three components are expressed within the experiment system software and control the presentation of trials to the subject. The last two are for logical organization purposes only.

Experiment Sequence Terminology

- *Trials* are the basic component of each experiment and consist of a subject performing the experiment task.
- *Blocks* are a grouped sequence of trials intended to be performed in succession with no planned breaks.
- *Sessions* are a sequence of blocks having the same overall viewing parameters (e.g. monoscopic and head-tracking).
- *Sittings* represent a group of sessions given on the same day. Some experiments can be completed within a single sitting, but longer experiments require multiple sittings on separate days.
- A *full session set* is used to describe the group of sessions administered to subjects across one or more sittings.

Pre-Experiment Activity

Before an experiment can begin, all potential subjects must read and sign a consent form. The consent form is required by the university ethics review board. The form gives the subject some basic information about the experiment, along with outlining the payment and withdrawal conditions. A sample consent form can be found in Appendix A.1. Once a subject has consented to perform the experiment, the screening tests described in Section 3.2 are performed. If successful on the screening tests, subjects proceed to the experiment.

Experiment Organization

Trials represent relatively short periods of activity, and are meant to be performed in fairly rapid succession. Trials are grouped together into a block. Each block was intended to be of five to ten minutes in duration followed by a break, to prevent subjects from becoming overly fatigued during the experiment. Short planned breaks are provided between blocks to allow subjects to rest their arms and their eyes. Blocks in turn are grouped into a session. Each session is performed under the same viewing conditions and subjects remain seated in front of the display during the entire session. Between sessions, subjects can get up and move about the room or go to the restroom.

At the end of each session, a short one page questionnaire regarding the difficulty of performing the task in the particular viewing condition is administered to the subject. The end-of-session questionnaires are presented in Appendices A.3 through A.6. At the completion of a full session set, when sessions for all the viewing conditions have been completed, a lengthier three page questionnaire dealing with all viewing conditions is administered. The end-of-experiment questionnaire can be found in Appendix A.7.

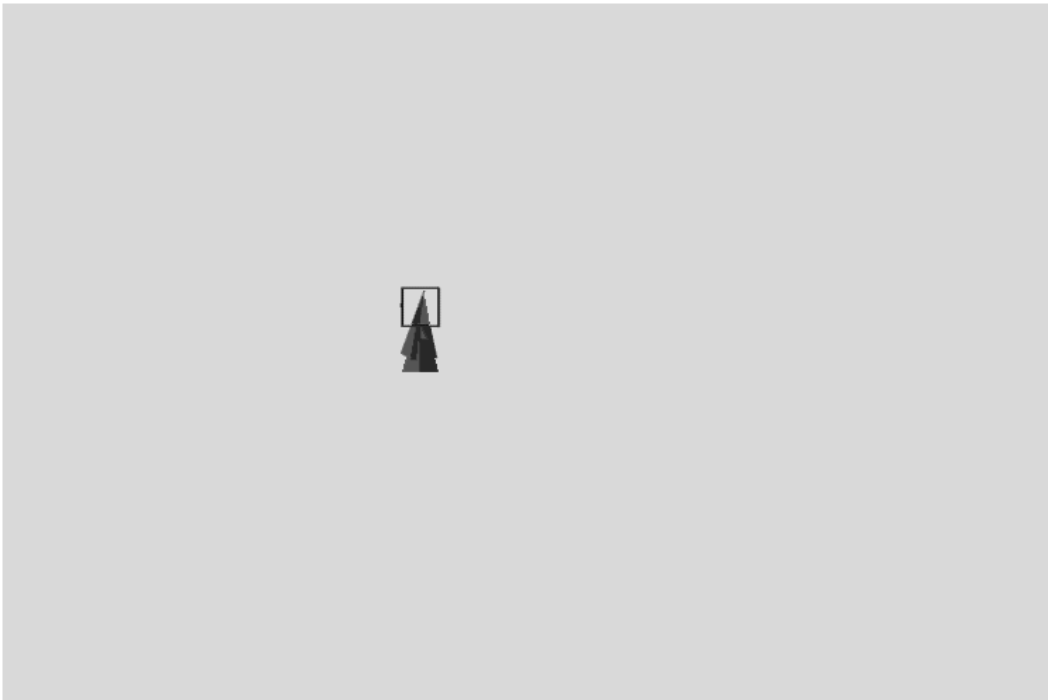


Figure 3.4: An image of the screen during the homing phase of a trial. The virtual pointer and homing target are in almost the same position and orientation. The red box indicating that the tip of the virtual pointer is within the homing tolerance is visible. The background of the screen was black. The background of the image has been lightened to allow better reproduction.

General Trial Operation

Subjects were required to perform a task in three dimensions. A subject manipulated the pointer via the Polhemus input device using his dominant right hand. The middle mouse button on the mouse held in the left hand was clicked to advance through the experiment trials.

A trial consisted of two phases, an initial homing phase during which the subject had to move the tip of the pointer to the tip of the homing target located at a fixed centre point in the virtual environment, and a targeting phase where the subject had to then move the pointer to the tip of the trial target. The homing target is also referred to as the start target and the

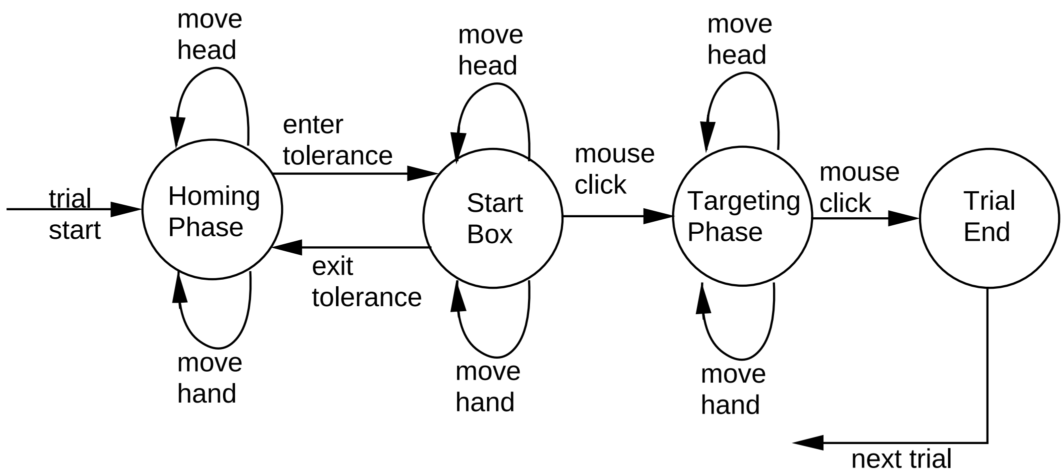


Figure 3.5: State transition diagram showing the stages of each trial.

trial target is referred to as the end target.

During the *homing phase*, subjects received visual feedback in the form of a red box that appeared around the tip of the homing target when the tip of the pointer was within a homing tolerance of 0.5cm from the homing target. When the red box appeared subjects could click the middle mouse button using their left hand to advance to the *targeting phase*; the homing target was removed and the trial target was displayed. Figure 3.4 shows an image of the screen during the homing phase of a trial.

Subjects could not advance to the targeting phase if the homing tolerance was not met. No feedback regarding proximity was given during the targeting phase. Subjects had to make their own determination of proximity using the visual cues available within the current experiment condition. Subjects clicked a second time with the mouse when they were satisfied with their performance during the targeting phase. Whatever stimuli had been present during the trial was then removed from the display and the position of the virtual pointer was frozen. After a short delay the system re-enabled the virtual pointer and advanced to the next trial. Figure 3.5 graphically represents the phases of a trial.

Experiment	1	2	3	4
Independent Variables				
Display Mode	X	X	X	X
Head-Tracking Mode	X		X	X
Target Position	X	X	X	X
Target Orientation		X		
Dependent Variables				
Trial Completion Time	X	X	X	X
Final Position Error	X	X	X	X
Final Position Axis Error	X	X	X	X
Final Orientation Error		X		
Final Orientation Axis Error		X		
RMS Error	X	X	X	X
RMS Axis Error	X	X	X	X
Path Length	X	X	X	X
Orientation Length		X		

Table 3.2: Summary of the independent variables that were manipulated for each experiment and the dependent variables that were analyzed for each experiment.

3.5 Design

The independent variables manipulated during each experiment and the organization of sessions was different for each experiment. Full descriptions appear in the chapter for each experiment. Table 3.2 provides a summary of the independent and dependent variables that were used in each experiment.

For compatibility between the stereoscopic and monoscopic conditions all target positions in the virtual environment were behind the screen surface and were displayed at a size appropriate to a perspective projection based on a nominal viewing position of 40cm in front of the screen surface for the fixed viewpoint condition. To mitigate the effects of the apparatus, subjects wore the stereo glasses and head-tracker (when appropriate) in all conditions.

3.6 Training

At the beginning of the first phase of the first experiment session subjects were given detailed instructions regarding the operation of the input device, the virtual environment and the task they were to perform. Other researchers (e.g. Hinckley et al. [53]) have chosen not to give detailed instructions to subjects in order to avoid possible influence by the experimenter. In my opinion this magnifies the already large differences between subjects. Subjects that may have had a passing encounter with a 3D interface or a similar input device are well ahead of those that have absolutely no experience. By providing detailed instructions to each subject I hoped to more quickly bring all subjects to a similar basic understanding of the operation of the input device and of the virtual environment.

Subjects are first trained in the operation of the pointer. The physical pointer is placed into their right hand and the mouse is placed into their left hand. They are then instructed using the following script that served as a guide for the experimenter, but was not read directly. Variations were made when necessary to accommodate individual subjects, but deviations were minimized.

> "On the screen you are seeing a 3D virtual environment. Notice the pointer on the screen and how it moves as you move the object in your right hand. As you move your hand left and right the pointer also moves left and right. As you move your hand up and down the pointer on the screen moves up and down. Now notice that as you move your hand away from you or toward you, the pointer on the screen also seems to move away from you or toward you. Finally, notice that you can also rotate the object in your hand and the pointer on the screen will rotate as well, allowing you to see other sides of the pointer."

Subjects are allowed to experiment with the pointer for a moment or two and then they are instructed about the primary depth cues in the system as follows.

> "Now let me tell you more about the depth cues available in the visual display. There is a perspective depth cue which is that objects get smaller when they are farther away from you and larger when they are closer to you. You'll notice that if you move the pointer away from you it gets smaller and if you move it towards you it gets larger. There is also an occlusion depth cue. Occlusion is what we call the situation that arises when an object closer to you blocks your view of an object that is father away. You will notice this if you move your pointer close to the stationary object on the screen. When the pointer is closer to you than the stationary object, the pointer blocks your view of the object and when the pointer moves farther away, it eventually moves behind the other object and blocks your view of the pointer."

The stereoscopic and motion parallax cues, when present, are described as follows.

> "In this mode the pointer should appear to move in space with a realistic sense of depth. We call this stereoscopic viewing. To get a feeling for how this mode works watch the pointer on the screen as you move it towards you and away from you."
>
> "In this mode you will notice that movement of your head changes the scene. As you move your head the view of the workspace is generated based on the estimated position of your eyes and should represent what you would see from that location if you were looking into a real environment."

After experimenting with the pointer and the viewing mode, the subject is asked if he has any questions. These are typically answered as long as the experimenter believes the answers will not bias the subject's performance in any of the visual modes. The experimenter then proceeds to train the subject to perform the task that he is required to perform in that particular experiment. Task specific training is described in the training section for each experiment.

3.7 Results

Each experiment was first screened for outliers. After outliers were removed the influence of the independent variables upon each of the dependent measures was analyzed using a repeated measures analysis of variance (ANOVA). Because of the number of factors involved in each of the analyses the ANOVA tables are large. The tables appear at the end of each chapter so they are easier to find and to facilitate better page layout.

Homogeneity of variance is an important assumption that underlies the analysis of variance technique used to analyze these and other experiments. The typical method of checking if this assumption is valid is to perform a plot of the model residuals against the predicted values. If the variance is relatively homogeneous, then the plot appears relatively uniform as predicted values increase. If the variance is not homogeneous, then the plot appears to spread out – indicating increasing variance – as the predicted values increase. A transformation of the data (log or square root) can be performed if the residual plot does not appear uniform. A second ANOVA can then be computed for the transformed data and used to make all statistical conclusions. Upon examination of the residuals, a log transformation was applied to all the dependent measures in all experiments.

A demonstration of this analysis appears in the Results section of the first experiment. While this analysis was carried out for each of the dependent variables in all of the experiments it is only presented once in full detail in Chapter 4.

Repeated measures ANOVA have an additional requirement of sphericity for the covariance matrix of the within subjects terms (Howell [54]). Violation of this assumption can inflate the Type I error rate. Adjustments to the degrees of freedom developed by Greenhouse and Geisser, and by Huynh and Feldt, can correct for violations of this assumption. To account for the possibility of violations of the sphericity requirement, the Huynh-Feldt adjusted degrees of freedom (F_{HF}) and probability are used in addition to the conventional F and p values and the adjusted p value is used to test for significance. It should be noted that when the numerator degrees of freedom is equal to one, no adjustment is necessary and thus no correction factor is computed.

3.8 Discussion

The results section provides only the statistical analysis of the data that was gathered for the experiment. Interpretation of the findings in the results section is performed in the discussion section for each experiment. The discussion section also draws upon related research and attempts to explain the possible causes of the observed results.

In a statistical sense a result is only significant if the probability that the observed data may be due to chance is less than some set threshold. However, in a logical sense, a finding that a particular independent variable did not influence task performance in some way is potentially just as interesting. These issues are explored in the experiments where certain independent variables failed to produce statistically significant results.

3.9 Conclusion

The conclusion section highlights a few of the most important findings for each experiment. Each of the experiments tended to produce a large number of significant results as well as an interesting non-significant result or two. While all of these findings are important, a few stand out because they can be more easily generalized to a broader class of tasks or be more readily applied within the context of systems currently under development.

Chapter 4

Experiment 1: Point Location

The point location experiment was the first experiment to be carried out. Point location is a 3D task even though a pointer making use of six degrees of freedom was used. In the point location task subjects have to match the position of the tip of a pointer to the position of the tip of a stationary target object. As one of the simplest possible tasks, the point location experiment served as a means of determining a baseline against which the results of future experiments could be compared.

4.1 Hypotheses

The experimental hypotheses of this experiment were that display mode (monoscopic versus stereoscopic), head-tracking mode (fixed viewpoint versus head-tracked viewpoint) and target position ($\pm$10cm along one of the X, Y or Z from a fixed starting location) have an effect upon performance in a 3D point location task. To evaluate the effect of the independent variables upon the task several different dependent measures were gathered.

The following six dependent measures were analyzed in this experiment: trial completion time, final position error, final position axis error, RMS error, RMS axis error and path length. A detailed description of these measures appears in Section 3.1. RMS error is the root mean square of the distance from the tip of the pointer to the closest point on the ideal path. For this experiment, the ideal path used to compute the RMS error is considered to be the straight line between the tip of the homing target and the tip of the trial target.

For statistical testing, an α level of 0.05 was used to determine whether or not the independent variables had a significant effect.

4.2 Participants

Subjects were recruited as specified in Chapter 3. One potential subject was excluded for failing to have normal colour vision. Eight subjects were accepted into the experiment.

4.3 Equipment

The standard equipment specified earlier in Section 3.3 earlier was used. The appearance of most of the equipment can be seen in Figure 3.1. Figure 4.1 shows an image of the screen from one of the trials in this experiment.

The pointer and target had identical appearances. Both were tetrahedron-like, in which the height (1.73cm) was twice the width (0.87cm). The difference in size along one dimension allowed one vertex to be easily identified as the tip and the opposite face to be identified as the base. The base was half intensity magenta and the other faces were full intensity red, blue and magenta. The shape of pointer can be seen in Figure 3.3. Note that in this experiment the front faces of the pointer and were a single solid colour as and not the checkerboard shown in Figure 3.3.

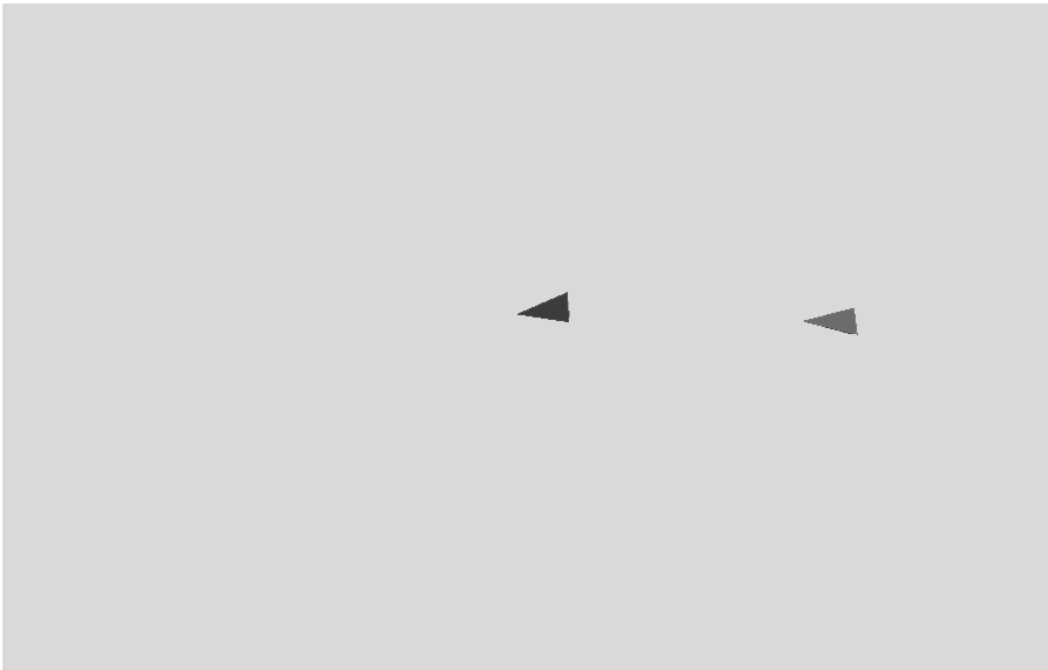


Figure 4.1: An image of the screen during a trial of the point location experiment. The target is visible to the right with the virtual pointer located near the centre of the image. The background of the screen was black. The background of the image has been lightened to allow better reproduction.

The colours used for the pointer and target were selected to avoid green because the decay rate of the green phosphor is slightly longer than a single frame time in the stereoscopic display, which results in bleeding between the images for each eye in the display, thereby impeding subjects' ability to correctly fuse the stereoscopic image.

4.4 Procedure

Subjects were required to perform a point location task in three dimensions. Subjects manipulated the pointer via the Polhemus input device using the dominant right hand. The middle mouse button on the mouse held in the left hand was clicked to advance through the experiment trials. A trial consisted of two phases, an initial homing phase in which the subject had to move the tip of the pointer to the tip of the homing target located at a fixed centre point in the virtual environment and a targeting phase where the subject had to then move the pointer to the tip of the trial target. During the homing phase, subjects received feedback in the form of a red box that appeared around the tip of the homing target when the tip of the pointer was within 0.5cm of the homing target. When the red box appeared subjects could click the middle mouse button using the left hand to advance to the targeting phase; the homing target was removed and the trial target was displayed. Subjects could not advance if the homing tolerance was not met. No feedback regarding proximity was given during the targeting phase. Subjects had to make their own determination of proximity using the visual cues available within the current experiment condition. Subjects clicked a second time with the mouse when they were satisfied with their performance. The trial target was then removed from the display and the position of the virtual pointer was frozen. After a short delay the system re-enabled the virtual pointer and advanced to the next trial.

4.5 Design

The experimental design consisted of four within-subject independent variables: display mode (2 levels, monoscopic or stereoscopic), head-tracking mode (2 levels, active or inactive), target position (6 levels, 10cm from the centre point along the positive and negative X, Y or Z axis) and block (4 levels, representing the successive blocks of trials within a particular session). For compatibility between the stereoscopic and monoscopic conditions, all target positions in the virtual environment were behind the plane of the screen. To mitigate the effects of the apparatus, subjects wore the head tracker and stereo glasses in all conditions even when head-tracking or stereo viewing were disabled. The receiver of the Polhemus was encased in modeling material that was shaped and coloured similar to the pointer in the virtual environment.

An experiment session comprised one of the four combinations of display and head-tracking modes. After each session, subjects completed a short questionnaire dealing only with that session. Each subject performed four sessions, thereby covering all four combinations of display and head-tracking modes. The combinations of display mode and head-tracking mode were counterbalanced across subjects according to a latin square where each condition was performed first, second, third or fourth an equal number of times. After the final session, each subject completed a longer questionnaire about all the conditions.

Each subject completed four consecutive sessions, one for each display condition. Each session was divided into five blocks, a training block followed by four experiment blocks. Each experiment block contained 30 trials consisting of five repetitions of each target position presented in random order. The training block was an abbreviation of an experiment block consisting of 15 trials. To minimize fatigue subjects were given a one minute break after each block.

Each subject performed a total of 480 experiment trials. Each subject spent approximately 90 minutes performing the experiment, including time for breaks and time to complete the questionnaires, but not including the time spent on the screening tasks.

4.6 Training

Task specific training was provided to subjects using the following script. As was the case for the general training in the use of the pointer and the environment, the the script served as a basis for the experimenter, but was not read directly.

> "Your task is to move your pointer from the starting point to the ending point as quickly and as accurately as possible. The starting point is located at the tip of the stationary object on the screen.

Move your pointer to the starting point now. Notice that when the tip of your pointer is close to the tip of the starting target a red box appears around the tip of the target."

[*Experimenter waits while subject moves pointer to start target*.]

"The red box must be present for you to be able to start a trial. To start a trial click the middle mouse button with your left hand."

[*Experimenter waits for subject to click with the middle button.*]

"The starting target has disappeared and another target has appeared somewhere in the virtual environment. Move your pointer to the target object so as to touch the tip of your pointer to the tip of the target."

[*Experimenter waits for subject to move pointer to target.*]

"Good."

"When you are satisfied that your pointer is as close as possible to the target, click the middle mouse button with your left hand. This will end the trial. The orientation of the pointer with respect to the target does not matter. You can have the pointer in any orientation you choose. All that matters is that you get the tip of the pointer as close as possible to the tip of the target."

"Please try to perform the task as quickly and as accurately as possible."

"Do you have any questions?"

The experimenter answered any questions that would not bias the subject's performance in any of the visual modes. Typical questions were, "Does the orientation of the pointer matter?" and "Does it matter what side of the pointer I have facing forward?" The answer to these questions was always, "No."

No instruction as to what was fast enough, or what was accurate enough was given, but the enforced accuracy requirement for the homing phase may have given subjects some guidance. If subjects asked whether they were to favour speed over accuracy or vice-versa they were told that they would have to decide for themselves how quickly or accurately they should perform the task.

The experimenter observed the subject during the training trials. When the training trials were over the experimenter asked and answered any questions using the same guideline stated previously. The experimenter then told the subject that he would be on his own during the actual trials. The experimenter then reminded the subject to perform the trials as quickly and as accurately as possible, left the room and started the first experiment block. The experimenter re-entered the room after each block to check on the subject and tell the subject to rest his arm for at least one minute. After waiting for a minute the subject was asked if he was ready to proceed. If the subject responded in the affirmative then the experimenter left the room again and initiated the next experiment block.

4.7 Results

The dependent variables stated earlier were gathered for each trial. Across all eight subjects a total of 3840 individual trials were completed.

Preliminary Screening

In an attempt to remove invalid trials, screening tests for outliers were performed. All trials with a final position error greater than twice the largest target dimension (3.46cm) were inspected. This yielded 21 trials (0.55%), which upon inspection revealed that most of the error was along the Z-axis. Because the Z-axis is the most difficult to visualize it was decided that it would be inappropriate to remove these trials. An attempt to screen out trials based on movement time, on a per subject basis, using a technique from Jolicoeur et al. [60] yielded 28 potential outlier trials (0.73%), however, because subjects were not given any guidance about how quickly to perform the task and because of the small number of trials involved, it was decided not to remove these trials either. Three trials were found to have a trial completion time under 0.5 seconds. Upon examination of logging data these trials were found to have been ended accidentally by subjects and were eliminated. Lastly, all trials with path lengths over 100cm were examined. This yielded a single trial with a 156cm path length. Upon examination of the video taped record of the session it was found that the subject proceeded to rub his nose during the trial using the hand that held the physical pointer. This trial was removed. The remainder of the analysis is based upon the 3836 remaining trials.

The experiment is a 2 (display mode) $\times$ 2 (head tracking mode) $\times$ 6 (target position) $\times$ 4 (session block) design with repeated measures on all factors. For the following analysis, the repetitions of each target position performed by a subject within a single block were averaged together to yield a single score. The result was a series of 768 measures, one for each subject in each of the conditions.

Analysis of Residuals

The residuals for trial completion time vs. predicted values are plotted in Figure 4.2. The plot shows how the spread of the residuals increases as the predicted values increase. A log

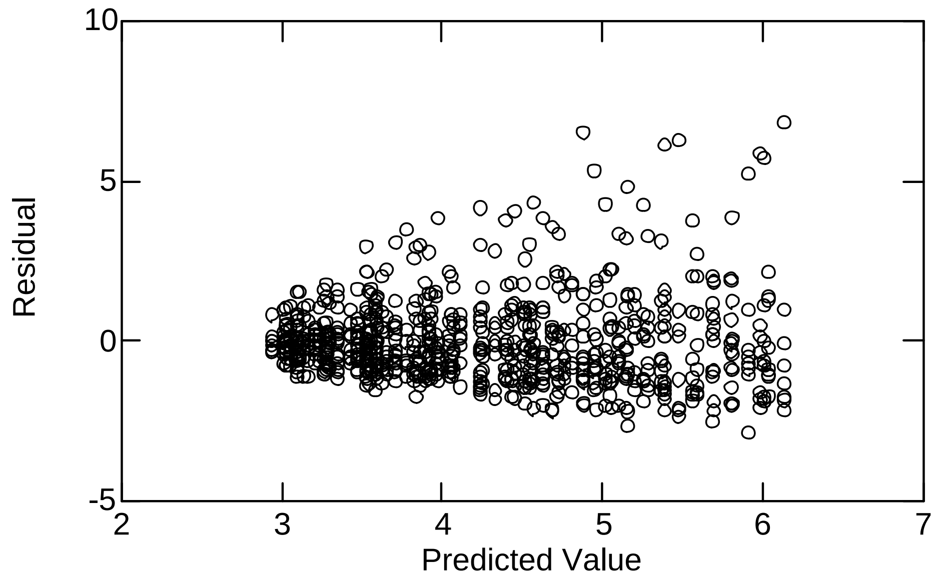


Figure 4.2: Predicted values vs. residuals for trial completion time.

transformation was thus performed on the data and the plot of the corresponding residuals can be seen in Figure 4.3. In this plot the residuals appear more uniform and are indicative of relatively homogeneous variance. Analysis of the residuals has been done for each of the ANOVAs, and a log transformation applied in every case. The plots and a detailed discussion are not included for the sake of brevity. It should be noted that whenever transformed data is used, all statistical procedures make use of the transformed data even though the non-transformed values are used in the interpretation of these results.

Degree of Freedom Adjustments

Repeated measures ANOVA have an additional requirement of sphericity for the covariance matrix of the within subjects terms (Howell [54]). Violation of this assumption can inflate the Type I error rate. Adjustments to the degrees of freedom developed by Greenhouse and Geisser, and Huynh and Feldt, can correct for violations of this assumption. To account for the possibility of violations of the sphericity requirement, the Huynh-Feldt (F_{HF}) adjusted

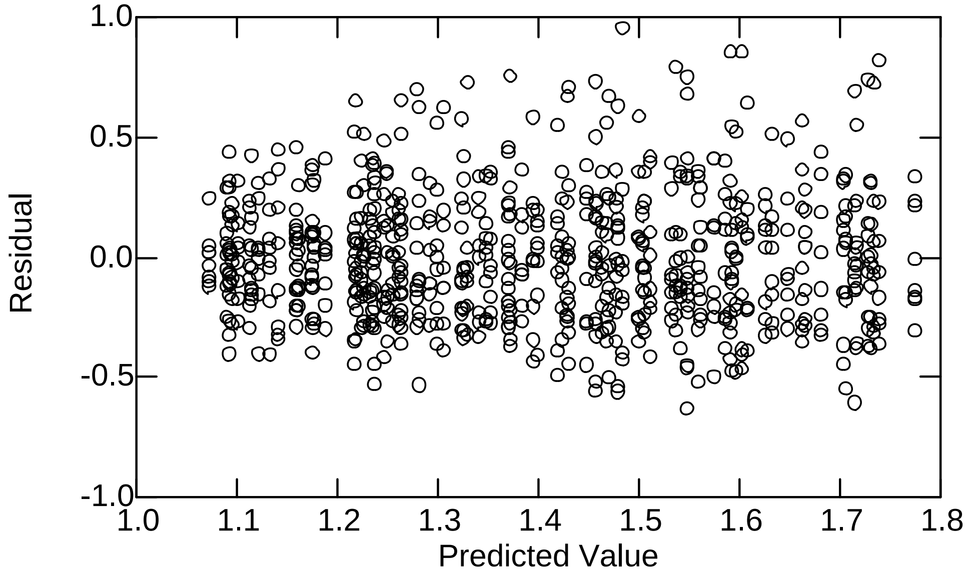


Figure 4.3: Predicted values vs. residuals for log of trial completion time.

degrees of freedom and probability are used in addition to the conventional F and p values and the adjusted p value is used to test for significance. The ANOVA tables at the end of the chapter provide both the conventional probability value (conv. p), the adjusted probability value (adj. p) and the the Huynh-Feldt epsilon value used to adjust the degrees of freedom (H-F ϵ). It should be noted that when the numerator degrees of freedom is equal to one, no adjustment is necessary.

Trial Completion Time

The repeated measures ANOVA for the log of trial completion time indicates significant main effects for display mode $\mathrm{F}(1,7) = 152.83$, $p < 0.0001$, target position $\mathrm{F}_{\mathrm{HF}}(3.7,25.9) = 11.32$, $p < 0.0001$, and block $\mathrm{F}_{\mathrm{HF}}(3,21) = 7.97$, $p = 0.0010$. Significant interaction effects were found for display mode $\times$ target position $\mathrm{F}_{\mathrm{HF}}(4.8,33.6) = 12.47$, $p < 0.0001$ and display mode $\times$ head-tracking mode $\times$ block $\mathrm{F}_{\mathrm{HF}}(3,21) = 8.39$, $p = 0.0007$. The remaining independent variable of head-tracking mode and all other interactions were not statistically

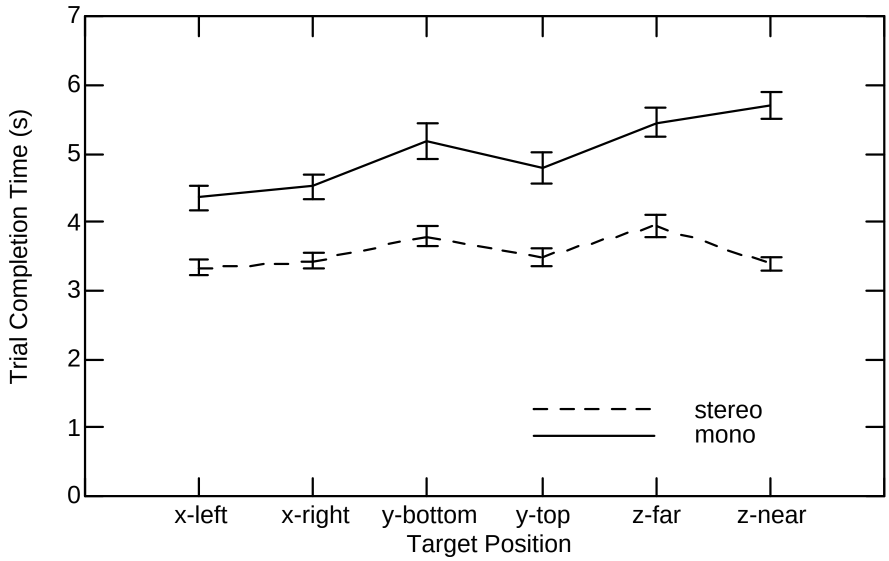


Figure 4.4: Mean of trial completion time in seconds for each of the target positions in both display modes.

significant ($p > 0.05$). Full results of the ANOVA can be seen in Table 4.18 at the end of this chapter.

The significant main effects for display mode and target position must be interpreted together in light of the significant display mode $\times$ target position interaction. Figure 4.4 shows the means for each display mode across the target positions. Post hoc analysis using Tukey's WSD procedure for the interaction reveals that changes in the pattern of trial com-

Target Position	**Mono**	**Stereo**
x-left	4.35	3.33
x-right	4.53	3.43
y-bottom	5.19	3.80
y-top	4.79	3.49
z-far	5.46	3.95
z-near	5.70	3.39
Overall	5.00	3.57

Table 4.1: Mean trial completion times in seconds for each target position in the two display modes.

	x-left	x-right	y-bottom	t-top	z-far	z-near
x-left			B		B	M
x-right			M		B	M
y-bottom						B
y-top					B	M
z-far						
z-near					S	

Table 4.2: Pairwise contrasts using Tukey's WSD between target positions in the monoscopic and stereoscopic display modes. B indicates that the target positions are significantly different in both display modes, M indicates that the target positions were significantly different only in the monoscopic display mode and S indicates a significant difference between target positions only in the stereoscopic display mode.

pletion time for two of the targets is responsible. The z-near target changes from having the slowest performance in the monoscopic condition to having the second fastest performance in the stereoscopic condition, and the bottom target changes from being significantly slower than the target to the right to no longer being different from that target. Table 4.1 shows the cell means for each target position in both display modes. Table 4.2 shows the details of all the pairwise contrasts between target position for the stereoscopic and monoscopic conditions.

The significant main effect for block must be interpreted in light of the significant display mode $\times$ head-tracking $\times$ block interaction. Figure 4.5 shows the mean trial completion time for each display and head-tracking mode combination across the four session blocks. In the monoscopic display mode, head-tracking appears to provide a slight benefit over the fixed viewpoint in the early blocks that disappears in the later blocks. In the stereoscopic display mode, head-tracking appears to hinder the early blocks when compared with the fixed viewpoint condition, but this difference is minimized in the later blocks.

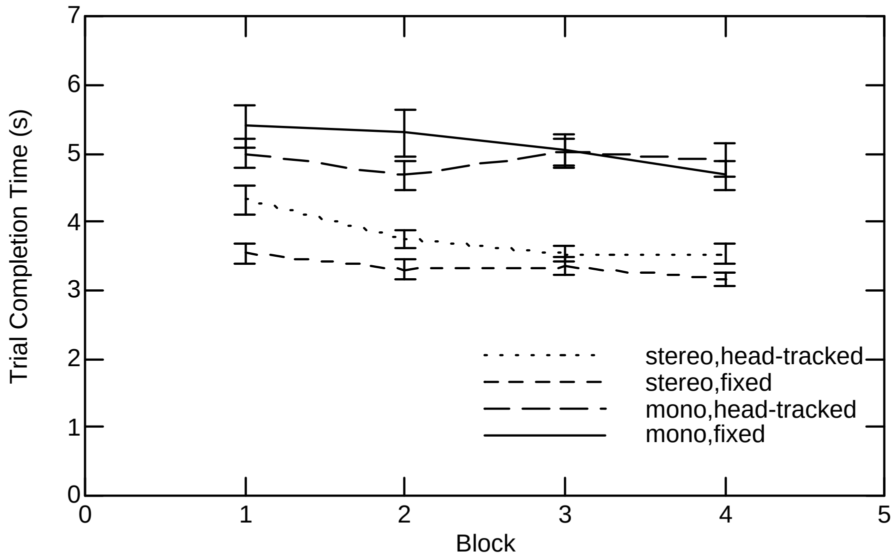


Figure 4.5: Plot of the mean trial completion time (seconds) vs. block for the display and head-tracking modes.

Final Position Error

The repeated measures ANOVA for log of final position error found a significant main effect for target position $F_{HF}(4.0,28.0) = 5.27$, $p = 0.0028$. A significant interaction effect was found for head-tracking mode $\times$ block $F_{HF}(3.0,20.8) = 6.03$, $p = 0.0041$. Table 4.19 at the end of this chapter provides the full results of the ANOVA for final position error.

Table 4.3 shows the final position error for each of the target positions. It is fairly clear that performance in the z-far target position is different from the other target positions.

Position	**Error**
x-left	0.34
x-right	0.35
y-bottom	0.33
y-top	0.35
z-far	0.43
z-near	0.35

Table 4.3: Mean of final position error in centimetres for each of the target positions.

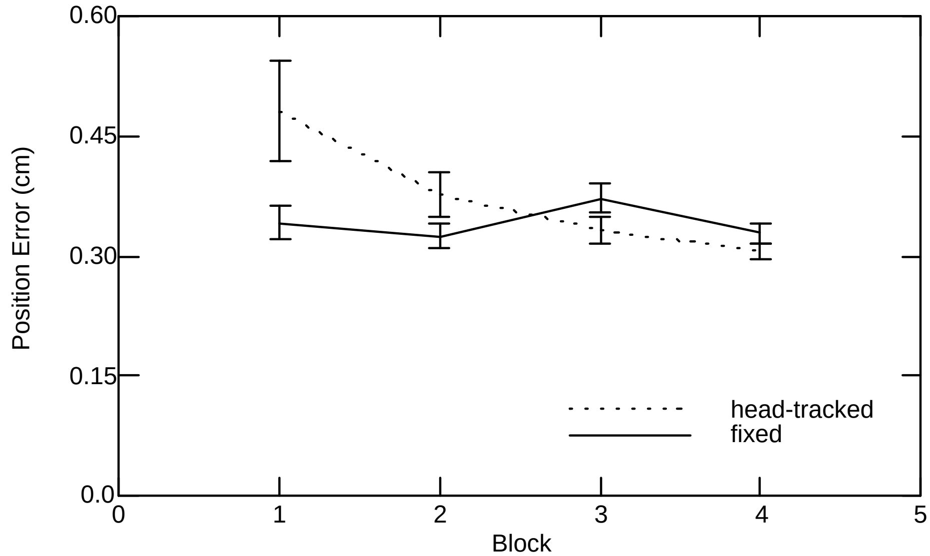


Figure 4.6: Plot of the mean final position error in centimetres vs. block for both head-tracking modes.

Table 4.4 provides the mean final position error for each of the blocks in both head tracking modes and Figure 4.6 illustrates the relationship between head-tracking mode and block. Final position error remains relatively constant across blocks in the non-head-tracked conditions. In the head-tracked conditions final position error starts out higher in the initial phases and then drops to the same level as the non-head-tracked condition in the later blocks.

Final Position Axis Error

A more detailed analysis of the positioning error can be obtained by considering the axis as an additional factor and performing a five factor ANOVA on the data. This ANOVA re-

Head-tracking mode	B 1	B 2	B 3	B 4
fixed	0.34	0.32	0.37	0.33
head-tracked	0.48	0.38	0.33	0.31

Table 4.4: Mean of final position error for both head-tracking modes in each of the four session blocks.

veals significant main effects for target position $F_{HF}(4.3,30.1) = 7.90$, $p = 0.0001$ and for axis $F_{HF}(2,14) = 27.99$, $p < 0.0001$. Significant two-way interactions were found for head-tracking mode $\times$ block $F_{HF}(3.0,21.0) = 4.75$, $p = 0.0111$, display $\times$ axis $F_{HF}(2.0,14.0) = 10.58$, $p = 0.0016$, head-tracking mode $\times$ axis $F_{HF}(1.8,12.5) = 4.25$, $p = 0.0430$, target position $\times$ axis $F_{HF}(6.0,42.0) = 3.88$, $p = 0.0036$ and block $\times$ axis $F_{HF}(6.0,42.0) = 3.35$, $p = 0.0086$. A single significant three-way interaction of head-tracking mode $\times$ target position $\times$ axis $F_{HF}(10.0,70.0) = 2.70$, $p = 0.0074$ was also found. Full results of the ANOVA can be found in Table 4.20 at the end of this chapter.

Interpretation of these significant effects can be divided into the following four groups where each group accounts for any significant effects wholly contained within that group.

- head-tracking mode $\times$ target position $\times$ axis
- head-tracking mode $\times$ block
- display mode $\times$ axis
- block $\times$ axis

Table 4.5 shows the mean final position axis error broken down by head-tracking mode and block. This is different from the final position error in that final position error represents the Euclidean distance between the pointer and the target, whereas the mean of the final position axis errors is the simple arithmetic mean of the final position axis error components. The mean of the final position axis errors remain fairly stable across the four blocks in the fixed viewpoint condition. In the head-tracked condition the mean of the final position axis error starts out higher in the first block and drops below the fixed viewpoint condition in the third and fourth blocks.

Table 4.6 shows the final position axis component errors in the two display modes. The source of the display mode $\times$ axis interaction is a result of the drop in final position axis

Block	Fixed	Head-tracked
1	0.17	0.21
2	0.16	0.17
3	0.18	0.16
4	0.16	0.15

Table 4.5: Mean component axis error in centimetres for the two head-tracking modes across the blocks of the experiment averaged over the other conditions.

error along the Z-axis between the monoscopic and stereoscopic display modes.

Axis	Mono	Stereo
X	0.13	0.16
Y	0.11	0.11
Z	0.32	0.20

Table 4.6: Mean error in cm along the X, Y and Z axes in the two display modes averaged over the other conditions.

Table 4.7 shows the axis component errors in each of the blocks. The source of the block $\times$ axis interaction stands out clearly. The Z-axis error improves from the first to the fourth block while the X- and Y-axis errors hardly vary at all.

Axis	B 1	B 2	B 3	B 4
X	0.14	0.14	0.15	0.14
Y	0.11	0.11	0.11	0.11
Z	0.32	0.25	0.25	0.21

Table 4.7: Mean error in centimetres along the X, Y and Z axes across the four blocks of the experiment averaged over the other conditions.

Finally, Table 4.8 shows the mean error along the X, Y and Z axes broken down by target position and head-tracking mode. The clear overall trend is for the Z-axis error to be larger than the X- and Y-axis errors. The source of the interaction effect is the large change in Z-axis error for the z-near and z-far targets between the non head-tracked and head-tracked modes.

	Fixed			Head-Tracked		
Pos.	**X**	**Y**	**Z**	**X**	**Y**	**Z**
left	0.19	0.10	0.20	0.15	0.10	0.24
right	0.15	0.10	0.25	0.12	0.12	0.27
top	0.14	0.13	0.22	0.11	0.11	0.25
bottom	0.16	0.11	0.24	0.12	0.11	0.26
z-near	0.12	0.09	**0.19**	0.12	0.11	**0.33**
far	0.17	0.13	**0.29**	0.14	0.12	**0.36**

Table 4.8: Mean error in centimetres along the X, Y and Z axes for each of the target positions in both head-tracking modes.

Path Length

A repeated measures ANOVA for the log of the path length found significant main effects for display mode $F(1,7) = 44.17$, $p = 0.0003$ and target position $F_{HF}(4.0,27.7) = 5.62$, $p = 0.0020$. A single significant two-way display mode $\times$ block interaction $F_{HF}(3.0,21.0) = 6.28$, $p = 0.0033$ was also found. Full results of the ANOVA are given in Table 4.21 at the end of the chapter.

Position	Length
x-left	16.78
x-right	18.65
y-bottom	19.37
y-top	19.95
z-far	17.89
z-near	17.38

Table 4.9: Mean path length in centimetres for each of the target positions.

Block	Mono	Stereo
1	19.92	17.88
2	19.34	17.08
3	20.47	16.33
4	19.73	15.93

Table 4.10: Mean path length in centimetres during each block in the two display modes.

The average path length in the stereoscopic display mode was 16.81cm, significantly shorter than the average path length of 19.86cm in the monoscopic mode. Table 4.9 shows the average path length for each of the target positions. The path length for the target to the left is shortest, the paths to the top and bottom targets are longest and the paths to the z-near and z-far targets fall in the middle. Table 4.10 shows the mean path length for each of the blocks of the experiment in the two display modes. The interaction effect is observable in the progressive reduction in path length for the stereoscopic display mode, while the path length remains relatively unchanged in the monoscopic display mode.

RMS Error

The ANOVA for RMS error of the the movement of the pointer revealed significant main effects for display mode $F(1,7) = 16.73$, $p = 0.0046$ and target position $F_{HF}(4.4,30.8) = 10.39$, $p < 0.0001$ as well as a single two-way display mode $\times$ position interaction $F_{HF}(3.6,25.2) = 3.22$, $p = 0.0328$. The results of the ANOVA for RMS error can be found in Table 4.22 at the end of the chapter.

Figure 4.7 illustrates the effects of target position and display mode upon RMS error. Overall the stereoscopic display mode has less error than the monoscopic display mode. The interaction effect is visible as the difference between the gap for the z-near and z-far target positions vs. the gap for the other target positions.

RMS Axis Error

As for final position error, the RMS error along the entire movement can be examined in terms of X, Y and Z components. Once again a five factor ANOVA was performed as for position error. Full results of this ANOVA are given in Table 4.23. Significant main effects were found for display mode $F(1,7) = 7.53$, $p = 0.0287$, target position $F_{HF}(4.6,32.2) =$

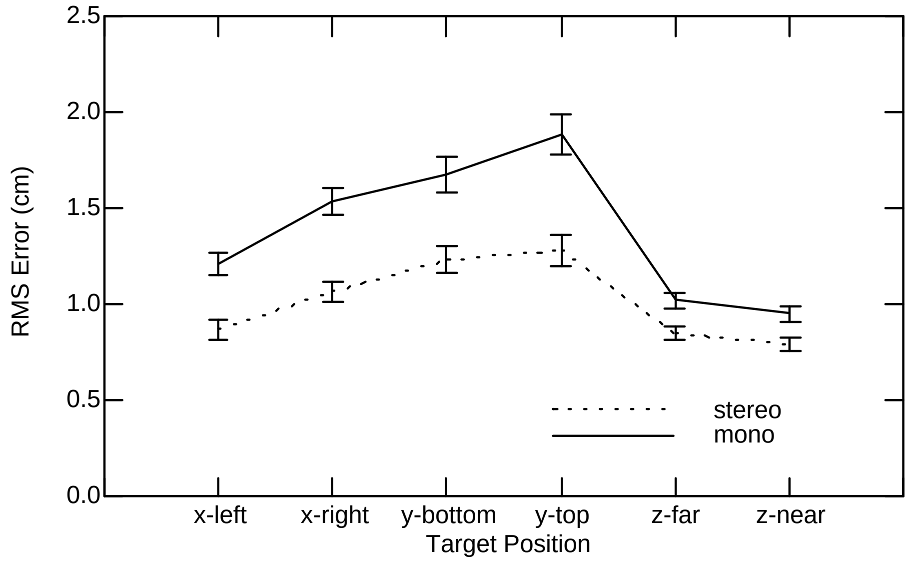


Figure 4.7: Plot of RMS error in centimetres for each of the target positions in both display modes.

10.78, $p < 0.0001$ and axis $F_{HF}(1.8,12.6) = 47.69$, $p < 0.0001$. Four significant two-way interactions were found for display mode $\times$ block $F_{HF}(2.1,14.7) = 3.78$, $p = 0.0457$, display mode $\times$ axis $F_{HF}(1.7,12.0) = 4.63$, $p = 0.0365$, target position $\times$ axis $F_{HF}(3.5,24.5) = 33.86$, $p < 0.0001$ and block $\times$ axis $F_{HF}(5.2,36.1) = 5.08$, $p = 0.0012$. One significant four-way interaction was found for display mode $\times$ head-tracking mode $\times$ block $\times$ axis $F_{HF}(6,42) = 2.62$, $p = 0.0300$.

Pos.	**X**	**Y**	**Z**
left	0.12	0.40	0.89
right	0.38	0.42	1.09
top	0.42	0.14	1.33
bottom	0.42	0.35	1.43
z-near	0.44	0.56	0.36
far	0.45	0.50	0.52

Table 4.11: Mean of RMS error in centimetres along the X, Y and Z axes across all target positions.

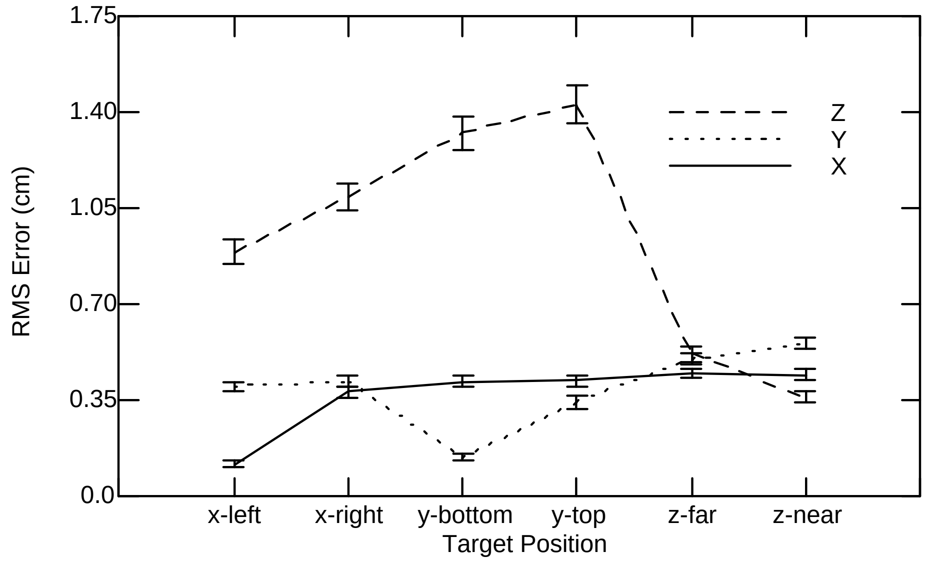


Figure 4.8: Plot of the mean RMS axis error in centimetres vs. target position for each of the axes.

Due to the significant target position $\times$ axis interaction, the effect of axis must be considered in conjunction with the effect for target position. Table 4.11 gives the RMS error along the X, Y and Z axes across all target positions and Figure 4.8 illustrates the relationship between these variables. The significant interaction is the product of three factors. First, the RMS error along the X-axis is relatively stable except for the left target position where it is much lower than the other cases. Second, the Y RMS error is relatively stable except for the bottom target position where it is much lower than the other positions. Third, the RMS error along the Z-axis is lower in the z-near and z-far target positions than in the other target positions.

Display mode	**X**	**Y**	**Z**
mono	0.41	0.41	1.12
stereo	0.33	0.38	0.75

Table 4.12: Mean of RMS error in centimetres along the X, Y and Z axes for both display modes.

The effects of display mode and axis are considered together because of the significant interaction between these two factors. Table 4.12 shows the mean RMS error between the stereoscopic and monoscopic display modes across the X, Y and Z axes. The interaction lies in the large effect that display mode has upon reducing the RMS error along the Z-axis in conjunction with the minimal effect in reducing the RMS error along the X and Y axes.

Questionnaires

At the end of each session subjects were required to answer four questions: two about task difficulty and two about the target positions. The two questions dealing with task difficulty were answered on a 5 point scale where 1 is "easy" and 5 is "hard." The first question asked subjects to rate the difficulty of determining the target position and the second question asked subjects to rate the difficulty of matching the target position. The two questions dealing with target positions asked subjects to indicate whether they found any target positions easier or harder to match and if so, which ones. A copy of the post-session questionnaire can be found in Appendix A.3.

At the end of the experiment subjects were asked to score all the display and head-tracking mode combinations according to task difficulty and ease of use. Subjects were also asked to rank the display mode and head-tracking mode combinations in order of their preference for the mode, the ease of use of the mode and their performance in the mode. A copy of the post-experiment questionnaire can be found in Appendix A.7.

In both the post-session and post-experiment questions, subjects clearly found the stereoscopic condition to be easier than the monoscopic condition, but did not indicate much difference between the head-tracked and non-head-tracked conditions. A summary of subjects' responses to the questions about the display mode and head-tracking combination is given in Table 4.13. Subject's responses regarding which targets were easier and which were

	Fixed		Head-tracked			
Question	**Mono**	**Stereo**	**Mono**	**Stereo**	**Friedman**	**p**
Determine position	3.0	1.9	3.1	2.0	7.09	0.07
Match position	3.1	1.3	2.5	1.8	9.86	0.02
Task difficulty	3.3	1.8	3.5	1.9	8.44	0.04
Mode usefulness	3.8	1.5	3.6	1.9	18.71	< 0.001

Table 4.13: Means of subjects' responses to the subjective questions regarding display mode and head tracking mode conditions. The first two questions were answered immediately after a particular condition, and the last two questions were answered after all conditions had been completed. All questions were answered on a scale of 1 to 5 where 1 indicated easy or useful and 5 indicated hard or not useful.

	Easier			
	Fixed		**Head-tracked**	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**
None	2	1	1	0
Z-Near	3	5	2	5
Far	0	0	0	0
XY axes	2	3	3	2
X-axis	2	0	1	0
Z-axis	0	0	0	1
Other			left	

Table 4.14: Summary of subjects' responses indicating which targets were easier across the head-tracking and display mode conditions. Column totals do not always add to eight because some subjects indicated multiple target positions. The row for other indicates a response given only once.

harder were coded by the experimenter into one of the target positions, target axes or combinations. In broad terms, subjects generally found the z-near targets and the X and Y axes to be easier and the z-far targets and the Z-axis to be harder across all display mode and head-tracking conditions. Subjects' coded responses are summarized in Table 4.14 and Table 4.15.

Figure 4.9 provides summaries of the rankings indicated by subjects for the various visual feedback modes according to their preference, ease of use and performance.

	Harder			
	Fixed		**Head-tracked**	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**
None	2	3	0	2
Z-Near	0	0	1	0
Far	5	5	4	4
XY axes	0	0	0	0
X-axis	1	0	0	1
Z-axis	0	0	3	1
Other		Top		Y-axis

Table 4.15: Summary of subjects' responses indicating which targets were harder across the head-tracking and display mode conditions. Column totals do not always add to eight because some subjects indicated multiple target positions. The row for other indicates a response given only once.

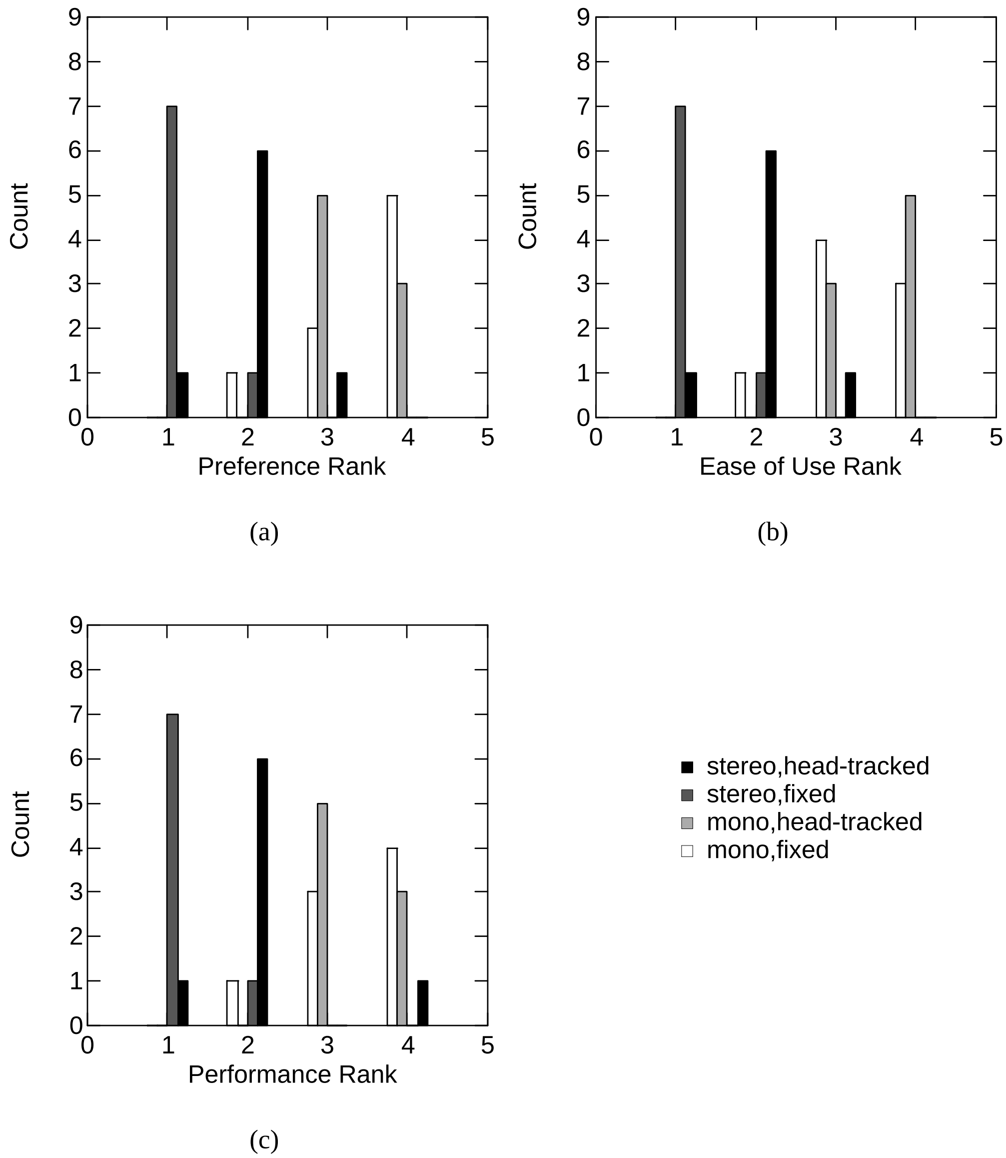


Figure 4.9: Rankings of subjects' preference, perceived ease of use and perceived performance for each of the viewing conditions. The height of a bar indicates the number of times a particular rank was assigned by a subject.

4.8 Discussion

The Results section showed that many of the independent variables had significant effects upon the dependent measures. These effects, their potential causes and their relationship to other work in the literature are explored in this section.

4.8.1 Speed

Stereoscopic viewing was faster than monoscopic viewing

The results for trial completion time clearly show a significant overall effect for display mode, with stereoscopic viewing superior to monoscopic viewing across all other conditions.

Trial completion time varied across target positions

The results also show a significant effect for target position in conjunction with a significant target position $\times$ display mode interaction indicating that the pattern of differences across target positions varies between display modes. The most noteworthy variation is the reversal for trial completion time for the z-near target position, from being slowest in the monoscopic condition to being comparable to the X-axis in the stereoscopic condition. This finding is somewhat surprising in light of work by Zhai [110], however, his studies did not consider the direction along each axis separately. Table 4.16 shows the normalized performance for each target position in the stereoscopic and monoscopic conditions.

Takemura et al. [93] had subjects perform a volume location task (1cm and 2cm cubes) using a field-sequential stereo display and a large number of target locations. They found no significant differences between any of the target positions. It may be that the reduced accuracy requirement of the volume location task resulted in a reduced sensitivity to position differences.

Ware and Balakrishnan [102] had subjects perform a volume location task (1cm cube) with targets to the left or behind the starting location of the pointer, using a head-tracked field-sequential stereo display. They found performance along the Z-axis to be 10% slower than performance along the X-axis.

Head-tracking had almost no influence

The results for trial completion time show no significant main effect for head-tracking. The only place where head-tracking had any effect upon trial completion time is in the head-tracking $\times$ display mode $\times$ block interaction. Upon inspection of the results it appears that in the stereoscopic display mode head-tracking appears to have degraded performance in three of the four blocks for this experiment condition.

4.8.2 Accuracy

Far target position had the largest position error

The results for final position error showed a significantly larger error for the z-far target position as compared with the other target positions. This may be a result of the difference in

Target position	**Mono**	**Stereo**
x-left	1.22	0.93
x-right	1.27	0.96
y-bottom	1.45	1.06
y-top	1.34	0.98
z-far	1.53	1.11
z-near	1.60	0.95
Overall	1.40	1.00

Table 4.16: Normalized trial completion time for each target position in the two display modes. Times have been normalized by dividing by the overall mean trial completion time for the stereo condition.

visual size of the target in the different positions due to perspective projection. Relative to the targets on the X and Y axes, the z-near target is bigger on the display and the z-far target is smaller on the display. In addition, as objects get farther away from the viewer, each individual pixel on the surface of the screen, represents a progressively larger area on the surface of the object. Subjects may have employed a pragmatic strategy of attempting to match the target position within a certain number of pixels. This would lead directly to an increase in error as measured in the physical workspace as any given pixel error represents a larger physical error as the target gets farther from the subject's viewpoint.

Head-tracking increased position error in early blocks

The head-tracking $\times$ block interaction shows that during the early blocks head-tracking significantly increased the final position error, but this effect disappears in the later blocks. There are a few possible explanations for this result. Subjects may have tired of the head-tracker after the first block or two and stopped making use of it. Alternately, subjects may have learned how to make use of the visual feedback afforded to them in the head-tracked mode and were able to reduce their error over time. However, as the improvement only allowed subjects to achieve the same level of performance as in the non head-tracked mode, the first explanation seems more plausible.

Considerations for head-tracking

Head-tracking used in conjunction with a stationary monitor is sometimes referred to as "Fish-tank Virtual Reality." An inherent property of this viewing model is that parallax shifts are induced in any points not located on the Z = 0 plane. You can simulate this effect yourself by looking at a distant object through a window and touching the projection of the object at the window's surface. If you then move your head left or right you will notice that the projection of the object does not stay in the same place, but moves with respect to your finger.

Sollenberger & Milgram [89] studied the ability of subjects to identify the visual network to which a specified point belonged in monoscopic and stereoscopic displays, with and without rotation of the network. They found that stereoscopic viewing and rotation improved performance both individually and in combination.

Arthur et al. [6] studied the same task, using monoscopic and stereoscopic displays with and without head-tracking. The motion parallax induced via head-tracking is similar to the effect of rotation. They found that the head-coupled stereoscopic mode resulted in the shortest response time and lowest error rate.

The lack of any significant effect (other than the interaction discussed above) involving head tracking is interesting. In the study by Arthur et al. [6] head movement was the only body movement required in the interaction task. I was unable to find many other studies that considered head-tracking within an interactive task other than the study by Ware and Balakrishnan [102]. Unfortunately, their study was primarily concerned with lag and frame rate and did not contrast fixed versus head-tracked viewpoints.

In this experiment, subjects had to co-ordinate any head movement used to obtain better visual information about the 3D environment with hand and arm movements used to perform the primary point location task. It may be that head movement was used in the initial phases of a monoscopic session to improve the discovery of target positions, but once those positions were known, head movement was no longer used because it interfered with the positioning task. A study where target position is allowed to vary more widely might be able to determine if this is indeed the case.

It is also possible that the mechanical nature of the ADL-1 head tracker discouraged head motion. The ADL-1 moves freely within its working volume, but the apparatus adds a small amount of inertia to head movements. While the ADL-1 is much less massive than most immersive virtual reality helmets, a device with less inertia might better afford head motion. If it is the case that the ADL-1 was somewhat restrictive of head motion, then it appears that the task demands were not sufficient to compel subjects to overcome this restriction. The ADL-1 is the same device used in the study by Arthur et al. [6] where a benefit was found for head-tracking.

Position error was largest along the Z-axis

The results of the analysis of final position axis error paints a complex picture. It is clear and not unexpected that the amount of error varies across the axes. Positions along the X and Y axes are fairly easy to visualize, but positions along the Z-axis are much more difficult. For both display modes the largest error is found in the Z-axis. Stereoscopic viewing is most influential in reducing the amount of Z-axis error. This is consistent with the findings of other researchers such as Zhai[111] and Massimino [65].

X-axis error was slightly higher in the stereoscopic display mode

While the error along the Y-axis is stable between display modes, the error along the X-axis is increased in the stereoscopic condition. This contrasts with the findings of Zhai [111] where error along the Y-axis was slightly worse than error along the X-axis. It should be noted that Zhai's study involved a different task (pursuit tracking), different input controllers (isometric rate and elastic rate), and only investigated a single display mode (stereoscopic). Furthermore, the stereoscopic display technique results in a loss of visual resolution along the Y-axis (but not along the X-axis) which could interfere with task performance. Additional work is clearly required in this area to shed more light on this finding.

Z-axis error was reduced across blocks

The significant effects for display, axis and display $\times$ axis show that the stereoscopic display mode significantly reduces the Z-axis error, but has no effect upon the X- and Y-axis errors. The significant block, axis and block $\times$ axis interaction show that over blocks subjects' Z-axis error was reduced, but X- and Y-axis error remained unchanged. This may indicate that subjects are operating at close to their optimal ability for the X and Y axes from the outset of the experiment. Further studies should investigate means of reducing Z-axis error as quickly as possible.

Head-tracking interfered with Z-axis position error reduction

The significant display mode $\times$ block and display mode $\times$ head-tracking mode $\times$ block $\times$ axis interaction shows that in both head-tracked and fixed viewpoint stereo modes, the Z-axis position error is reduced from the first block to the last block, although the Z-axis position error is consistently higher in the head-tracked mode. The two monoscopic conditions do not exhibit this steady reduction, but instead exhibit much more variability across blocks. This may indicate that subjects were unable to make effective use of the head-tracker in the monoscopic mode. Even though the head-tracker was not beneficial in the stereoscopic mode it did not interfere with subjects' ability to steadily reduce their Z-axis error.

Head-tracking increased Z-axis error for Z-axis targets

The main effects for target position and axis in conjunction with the two-way target position $\times$ axis and head-tracking mode $\times$ axis interactions and three-way target position $\times$ head-tracking mode $\times$ axis interaction must be considered together. In doing so it becomes apparent that across both display modes, the largest detrimental effect of head-tracking was upon the Z-axis error for the z-near and z-far target positions in the head-tracked conditions.

Z-axis targets had the lowest RMS error

Display mode and target position have a joint effect upon RMS error. There is a clear difference in the RMS error between display modes for the target positions along either the X or Y axes. This difference becomes non-significant in the two Z-axis target positions. This result and the result for path length reveal that there are some elements of the task that benefit from the Z-axis rather than being hindered by it. It is important to note that in this task, RMS error is being computed against the straight line path from the start position to the target position. This causes the metric itself to be somewhat forgiving along the axis upon which the path is located. So in the case of straight lines, path length and path error can be reduced by arranging targets along the Z-axis if one is willing to accept more end point error.

RMS error revealed weakness for certain target positions

To obtain a detailed view of how the actual movement varied from the optimal movement, RMS error along the X, Y and Z axes was investigated. There were many significant effects and interactions, making it somewhat difficult to draw conclusions. However, a few things did stand out.

The effects for axis, position and axis $\times$ position upon RMS error indicate that the left target had the smallest X-axis error, the top position had the smallest Y-axis error and the z-near and z-far targets had the smallest Z-axis errors. However, this is not too surprising given the nature of the task and how X, Y and Z errors with respect to the optimal path are computed. What is surprising is that the X-axis error for the right target and the Y-axis error for the bottom target are comparable to the errors along other axes. Rather than pointing out the advantage of certain positions, this result reinforces the weakness of certain positions, in particular the right target and the bottom target. The two Y-axis targets exhibit the largest Z-axis errors indicating that these positions seem to make Z-axis control more difficult.

Pos.	**Time mono**	**Time stereo**	**Path**
x-left	1	1	1
x-right	2	3	4
y-bottom	4	5	5
y-top	3	4	6
z-far	5	6	3
z-near	6	2	2

Table 4.17: Rank ordering of trial completion time (in both display modes) and path length.

4.8.3 Conciseness

Stereoscopic viewing produced a reduction in path length over blocks

In terms of the overall path length, stereoscopic viewing was significantly better than monoscopic viewing. When the display mode $\times$ block interaction is examined it becomes evident that there is a block to block improvement in path length for stereo viewing, but monoscopic viewing is mostly unchanged across the blocks.

Z-axis targets had short path lengths but long trial completion times

The effect of target position upon path length is interesting, especially when compared with the effects of target position upon trial completion time. The movement path to the left target was shortest and is consistent with trial completion time in that it was also the fastest. However, the z-near and z-far targets have relatively short movement paths in contrast to their trial completion times that were generally high. The longest paths belong to the top and bottom target positions. Table 4.17 compares the rankings of trial completion time and path length for the target positions.

Considerations for path length

Zhai [112] has recently suggested the use of the ratio of total movement to shortest possible movement as a measure of the coordination of the movement. In this sense the most coordinated movement will be the one with values approaching 1.0 and less coordinated movements will have larger values. The basic idea is that simultaneous movement along all three axes will minimize the total path length. Non-simultaneous movement along the axes will increase the total path length. Unfortunately, this may not provide sufficient insight into where any strengths or weaknesses lie.

4.8.4 Felicity

The subjective responses made by subjects regarding the effectiveness of the display modes and head-tracking modes is clearly in line with their measured performance. Subjects found the stereoscopic condition easier and more useful than the monoscopic condition, while they found little difference between the head-tracked and non-head-tracked conditions.

Subjects felt the near target was easier even when measured performance indicated it was harder

When required to specify which target positions were easier or harder to match, subject's responses were generally consistent with measured performance, but there were a few surprising exceptions. Some subject's found the z-near target to be easier in both the stereoscopic and monoscopic display modes and seldom singled it out as being harder. This is consistent with their performance in the stereoscopic condition, but contrasts sharply with performance in the monoscopic condition where the z-near target was the most difficult.

The match or mismatch between actual measured performance and subjective perceptions is important. When subjects perceive a task to be more difficult they may avoid the task, or they may expend more effort on the task if they are trying to meet some speed or accuracy requirement. If the task is considered to be less difficult, subjects might show a preference for the task, or they might feel that less attention is required to meet a speed or accuracy requirement. In interactive tasks it may be more important to focus on potential mismatches between performance and perception than simply trying to improve performance. This experiment has shown that a potential mismatch between performance and perception exists in a 3D point location task. Further investigation is required to determine whether such a mismatch is significant and if it has an adverse effect upon performance.

4.9 Conclusions

The purpose of this experiment was to determine what effect, if any, stereoscopic viewing, head-tracking and target position would have on performance in a point location task. As expected we found that stereoscopic viewing is superior to monoscopic viewing. No significant main effect was found for head-tracking, but there were several situations in which head-tracking interacted with some other variable. These cases tended to indicate that head-tracking resulted in somewhat poorer performance, generally through increased error or reduced skill improvement over time. These findings for head-tracking stand in contrast to other studies that made use of head-tracking where it was found to provide significant benefit to subjects. This indicates that the benefit of head-tracking should not be considered a foregone conclusion and raises many opportunities for further study. Some of the possible questions include: Was head-tracking detrimental because of the mechanical nature of the device used? Were there insufficient cues in the environment from which to obtain depth information via motion parallax?

Performance was found to vary across target positions both in terms of the time to complete the trial, final position error, final position axis error, path length, total RMS error and axis RMS error. In fact, target position was the only independent variable that had a significant effect upon every dependent measure that was used. It should be clear that the position of the target in the workspace had a significant influence upon subjects' ability to perform the task. When all the dependent measures are considered, it is clear that no single target position offered the best or worst of everything. Nonetheless, the target to the left had good performance along many of the measures used, and the bottom target had poor performance along many of the measures. Interestingly, the z-near and far targets had shorter path lengths and tended to minimize the RMS error across all three axes.

The difference in error across axes is not in itself that surprising or interesting. However, when combined with other independent variables, it is possible to see how different target positions or display modes influence error levels. One of the most interesting results is the detrimental effect that head-tracking has, either increasing Z-axis error in some cases, or inhibiting improvement in others.

Finally, we found that while subjective perceptions of the effectiveness of the different display conditions matched objective performance measures, perceptions of the ease and/or difficulty associated with different target positions did not always match objective performance measures.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	20.069	1	20.069	152.83	< 0.0001		
Error	0.919	7	0.131				
head-track	0.419	1	0.419	0.33	0.5820		
Error	8.801	7	1.257				
d×h	0.911	1	0.911	2.59	0.1516		
Error	2.463	7	0.352				
position	3.613	5	0.723	11.32	< 0.0001	< 0.0001	0.74
Error	2.235	35	0.064				
d×p	1.332	5	0.266	12.47	< 0.0001	< 0.0001	0.96
Error	0.747	35	0.021				
h×p	0.019	5	0.004	0.18	0.9689	0.9534	0.84
Error	0.745	35	0.021				
d×h×p	0.117	5	0.023	1.01	0.4262	0.4156	0.73
Error	0.811	35	0.023				
block	1.299	3	0.433	7.97	0.0010	0.0010	1.00
Error	1.141	21	0.054				
d×b	0.219	3	0.073	2.52	0.0855	0.0938	0.90
Error	0.608	21	0.029				
h×b	0.075	3	0.025	0.74	0.5419	0.5419	1.00
Error	0.716	21	0.034				
d×h×b	0.366	3	0.122	8.39	0.0007	0.0007	1.00
Error	0.305	21	0.015				
p×b	0.272	15	0.018	1.40	0.1609	0.1627	0.98
Error	1.358	105	0.013				
d×p×b	0.214	15	0.014	1.19	0.2935	0.3031	0.84
Error	1.264	105	0.012				
h×p×b	0.185	15	0.012	0.85	0.6215	0.6067	0.86
Error	1.526	105	0.015				
d×h×p×b	0.229	15	0.015	0.99	0.4762	0.4687	0.77
Error	1.625	105	0.015				
Total	84.468						

Table 4.18: Experiment 1 - Repeated measures ANOVA table for log trial completion time.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	2.341	1	2.341	1.22	0.3063		
Error	13.456	7	1.922				
head-track	0.156	1	0.156	0.09	0.7709		
Error	11.919	7	1.703				
d$\times$h	0.059	1	0.059	0.06	0.8166		
Error	7.154	7	1.022				
position	5.579	5	1.116	5.27	0.0010	0.0028	0.80
Error	7.415	35	0.212				
d$\times$p	0.511	5	0.102	0.68	0.6392	0.6305	0.94
Error	5.231	35	0.149				
h$\times$p	0.786	5	0.157	2.21	0.0753	0.0808	0.93
Error	2.488	35	0.071				
d$\times$h$\times$p	0.367	5	0.073	0.57	0.7235	0.7235	1.00
Error	4.522	35	0.129				
block	1.069	3	0.356	2.06	0.1357	0.1432	0.91
Error	3.626	21	0.173				
d$\times$b	0.166	3	0.055	0.29	0.8286	0.8286	1.00
Error	3.949	21	0.188				
h$\times$b	2.633	3	0.878	6.03	0.0040	0.0041	0.99
Error	3.055	21	0.145				
d$\times$h$\times$b	0.473	3	0.158	0.95	0.4336	0.4109	0.68
Error	3.477	21	0.166				
p$\times$b	1.358	15	0.091	1.11	0.3554	0.3554	1.00
Error	8.551	105	0.081				
d$\times$p$\times$b	1.042	15	0.069	0.82	0.6503	0.6069	0.66
Error	8.864	105	0.084				
h$\times$p$\times$b	1.447	15	0.096	1.32	0.2020	0.2020	1.00
Error	7.661	105	0.073				
d$\times$h$\times$p$\times$b	1.669	15	0.111	1.56	0.0977	0.1274	0.74
Error	7.485	105	0.071				
Total	142.679						

Table 4.19: Experiment 1 – Repeated measures ANOVA for log of final position error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.174	1	0.174	0.04	0.8504		
Error	31.735	7	4.534				
head-track	0.018	1	0.018	0.01	0.9398		
Error	20.422	7	2.917				
d×h	0.250	1	0.250	0.21	0.6609		
Error	8.334	7	1.191				
position	18.112	5	3.622	7.90	$<$ 0.0001	0.0001	0.86
Error	16.043	35	0.458				
d×p	1.381	5	0.276	1.08	0.3889	0.3889	1.00
Error	8.958	35	0.256				
h×p	1.987	5	0.397	2.22	0.0744	0.0899	0.83
Error	6.273	35	0.179				
d×h×p	0.580	5	0.116	0.35	0.8780	0.8780	1.00
Error	11.553	35	0.330				
block	0.714	3	0.238	0.59	0.6307	0.6280	0.98
Error	8.525	21	0.406				
d×b	0.142	3	0.047	0.13	0.9402	0.9402	1.00
Error	7.576	21	0.361				
h×b	3.397	3	1.132	4.75	0.0111	0.0111	1.00
Error	5.002	21	0.238				
d×h×b	1.362	3	0.454	1.77	0.1839	0.1945	0.84
Error	5.390	21	0.257				
p×b	1.578	15	0.105	0.62	0.8517	0.8517	1.00
Error	17.777	105	0.169				
d×p×b	2.120	15	0.141	0.80	0.6730	0.6435	0.77
Error	18.501	105	0.176				
h×p×b	2.779	15	0.185	1.25	0.2461	0.2461	1.00
Error	15.536	105	0.148				
d×h×p×b	3.383	15	0.226	1.66	0.0703	0.0829	0.87
Error	14.253	105	0.136				

Table 4.20: Experiment 1 – Repeated measures ANOVA for log of final position error along axes (continued on next page).

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
axis	211.819	2	105.909	27.99	< 0.0001	< 0.0001	1.00
Error	52.977	14	3.784				
d×a	25.900	2	12.950	10.58	0.0016	0.0016	1.00
Error	17.130	14	1.224				
h×a	11.740	2	5.870	4.25	0.0361	0.0430	0.89
Error	19.337	14	1.381				
d×h×a	0.910	2	0.455	0.66	0.5305	0.5305	1.00
Error	9.602	14	0.686				
p×a	11.918	10	1.192	3.88	0.0003	0.0036	0.60
Error	21.489	70	0.307				
d×p×a	2.307	10	0.231	1.04	0.4239	0.4221	0.85
Error	15.604	70	0.223				
h×p×a	4.151	10	0.415	2.70	0.0074	0.0074	1.00
Error	10.760	70	0.154				
d×h×p×a	0.616	10	0.062	0.33	0.9713	0.9713	1.00
Error	13.195	70	0.189				
b×a	3.257	6	0.543	3.35	0.0086	0.0086	1.00
Error	6.799	42	0.162				
d×b×a	1.248	6	0.208	1.00	0.4374	0.4294	0.78
Error	8.725	42	0.208				
h×b×a	1.798	6	0.300	1.71	0.1432	0.1518	0.91
Error	7.375	42	0.176				
d×h×b×a	0.740	6	0.123	0.45	0.8389	0.8389	1.00
Error	11.440	42	0.272				
p×b×a	3.770	30	0.126	0.97	0.5147	0.5071	0.80
Error	27.183	210	0.129				
d×p×b×a	3.738	30	0.125	0.91	0.5997	0.5997	1.00
Error	28.635	210	0.136				
h×p×b×a	3.748	30	0.125	0.88	0.6512	0.6512	1.00
Error	29.841	210	0.142				
d×h×p×b×a	3.481	30	0.116	0.83	0.7257	0.7061	0.85
Error	29.454	210	0.140				
Total	903.605						

Table 4.20: Experiment 1 – Repeated measures ANOVA for log of final position error along axes (continued from previous page).

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	5.484	1	5.484	44.17	0.0003		
Error	0.869	7	0.124				
head-track	0.196	1	0.196	0.70	0.4311		
Error	1.963	7	0.280				
d×h	0.048	1	0.048	0.42	0.5360		
Error	0.801	7	0.114				
position	2.414	5	0.483	5.62	0.0007	0.0020	0.79
Error	3.009	35	0.086				
d×p	0.107	5	0.021	1.31	0.2826	0.2906	0.80
Error	0.572	35	0.016				
h×p	0.031	5	0.006	0.18	0.9678	0.8800	0.51
Error	1.196	35	0.034				
d×h×p	0.114	5	0.023	0.70	0.6253	0.5301	0.46
Error	1.132	35	0.032				
block	0.279	3	0.093	2.44	0.0928	0.0928	1.00
Error	0.800	21	0.038				
d×b	0.424	3	0.141	6.28	0.0033	0.0033	1.00
Error	0.473	21	0.023				
h×b	0.041	3	0.014	0.60	0.6200	0.6006	0.88
Error	0.478	21	0.023				
d×h×b	0.011	3	0.004	0.25	0.8619	0.8619	1.00
Error	0.297	21	0.014				
p×b	0.128	15	0.009	0.68	0.7973	0.6869	0.47
Error	1.310	105	0.012				
d×p×b	0.146	15	0.010	0.92	0.5438	0.5237	0.72
Error	1.111	105	0.011				
h×p×b	0.124	15	0.008	0.63	0.8414	0.7817	0.67
Error	1.373	105	0.013				
d×h×p×b	0.124	15	0.008	0.70	0.7811	0.7678	0.91
Error	1.247	105	0.012				
Total	34.388						

Table 4.21: Experiment 1 – Repeated measures ANOVA for log of path length.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	16.173	1	16.173	16.73	0.0046		
Error	6.768	7	0.967				
head-track	1.519	1	1.519	1.14	0.3216		
Error	9.346	7	1.335				
d×h	0.160	1	0.160	0.44	0.5265		
Error	2.524	7	0.361				
position	32.655	5	6.531	10.39	< 0.0001	< 0.0001	0.88
Error	22.000	35	0.629				
d×p	1.619	5	0.324	3.22	0.0171	0.0328	0.72
Error	3.522	35	0.101				
h×p	1.059	5	0.212	1.13	0.3606	0.3607	0.85
Error	6.534	35	0.187				
d×h×p	0.498	5	0.100	0.63	0.6786	0.6638	0.91
Error	5.543	35	0.158				
block	0.507	3	0.169	1.95	0.1519	0.1519	1.00
Error	1.818	21	0.087				
d×b	1.258	3	0.419	2.80	0.0652	0.0652	1.00
Error	3.148	21	0.150				
h×b	0.356	3	0.119	1.38	0.2754	0.2774	0.93
Error	1.801	21	0.086				
d×h×b	0.291	3	0.097	0.84	0.4848	0.4848	1.00
Error	2.411	21	0.115				
p×b	0.879	15	0.059	1.13	0.3420	0.3597	0.54
Error	5.462	105	0.052				
d×p×b	0.783	15	0.052	0.95	0.5150	0.5052	0.81
Error	5.780	105	0.055				
h×p×b	0.606	15	0.040	0.61	0.8641	0.7816	0.58
Error	6.995	105	0.067				
d×h×p×b	0.457	15	0.030	0.64	0.8394	0.7470	0.54
Error	5.034	105	0.048				
Total	166.099						

Table 4.22: Experiment 1 – Repeated measures ANOVA for log of RMS error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	29.609	1	29.609	7.53	0.0287		
Error	27.518	7	3.931				
head-track	4.722	1	4.722	1.91	0.2094		
Error	17.304	7	2.472				
d×h	1.107	1	1.107	0.87	0.3820		
Error	8.908	7	1.273				
position	86.635	5	17.327	10.78	<0.0001	<0.0001	0.92
Error	56.236	35	1.607				
d×p	3.530	5	0.706	1.96	0.1087	0.1087	1.00
Error	12.589	35	0.360				
h×p	3.289	5	0.658	1.60	0.1869	0.1869	1.00
Error	14.426	35	0.412				
d×h×p	1.793	5	0.359	0.80	0.5594	0.5399	0.82
Error	15.751	35	0.450				
block	0.353	3	0.118	0.64	0.5972	0.5830	0.90
Error	3.851	21	0.183				
d×b	3.053	3	1.018	3.78	0.0260	0.0457	0.70
Error	5.660	21	0.270				
h×b	0.833	3	0.278	1.08	0.3773	0.3773	1.00
Error	5.377	21	0.256				
d×h×b	1.057	3	0.352	1.51	0.2422	0.2422	1.00
Error	4.915	21	0.234				
p×b	1.566	15	0.104	0.71	0.7673	0.6257	0.35
Error	15.392	105	0.147				
d×p×b	2.038	15	0.136	1.11	0.3563	0.3563	1.00
Error	12.846	105	0.122				
h×p×b	1.596	15	0.106	0.64	0.8322	0.7835	0.72
Error	17.352	105	0.165				
d×h×p×b	0.864	15	0.058	0.50	0.9361	0.8654	0.58
Error	12.104	105	0.115				

Table 4.23: Experiment 1 – Repeated measures ANOVA for log of RMS error along component axes (continued on next page)

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
axis	408.225	2	204.112	47.69	< 0.0001	< 0.0001	0.90
Error	59.919	14	4.280				
d$\times$a	11.752	2	5.876	4.63	0.0286	0.0365	0.86
Error	17.764	14	1.269				
h$\times$a	2.223	2	1.112	1.48	0.2612	0.2612	1.00
Error	10.518	14	0.751				
d$\times$h$\times$a	0.979	2	0.490	0.59	0.5696	0.5696	1.00
Error	11.694	14	0.835				
p$\times$a	614.963	10	61.496	33.86	< 0.0001	< 0.0001	0.35
Error	127.121	70	1.816				
d$\times$p$\times$a	6.890	10	0.689	1.89	0.0611	0.0611	1.00
Error	25.521	70	0.365				
h$\times$p$\times$a	2.169	10	0.217	0.49	0.8879	0.7998	0.57
Error	30.684	70	0.438				
d$\times$h$\times$p$\times$a	1.449	10	0.145	0.59	0.8156	0.7015	0.48
Error	17.150	70	0.245				
b$\times$a	3.205	6	0.534	5.08	0.0005	0.0012	0.86
Error	4.414	42	0.105				
d$\times$b$\times$a	0.871	6	0.145	1.37	0.2489	0.2653	0.73
Error	4.450	42	0.106				
h$\times$b$\times$a	0.638	6	0.106	1.22	0.3147	0.3255	0.59
Error	3.658	42	0.087				
d$\times$h$\times$b$\times$a	1.367	6	0.228	2.62	0.0300	0.0300	1.00
Error	3.649	42	0.087				
p$\times$b$\times$a	2.090	30	0.070	0.81	0.7432	0.6916	0.66
Error	17.966	210	0.086				
d$\times$p$\times$b$\times$a	1.683	30	0.056	0.70	0.8779	0.8213	0.67
Error	16.837	210	0.080				
h$\times$p$\times$b$\times$a	1.488	30	0.050	0.52	0.9828	0.9436	0.59
Error	20.094	210	0.096				
d$\times$h$\times$p$\times$b$\times$a	2.165	30	0.072	0.93	0.5734	0.5527	0.71
Error	16.268	210	0.077				
Total	1896.938						

Table 4.23: Experiment 1 – Repeated measures ANOVA for log of RMS error along component axes (continued from previous page)

Chapter 5

Experiment 2: Docking

The docking experiment was the second of the four experiments investigating 3D interaction to be carried out. Docking is a task that makes use of all six degrees of freedom present in the input control. In the docking task subjects have to match the position and the orientation of a pointer to the position and orientation of a stationary target. The docking experiment was the only one of the four experiments that explicitly made use of the orientation degrees-of-freedom and the only one that did not make use of the head-tracker. Figure 5.1 shows the equipment setup that was used in this experiment.

5.1 Hypotheses

The experimental hypotheses of this experiment are that display mode (monoscopic versus stereoscopic), target position ($\pm$10cm along one of the X, Y or Z axes from a fixed starting location) and target orientation (a default orientation and $\pm 45°$ from the default orientation around one of the X, Y, or Z axes) have an effect upon performance in a six degree of free-

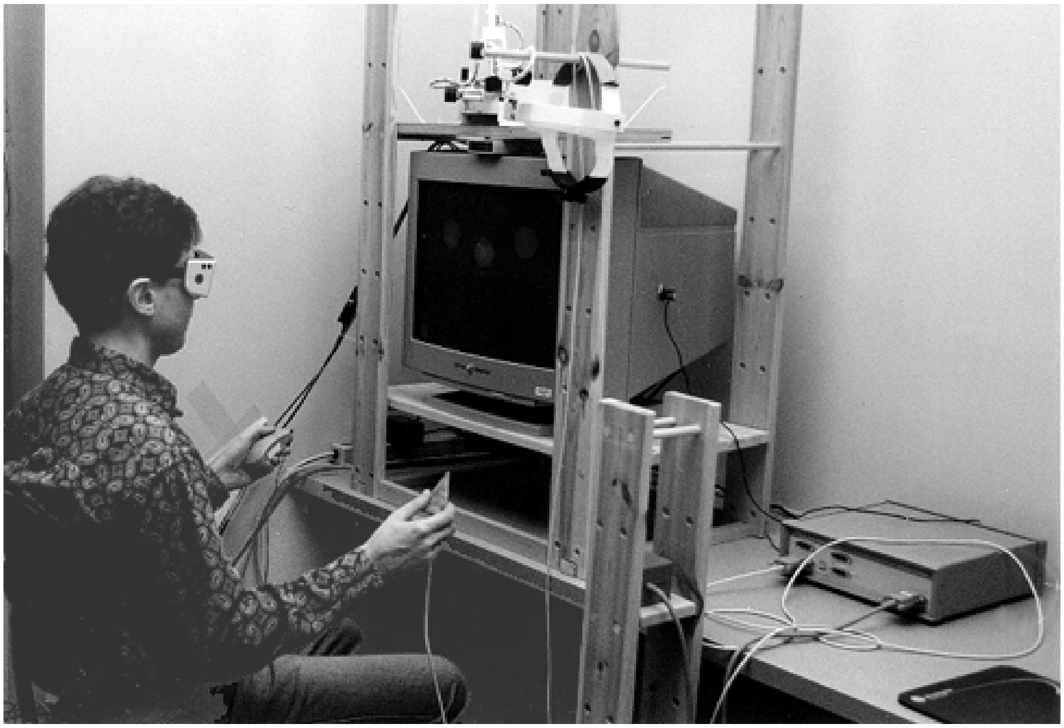


Figure 5.1: Photograph of the experiment system setting used in this experiment. The subject can be seen seated in front of the monitor wearing the LCD glasses. The pointer can be seen in the right hand and the mouse in the left hand. The Fastrak controller and transmitter can be seen in the lower right.

dom docking task. To evaluate the effect of the independent variables upon the task several different dependent measures were gathered.

The following eight dependent measures were analyzed in this experiment: trial completion time, final position error, final position axis error, RMS error, path length, final orientation error, orientation length and final orientation axis error. A detailed description of these measures appears in Section 3.1. RMS error is the root mean square of the distance from the tip of the pointer to the closest point on the ideal path. For this experiment the ideal path used to compute the RMS error is considered to be a straight line between the tip of the homing target and the tip of the trial target. Orientation is not considered in the determination of the ideal path.

For statistical testing, an α level of 0.05 was used to determine whether or not the independent variables had a significant effect.

5.2 Participants

Subjects were recruited as specified in Chapter 3. During the screening process two potential subjects were excluded because they were unable to make correct depth judgments during the screening test. Eight subjects were accepted into the experiment, none of whom participated in the other experiments.

5.3 Equipment

The standard equipment specified in Chapter 3 was used with the exception that the ADL-1 head-tracker was not worn by subjects in this experiment. Figure 5.1 shows the the hardware used in this study. Figure 5.2 shows an image of the screen from one of the trials in this experiment.

The pointer and target in this task were identical tetrahedron-like objects, in which the height (1.73cm) was twice the width (0.87cm), and one face was perpendicular to the base. The difference in size along one dimension allows one vertex to be easily identified as the tip and one face to be identified as the base.

In the default orientation the tip points upward along the positive Y axis, the base was parallel to the XZ plane and the perpendicular face was to the back and parallel to the XY plane. The base was half intensity magenta, the back face was full intensity magenta and the two front faces used both full intensity red and full intensity blue. The two front faces had a different colour in their top and bottom halves. The top half of one face was red and the bottom half of the same face was blue. The opposite face had the reverse colouration.

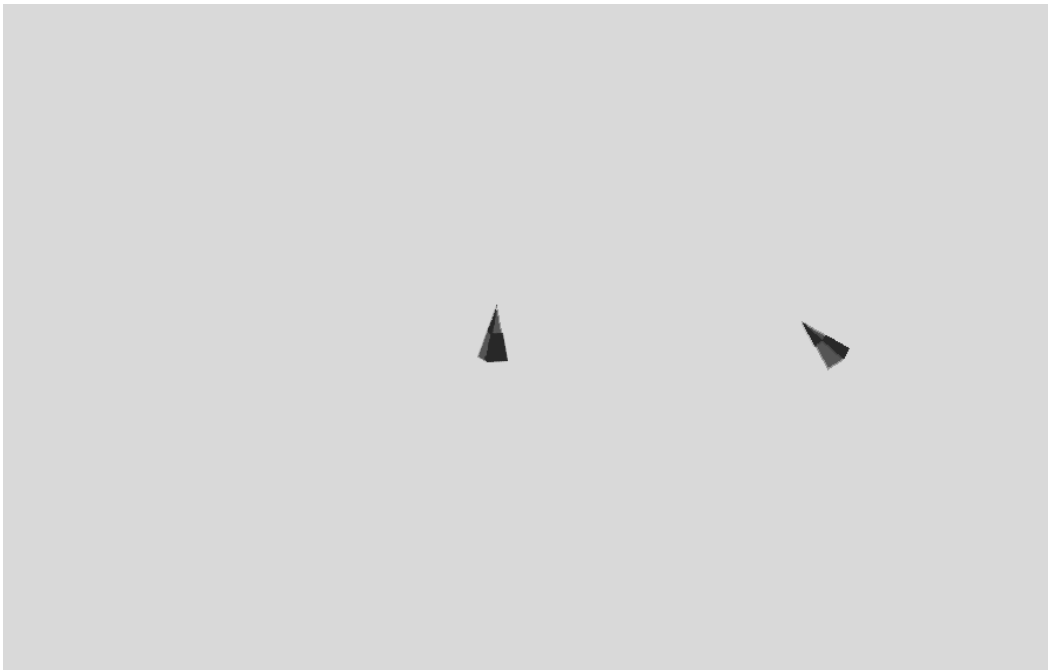


Figure 5.2: An image of the screen during a trial of the docking experiment. The target is visible to the right in the +45° Z orientation with the virtual pointer located near the centre of the image. The background of the screen was black. The background of the image has been lightened to allow better reproduction.

Figure 3.3 illustrates the appearance of the pointer and target in different orientations.

5.4 Procedure

Subjects were required to perform a 6 DOF docking task. Subjects manipulated the pointer via the Polhemus input device using the dominant right hand, and used the middle mouse button with the mouse held in the left hand to advance through the experiment trials. A trial consisted of two phases, an initial homing phase and a docking phase.

During the homing phase the subject had to move the tip of the pointer to the tip of the homing target. The homing target was located at a fixed centre point and was always in the default orientation. Subjects received feedback in the form of a red box that appeared around the tip of the homing target when the tip of the pointer was within 0.5cm of the homing target. When the red box appeared subjects could click the middle mouse button using the left hand to advance to the docking phase; the homing target was removed and the trial target was displayed at one of the six positions in one of the seven orientations. Subjects could not advance to the docking phase if the homing tolerance was not met. The pointer was not required to be in any specific orientation for the homing phase.

During the docking phase subjects were instructed to match the position and orientation of the pointer with the position and orientation of the trial target. No feedback regarding proximity (position or orientation) was given during the docking phase. Subjects had to make their own determination of proximity using the visual cues available within the particular experiment condition. Subjects middle clicked with the mouse when they were satisfied with their performance. After a short delay the system advanced to the next trial.

5.5 Design

The experimental design consisted of three within-subject independent variables: display mode, target position and target orientation. For compatibility between the stereoscopic and monoscopic conditions all target positions in the virtual environment were behind the screen surface and were displayed at a size appropriate to a perspective projection based on a nominal viewing position of 40cm in front of the screen surface. To mitigate the effects of the apparatus, subjects wore the stereo glasses in all conditions. The receiver of the Polhemus was encased in modeling material that was shaped and coloured to resemble the pointer in the virtual environment. The physical pointer is shown in Figure 3.2.

Each subject participated in two sessions conducted on separate days. One session was conducted using the monoscopic display mode and the other session was conducted using the stereoscopic display mode. After each session, subjects completed a short questionnaire dealing only with that session. The selection of initial display mode was counterbalanced so that half the subjects were given the monoscopic display first and half the subjects were given the stereoscopic display first. After the second session, each subject also completed a longer questionnaire about both conditions.

Each session was divided into seven blocks, a training block followed by six experiment blocks. Each experiment block contained 42 trials consisting of all combinations of target position and target orientation presented in random order. The training block was an abbreviation of an experiment block consisting of 20 randomly selected trials. During the first 10 trials of the training phase subjects were given feedback telling them when they were within 1cm and/or 10° of the target. No feedback was given for the remaining 10 trials of the training block. To minimize fatigue subjects were given a one minute break after each block.

Each subject performed a total of 504 experiment trials. Each subject spent between 60 and 90 minutes to complete each session including time for breaks and time to complete the questionnaires. Prior to the experiment session an additional 15 minutes was spent on the screening tasks.

5.6 Training

The following script served as the basis for training subjects on this task.

> "Your task is to move the pointer from the starting position to a target, matching both the position and orientation of the target. You should perform the task as quickly and as accurately as possible. The starting point is located at the tip of the stationary object on the screen. Move your pointer to the starting point now. Notice

that when the tip of your pointer is close to the tip of the starting target a red box appears around the tip of the target. The red box must be present for you to be able to start a trial. To start a trial click the middle mouse button with your left hand."

[*Experimenter waits for subject to start trial.*]

"The starting target has disappeared and another target has appeared somewhere in the virtual environment. Move your pointer to the target object so that your pointer is in the same location and same orientation as the target. Notice that when the tip of the pointer is close to the tip of the target a red box appears around the tip of the target. When the orientation of the pointer is close to the orientation of the target, a red cross-hair appears at the tip of the target. When both the position and orientation are close the red box and the red cross-hair will be present. When you are satisfied with your match click the middle mouse button with your left hand to end the trial."

[*Experimenter allows subject to perform a few trials.*]

"The feedback you are getting when matching the position and orientation of the target will only last for another few trials. During the actual experiment trials you will not receive any feedback regarding your match."

[*After first ten training trials.*]

"Now feedback has been removed. Try to match the position and orientation of the target to the best of your ability. When you are satisfied with your match click on the middle mouse button to end the trial."

"Please try to perform the task as quickly and as accurately as possible."

"Do you have any questions?"

The experimenter answered any questions that would not bias the subject's performance in any of the visual modes. A typical question was, "Do I have to match the orientation of the start target?" Subjects were informed that they do not have to match the orientation of the start target but they could do so if they desired.

The experimenter observed the subject during the training trials. When the training trials were over the experimenter asked and answered any questions using the same guideline stated previously. The experimenter then told the subject that he would be on his own during the actual trials. The experimenter then reminded the subject to perform the trials as quickly and as accurately as possible, left the room and started the first experiment block. The experimenter re-entered the room after each block to check on the subject and tell the subject to rest his arm for at least a minute. After waiting for a minute the subject was asked if he was ready to proceed. If the subject responded in the affirmative then the experimenter left the room and initiated the next experiment block.

5.7 Results

The dependent variables were computed for each trial. Across all eight subjects a total of 4032 individual trials were completed. The mean trial completion time was 8.00s with a standard deviation of 4.66s.

Preliminary Screening

The data was screened for outliers by removing all trials with a trial completion time more than three standard deviations above the mean (75 trials, 1.9%), any trials that were ended accidentally (1 additional trial), any trials with a path length over 150cm (3 additional trials) and any trials with more than 900°of rotation (3 additional trials). The trials three standard deviations above the mean were examined in more detail and 68 of these trials (91% of those that were three standard deviations above the mean) were in the monoscopic condition. Note that many of the trials excluded because they were three standard deviations above the mean for trial completion time, also fell into one of the other exclusion categories. All further analysis was conducted using the 3950 remaining trials.

The experiment was a 2 (display mode) $\times$ 6 (target position) $\times$ 7 (target orientation) design with repeated measures on all factors. For the remaining analysis, the repetitions for each target position and orientation condition performed by a subject within a single session were averaged together to yield a single score. The result is a series of 672 measures per dependent variable, one for each subject in each of the conditions.

Residual Analysis and Degree of Freedom Adjustments

Residuals were analysed for homogeneity of variance as described in Chapter 3 and demonstrated in Chapter 4. A log transform was applied to all of the dependent variable to make the variance more uniform and suitable for analysis. To account for violations of the sphericity assumption, the Huynh-Feldt adjusted degrees of freedom were used to test for significance.

Trial Completion Time

The repeated measures ANOVA for trial completion time indicates significant main effects for display mode $\mathrm{F}(1,7) = 5.84$, $p = 0.0463$ target position $\mathrm{F}_{\mathrm{HF}}(4.3,29.8) = 4.26$, $p = 0.0068$

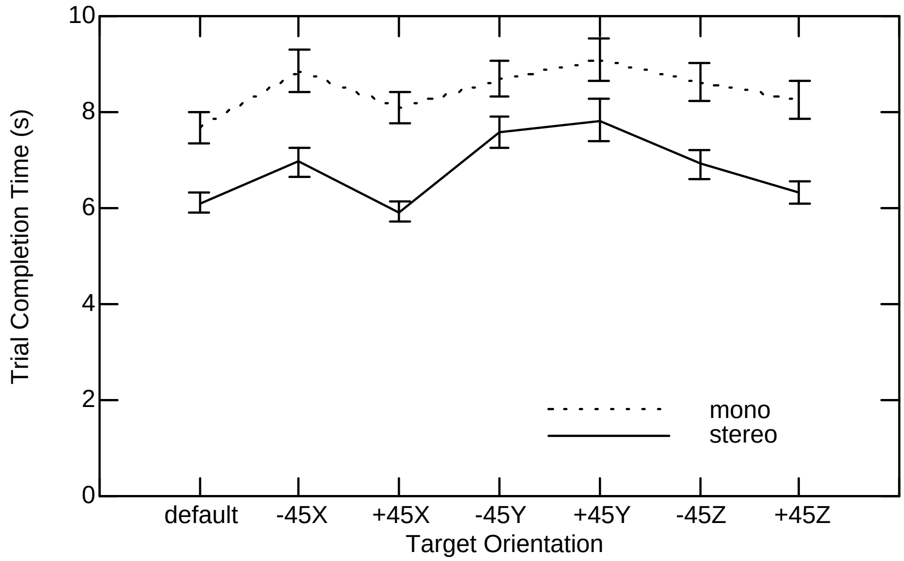


Figure 5.3: Mean of trial completion time in seconds for each of the target orientation conditions in both display modes.

Orientation	Mono	Stereo
default	7.70	6.11
−45°X	8.87	6.96
+45°X	8.10	5.94
−45°Y	8.69	7.58
+45°Y	9.10	7.83
−45°Z	8.63	6.92
+45°Z	8.27	6.35

Table 5.1: Mean trial completion time in seconds for all target orientations in both display modes.

and target orientation $F_{HF}(5.8,40.3) = 6.93$, $p < 0.0001$. Significant two-way interactions were found for display mode × target orientation $F_{HF}(6,42) = 2.64$, $p = 0.0291$ and target position × target orientation $F_{HF}(29.4,205.8) = 3.51$, $p < 0.0001$. Table 5.16 at the end of the chapter provides full details of the ANOVA results.

The mean trial completion time was 8.48s in the monoscopic display mode and 6.81s in the stereoscopic display mode. However, due to the significant display mode × target ori-

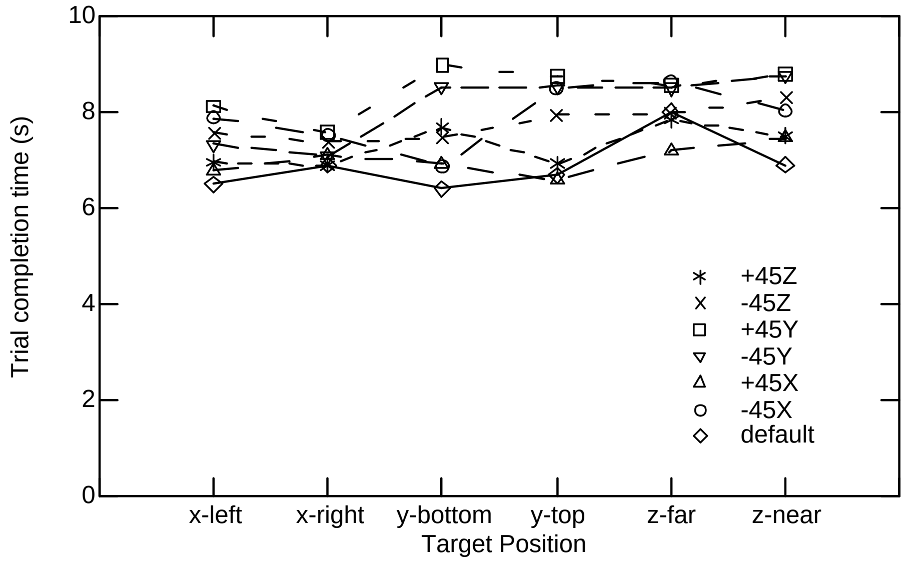


Figure 5.4: Graph of trial completion time in seconds for each target orientation condition across target positions. Error bars are removed to improve legibility

entation interaction these factors must be examined together. Figure 5.3 shows the relationship in trial completion time for the two display modes. Stereoscopic viewing is superior to monoscopic viewing across all target orientations. The significant interaction is a result of the difference in improvement for the stereoscopic display mode across the orientations that ranges from a low of 1.11s between the two $-45°$Y conditions to a high of 2.16s in the two $+45°$X conditions. Table 5.1 gives the trial completion time for the different orientations in the two display modes.

The mean trial completion time for the target position and target orientation conditions must be examined together in light of the significant interaction between these effects. Table 5.2 shows the mean trial completion time for each of the target position and target orientation conditions and Figure 5.4 illustrates these conditions.

Closer examination reveals the source of the significant interaction effect. A post-hoc analysis using Tukey's HSD as modified by Cicchetti [29] to allow for only row and

Cond.	**def.**	**−45X**	**+45X**	**−45Y**	**+45Y**	**−45Z**	**+45Z**	**all**
x-left	6.52	7.89	6.78	7.34	8.13	7.58	6.97	7.32
x-right	6.91	7.52	7.12	7.08	7.58	7.38	6.90	7.21
y-bottom	6.41	6.88	6.92	8.53	8.99	7.47	7.69	7.56
y-top	6.69	8.50	6.59	8.55	8.74	7.95	6.93	7.71
z-far	8.01	8.65	7.21	8.53	8.56	7.98	7.85	8.11
z-near	6.90	8.04	7.47	8.77	8.79	8.29	7.51	7.97
all	6.91	7.91	7.02	8.13	8.47	7.78	7.31	

Table 5.2: Mean trial completion time in seconds for all combinations of the target position and target orientation conditions. The row and column labeled all indicates the overall mean for the orientation or position condition respectively.

column contrasts was conducted. The analysis indicates that where an effect exists it is between the extreme values within a given target orientation (column) or target position (row). There was no significant effect for orientation in the x-right target position. There was no significant effect for target position in the +45°X, +45°Y or ±45°Z target orientation conditions.

Final Position Error

The repeated measures ANOVA for final position error revealed significant main effects for display mode $F(1,7) = 10.56$, $p = 0.0141$, target position $F_{HF}(2.9,20.3) = 4.53$, $p = 0.0143$ and target orientation $F_{HF}(4.4,31.1) = 3.58$, $p = 0.0140$. There were no significant interactions. Summary results of of the ANOVA for total position error are given in Table 5.17 at the end of the chapter.

The final position error in the monoscopic display mode was 0.45cm and the final position error in the stereoscopic display mode was 0.24cm. Table 5.3 gives the final position error in each of target position averaged over the other conditions and Table 5.4 gives the final position error in each of the target orientations averaged over the other conditions. We

Pos.	Position Error
x-left	0.32
x-right	0.33
y-bottom	0.35
y-top	0.33
z-far	0.39
z-near	0.35

Table 5.3: Mean final position error in centimetres for each of the target positions averaged over the other conditions.

Orientation	Position Error
default	0.32
$-45°$X	0.36
$+45°$X	0.37
$-45°$Y	0.38
$+45°$Y	0.37
$-45°$Z	0.32
$+45°$Z	0.28

Table 5.4: Mean of final position error in centimetres for each of the target orientations averaged over the other conditions.

can see that the position to the left has the lowest error while the z-far position has the highest error. For the orientation conditions, the $-45°$Y condition had the highest final position error while the $+45°$Z condition had the lowest final position error.

Final Orientation Error

A repeated measures ANOVA for orientation error reveals significant main effects for target orientation $F_{HF}(3.1,21.4) = 15.05$, $p < 0.0001$ and target position $F_{HF}(3.4,23.5) = 4.08$, $p = 0.0156$ and a significant display mode $\times$ target position interaction $F_{HF}(4.6,32.2) = 2.81$, $p = 0.0360$. Results of the ANOVA appear in Table 5.18 at the end of the chapter.

Table 5.5 shows the amount of orientation error in each of target orientations. The $\pm 45°$Y conditions clearly stand out as having more error than the other conditions. The effect of target position must be considered together with the effect for display mode because

Orientation	Orientation Error
default	6.55
$-$**45°X**	5.51
+**45°X**	8.70
$-$**45°Y**	14.02
+**45°Y**	14.41
$-$**45°Z**	8.71
+**45°Z**	7.84

Table 5.5: Mean of final orientation error in degrees for each of the target orientation conditions averaged over the other conditions.

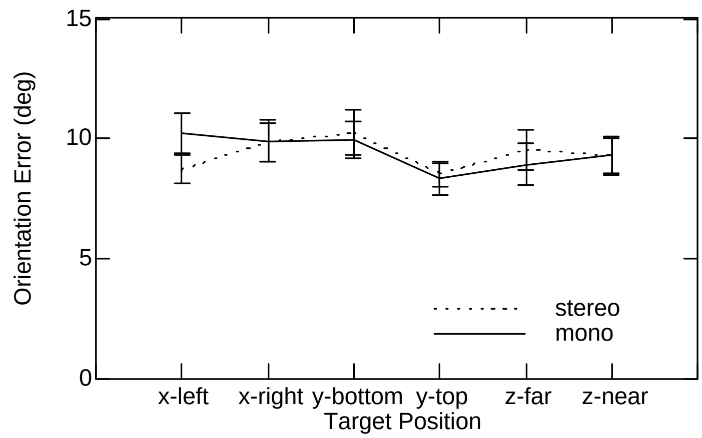


Figure 5.5: Mean orientation error in degrees vs. target position for both display modes.

of the significant interaction. Figure 5.5 shows the orientation error across target positions for both display modes. The divergence between the stereoscopic and monoscopic conditions for the left target position is the source of the interaction effect. Table 5.6 provides the total orientation error in each target position condition for both display modes.

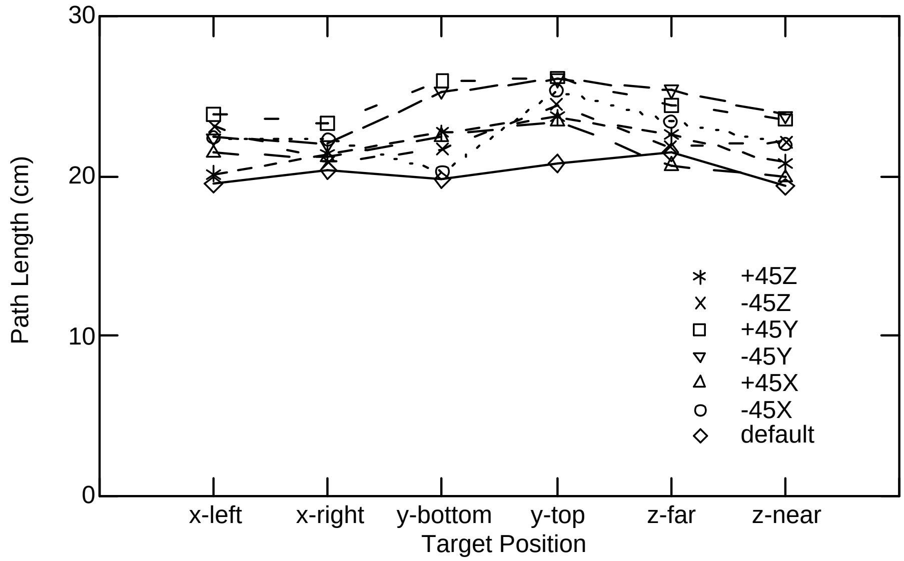


Figure 5.6: Mean path length in centimetres across target positions for each of the target orientation conditions.

Path Length

The repeated measures ANOVA for path length revealed significant main effects for display mode $F(1,7) = 8.27$, $p = 0.0238$ and target orientation $F_{HF}(3.3,23.1) = 6.49$, $p = 0.0019$ as well as a significant two-way target position $\times$ target orientation interaction $F_{HF}(12.3,86.1) = 1.97$, $p = 0.0344$. The full ANOVA table appears in Table 5.19 at the end of the chapter.

The mean path length in the monoscopic condition was 24.90cm and the mean path

Condition	Mono	Stereo
x-left	10.21	8.70
x-right	9.91	9.86
y-bottom	9.93	10.24
y-top	8.31	8.52
z-far	8.91	9.53
z-near	9.29	9.28

Table 5.6: Mean orientation error in degrees for all target positions in both display modes.

Cond.	def.	−45X	+45X	−45Y	+45Y	−45Z	+45Z	all
x-left	19.52	22.42	21.45	22.44	23.91	23.12	20.15	21.82
x-right	20.40	22.35	21.19	22.02	23.30	20.89	21.42	21.65
y-bottom	19.79	20.28	22.52	25.30	25.96	21.69	22.74	22.61
y-top	20.82	25.38	23.47	26.09	26.18	24.46	23.72	24.30
z-far	21.52	23.44	20.69	25.41	24.49	21.81	22.60	22.85
z-near	19.36	22.03	19.92	23.84	23.54	22.20	20.81	21.67
all	20.23	22.65	21.54	24.18	24.56	22.36	21.91	

Table 5.7: Mean path length in centimetres for all combinations of the target position and target orientation conditions. The row and column labeled all indicates the overall mean for the orientation or position condition respectively.

length in the stereoscopic condition was 20.09cm.

The effect of target orientation must be considered together with the effect for target position because of the significant two-way interaction. Figure 5.6 illustrates the relationship between target position and path length for each of the target orientations. Table 5.7 gives the path length for each of the target position and target orientation conditions. The key things to notice are:

- the default target orientation generally yielded the shortest movement paths,
- the $\pm 45°$Y orientations generally yielded the longest movement paths,
- some target position and target orientation combinations had movement paths that were comparable to the default orientation.

Orientation Length

The repeated measures ANOVA for orientation length revealed a significant main effect for target orientation $F_{HF}(4.7,33.2) = 24.13$, $p < 0.0001$ and a significant target position x target orientation interaction $F_{HF}(21.6,151.2) = 2.56$, $p = 0.0005$. Because of the interaction these

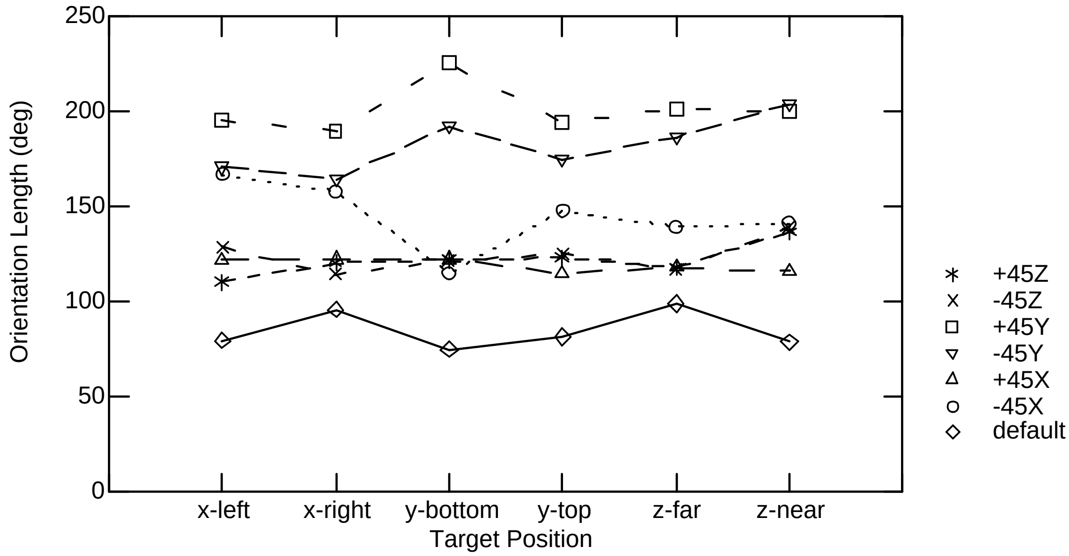


Figure 5.7: Mean orientation length in degrees across target positions for each of the target orientation conditions.

Cond.	**def.**	**−45X**	**+45X**	**−45Y**	**+45Y**	**−45Z**	**+45Z**
x-left	79.42	166.96	121.83	171.01	195.35	128.72	110.48
x-right	95.76	158.32	122.37	164.46	190.09	114.30	120.35
y-bottom	74.87	115.28	122.44	192.39	225.39	121.39	122.26
y-top	81.59	148.10	114.90	174.79	194.50	125.28	123.18
z-far	99.05	139.46	118.02	186.68	201.57	117.40	118.32
z-near	79.01	141.28	115.92	204.24	200.31	139.41	136.85
all	84.95	144.90	119.25	182.26	201.20	124.42	121.91

Table 5.8: Mean orientation length in degrees for all combinations of the target position and target orientation conditions. The row labeled all indicates the overall mean for the orientation condition.

factors are considered together. The full ANOVA results appear in Table 5.20 at the end of the chapter.

Figure 5.7 illustrates the effect of target position on orientation length for each of the target orientation conditions. Table 5.8 provides the mean orientation length in the different target position and target orientation conditions. The key things to notice are:

Condition	Mono	Stereo
x-left	1.51	0.82
x-right	1.85	1.12
y-bottom	1.93	1.29
y-top	2.26	1.15
z-far	1.09	0.96
z-near	0.86	0.86

Table 5.9: Mean RMS error in centimetres for the target positions in both display modes, averaged over the other conditions.

- the default orientation generally required the least amount of reorientation,
- the $\pm 45°$Y target orientations generally required the most reorientation,
- the amount of reorientation is generally consistent across target positions for each of the target orientation conditions, with the noticeable exception of the $-45°$X condition in the y-bottom target position.

RMS Error

The repeated measures ANOVA for RMS error revealed significant main effects for display mode $F(1,7) = 19.30$, $p = 0.0032$, target position $F_{HF}(5,35) = 11.68$, $p < 0.0001$ and target orientation $F_{HF}(4.5,31.5) = 4.00$, $p = 0.0077$. Significant two-way interactions were found for display mode $\times$ target position $F_{HF}(3.8,26.6) = 7.06$, $p = 0.0006$ and target position $\times$ target orientation $F_{HF}(17.7,123.9) = 1.77$, $p = 0.0363$. Full results of the ANOVA for RMS error appear in Table 5.21 at the end of the chapter.

Table 5.9 provides the RMS error for each of the target positions in both display modes. Stereoscopic viewing reduces RMS error in the x-left, x-right, y-bottom and y-top target positions. RMS error is not significantly reduced by stereoscopic viewing for the z-near and z-far target positions.

Figure 5.8 illustrates the RMS error across target positions for each target orienta-

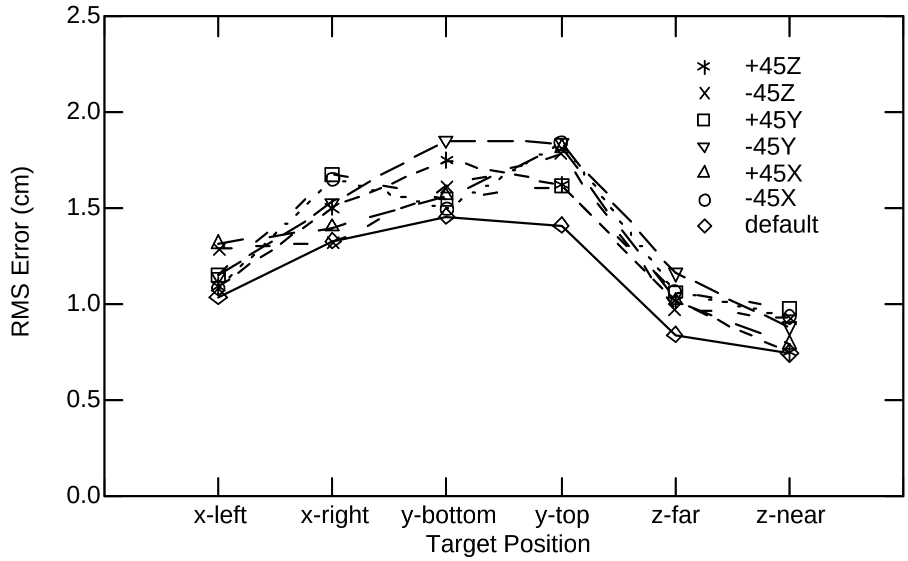


Figure 5.8: Graph of target position vs. RMS error in centimetres for each of the target orientation conditions. Error bars are removed to improve legibility.

tion. The general pattern for RMS error remains consistent for each of the target orientation conditions. The interaction effect is evident in the increased variation in RMS error for the y-top target position and the reduced variation in RMS error for the z-near target position.

Final Position Axis Error

To gain an understanding of how both position and orientation error are distributed across the X, Y and Z axes, axis was introduced as a factor and a repeated measures ANOVA was performed with this additional factor.

The repeated measures ANOVA for final position axis error identified significant main effects for display mode $F(1,7) = 6.66$, $p = 0.0365$, target position $F_{HF}(3.1,21.4) = 20.47$, $p < 0.0001$, target orientation $F_{HF}(4.0,28.1) = 4.00$, $p = 0.0109$, and axis $F_{HF}(1.1,7.4) = 116.36$, $p < 0.0001$. Significant two-way interactions were found for display mode $\times$ target position $F_{HF}(4.0,27.7) = 2.93$, $p = 0.0394$, display mode $\times$ axis $F_{HF}(1.1,7.8) = 19.10$,

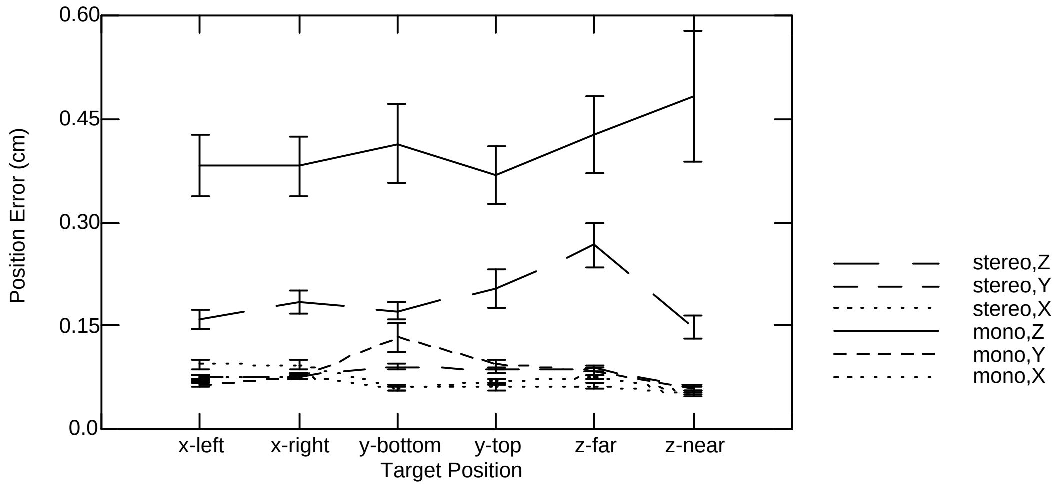


Figure 5.9: Means of final position error in centimetres along each of the axes across all target positions for both display modes.

Pos.	**Mono**			**Stereo**		
	X	**Y**	**Z**	**X**	**Y**	**Z**
x-left	0.09	0.06	0.38	0.08	0.08	0.16
x-right	0.09	0.08	0.38	0.08	0.08	0.19
y-bottom	0.06	0.13	0.41	0.06	0.09	0.17
y-top	0.06	0.09	0.37	0.07	0.09	0.20
z-far	0.06	0.08	0.43	0.08	0.09	0.27
z-near	0.05	0.06	0.48	0.05	0.06	0.15

Table 5.10: Mean position error in centimetres along each axis for each target position in both display modes.

P = 0.0021, target position $\times$ axis $\mathrm{F_{HF}}(6.8,47.6) = 8.74$, $p < 0.0001$ and target orientation $\times$ axis $\mathrm{F_{HF}}(9.7,68.0) = 3.68$, p = 0.0006.

A single three-way interaction of display mode $\times$ target position $\times$ axis $\mathrm{F_{HF}}(8.8,61.6) = 2.84$, p = 0.0076 was also found. Full results of the ANOVA appear in Table 5.22 at the end of the chapter.

Table 5.10 provides the position error across each axis for both display modes and all target positions. Figure 5.9 illustrates these results. This table and figure can be used to

Condition	X	Y	Z
default	0.06	0.08	0.28
$-45°$X	0.07	0.10	0.30
$+45°$X	0.07	0.09	0.33
$-45°$Y	0.07	0.08	0.34
$+45°$Y	0.07	0.08	0.33
$-45°$Z	0.07	0.08	0.28
$+45°$Z	0.07	0.07	0.24

Table 5.11: Position error in centimetres along the X, Y and Z axes for each of the target orientation conditions.

understand the three-way display mode $\times$ target position $\times$ axis interaction, the two-way display mode $\times$ target position, display mode $\times$ axis and target position $\times$ axis interactions, and the main target position and display mode effects. It is clear that there is more error along the Z-axis than along the X or Y axes. Stereoscopic viewing has a significant impact in reducing the amount of error along the Z-axis. The largest reduction in Z-axis error is for the z-near target and the smallest reduction in Z-axis error is for the z-far target.

Table 5.11 provides the position error across each axis for each of the target orientations. This table helps explain the source of the target orientation $\times$ axis interaction as well as the target orientation main effect. Once again it is clear that there is more error along the Z-axis. Target orientation has practically no effect upon the position error along the X and Y axes, but there is a relatively large difference in Z-axis error across the target orientations.

Final Orientation Axis Error

The repeated measures ANOVA for final orientation axis error identified a significant main effect for axis $F_{HF}(2,14) = 132.13$, $p < 0.0001$. Significant two-way interactions were found for display mode $\times$ orientation $F_{HF}(3.2,22.3) = 3.60$, $p = 0.0274$, display mode $\times$ axis $F_{HF}(2,14) = 6.92$, $p = 0.0081$, and target orientation $\times$ axis $F_{HF}(12,84) = 18.78$, $p < 0.0001$.

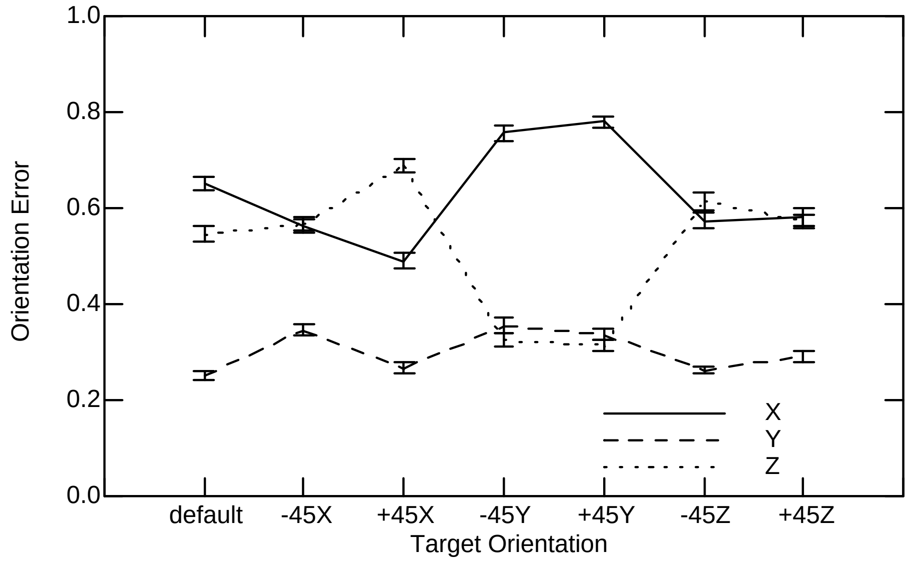


Figure 5.10: Orientation error for each of the target orientation conditions about the X, Y and Z axes. The values represent the projection of the unit axis of rotation for the orientation error onto the X, Y and Z axes.

A single three-way interaction of target position × target orientation x axis $F_{HF}(51.6,361.2) = 1.42$, $p = 0.0375$, was also found. Full results of the ANOVA appear in Table 5.23 at the end of the chapter. Because the purpose of this analysis is to examine the effect of axis, only the effect of axis and its interactions are analyzed further.

Display Mode	**X**	**Y**	**Z**
mono	0.61	0.30	0.53
stereo	0.64	0.30	0.51

Table 5.12: Mean of orientation error along each axis. The values represent the projection of the unit axis of rotation onto the X, Y and Z axes.

The display mode and display mode × axis effects are considered together because of the significant interaction. Table 5.12 shows the amount of orientation error about the X, Y and Z axes in both display modes. The Y-axis orientation error remains stable between the display modes, but there is a slight shift of orientation error from the Z-axis in the mono-

Condition	**X**	**Y**	**Z**
default	0.65	0.25	0.55
$-45°$X	0.56	0.35	0.57
$+45°$X	0.49	0.27	0.69
$-45°$Y	0.76	0.35	0.33
$+45°$Y	0.78	0.34	0.31
$-45°$Z	0.58	0.26	0.62
$+45°$Z	0.58	0.29	0.57

Table 5.13: Mean of orientation error along each axis for each of the target orientations. The values represent the projection of the unit axis of rotation onto the X, Y and Z axes.

scopic display mode to the X-axis in the stereoscopic display mode.

Figure 5.10 illustrates the relationship between the amount of orientation error about the X, Y and Z axes and the target orientation. Once again, the Y-axis error remains relatively stable across target orientations. There is a shift of error from the Z-axis to the X-axis in the $\pm 45°$Y target orientation conditions. Table 5.13 provides the amount of orientation error about the X, Y and Z axes in each of the target orientations.

Figure 5.11 illustrates the amount of orientation error for each of the target orientations across target positions for each of the X, Y and Z axes. This figure can be used to help interpret the target position $\times$ target orientation $\times$ axis interaction. The interaction was due to the substantial amount of variation in orientation error about the axes across the various target position and target orientation conditions. The Y-axis orientation error was generally lower compared to the other axes, the X-axis orientation error was generally higher and the Z-axis orientation error was spread out with some conditions having had a consistently lower Z-axis orientation error and some having had a consistently higher z-axis orientation error. The $\pm 45°$Y target orientation conditions stand out as having their error biased away from the Z-axis and towards the X-axis.

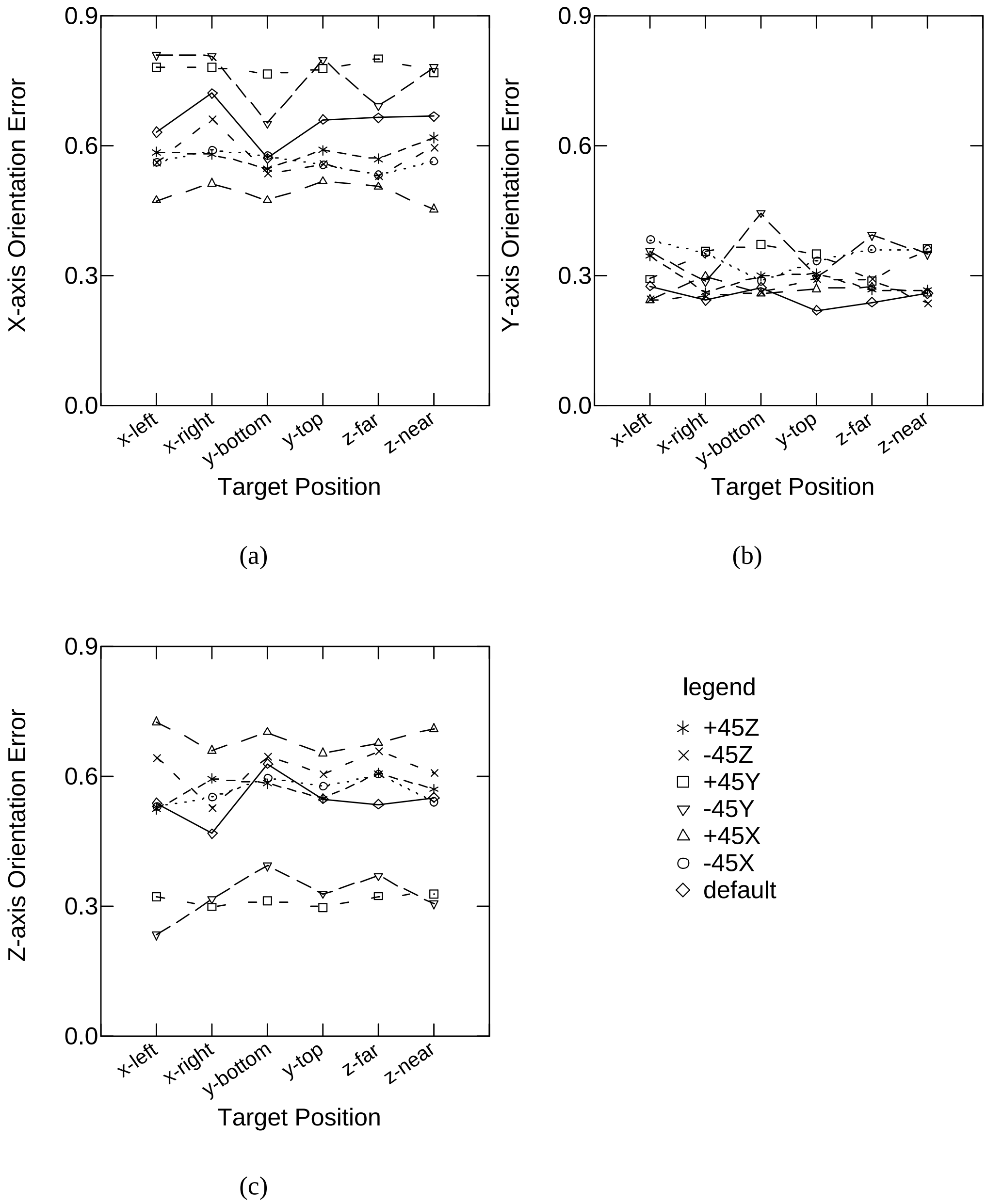


Figure 5.11: Orientation error for each of the target position and target orientation conditions about the X, Y or Z axis. The values represent the projection of the unit axis of rotation for the orientation error onto the X, Y and Z axes.

Questionnaires

At the end of each display mode session, subjects were asked to complete a questionnaire consisting of eight questions: two about positioning difficulty, two about orientation difficulty, two about target positions and two about target orientations. The questions relating to difficulty were answered on a five point scale where 1 was "easy" and 5 was "hard". The difficulty questions asked subjects to separately rate the difficulty of determining or matching the position or orientation of the targets. The questions about target positions and orientations asked subjects to indicate which if any target positions or orientations were easier or harder to match. A copy of the post-session questionnaire is located in Appendix A.4.

At the end of the experiment subjects were asked to score both display modes according to task difficulty and usefulness and to rank the display modes according to preference, ease of use and performance. The same post-experiment questionnaire was used in all the experiments. Subjects were asked to disregard any portions of the questionnaire dealing with head-tracking. A copy of the post-experiment questionnaire can be found in Appendix A.7.

A summary of subjects' responses to the post-session and post-experiment questions in which they were asked to score the display modes is given in Table 5.14.

In the post-session questionnaires that asked subjects to separately consider the position and orientation components of the task, subjects stated that stereoscopic viewing made it easier to determine the position of the target. The viewing condition had an almost significant effect upon the ease of matching the position of the target. Subjects indicated that there was very little difference between the display modes in terms of determining and matching the orientation of the target. Even though the mean scores of subjects comparing the task difficulty and display mode usefulness indicated that the stereoscopic display mode was easier and more useful, the Friedman test indicated that the differences were not significant.

Question	Mono	Stereo	Friedman	p
det. target pos.	3.1	1.5	8.00	0.005
match target pos.	2.6	1.9	3.13	0.077
det. orientation	2.0	1.9	0.00	1.000
match orientation	2.3	2.3	0.13	0.724
task difficulty	2.8	1.9	2.00	0.157
mode usefulness	2.8	1.9	2.00	0.157

Table 5.14: Means of subjects' responses to the subjective questions regarding display mode conditions. The first four questions were answered immediately after a particular condition, and the last two questions were answered after all conditions had been completed. All questions were answered on a scale of 1 to 5 where 1 indicated easy or useful and 5 indicated hard or not useful.

Subjects' responses regarding which targets were easier or harder were coded by the experimenter into a few categories. Summaries of subjects' coded responses can be found in Table 5.15. While some subjects considered that the orientation condition made the positioning component of the task more difficult, no subjects indicated that the position made orienting more difficult. The z-far target was generally considered the hardest and the z-near target the easiest. Subjects did not single out any other individual targets as being easier or harder than the others.

Seven of the eight subjects indicated that they thought that the orientation was harder to match when only one face was visible in the stereoscopic display mode. Having only a single face showing corresponds to the $\pm 45°$orientation conditions. Five of the eight subjects indicated that having more than one face visible made the orientation match easier in the stereoscopic display mode.

At the end of the experiment subjects were asked to rank the two display modes that they used to perform the task according to preference, ease of use and performance. Figure 5.12 provides a summary of the rankings. Two out of eight (Friedman = 2.0, p = 0.157) stated that they preferred monoscopic viewing; three out of eight (Friedman = 0.50, p = 0.480) stated that they found the monoscopic display easier to use; and two out of eight (Friedman

	Position				Orientation			
	Easier		Harder		Easier		Harder	
Response	**M**	**S**	**M**	**S**	**M**	**S**	**M**	**S**
None	1	3		3	3	2	2	
z-near	3	4	1					
z-far			5	4				
1 face			3	3			5	7
$>$ 1 face	3	2			3	5		
default		2			1	1		
XY axes	1							
$-45°$X				1	1		1	1
$+45°$Y					1			

Table 5.15: Summary of subjects' responses indicating which target positions and orientations they thought were easier or harder to match in the monoscopic(M) and stereoscopic (S) display modes. Column totals do not always sum to eight because some subjects indicated multiple conditions.

= 2.0, p = 0.157) stated that their performance was better in the monoscopic condition.

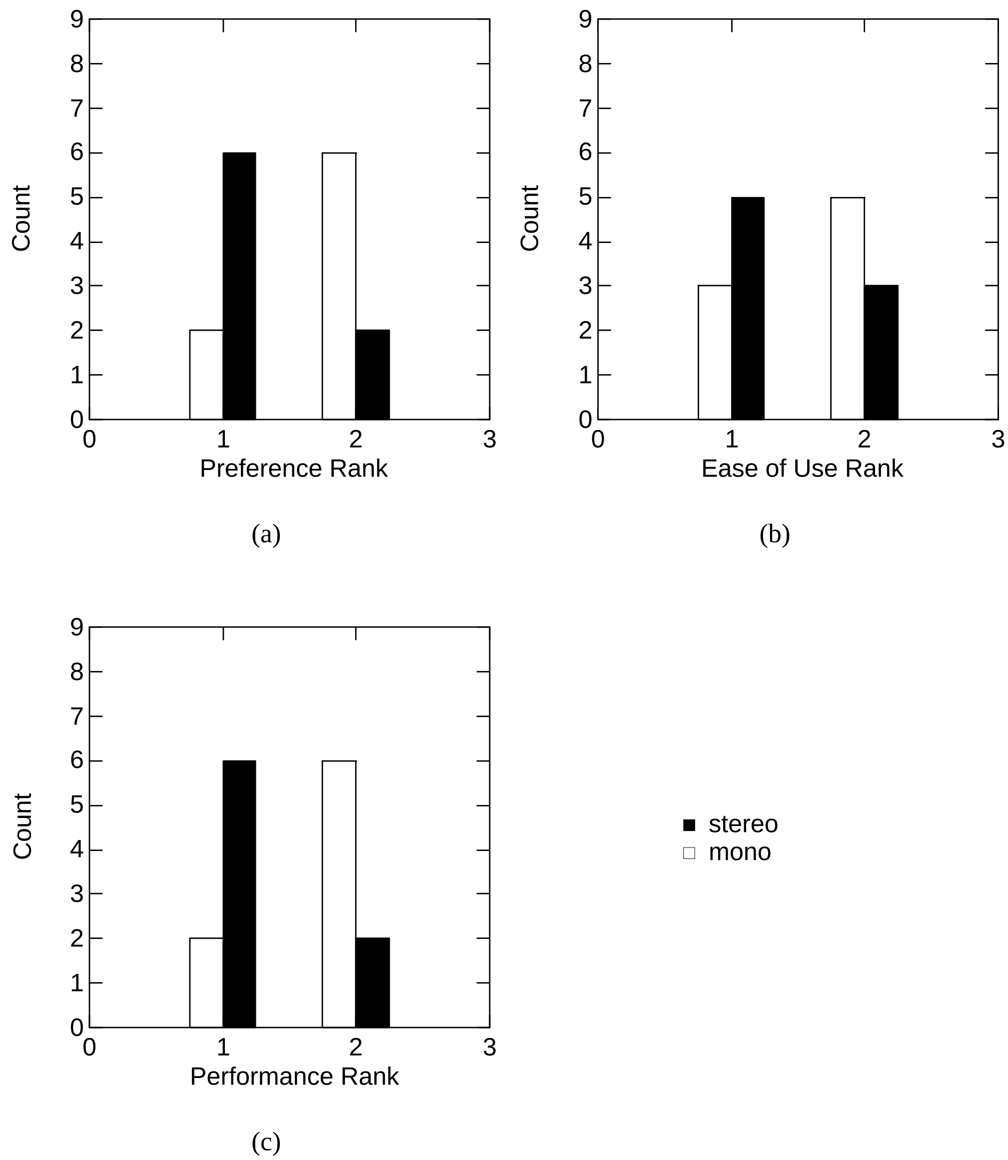


Figure 5.12: Rankings of subjects' preference, perceived ease of use and perceived performance for each of the viewing conditions. The height of a bar indicates the number of times a particular rank was assigned by a subject.

5.8 Discussion

This docking experiment made use of a large number of dependent variables each of which was affected in some way by the independent variables that were manipulated. The large set of dependent variables provides a great deal of insight into the mechanics of the docking task.

5.8.1 Speed

The results for trial completion time showed significant main effects for display mode, target position and target orientation as well as several significant two-way interactions.

Stereoscopic viewing improves trial completion time

The significant main effect for display mode showed that when averaged over the other factors, docking in the stereoscopic viewing mode was faster than in the monoscopic viewing mode. This is not surprising given the body of previous work in this area.

Position and orientation of targets affects trial completion time

The results for trial completion time also showed significant main effects for target position and target orientation along with a significant interaction between these variables. Because of the significant interaction, target position and target orientation must be considered together.

The joint effect of target position and target orientation is new and interesting. From Table 5.2 it is apparent that there are many position and orientation combinations that are significantly different from each other, however, the results are more interesting when considered as either orientation groupings or position groupings. Where a significant difference exists we see a position or orientation condition where performance is not uniform across the other factor. The lack of a significant effect for certain position or orientation groupings may suggest that certain positions are immune to the effects of orientation and certain orientations are immune to the effects of position. An alternate possibility is that certain positions may impede performance across all orientations, in effect making all orientations equally difficult. Considered from the view of orientations, certain orientations may make all positions equally difficult.

X-axis performance generally fastest, Z-axis performance generally slowest

Some general performance trends are clearly visible. The target to the x-right is fastest in four of seven orientation conditions and the target to the x-left is second fastest in four of seven conditions. This is similar to Experiment 1 where the left target was fastest and the right target was second fastest. Performance for the z-far target was generally poor; however, for the two orientation conditions that generally exhibited the poorest performance, the z-far target ranks higher. Closer examination of Table 5.2 shows that this is likely due to a degradation in performance for the other conditions rather than a genuine improvement in performance for the z-far position.

Y-axis orientations slowest

The $\pm 45°$Y conditions generally exhibit the poorest performance in terms of trial completion time. The default target orientation that had the same orientation as the start target generally had the best performance, but this performance was degraded substantially for the z-far target position. The poor orientation match accuracy for the $\pm 45°$Y conditions suggests that trial completion time may understate the difficulty of these two conditions.

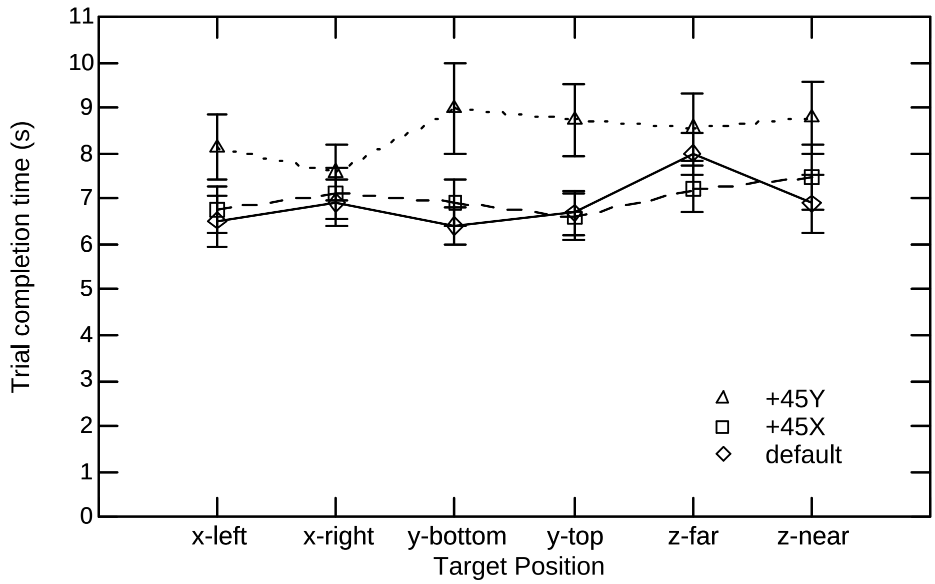


Figure 5.13: Graph of trial completion time in the default, $+45°$X and $+45°$Y target orientation conditions across all target positions with error bars.

Some orientations match the performance of the default orientation

One might easily conclude that the default orientation would directly lead to the best performance of all the target orientation conditions. While this seems to be the case in many of the target positions, it is certainly not true overall. Figure 5.13 illustrates trial completion time for three orientation conditions, the default (best overall), $+45°$Y (worst overall) and $+45°$X (2nd best overall). It is fairly that performance across the default orientation is roughly similar to performance across the $+45°$X condition and that both these conditions are mostly different from performance in the $+45°$Y condition. The exception that stands out is that for the “right” target position there is little difference between any of these conditions.

Evidence that orientation occurs during translation

Jacob et al. [57] suggest that input devices should be matched to the task based upon the task's integrability or separability. It is important to determine whether the position and orientation components that make up the docking task are integrable or separable. If they are integrable they benefit from being considered as a single 6 DOF task and justify the use of a 6 DOF input control. If they are separable, then an input control with fewer degrees of freedom might be used instead. Evidence that positioning and orientation occur in parallel rather than serially might suggest that these two subtasks are integrable.

Wang et al. [99] conducted a study of transportation and orientation in an augmented virtual environment. In their study they specifically looked at the overlap between orientation time and transportation time. They found that the orientation component occurred within the transportation component and the transportation component was the limiting factor.

Ware [103] conducted a study of object placement in a virtual environment. He investigated position only, orientation only and combined position and orientation tasks. He found that position and orientation could be performed in roughly the same amount of time as orientation alone indicating that the orientation component seemed to be the limiting factor. An interesting aspect of Ware's tasks was that he allowed subjects to manipulate their view of the scene during the course of each trial.

When we compare this experiment to Experiment 1, which only looked at positioning, we find that trial completion time in this experiment increased by 3.48s in the monoscopic condition and 3.24s in the stereoscopic condition over the times reported in that experiment. Clearly the docking task is harder, but if we believe that orientation matching should take more than the average of 3.36s, it suggests at least some overlap in the positioning and orientation activities. Hinckley et al. [53] report a rotation matching time of about

15.4s for males using a similar input device, although the rotations were likely of a larger magnitude. Further analysis might allow us to better understand the relationship between these two operations.

5.8.2 Accuracy

Position error reveals results that are generally similar to the trial completion time results. As was expected, stereoscopic viewing generally reduced the overall error by 47% as compared with monoscopic viewing. The z-far target position has the greatest error and the x-left, x-right and y-top positions have the least error.

Target orientation influences translation accuracy

The effect of target orientation on position error is interesting. It indicates that certain target orientations have an effect upon subjects' ability to match tip positions. The $\pm 45°$Y orientations have the largest position error, but interestingly, the $\pm 45°$X orientations have similar position error levels. The smallest position error levels are exhibited by the default and $\pm 45°$Z orientations.

Rotations about the X-axis change the depth of the pointer tip, while rotations about the Y-axis or Z-axis do not. Performance for the $\pm 45°$Y orientations is generally poor and so it is not surprising that they also exhibit a large position error. The increased position error measures for the $\pm 45°$X orientations may be a result of the depth changes for the tip of the target.

Z-near target shows the largest improvement in Z-axis error

Examination of the results for position error across the X, Y and Z axes highlights the Z-axis as the source of much of the positioning error. Stereoscopic viewing significantly reduces the amount of Z-axis error. The largest improvement is for the z-near target position and the smallest improvement is for the z-far target. These two target positions had a somewhat higher Z-axis error level than the other target positions in the monoscopic display mode. In the stereoscopic display mode the Z-axis error for the z-near target is comparable to most of the other target positions, but the z-far target still shows a higher Z-axis error level.

Final Orientation Error

The results for final orientation error highlight the poor performance for the $\pm 45°$Y orientations. The $\pm 45°$Y orientations have an average rotation error of 14.23°while the average rotation error for the other conditions is 7.46°. Display mode and target position combine in their effect upon orientation error. The significant interaction for these factors indicates that orientation error is not consistent across the combinations of these two factors. Stereoscopic viewing had a lower orientation error for the left target position, but a higher orientation error for the z-far target.

Considerations for orientation error

Several issues must be considered carefully when discussing orientation error. The nature of rotations limits the maximum orientation error to 180°. In contrast, the maximum position error is unlimited.

Rotations about one axis can affect rotations about other axes. For example, consider an object aligned with the X, Y and Z axes that is rotated by 30°about the object's local Y axis, corresponding to a rotation of 30°about the global Y axis. Now consider the same object first rotated by 90°about the global X axis. Now in order to obtain the same 30°local Y axis rotation we must rotate by 30°about the global Z axis. Translations along one axis can never affect translations along another axis.

The order of rotations is important. For example, a 45°rotation about the Y-axis followed by a 90°rotation about the X-axis is equivalent to a 90°rotation about the X-axis followed by a 45°rotation about the Z-axis. Mathematically, the computation of the X, Y and Z axis rotations to obtain a given orientation requires that a specific ordering be imposed. A subject might chose any order. From a conceptual point of view this means that there may not be an exact correspondence between the rotations a subject might make and the mathematically determined rotations. In the case of position errors, the order of translation is irrelevant and can never cause any confusion.

Any combination of rotations about the X, Y and Z axes can be represented as a single arc rotation about some axis. In this representation the rotation is not disproportionately biased toward any of the axes. The mathematical computation that determines the X, Y and Z axis rotations that would produce the same final orientation as the single arc rotation can produce a disproportionate bias. In this computation, the maximum Y-axis rotation that can result is 90°. For the X and Z axes, the maximum rotations are 180°.

In order to analyze rotation error for some sort of X, Y, or Z axis tendency, an unbiased mechanism is needed for determining the amount of axis rotation error. The mechanism used here is to consider the axis of the single arc rotation as a unit vector and then compare the magnitudes of the vector components.

Y-axis orientations were much harder to match

The analysis for total rotation error shows that two target orientations are much more difficult to match than any of the others. These two orientation conditions are the ±45°rotations about the Y axis. This happens to be the axis of symmetry of the object. However, this is unlikely to be the primary reason why these conditions resulted in the poorest performance. These two particular orientations result in only a single face of the target being visible to the subject, perhaps making it difficult to find any landmark on the target to aid in an orientation match. This is consistent with subjective responses where almost all subjects indicated that orientations with only a single face visible were the most difficult to match.

Overall, the total rotation error was 9.39°. When the $\pm 45°$Y conditions are removed the total rotation error drops to 7.46°. This is consistent with a study of 2D rotation techniques by Chen [26] and a recent study comparing 2D and 3D techniques by Hinckley et al. [53]. It should be noted that in the study by Hinckley et al. subjects were given feedback after each match, while in our study no feedback was provided.

It is foreseeable that in a typical interactive setting there are likely to be orientation conditions that need to be matched where the visual appearance makes this difficult. In a study of object placement by Ware [103], subjects were allowed to orient the scene to find a more favourable viewing condition. A study of 3D manipulation techniques by Phillips et al. [75] investigated automatic scene rotation to afford a better viewing angle. While these approaches may be workable in some situations, there are likely to be conditions where subjects are unwilling to accept or unable to make viewpoint changes. Further work is needed to determine a means of improving orientation matching performance in such circumstances.

Orientation error biased toward the X-axis

When we look at the individual axis errors for the $\pm 45°$Y orientations we find that the majority of the error is about the X axis and not about the axis of rotation. An examination of the overall rotation errors across all target orientations indicates that the Y axis had the least amount of error while the X axis had the most error.

In the case of position error, Z axis performance is poorest, because it relies on depth cues rather than position cues. Zhai argues that for rotation errors performance about the Z axis should be best because these rotations do not change the depth of any portion of the target, while rotations about the X and Y axes do alter the depth of portions of the target. The findings presented here are somewhat consistent with his in the sense that the X-axis orientation error was found to be the largest. However, whereas Zhai found the Z rotation error to be smallest, the Y-axis orientation error was found to be smallest in this experiment. It should be noted that Zhai used a tracking task with wireframe objects and semi-transparency, while our task was docking with opaque objects. Great care must be taken in interpreting this result because the X rotation error is heavily biased by performance in the $\pm 45°$Y conditions. In the other orientation conditions, the X and Z rotation errors are roughly equivalent, with a clear trend towards the Y rotation error being smallest.

RMS Error

RMS error was reduced in the stereoscopic display mode. This was expected in light of the strong influence that stereoscopic viewing has upon the task. Aside from the two Z-axis target positions, the left target position showed the least RMS error and thus indicates the closest adherence to an imaginary line segment between the start and end positions. The default orientation had the least amount of RMS error (it also had the shortest path length) and this would seem to indicate that subjects were best able to stick to an optimal movement path when little orientation change was required.

Considerations for RMS error

When examining the effect of target position upon RMS error it is important to consider how the RMS error is computed. To compute RMS error it is necessary to compute the position error for each individual point in the movement. This is accomplished by finding the closest point to the pointer on the line connecting the start position to the end position. The distance between the pointer and the closest point represents the error for that point in the movement.

For the two targets along the X-axis the amount of X error for any pointer location between the start and end point will be zero. The X error can only be non-zero if the pointer goes beyond the start or end points of the line segment. For the two Z-axis targets this means that the Z-axis error will be zero as long as the Z coordinate of the pointer is between the Z coordinates of the start and end positions. As the Z-axis error is generally much larger than the X and Y axis errors, it follows that in computation of the RMS error based upon target position the two Z-axis target positions will benefit.

5.8.3 Conciseness

The average path length in the monoscopic display mode was 24.90 cm. The stereoscopic display mode reduced the average path length to 20.09cm, a reduction of 4.81cm or 19% of movement. As was the case for trial completion time, the significant target position $\times$ target orientation interaction revealed a complex pattern of influence for target positions and target orientations.

Increased orientation difficulty corresponds with increased pointer movement

In general the pattern of path lengths across orientation conditions appears similar to the pattern for trial completion time across orientation conditions. The default orientation generally had the shortest path length. The longest path lengths appeared in the $\pm 45°$Y conditions. It appears that greater difficulty in matching the target orientation results in greater translation of the pointer. During the training sessions, subjects were observed to have moved the pointer away from the target to check their orientation match and then move the pointer back to the target when they were satisfied with their orientation to match the target position.

Amount of pointer orientation varied across target orientations

The results for orientation length showed that the least amount of rotation generally occurred for the default target orientation. The most rotation generally occurred for the $\pm$ 45°Y conditions. The other target orientation conditions range between these extremes. Once again the significant interaction indicated a complex pattern of influence for the target position and target orientation conditions.

The reduced orientation length for the default target position suggests that subjects brought the orientation of the pointer into close correspondence with the orientation of the homing target even though they were not required to do so.

Stereoscopic viewing provides little benefit to rotation

Examining all the dependent variables related to orientation showed almost no effect for display mode. The only significance of any kind is in the display mode $\times$ target position interaction in the analysis for orientation error. This would indicate that stereoscopic viewing was of little benefit for orientation based tasks. The improvement in trial completion time for stereoscopic viewing was a result of the strong effect upon the positioning component of the task. Orientation error was mostly influenced by the orientation of the target and to some degree by the position of the target.

5.8.4 Felicity

Subjective responses regarding the effectiveness of the display mode for the task were generally in line with measured performance although a few interesting results did appear. A number of subjects (38%) felt that the monoscopic condition was easier to use than the stereoscopic condition in contrast with overall performance. A few subjects indicated on their questionnaires that the orientation component of the task was more difficult than the position component and that stereo was not helpful for the orientation component. If subjects based their determination for ease of use upon the component of the task they felt to be more difficult, then it would follow that those who thought the orientation component was more difficult would see little benefit for stereoscopic viewing.

Subjects became more definitive about the target orientations that were easier and harder in the stereoscopic display mode. This might indicate that stereoscopic viewing made the positioning component of the task easier and allowed subjects to focus more attention on the orientation component of the task. Subjects felt that certain orientations made position matching easier, but there were no positions that made orientation matching easier.

5.9 Conclusions

The goal of this experiment was to achieve a better understanding of the effects of display mode, target position and target orientation on a docking task. The most important findings are:

- Stereoscopic viewing appears to have little influence upon the orientation component of the task.
- Target position and target orientation have a significant effect upon performance. This suggests that hidden biases might exist within studies that do not explicitly consider the position and orientation of the target.
- Differences in performance by orientation may be more closely coupled to the object than the axis of rotation. On a per axis basis orientation errors seem unrelated to the axis of rotation of the object.

Finally, it seems that even though overall subjective measures are generally consistent with objective measures, stark inconsistencies exist. One of the most noteworthy in this experiment being that three of eight subjects found the monoscopic display mode easier to use than the stereoscopic display mode.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	7.457	1	7.457	5.84	0.0463		
Error	8.935	7	1.276				
position	1.091	5	0.218	4.26	0.0039	0.0068	0.85
Error	1.793	35	0.051				
d$\times$p	0.173	5	0.035	1.50	0.2134	0.2512	0.47
Error	0.806	35	0.023				
orientation	2.895	6	0.482	6.93	< 0.0001	< 0.0001	0.96
Error	2.922	42	0.070				
d$\times$o	0.372	6	0.062	2.64	0.0291	0.0291	1.00
Error	0.987	42	0.024				
p$\times$o	1.379	30	0.046	3.51	< 0.0001	< 0.0001	0.98
Error	2.751	210	0.013				
d$\times$p$\times$o	0.614	30	0.020	1.28	0.1605	0.1991	0.68
Error	3.354	210	0.016				
Total	76.766						

Table 5.16: Experiment 2 – Repeated measures ANOVA for log of trial completion time.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	37.559	1	37.559	10.56	0.0141		
Error	24.906	7	3.558				
position	5.349	5	1.070	4.53	0.0028	0.0143	0.58
Error	8.272	35	0.236				
d$\times$p	2.253	5	0.451	2.16	0.0813	0.1157	0.66
Error	7.309	35	0.209				
orientation	7.206	6	1.201	3.58	0.0059	0.0140	0.74
Error	14.088	42	0.335				
d$\times$o	2.870	6	0.478	1.94	0.0966	0.1496	0.53
Error	10.359	42	0.247				
p$\times$o	5.628	30	0.188	1.49	0.0570	0.0570	1.00
Error	26.452	210	0.126				
d$\times$p$\times$o	4.953	30	0.165	1.18	0.2447	0.2451	1.00
Error	29.290	210	0.139				
Total	280.521						

Table 5.17: Experiment 2 – Repeated measures ANOVA for log of final position error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.035	1	0.035	0.24	0.6399		
Error	1.030	7	0.147				
position	1.864	5	0.373	4.08	0.0051	0.0156	0.67
Error	3.200	35	0.091				
d×p	0.694	5	0.139	2.81	0.0309	0.0360	0.92
Error	1.729	35	0.049				
orientation	63.870	6	10.645	15.05	< 0.0001	< 0.0001	0.51
Error	29.716	42	0.708				
d×o	1.384	6	0.231	1.99	0.0884	0.0884	1.00
Error	4.863	42	0.116				
p×o	4.638	30	0.155	1.60	0.0308	0.0614	0.66
Error	20.284	210	0.097				
d×p×o	1.980	30	0.066	1.29	0.1564	0.1740	0.84
Error	10.767	210	0.051				
Total	173.863						

Table 5.18: Experiment 2 – Repeated measures ANOVA for log of final orientation error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	6.960	1	6.960	8.27	0.0238		
Error	5.894	7	0.842				
position	0.984	5	0.197	2.65	0.0391	0.0532	0.81
Error	2.597	35	0.074				
d×p	0.286	5	0.057	2.60	0.0424	0.0573	0.80
Error	0.772	35	0.022				
orientation	2.390	6	0.398	6.49	0.0001	0.0019	0.55
Error	2.579	42	0.061				
d×o	0.184	6	0.031	1.84	0.1148	0.1265	0.88
Error	0.701	42	0.017				
p×o	0.846	30	0.028	1.97	0.0031	0.0344	0.41
Error	3.000	210	0.014				
d×p×o	0.338	30	0.011	1.16	0.2720	0.3086	0.58
Error	2.047	210	0.010				
Total	46.559						

Table 5.19: Experiment 2 – Repeated measures ANOVA for log of path length.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.398	1	0.398	0.33	0.5844		
Error	8.486	7	1.212				
position	0.083	5	0.017	0.15	0.9777	0.9777	1.00
Error	3.775	35	0.108				
d×p	0.331	5	0.066	2.38	0.0587	0.0705	0.86
Error	0.974	35	0.028				
orientation	58.470	6	9.745	24.13	$<$ 0.0001	$<$ 0.0001	0.79
Error	16.961	42	0.404				
d×o	0.782	6	0.130	1.68	0.1503	0.1741	0.75
Error	3.261	42	0.078				
p×o	3.406	30	0.114	2.56	0.0001	0.0005	0.72
Error	9.295	210	0.044				
d×p×o	0.907	30	0.030	1.02	0.4425	0.4413	0.61
Error	6.217	210	0.030				
Total	177.674						

Table 5.20: Experiment 2 – Repeated measures ANOVA for log of orientation length.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	22.415	1	22.415	19.30	0.0032		
Error	8.129	7	1.161				
position	31.393	5	6.279	11.68	$<$ 0.0001	$<$ 0.0001	1.00
Error	18.809	35	0.537				
d×p	8.423	5	1.685	7.06	0.0001	0.0006	0.76
Error	8.356	35	0.239				
orientation	2.915	6	0.486	4.00	0.0030	0.0077	0.75
Error	5.103	42	0.122				
d×o	1.082	6	0.180	2.31	0.0516	0.0645	0.84
Error	3.283	42	0.078				
p×o	3.096	30	0.103	1.77	0.0112	0.0363	0.59
Error	12.234	210	0.058				
d×p×o	0.903	30	0.030	0.65	0.9170	0.8187	0.49
Error	9.669	210	0.046				
Total	210.690						

Table 5.21: Experiment 2 – Repeated measures ANOVA for log of RMS error.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	21.337	1	21.337	6.66	0.0365		
Error	22.441	7	3.206				
position	22.901	5	4.580	20.47	< 0.0001	< 0.0001	0.61
Error	7.831	35	0.224				
d×p	3.037	5	0.607	2.93	0.0260	0.0394	0.79
Error	7.261	35	0.207				
orientation	9.508	6	1.585	4.00	0.0030	0.0109	0.67
Error	16.646	42	0.396				
d×o	3.596	6	0.599	2.04	0.0810	0.0810	1.00
Error	12.324	42	0.293				
p×o	6.730	30	0.224	1.09	0.3470	0.3470	1.00
Error	43.107	210	0.205				
d×p×o	6.323	30	0.211	1.03	0.4316	0.4328	0.78
Error	43.023	210	0.205				
axis	577.772	2	288.886	116.36	< 0.0001	< 0.0001	0.53
Error	34.759	14	2.483				
d×a	46.542	2	23.271	19.10	0.0001	0.0021	0.56
Error	17.058	14	1.218				
p×a	21.062	10	2.106	8.74	< 0.0001	< 0.0001	0.68
Error	16.875	70	0.241				
d×p×a	6.484	10	0.648	2.84	0.0051	0.0076	0.88
Error	15.984	70	0.228				
o×a	10.751	12	0.896	3.68	0.0002	0.0006	0.81
Error	20.435	84	0.243				
d×o×a	2.686	12	0.224	1.06	0.4002	0.3961	0.40
Error	17.663	84	0.210				
p×o×a	10.174	60	0.170	1.39	0.0344	0.0808	0.57
Error	51.093	420	0.122				
d×p×o×a	9.602	60	0.160	1.10	0.2938	0.2938	1.00
Error	61.099	420	0.145				
Total	1302.071						

Table 5.22: Experiment 2 – Repeated measures ANOVA for log of final position axis error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.000	1	0.000	0.01	0.9449		
Error	0.304	7	0.043				
position	0.254	5	0.051	1.10	0.3800	0.3796	0.95
Error	1.620	35	0.046				
d×p	0.341	5	0.068	2.50	0.0487	0.0523	0.95
Error	0.952	35	0.027				
orientation	2.602	6	0.434	1.85	0.1132	0.1511	0.64
Error	9.867	42	0.235				
d×o	1.358	6	0.226	3.60	0.0057	0.0274	0.53
Error	2.637	42	0.063				
p×o	2.247	30	0.075	1.32	0.1366	0.1459	0.92
Error	11.950	210	0.057				
d×p×o	1.461	30	0.049	1.17	0.2546	0.2599	0.94
Error	8.713	210	0.041				
axis	211.742	2	105.871	132.13	$<$ 0.0001	$<$ 0.0001	1.00
Error	11.218	14	0.801				
d×a	1.351	2	0.675	6.92	0.0081	0.0081	1.00
Error	1.365	14	0.098				
p×a	2.598	10	0.260	1.43	0.1851	0.1851	1.00
Error	12.708	70	0.182				
d×p×a	0.913	10	0.091	0.72	0.6997	0.6997	1.00
Error	8.832	70	0.126				
o×a	91.690	12	7.641	18.78	$<$ 0.0001	$<$ 0.0001	1.00
Error	34.172	84	0.407				
d×o×a	2.948	12	0.246	1.58	0.1126	0.1206	0.92
Error	13.045	84	0.155				
p×o×a	11.333	60	0.189	1.42	0.0277	0.0375	0.86
Error	55.927	420	0.133				
d×p×o×a	6.678	60	0.111	1.17	0.1928	0.2249	0.73
Error	39.958	420	0.095				
Total	554.402						

Table 5.23: Experiment 2 – Repeated measures ANOVA for log of final orientation axis error.

Chapter 6

Experiment 3: Line Tracing

The line tracing experiment was the third of four experiments to be carried out. The line tracing task required subjects to move the tip of a pointer from a starting position to an end position while keeping the tip of the pointer as close to a line (shown on the display as a narrow cylinder) as possible. In the line tracing task subjects were free to reorient the pointer in any manner they might choose.

6.1 Hypotheses

The experimental hypotheses of this experiment are that display mode (monoscopic versus stereoscopic), head-tracking mode (fixed viewpoint versus head-tracked viewpoint), target position ($\pm$10cm along one of the X, Y or Z axes from a fixed starting location) and block (four levels, representing the successive blocks of trials within a particular session) have an effect upon performance in a line tracing task. To evaluate the effect of the independent variables upon the task several different dependent measures were gathered.

The following six dependent measures were analyzed in this experiment: trial completion time, final position error, final position axis error, RMS error, RMS axis error and path length. A detailed description of these measures appears in Section 3.1. RMS error is the root mean square of the distance from the tip of the pointer to the closest point on the ideal path. For this experiment the ideal path used to compute the RMS error is considered to be the straight line from the starting position to the target position.

For statistical testing, an α level of 0.05 was used to determine whether or not the independent variables had a significant effect.

6.2 Participants

Subjects were recruited as specified in Chapter 3. One potential subject was excluded because he could not fuse the stereo pairs in the stereo screening test. Eight subjects were accepted into the experiment, none of whom had participated in the earlier experiments.

6.3 Equipment

The standard equipment described earlier in Section 3.3 was used, including the head-tracker. Figure 3.1 shows the hardware used in this experiment.

A long narrow cylinder was used to connect the start and end points rather than a single pixel straight line. In order to make the lighting more natural a white light source was used to illuminate the cylinder. Because of the lighting a small amount of green may appear on the display. To mask potential bleeding a dark gray background was used rather than the black background used in earlier experiments. Figure 6.1 shows an image of the screen from one of the trials of this experiment.

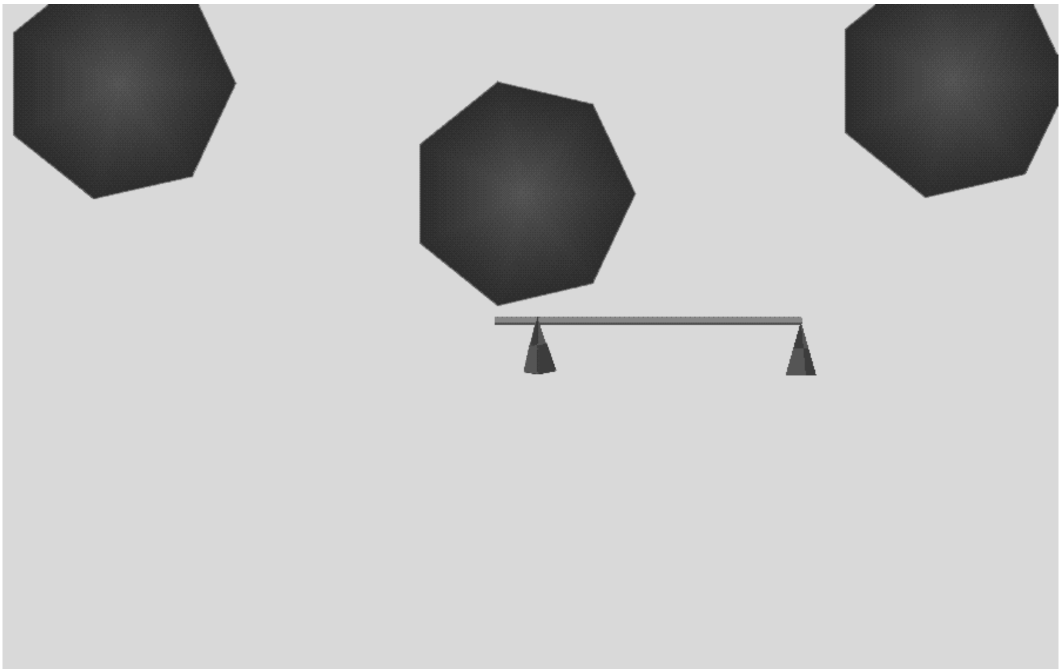


Figure 6.1: An image of the screen during a trial of the line tracing experiment. The cylinder representing the line to be traced is visible with the trial target at the right-end and the pointer moving along the line from the left-end. The large shaded objects toward the top of the image are a part of the vection background.

6.4 Procedure

Subjects were required to perform a line tracing task in a 3D environment. The pointer had six degrees of freedom, but to accomplish the task, only the position of the tip of the pointer was important. The pointer and target in this task were identical tetrahedron-like objects, in which the height (1.73cm) was twice the width (0.87cm) and one face was perpendicular to the base. The difference in size along one dimension allows one vertex to be easily identified as the tip and one face to be identified as the base. This is the same pointer used in the docking experiment, and only differs from the pointer in the point location experiment in terms of colouration as described in Section 3.3.

In the default orientation the tip points upward along the positive Y-axis, the base was parallel to the XZ plane and the perpendicular face was to the back and parallel to the XY plane. The base was half intensity magenta, the back face was full intensity magenta and the two front faces used both full intensity red and full intensity blue. The two front faces had a different colour in their top and bottom halves. The top half of one face was red and the bottom half of the same face was blue. The opposite face had the reverse colouration. Figure 3.3 illustrates the appearance of the pointer and target in different orientations.

Subjects manipulated the pointer via the Polhemus input device using the dominant right hand, and used the middle mouse button with the mouse held in the left hand to advance through the experiment trials. A trial consisted of two phases, an initial homing phase and a tracing phase.

During the homing phase the subject had to move the tip of the pointer to the tip of the homing target. The homing target was located at a fixed centre point and was always in the default orientation. Subjects received feedback in the form of a red box that appeared around the tip of the homing target when the tip of the pointer was within 0.5cm of the homing target. When the red box appeared subjects could click the middle mouse button using the left hand to advance to the tracing phase; the homing target was removed and the trial target was displayed at one of the six positions. A narrow cylinder was drawn from the start position (the tip of the homing target) to the tip of the trial target. Subjects could not advance to the docking phase if the homing tolerance was not met. The pointer was not required to be in any specific orientation for the homing phase.

During the tracing phase subjects were instructed to move the pointer to the trial target while keeping the tip of the pointer as close to the line (cylinder) as possible. No feedback regarding proximity to either the line or trial target was given during the tracing phase. Subjects had to make their own determination of proximity using the visual cues available within the particular experiment condition. Subjects middle clicked with the mouse when they were satisfied with their performance. After a short delay the system advanced to the next trial.

Small viewing shifts in the spirit of Phillips et al. [75] to the left or right were used for the two Z-axis target positions (the z-near and z-far positions) when the viewpoint was fixed. The direction of the shift to the left or right was chosen at random. Without these view shifts the cylinder would appear as a small round circle in the two Z-axis target positions.

An additional change to the display was to introduce a vection background similar to the one used by Arthur et al. [6]. The goal was to increase the perceived motion parallax in the head-tracked viewing conditions.

6.5 Design

The experimental design consisted of four within-subject independent variables: display mode, head-tracking mode, target position and session block. For compatibility between the stereoscopic and monoscopic conditions all target positions in the virtual environment were behind the screen surface and were displayed at a size appropriate to a perspective projection based on a nominal viewing position of 40cm in front of the screen surface in the fixed viewpoint condition. To mitigate the effects of the apparatus, subjects wore the stereo glasses and head-tracker in all conditions. The receiver of the Polhemus was encased in modeling material that was shaped and coloured to resemble the pointer in the virtual environment. The physical pointer is shown in Figure 3.2.

An experiment session was comprised of one of the four combinations of display and head-tracking modes. After each session, subjects completed a short questionnaire dealing only with that session. Each subject performed four sessions, thereby covering all four combinations of display and head-tracking modes. The combinations of display mode and head-tracking mode were counterbalanced across subjects according to a latin square where each condition was performed first, second, third or fourth an equal number of times. After the final session, each subject completed a longer questionnaire about all the conditions.

Each subject completed four consecutive sessions, one for each display condition. Each session was divided into five blocks, a training block followed by four experiment blocks. Each experiment block contained 30 trials consisting of five repetitions of each target position presented in random order. The training block was an abbreviation of an experiment block consisting of 20 trials. During the first 10 trials of the training block subjects were given feedback telling them when they were within 1cm of the line or trial target. No feedback was given for the remaining 10 trials of the training block. To minimize fatigue subjects were given a one minute break after each block.

Each subject performed a total of 480 experiment trials. Each subject spent approximately 90 minutes performing the experiment including time for breaks and time to complete the questionnaires, but not including the time spent on the screening tasks.

6.6 Training

Subjects were trained using the following script. The script served as a basis for the experimenter, but was not read directly.

> “Your task is to move your pointer from the starting point to the ending point as quickly and as accurately as possible while attempting to keep the tip of your pointer as close as possible to the line. The starting point is located at the tip of the stationary object on the screen. Move your pointer to the starting point now. Notice that when the tip of your pointer is close to the tip of the starting target a red box appears around the tip of the target.”

[*Experimenter waits while subject moves pointer to start target.*]

"The red box must be present for you to be able to start a trial. To start a trial click the middle mouse button with your left hand."

[*Experimenter waits for subject to click with the middle button.*]

"The starting target has disappeared and another target has appeared somewhere in the virtual environment along with a line that goes from where the starting target was to the tip of the target now visible. Move your pointer to the target while trying to keep the tip of your pointer as close to the line as you can."

"If you find that you have moved away from the line, try to move the pointer closer to the line. Don't move the pointer back to the start of the line."

[*Experimenter waits for subject to move pointer to target.*]

"When you are satisfied that you have reached the target with your pointer click the middle mouse button with your left hand. This will end the trial. The orientation of the pointer with respect to the line or target does not matter. You can have the pointer in any orientation you choose. All that matters is that you get the tip of the pointer as close as possible to the tip of the target while following the line."

"Please try to perform the task as quickly and as accurately as possible."

"Do you have any questions?"

The experimenter answered any questions that would not bias the subject's performance in any of the visual modes. Typical questions were, "Does the orientation of the pointer matter?" and "Does it matter what side of the pointer I have facing forward?" The answer to these questions was always no.

No instruction as to what was fast enough, or what was accurate enough was given, but the enforced accuracy requirement for the homing phase may have given subjects some guidance. If subjects asked whether they were to favour speed over accuracy or vice-versa they were told that they would have to decide for themselves how quickly or accurately they should perform the task.

The experimenter observed the subject during the training trials. When the training trials were over the experimenter asked for and answered any questions using the same guideline stated previously. The experimenter then told the subject that he would be on his own during the actual trials. The experimenter then reminded the subject to perform the trials as quickly and as accurately as possible, left the room and started the first experiment block. The experimenter re-entered the room after each block to check on the subject and tell the subject to rest his arm for at least a minute. After waiting for a minute the subject was asked if he was ready to proceed. If the subject responded in the affirmative then the experimenter left the room and initiated the next experiment block.

6.7 Results

The dependent variables stated earlier were gathered for each trial. Across all eight subjects a total of 3840 individual trials were completed.

Preliminary Screening

In an attempt to remove invalid trials, screening tests for outliers were performed. One trial had a trial completion time below 0.5s and a path length below 1cm. This trial was considered to be an erroneous click by the subject. There were no trials requiring more than 30s and only 11 trials required more than 20s. One trial had a path length over 60cm (81.46cm). Based on the results of the point location experiment all these trials were considered valid and only the trial with the erroneous click was excluded. The remainder of the analysis is based upon the 3839 remaining trials.

The experiment is a 2 (display mode) $\times$ 2 (head-tracking mode) $\times$ 6 (target position) $\times$ 4 (session block) design with repeated measures on all factors. For the remaining analysis, the repetitions of each target position performed by a subject within a single block are averaged together to yield a single score. The result is a series of 768 measures, one for each subject in each of the conditions.

Residual Analysis and Degree of Freedom Adjustments

Residuals were analysed for homogeneity of variance as described in Chapter 3 and demonstrated in Chapter 4. A log transform was applied to each dependent variable to make the variance more uniform and suitable for analysis. To account for violations of the sphericity assumption, the Huynh-Feldt adjusted degrees of freedom are used to test for significance.

Trial Completion Time

The repeated measures ANOVA for trial completion time revealed significant main effects for target position $F_{HF}(3.4,23.5) = 32.21$, $p < 0.0001$ and block $F_{HF}(2.5,17.2) = 7.15$, $p = 0.0038$ as well as a significant head-tracking mode $\times$ target position interaction $F_{HF}(5,35) = 3.10$, $p = 0.0201$. Full results of the ANOVA appear in Table 6.16 located at the end of the

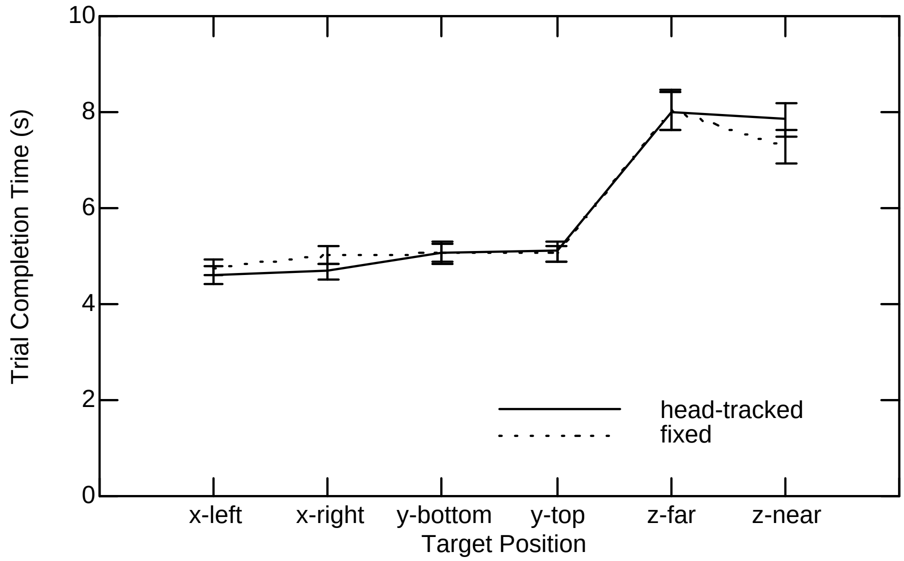


Figure 6.2: Mean of trial completion time in seconds for each target position in the two head-tracking modes.

chapter.

The effect for block indicates that there is a significant improvement in trial completion time for the task averaged over the other conditions. The average movement time for the first block is 6.22 seconds, dropping to 6.03s in the second block, 5.68s in the third block and 5.60s in the final block.

The significant effect for target position is considered together with the effect of head-tracking mode because of the significant interaction. Table 6.1 shows the trial completion time in each of the target positions for both head-tracking modes and Figure 6.2 illustrates the relationship between these factors. Many conditions show almost no difference between the head-tracked and fixed viewpoint conditions, however, there is a decrease in trial completion time for the head-tracked condition for the x-right target position and an increase in trial completion time for the z-near target position.

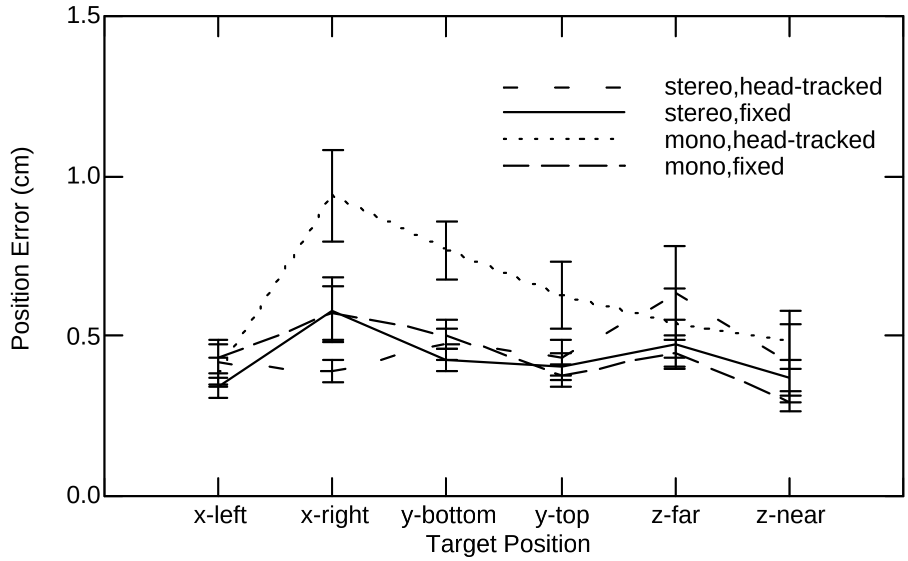


Figure 6.3: Mean position error in centimetres for each of the target positions in all the display mode and head-tracking mode combinations.

Final Position Error

The repeated measures ANOVA for final position error shows a significant main effect for target position $F_{HF}(5,35) = 5.90$, $p = 0.0005$, a significant two-way interaction for display mode $\times$ head-tracking mode $F_{HF}(1,7) = 12.11$, $p = 0.0103$ and a significant three-way interaction for display mode $\times$ head-tracking mode $\times$ target position $F_{HF}(5,35) = 4.09$, $p = 0.0050$. Full results of the ANOVA appear in Table 6.17 located at the end of the chapter.

Pos.	**Fixed**	**Head-tracked**	**Overall**
x-left	4.77	4.62	4.69
x-right	5.03	4.69	4.86
y-bottom	5.07	5.08	5.08
y-top	5.06	5.11	5.08
z-far	8.05	8.02	8.03
z-near	7.28	7.85	7.57

Table 6.1: Mean trial completion time in seconds for the target positions in both head-tracking modes and overall.

Due to the significant three-way interaction the display mode, head-tracking mode and target position factors are considered together. Table 6.2 gives the mean final position error for these factors and Figure 6.3 illustrates these factors. We can see that in the monoscopic conditions, head-tracking generally results in higher final position error, especially in the x-right target position. In the stereoscopic display mode head-tracked and fixed viewpoint performance is roughly equivalent except for the x-right target position that shows a marked improvement in the head-tracked condition and the z-far target position that shows a degradation in the head-tracked condition.

Final Position Axis Error

The repeated measures ANOVA for final position axis error reveals significant main effects for position $F_{HF}(5,35) = 4.50$, $p = 0.0028$ and axis $F_{HF}(1.5,10.4) = 36.50$, $p < 0.0001$. Significant two-way effects were found for display mode $\times$ head-tracking mode $F_{HF}(1,7) = 9.39$, $p = 0.0182$, display mode $\times$ axis $F_{HF}(1.2,8.3) = 9.57$, $p = 0.0121$ and position $\times$ axis $F_{HF}(10,70) = 8.15$, $p < 0.0001$. A significant three-way interaction was found for display mode $\times$ head-tracking mode $\times$ target position $F_{HF}(5,35) = 4.51$, $p = 0.0028$. Because

	Mono		**Stereo**		
Pos.	**Fixed**	**HT**	**Fixed**	**HT**	**All**
x-left	0.43	0.39	0.34	0.42	0.39
x-right	0.57	0.94	0.58	0.39	0.62
y-bottom	0.51	0.77	0.43	0.48	0.55
y-top	0.38	0.63	0.41	0.43	0.46
z-far	0.45	0.54	0.48	0.63	0.53
z-near	0.30	0.49	0.37	0.42	0.39
all	0.44	0.63	0.43	0.46	

Table 6.2: Mean of final position error in centimetres for each target position in both head-tracking modes and both display modes. The column labeled all provides the mean position error for each target position over all other factors and the row labeled all provides the mean final position error for the specific viewing condition.

the purpose of this analysis is to discover the possible effects of axis, only main effects or interactions involving the axis factor will be examined in detail. Full results of the ANOVA appear in Table 6.18 located at the end of the chapter.

Table 6.3 shows the final position error along each axis and Table 6.4 provides the final position error for each display mode along each axis. In each case the Z-axis error is clearly much higher than the X- and Y-axis errors which are roughly equivalent. From the display mode $\times$ axis interaction it is evident that the Z-axis error shows an improvement in the stereoscopic display mode, while there is almost no change in the X- and Y-axis errors.

Axis	**X**	**Y**	**Z**
Error	0.12	0.14	0.41

Table 6.3: Mean of final position error along the X, Y and Z axes.

Display Mode	**X**	**Y**	**Z**
mono	0.12	0.13	0.46
stereo	0.12	0.15	0.35

Table 6.4: Mean of final position error in centimetres along the X, Y and Z axes for both display modes.

Table 6.5 provides the final position error for each position along each axis and the relationship between these factors is illustrated in Figure 6.4. The Z-axis error is clearly greater than the X- and Y-axis errors in every case. Interestingly, while the x-left target po-

Pos.	**X**	**Y**	**Z**
x-left	0.12	0.11	0.32
x-right	0.16	0.12	0.54
y-bottom	0.10	0.20	0.44
y-top	0.09	0.17	0.44
z-far	0.14	0.13	0.45
z-near	0.13	0.11	0.38

Table 6.5: Mean of final position error in centimetres along the X, Y and Z axes for each of the target positions.

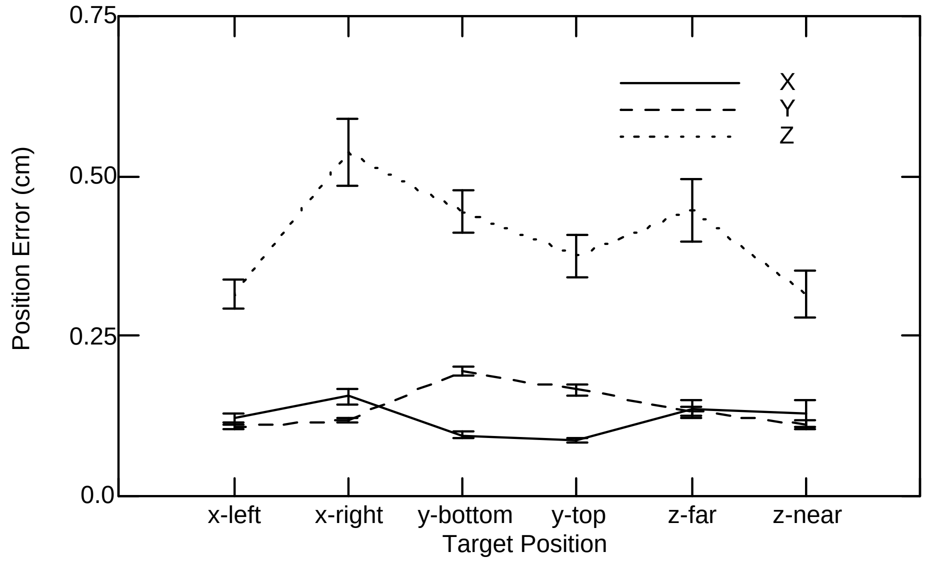


Figure 6.4: Mean of final position axis error in centimetres for each axis across all target positions.

sition has the least amount of Z-axis error, the x-right target position has the most Z-axis error. It is also evident while the amount of X- and Y-axis error appears roughly for the X- and Z-axis targets, the two Y-axis targets (y-bottom and y-top) exhibit a larger amount of Y-axis error.

Path Length

The repeated measures ANOVA for path length reveal significant main effects for target position $F_{HF}(4.5,31.5) = 23.87$, $p < 0.0001$ and block $F_{HF}(3,21) = 13.87$, $p < 0.0001$ as well as a significant head-tracking mode $\times$ target position interaction $F_{HF}(5,35) = 5.90$, $p = 0.0005$. Full results of the ANOVA appear in Table 6.19 located at the end of the chapter.

The significant main effect for block indicates a significant decrease in path length across the blocks. Table 6.6 shows the mean path length in each of the blocks.

The significant target position and head-tracking mode $\times$ target position interaction

Block	**1**	**2**	**3**	**4**
Path Length	14.62	14.29	13.96	13.84

Table 6.6: Mean path length in centimetres in each of the session blocks.

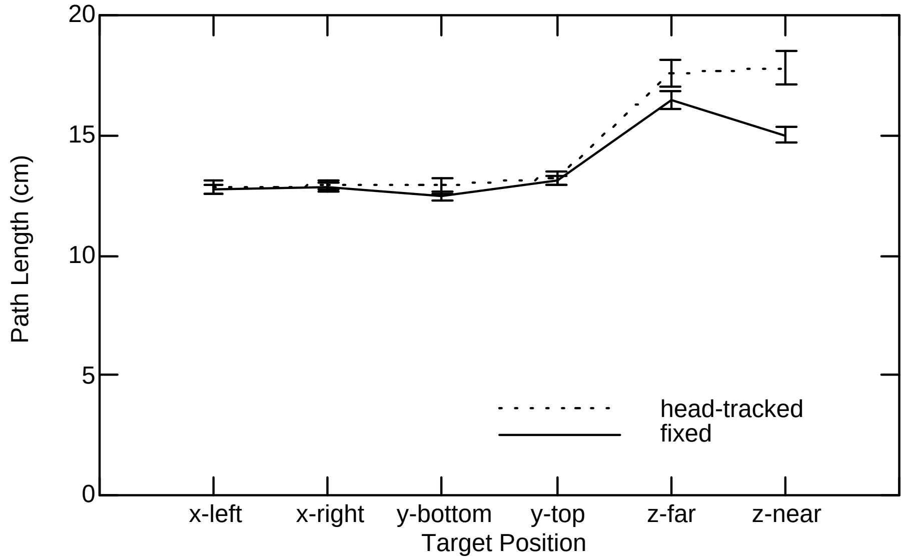


Figure 6.5: Mean path length in centimetres for each of the target positions in the two head-tracking modes.

are considered together. Table 6.7 provides the mean path length for each target position in both head-tracking modes and Figure 6.5 illustrates these values. We can see that the path length for the z-near and z-far target positions is higher than for the other target positions. In the head-tracked mode the path lengths for the z-near and z-far target positions increases dramatically, while remaining essentially unchanged for the other target positions.

RMS Error

The repeated measures ANOVA for RMS error indicates significant main effects for head-tracking mode $F_{HF}(1,7) = 7.62$, $p = 0.0281$ and target position $F_{HF}(5,35) = 3.44$, $p = 0.0124$ as well as a significant display mode $\times$ head-tracking mode interaction $F_{HF}(1,7) = 10.20$, p

= 0.0152 . Full results of the ANOVA appear in Table 6.20 located at the end of the chapter.

The main effect for head-tracking mode and the display mode $\times$ head tracking mode interaction are considered together. Table 6.8 gives the RMS error for each combination of display and head-tracking modes. From the table it is evident that the RMS error in the two fixed viewpoint modes is essentially equivalent. RMS error is larger in both head-tracked conditions than in the corresponding fixed viewpoint conditions. The interaction is a result of the smaller increase in RMS error for head-tracking in the stereoscopic display mode as compared to the monoscopic display mode.

Table 6.9 gives the RMS error in each of the target positions. From the table it is evident that the x-right target position has a much higher RMS error than any of the other target positions.

RMS Axis Error

The repeated measures ANOVA for RMS error along the axes reveals significant main effects for head-tracking mode $F_{HF}(1,7) = 5.59$, $p = 0.0500$, target position $F_{HF}(4.4,30.5) = 16.00$, $p < 0.0001$ and axis $F_{HF}(1.2,8.1) = 91.24$, $p < 0.0001$. Significant two-way interactions were found for display mode $\times$ head-tracking mode $F_{HF}(1,7) = 5.72$, $p = 0.0481$, display mode $\times$ axis $F_{HF}(1.4,9.9) = 8.15$, $p = 0.0116$ and target position $\times$ axis $F_{HF}(7.9,55.3)$

Pos.	**Fixed**	**HT**	**All**
x-left	12.77	12.84	12.80
x-right	12.88	12.92	12.90
y-bottom	12.50	12.93	12.72
y-top	13.11	13.25	13.18
z-far	16.50	17.60	17.05
z-near	15.03	17.82	16.42

Table 6.7: Mean path length in centimetres for each target position in both the head-tracked and fixed viewpoint conditions. The column labeled all provides the mean for each target position averaged over the other conditions.

Head-tracking Mode	Mono	Stereo
Fixed	0.54	0.55
Head-tracked	0.77	0.62

Table 6.8: Mean RMS error in centimetres for all combinations of display mode and head-tracking mode.

Pos.	RMS error
x-left	0.59
x-right	0.77
y-bottom	0.58
y-top	0.61
z-far	0.58
z-near	0.60

Table 6.9: Mean RMS error in centimetres for each of the target positions.

= 110.64, $p < 0.0001$. Because the purpose of this analysis is to detect how the RMS error differs across the axes, only the axis factor and its interactions will be examined in detail. Full results of the ANOVA appear in Table 6.21 located at the end of the chapter.

Table 6.10 gives the RMS error along each axis. The RMS error along the Z-axis is clearly much higher than along either the X- or Y-axes.

Axis	X	Y	Z
RMS Error	0.19	0.20	0.48

Table 6.10: Means of RMS error in centimetres along the X, Y and Z axes.

Table 6.11 gives the RMS error along each axis in both display modes. Stereoscopic viewing does not significantly change the RMS error for the X- and Y-axes. Stereoscopic viewing does reduce the RMS error along the Z-axis.

Table 6.12 gives the RMS error along each axis for all the target positions and Figure 6.6 illustrates the relationship of these factors. It is evident that for targets positions on a particular axis, the error along that axis is reduced. Thus there is a lower RMS error for the x-left and x-right target positions along the X-axis, the y-bottom and y-top target positions

Display Mode	X	Y	Z
mono	0.18	0.19	0.53
stereo	0.20	0.21	0.43

Table 6.11: Mean of RMS error in centimetres along the X, Y and Z axes for both display modes.

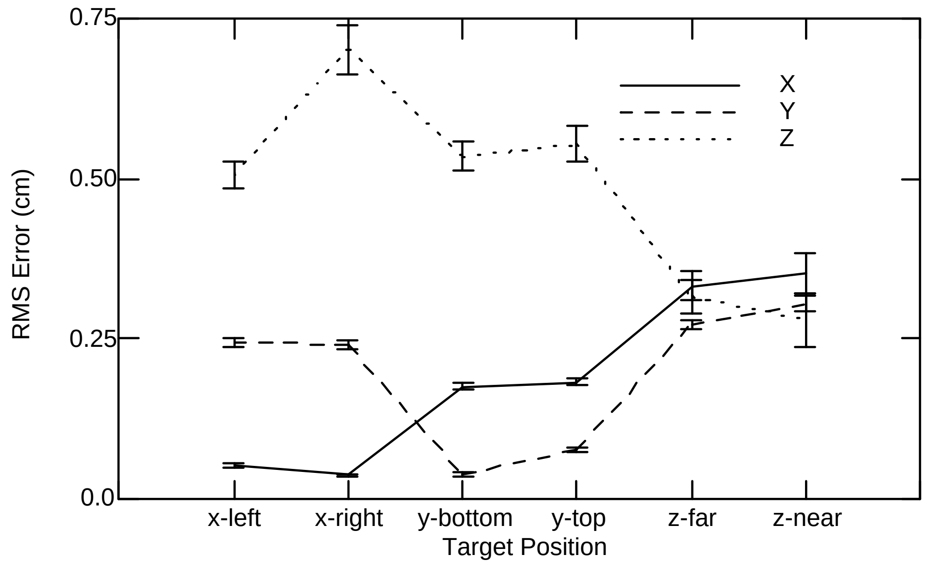


Figure 6.6: Mean of RMS axis error in centimetres for each of the axes across all the target positions.

along the Y-axis and the z-near and z-far targets along the Z-axis. For the two Z-axis targets, the X, Y and Z RMS error levels appear roughly equivalent. The x-right target position exhibits a Z-axis RMS error that is higher than any of the other target positions.

Questionnaires

At the end of each session subjects were required to answer a series of questions about different aspects of the task. Four questions required the subject to rate the difficulty of different components of the task. They were asked to indicate the difficulty of determining the position of the target and the line. They were also asked to indicate the difficulty of matching the target position and tracing the line. Each of these questions was answered on a 5 point scale where 1 is “easy” and 5 is “hard.” Four additional questions asked subjects to indicate if they found any target positions easier or harder to match, or whether they found any lines easier or harder to trace. If a subject indicated that a component was easier or harder, a free form answer was used to describe the particular conditions. A copy of the post-session questionnaire is located in Appendix A.5.

At the end of the experiment, subjects were asked to score all the display and head-tracking mode combinations according to task difficulty and ease of use. Additionally, at the end of the experiment subjects were asked to conduct three rankings of the four different viewing conditions. Subjects were asked to rank the viewing modes based on their preference, ease of use and their performance. Low rankings were best and high rankings were worst. A copy of the post-experiment questionnaire can be found in Appendix A.7.

Friedman rank tests failed to find any significant differences between the subjects' scores for any of the viewing conditions. Looking at the scores themselves for any possible trends, subjects indicated almost no difference between head-tracked and fixed viewpoint monoscopic conditions in the target related elements of the task, but indicated that head-tracking made the line related aspects of the task easier. This was also the case to a somewhat lesser degree in the stereoscopic conditions. Table 6.13 gives the mean scores for the post

Pos.	**X**	**Y**	**Z**
x-left	0.05	0.25	0.51
x-right	0.04	0.24	0.70
y-bottom	0.18	0.04	0.54
y-top	0.18	0.08	0.55
z-far	0.33	0.27	0.32
z-near	0.35	0.31	0.28

Table 6.12: Mean of RMS error in centimetres along the X, Y and Z axes for each of the target positions.

	Fixed		Head-tracked			
Question	**Mono**	**Stereo**	**Mono**	**Stereo**	**Friedman**	**p**
det. target pos.	2.4	2.1	2.4	2.0	2.36	0.501
match target pos.	2.8	2.8	2.9	2.5	1.61	0.657
det. line pos.	3.3	2.9	2.6	2.3	3.15	0.369
trace line	3.8	3.3	3.0	3.0	2.74	0.434
task difficulty	3.3	2.9	2.8	2.4	2.36	0.501
mode usefulness	3.1	2.8	2.1	2.3	1.99	0.575

Table 6.13: Means of subjects' responses to the subjective questions regarding display mode and head tracking mode conditions. The first four questions were answered immediately after a particular condition, and the last two questions were answered after all conditions had been completed. All questions were answered on a scale of 1 to 5 where 1 indicated easy or useful and 5 indicated hard or not useful.

session and post experiment questions.

For the post-experiment questions subjects indicated that the task was easiest in the head-tracked stereoscopic condition and hardest in the fixed viewpoint monoscopic. In terms of the usefulness of the display mode, subjects found the fixed viewpoint stereoscopic condition to be slightly more useful than the head-tracked stereoscopic condition.

The post-experiment rankings were generally inconclusive. Friedman's rank test did not indicate significance for any of the rankings. The counts for preference, ease of use and performance do not clearly indicate any clear bias for or against any of the viewing conditions. The strongest response was that five subjects felt that their performance was worst in the fixed viewpoint monoscopic viewing condition. Four subjects felt that their performance was best in the head-tracked stereoscopic viewing condition. Figure 6.7 provides summary information for subjects rankings.

Subjects' free form answers to the questions about the type of targets and lines were classified into general categories. Subjects tended to indicate that the targets in the plane parallel to the screen (XY plane) tended to be easier. Both targets and lines along the Z-axis were considered to be harder. The categorized summaries of subjects' answers can be found

Easier								
	Target				Line			
	Fixed		Head-tracked		Fixed		Head-tracked	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**
none	2	3	2	3	4	5	4	5
XY plane	4	2	3	3	3	1	2	1
X-axis	1	2	1	1	1	1	1	1
Y-axis	1		2	1			1	1
Z-axis		1				1		

Table 6.14: Coded summary of subjects' responses to the questions about which targets and lines were easier to match the position of or trace.

Harder								
	Target				Line			
	Fixed		Head-tracked		Fixed		Head-tracked	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**
none	2	4	2	3	2	3	4	4
XY plane			1					
X-axis	1							
Y-axis								
Z-axis	5	3	5	3	6	5	4	3
z-far		1		2				
z-near								1

Table 6.15: Coded summary of subjects' responses to the questions about which targets and lines were harder to match the position of or trace.

in Tables 6.14 and 6.15.

6.8 Discussion

In contrast with the point location and docking studies, this study failed to show a strong effect for stereoscopic viewing. Rather than a task where head-tracking would be beneficial, a task that had little benefit from stereoscopic viewing was found.

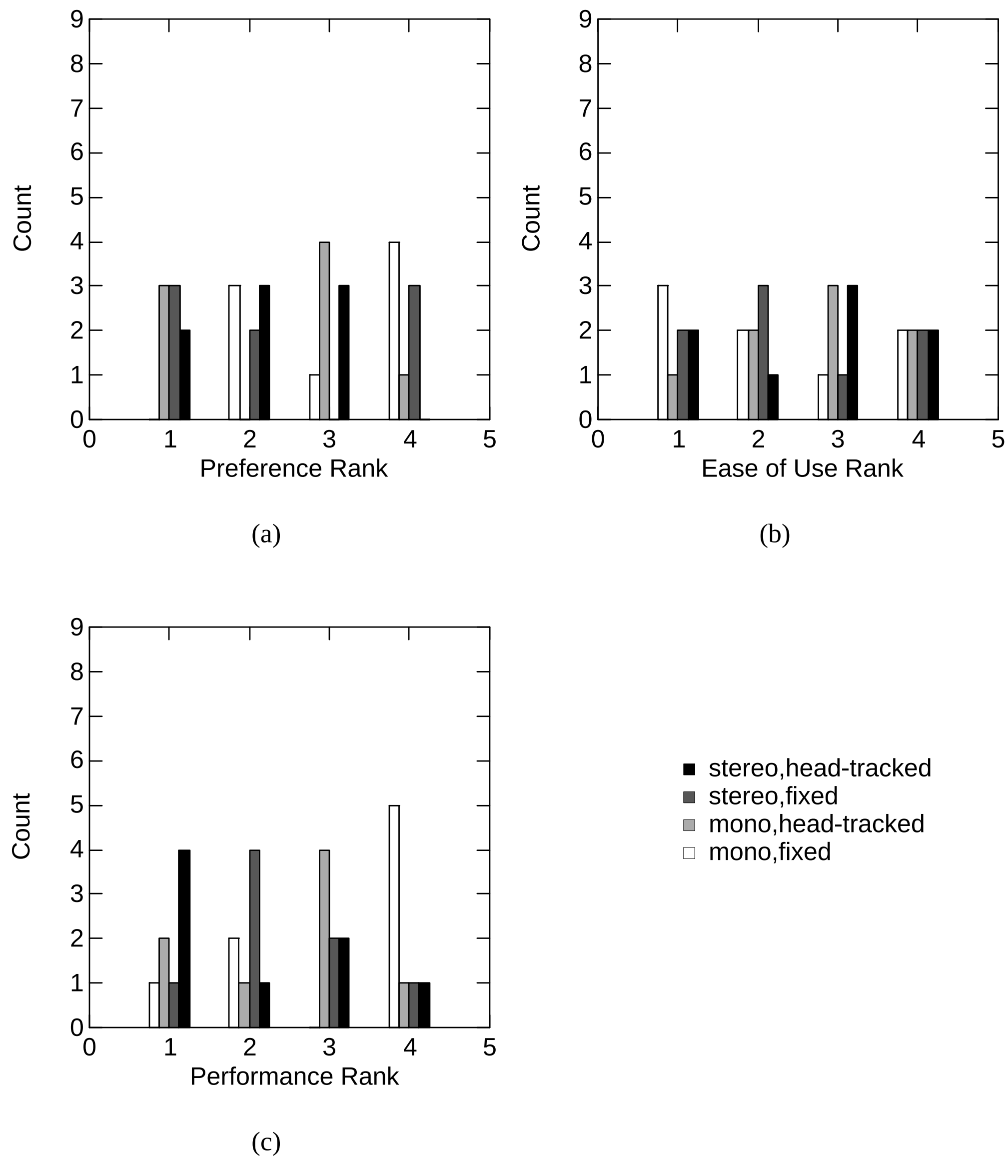


Figure 6.7: Rankings of subjects' preference, perceived ease of use and perceived performance for each of the viewing conditions. The height of a bar indicates the number of times a particular rank was assigned by a subject.

6.8.1 Speed

Z-axis target positions had slower performance

The results for trial completion time shows a strong difference for target position with the z-near and z-far targets clearly standing out as the poorer performers.

Display mode had no effect upon trial completion time

The surprising thing is the lack of an effect for display mode. One possible explanation is the presence of the line in the display. The line allows a subject to obtain occlusion based depth information throughout the task. The occlusion cue is a strong source of relative depth information. According to Wickens, [106] the occlusion cue is stronger than the stereoscopic depth cue.

Small effect for head-tracking was mixed

There is a small influence for head-tracking in the interaction of the head-tracking factor with target position. For the majority of target positions head-tracking provides no advantage or disadvantage. In the x-right target position head-tracking provides a slight benefit, but in the z-near target position head-tracking appears to be detrimental.

6.8.2 Accuracy

Faster performance for the x-right target may have come at the expense of accuracy

Final position error provides some additional insight here. The x-right target position in the head-tracked monoscopic condition has a much higher final position error. The implication is that there is not really much benefit provided by head-tracking for the x-right target position. With the only source of potential improvement for head-tracking eliminated the conclusion seems to be that head-tracking is either neutral or detrimental to the performance of the task.

Head-tracking impaired monoscopic accuracy, stereoscopic accuracy remained unchanged

Examination of the display mode $\times$ head-tracking mode interaction upon final position error shows that head-tracking is clearly detrimental in the monoscopic display mode and essentially neutral in the stereoscopic display mode. The positive outlook is that if stereoscopic viewing is used and subjects consider head-tracking to be beneficial then it can be used without degrading performance.

X-right target position showed higher Z-axis error

When axis is considered, it is clear that the Z-axis is the source of the majority of the position error. The stereoscopic display mode has a significant effect in reducing the amount of Z-axis error. An interesting anomaly is the high Z-axis error in the x-right target position. The x-right target position also had a higher Z-axis RMS error.

Head-tracking made it harder to stay close to the line

The results for RMS error support the findings of the other variables. Head-tracking is detrimental in both display modes. In the stereoscopic display mode the gap between the head-tracked and fixed viewpoint conditions was decreased, but head-tracking still increased RMS error. The x-right target position has a higher RMS error than the other target positions. Once again this target position stands out as being more difficult than the others.

X-right target position showed higher Z-axis RMS error

Examination of the axis RMS errors shows that the stereoscopic display mode reduces the RMS error along the Z-axis. In addition we can see that the x-right target position in the monoscopic condition stands out with much more Z-axis RMS error than the other target positions. Some property of moving to the x-right target position made it harder to control Z-axis movement in contrast to the x-left, y-bottom and y-top target positions.

6.8.3 Conciseness

As in earlier experiments path length indicates the increased difficulty of the z-near and z-far target positions. Head-tracking significantly degrades performance for these two positions.

Head-tracking hinders performance

The results from several dependent variables clearly indicate that head-tracking is generally detrimental to performance. This is the opposite of findings by Sollenberger & Milgram [89] and Arthur et al. [6]. Sollenberger and Milgram found a rotational display to be beneficial and the combination of stereoscopic viewing and rotational display to be even more beneficial. Arthur et al. duplicated these results but used head-tracking in place of a rotational display.

There are two principal differences between this work and the earlier work of these other researchers. One is the complexity of the display. The displays in this work are relatively simple with only a few objects (the pointer, the line to trace and the target) appearing in the display. The other difference is that in the earlier work subjects only had to view stimuli on the display and draw a conclusion. In this work subjects have to move a pointer around within the virtual environment in addition to visualizing the stimuli. One difficulty that subjects tend to have in this setting is keeping their hand steady while changing their view. In general, a fairly large head movement is needed in order to noticeably change the view. This causes a sizeable movement of a subject's shoulders and requires subjects to counteract this movement in order to keep the pointer steady. Both of these possibilities hold potential for future research. One avenue of inquiry would be to investigate what level of scene complexity is needed to make head-tracking either essential or beneficial. An alternative course of action would be to investigate at what level of scene complexity (if any) the benefit provided by head-tracking can overcome the need to coordinate head and hand movements.

6.8.4 Felicity

In contrast with the findings of Experiments one and two, none of the scores or rankings in this experiment were found to be significant. The scores and rankings focused upon the display and head-tracking modes. It is not too surprising that the subjective measures did not indicated a clear bias in light of the other results which showed little or no effect for the independent variables involved.

Examination of the scores in Table 6.13 shows a very slight bias towards the head-tracked stereoscopic condition. It seems likely that subjects exhibited some sort of bias towards the head-tracked viewpoint and stereoscopic display mode because of their novelty. If this was indeed the case then it indicates that subjects were able to determine during the course of the experiment that head-tracking and stereoscopic viewing provided little benefit.

6.9 Conclusion

One inevitable conclusion based upon several of the dependent measures is that head-tracking is more often detrimental to the performance of the line tracing task than it is beneficial. While the effect upon trial completion time is minimal, the effect on several of the error measures is significant.

Stereoscopic viewing had been expected to improve task performance for all the dependent measures, but based on the results, stereoscopic viewing is mostly effective at reducing the error level for the path related dependent measures.

The x-right target position appears to have been more difficult in this task than in the pointing and docking tasks studied earlier, although there were certain trends indicating weakness for this target position. This is an unusual finding and might indicate a task related weakness for the x-right target position. To verify this finding it would be useful to conduct a single experiment that focuses specifically upon task differences.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	2.709	1	2.709	4.10	0.0825		
Error	4.625	7	0.661				
head-track	0.019	1	0.019	0.01	0.9116		
Error	9.880	7	1.411				
d×h	0.114	1	0.114	0.25	0.6340		
Error	3.218	7	0.460				
position	33.326	5	6.665	32.21	< 0.0001	< 0.0001	0.67
Error	7.242	35	0.207				
d×p	0.307	5	0.061	0.77	0.5788	0.4774	0.38
Error	2.796	35	0.080				
h×p	0.444	5	0.089	3.10	0.0201	0.0201	1.00
Error	1.002	35	0.029				
d×h×p	0.156	5	0.031	1.27	0.2978	0.2986	0.98
Error	0.860	35	0.025				
block	1.279	3	0.426	7.15	0.0017	0.0038	0.82
Error	1.252	21	0.060				
d×b	0.216	3	0.072	1.38	0.2774	0.2837	0.73
Error	1.101	21	0.052				
h×b	0.195	3	0.065	1.02	0.4049	0.4032	0.96
Error	1.340	21	0.064				
d×h×b	0.059	3	0.020	0.25	0.8638	0.7903	0.68
Error	1.673	21	0.080				
p×b	0.273	15	0.018	1.42	0.1517	0.1930	0.64
Error	1.347	105	0.013				
d×p×b	0.077	15	0.005	0.43	0.9675	0.9291	0.67
Error	1.265	105	0.012				
h×p×b	0.232	15	0.016	1.30	0.2143	0.2620	0.53
Error	1.250	105	0.012				
d×h×p×b	0.156	15	0.010	1.09	0.3755	0.3755	1.00
Error	1.004	105	0.010				
Total	105.771						

Table 6.16: Experiment 3 – Repeated measures ANOVA for log of trial completion time.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	3.497	1	3.497	1.89	0.2113		
Error	12.931	7	1.847				
head-track	2.634	1	2.634	1.53	0.2559		
Error	12.046	7	1.721				
d×h	2.684	1	2.684	12.11	0.0103		
Error	1.551	7	0.222				
position	18.787	5	3.757	5.90	0.0005	0.0005	1.00
Error	22.305	35	0.637				
d×p	3.496	5	0.699	1.92	0.1157	0.1617	0.56
Error	12.742	35	0.364				
h×p	0.802	5	0.160	0.41	0.8378	0.8378	1.00
Error	13.652	35	0.390				
d×h×p	3.235	5	0.647	4.09	0.0050	0.0050	1.00
Error	5.543	35	0.158				
block	0.704	3	0.235	0.87	0.4715	0.4280	0.58
Error	5.652	21	0.269				
d×b	0.890	3	0.297	1.35	0.2858	0.2906	0.75
Error	4.624	21	0.220				
h×b	1.023	3	0.341	1.67	0.2033	0.2033	1.00
Error	4.280	21	0.204				
d×h×b	0.306	3	0.102	0.74	0.5378	0.5057	0.75
Error	2.876	21	0.137				
p×b	1.809	15	0.121	1.02	0.4384	0.4310	0.56
Error	12.381	105	0.118				
d×p×b	1.993	15	0.133	0.96	0.4989	0.4849	0.70
Error	14.488	105	0.138				
h×p×b	2.456	15	0.164	1.26	0.2405	0.2849	0.51
Error	13.641	105	0.130				
d×h×p×b	0.880	15	0.059	0.45	0.9597	0.8728	0.49
Error	13.717	105	0.131				
Total	338.773						

Table 6.17: Experiment 3 – Repeated measures ANOVA table for log of final position error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	1.166	1	1.166	0.45	0.5220		
Error	17.961	7	2.566				
head-track	7.272	1	7.272	2.00	0.2003		
Error	25.469	7	3.638				
d×h	3.031	1	3.031	9.39	0.0182		
Error	2.258	7	0.323				
position	30.211	5	6.042	4.50	0.0028	0.0028	1.00
Error	46.990	35	1.343				
d×p	4.130	5	0.826	1.47	0.2235	0.2239	0.99
Error	19.626	35	0.561				
h×p	1.672	5	0.334	0.39	0.8503	0.7979	0.74
Error	29.789	35	0.851				
d×h×p	7.874	5	1.575	4.51	0.0028	0.0028	1.00
Error	12.219	35	0.349				
block	0.792	3	0.264	0.51	0.6771	0.5848	0.58
Error	10.783	21	0.513				
d×b	0.936	3	0.312	0.89	0.4633	0.4451	0.78
Error	7.372	21	0.351				
h×b	1.945	3	0.648	1.43	0.2627	0.2627	1.00
Error	9.536	21	0.454				
d×h×b	0.180	3	0.060	0.19	0.9027	0.8901	0.92
Error	6.679	21	0.318				
p×b	3.980	15	0.265	0.88	0.5868	0.5696	0.81
Error	31.626	105	0.301				
d×p×b	3.991	15	0.266	0.93	0.5321	0.5321	1.00
Error	29.984	105	0.286				
h×p×b	4.310	15	0.287	1.05	0.4099	0.4104	0.63
Error	28.679	105	0.273				
d×h×p×b	2.631	15	0.175	0.59	0.8794	0.7947	0.57
Error	31.396	105	0.299				

Table 6.18: Experiment 3 – Repeated measures ANOVA for log of final position error along axes (continued on next page)

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
axis	494.703	2	247.352	36.50	$<$ 0.0001	$<$ 0.0001	0.74
Error	94.873	14	6.777				
d×a	11.509	2	5.754	9.57	0.0024	0.0121	0.59
Error	8.421	14	0.601				
h×a	1.945	2	0.973	1.61	0.2346	0.2355	0.97
Error	8.453	14	0.604				
d×h×a	1.096	2	0.548	1.51	0.2546	0.2570	0.89
Error	5.079	14	0.363				
p×a	43.420	10	4.342	8.15	$<$ 0.0001	$<$ 0.0001	1.00
Error	37.313	70	0.533				
d×p×a	4.810	10	0.481	1.49	0.1614	0.1896	0.73
Error	22.591	70	0.323				
h×p×a	3.752	10	0.375	1.37	0.2137	0.2137	1.00
Error	19.212	70	0.274				
d×h×p×a	1.969	10	0.197	0.97	0.4781	0.4781	1.00
Error	14.225	70	0.203				
b×a	0.748	6	0.125	0.63	0.7047	0.7047	1.00
Error	8.305	42	0.198				
d×b×a	1.000	6	0.167	0.85	0.5425	0.5425	1.00
Error	8.280	42	0.197				
h×b×a	1.112	6	0.185	0.81	0.5663	0.5523	0.86
Error	9.578	42	0.228				
d×h×b×a	0.704	6	0.117	0.62	0.7158	0.7158	1.00
Error	7.992	42	0.190				
p×b×a	3.939	30	0.131	0.92	0.5837	0.5837	1.00
Error	29.820	210	0.142				
d×p×b×a	4.977	30	0.166	1.08	0.3598	0.3598	1.00
Error	32.179	210	0.153				
h×p×b×a	4.599	30	0.153	0.98	0.4961	0.4926	0.88
Error	32.728	210	0.156				
d×h×p×b×a	3.083	30	0.103	0.72	0.8526	0.8526	1.00
Error	29.774	210	0.142				
Total	1612.315						

Table 6.18: Experiment 3 – Repeated measures ANOVA for log of final position error along axes (continued from previous page)

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	0.375	1	0.375	5.00	0.0605		
Error	0.525	7	0.075				
head-track	0.300	1	0.300	2.04	0.1960		
Error	1.027	7	0.147				
d×h	0.026	1	0.026	0.33	0.5849		
Error	0.561	7	0.080				
position	10.528	5	2.106	23.87	< 0.0001	< 0.0001	0.90
Error	3.088	35	0.088				
d×p	0.043	5	0.009	0.35	0.8811	0.7245	0.42
Error	0.863	35	0.025				
h×p	0.509	5	0.102	5.90	0.0005	0.0005	1.00
Error	0.604	35	0.017				
d×h×p	0.071	5	0.014	0.99	0.4396	0.4256	0.71
Error	0.504	35	0.014				
block	0.272	3	0.091	13.87	< 0.0001	< 0.0001	1.00
Error	0.137	21	0.007				
d×b	0.096	3	0.032	2.04	0.1388	0.1388	1.00
Error	0.328	21	0.016				
h×b	0.013	3	0.004	0.17	0.9172	0.8209	0.58
Error	0.539	21	0.026				
d×h×b	0.030	3	0.010	0.44	0.7296	0.7026	0.86
Error	0.486	21	0.023				
p×b	0.198	15	0.013	1.39	0.1639	0.2179	0.54
Error	0.993	105	0.009				
d×p×b	0.090	15	0.006	0.77	0.7109	0.6386	0.56
Error	0.821	105	0.008				
h×p×b	0.163	15	0.011	1.02	0.4378	0.4255	0.43
Error	1.118	105	0.011				
d×h×p×b	0.134	15	0.009	0.95	0.5140	0.4797	0.48
Error	0.985	105	0.009				
Total	31.270						

Table 6.19: Experiment 3 – Repeated measures ANOVA for log of path length.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	1.957	1	1.957	3.91	0.0886		
Error	3.507	7	0.501				
head-track	5.345	1	5.345	7.62	0.0281		
Error	4.908	7	0.701				
d×h	2.366	1	2.366	10.20	0.0152		
Error	1.624	7	0.232				
position	7.555	5	1.511	3.44	0.0124	0.0124	1.00
Error	15.365	35	0.439				
d×p	3.505	5	0.701	2.85	0.0290	0.0552	0.66
Error	8.597	35	0.246				
h×p	1.916	5	0.383	1.38	0.2539	0.2539	1.00
Error	9.689	35	0.277				
d×h×p	0.358	5	0.072	0.54	0.7466	0.7273	0.89
Error	4.660	35	0.133				
block	0.586	3	0.195	1.39	0.2747	0.2829	0.49
Error	2.959	21	0.141				
d×b	0.355	3	0.118	1.82	0.1747	0.1946	0.72
Error	1.368	21	0.065				
h×b	0.201	3	0.067	1.10	0.3717	0.3698	0.92
Error	1.280	21	0.061				
d×h×b	0.037	3	0.012	0.12	0.9463	0.9463	1.00
Error	2.147	21	0.102				
p×b	0.575	15	0.038	0.65	0.8305	0.7312	0.52
Error	6.236	105	0.059				
d×p×b	0.739	15	0.049	0.96	0.5006	0.4835	0.65
Error	5.382	105	0.051				
h×p×b	1.549	15	0.103	1.61	0.0820	0.1210	0.66
Error	6.713	105	0.064				
d×h×p×b	0.808	15	0.054	0.86	0.6046	0.5494	0.52
Error	6.538	105	0.062				
Total	140.166						

Table 6.20: Experiment 3 – Repeated measures ANOVA for log of RMS error.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.260	1	0.260	0.29	0.6074		
Error	6.300	7	0.900				
head-track	9.955	1	9.955	5.59	0.0500		
Error	12.460	7	1.780				
d×h	4.874	1	4.874	5.72	0.0481		
Error	5.967	7	0.852				
position	137.724	5	27.545	16.00	< 0.0001	< 0.0001	0.87
Error	60.273	35	1.722				
d×p	2.195	5	0.439	1.27	0.2980	0.3032	0.84
Error	12.084	35	0.345				
h×p	5.562	5	1.112	1.62	0.1810	0.1810	1.00
Error	24.062	35	0.687				
d×h×p	1.463	5	0.293	1.18	0.3380	0.3382	0.99
Error	8.668	35	0.248				
block	0.654	3	0.218	0.77	0.5256	0.4537	0.50
Error	5.976	21	0.285				
d×b	0.834	3	0.278	1.15	0.3512	0.3512	1.00
Error	5.064	21	0.241				
h×b	0.491	3	0.164	0.54	0.6619	0.6464	0.91
Error	6.395	21	0.305				
d×h×b	0.783	3	0.261	0.83	0.4906	0.4878	0.97
Error	6.580	21	0.313				
p×b	1.606	15	0.107	0.67	0.8066	0.8066	1.00
Error	16.739	105	0.159				
d×p×b	1.153	15	0.077	0.51	0.9293	0.9293	1.00
Error	15.768	105	0.150				
a h×p×b	4.140	15	0.276	1.34	0.1912	0.2200	0.72
Error	21.604	105	0.206				
d×h×p×b	1.781	15	0.119	0.62	0.8536	0.7961	0.68
Error	20.142	105	0.192				

Table 6.21: Experiment 3 – Repeated measures ANOVA for log of RMS axis error (continued on next page)

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
axis	564.585	2	282.293	91.24	$<$ 0.0001	$<$ 0.0001	0.58
Error	43.316	14	3.094				
d×a	8.656	2	4.328	8.17	0.0044	0.0116	0.71
Error	7.412	14	0.529				
h×a	3.420	2	1.710	3.46	0.0601	0.0601	1.00
Error	6.916	14	0.494				
d×h×a	0.690	2	0.345	1.87	0.1909	0.2015	0.79
Error	2.587	14	0.185				
p×a	1226.584	10	122.658	110.64	$<$ 0.0001	$<$ 0.0001	0.79
Error	77.603	70	1.109				
d×p×a	7.194	10	0.719	2.42	0.0153	0.0864	0.33
Error	20.777	70	0.297				
h×p×a	3.339	10	0.334	1.12	0.3596	0.3596	1.00
Error	20.855	70	0.298				
d×h×p×a	0.876	10	0.088	0.48	0.8950	0.8282	0.64
Error	12.668	70	0.181				
b×a	0.370	6	0.062	0.52	0.7872	0.7156	0.65
Error	4.946	42	0.118				
d×b×a	1.032	6	0.172	1.26	0.2978	0.2978	1.00
Error	5.749	42	0.137				
h×b×a	0.332	6	0.055	0.40	0.8773	0.8747	0.98
Error	5.862	42	0.140				
d×h×b×a	0.312	6	0.052	0.49	0.8119	0.7660	0.76
Error	4.460	42	0.106				
p×b×a	2.819	30	0.094	1.00	0.4669	0.4651	0.86
Error	19.655	210	0.094				
d×p×b×a	2.583	30	0.086	0.84	0.7056	0.6704	0.73
Error	21.485	210	0.102				
h×p×b×a	4.602	30	0.153	1.48	0.0587	0.0989	0.65
Error	21.710	210	0.103				
d×h×p×b×a	2.299	30	0.077	0.75	0.8290	0.6840	0.35
Error	21.570	210	0.103				
Total	2569.153						

Table 6.21: Experiment 3 – Repeated measures ANOVA for log of RMS axis error (continued from previous page)

Chapter 7

Experiment 4: Curve Tracing

The curve tracing experiment was the last of four experiments to be carried out to investigate the effects of display mode, head-tracking and target position on a 3D interactive task. The curve tracing task required subjects to move the tip of a pointer from a starting position to an end position while keeping the tip of the pointer as close to a two dimensional planar curve as possible. The curve was always a single wavelength of a sine curve. The curve was approximated for rendering purposes by a series of connected cylindrical segments. In the curve tracing task as in the line tracing task subjects were free to reorient the pointer.

7.1 Hypotheses

The experimental hypotheses of this experiment are that display mode (monoscopic versus stereoscopic), head-tracking mode (fixed viewpoint versus head-tracked viewpoint), target position ($\pm$10cm along one of the X, Y or Z axes from a fixed starting location) and block (four levels, representing the successive blocks of trials within a particular session) have an effect upon performance in a curve tracing task. To evaluate the effect of the independent variables upon the task several different dependent measures were gathered.

The following six dependent measures were analyzed in this experiment: trial completion time, final position error, final position axis error, RMS error, RMS axis error and path length. A detailed description of these measures appears in Section 3.1. RMS error is the root mean square of the distance from the tip of the pointer to the closest point on the ideal path. For this experiment the ideal path used to compute the RMS error is considered to be the sine wave from the starting position to the target position.

For statistical testing, an α level of 0.05 was used to determine whether or not the independent variables had a significant effect.

7.2 Participants

Subjects were recruited as specified in Chapter 3. One potential subject was excluded for failing the handedness screening. A second potential subject had to be replaced because of a fire alarm that occurred part way through a session. A third potential subject had to be replaced because of a data loss. Eight subjects were accepted into the experiment, none of whom participated in any of the earlier experiments.

7.3 Equipment

The standard equipment described earlier in Section 3.3 was used, including the head-tracker. Figure 3.1 shows the hardware used in this experiment.

A sine wave composed of 12 narrow cylindrical segments was used to connect the start and end points rather than a single pixel curve. In order to make the lighting more natural a white light source was used to illuminate the curve. Because of the lighting a small

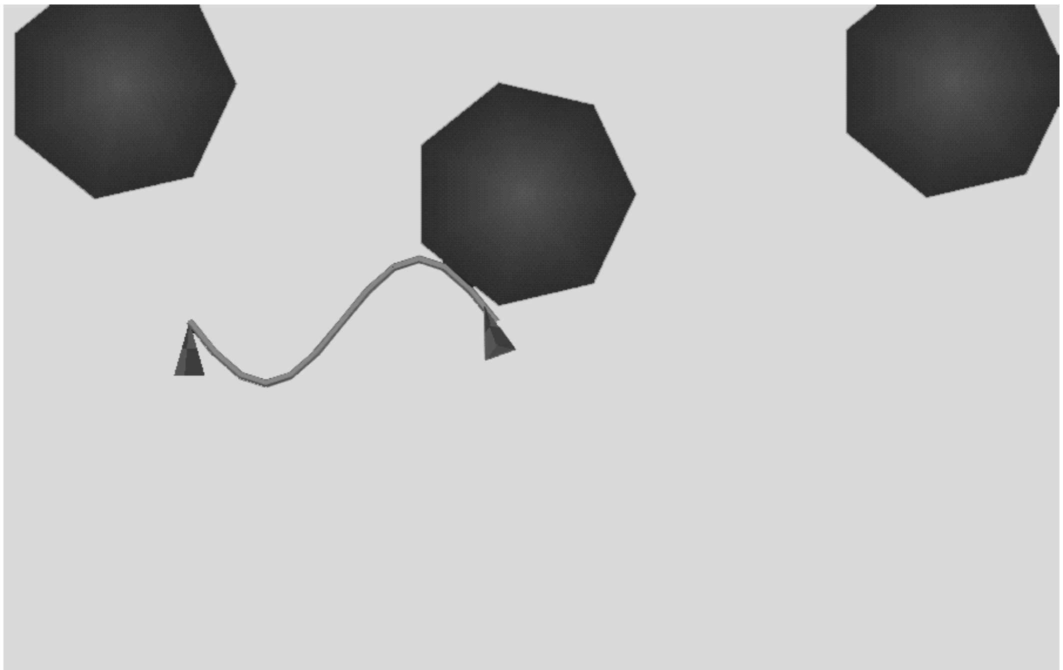


Figure 7.1: An image of the screen during a trial of the curve tracing experiment. The sine curve is visible with the trial target at the left-end and the pointer moving along the curve from the right-end. The large shaded objects toward the top of the image are a part of the vection background.

amount of green may appear on the display. To mask potential bleeding a dark gray background was used rather than the black background used in earlier experiments. Figure 7.1 is an image of the screen during a trial of this experiment.

7.4 Procedure

Subjects were required to perform a curve tracing task in a 3D environment. The pointer had six degrees of freedom, but to accomplish the task, only the position of the tip of the pointer was important. The pointer and target in this task were identical tetrahedron-like objects, in which the height (1.73cm) was twice the width (0.87cm) and one face was perpendicular to the base. The difference in size along one dimension allows one vertex to be easily identified as the tip and one face to be identified as the base. This is the same pointer used in the docking and line tracing experiments.

In the default orientation the tip points upward along the positive Y-axis, the base was parallel to the XZ plane and the perpendicular face was to the back and parallel to the XY plane. The base was half intensity magenta, the back face was full intensity magenta and the two front faces used both full intensity red and full intensity blue. The two front faces had a different colour in their top and bottom halves. The top half of one face was red and the bottom half of the same face was blue. The opposite face had the reverse colouration. Figure 3.3 illustrates the appearance of the pointer and target in different orientations.

Subjects manipulated the pointer via the Polhemus input device using the dominant right hand, and used the middle mouse button with the mouse held in the left hand to advance through the experiment trials. A trial consisted of two phases, an initial homing phase and a tracing phase.

During the homing phase the subject had to move the tip of the pointer to the tip of the homing target. The homing target was located at a fixed centre point and was always in the default orientation. Subjects received feedback in the form of a red box that appeared around the tip of the homing target when the tip of the pointer was within 0.5cm of the homing target. When the red box appeared subjects could click the middle mouse button using the left hand to advance to the tracing phase; the homing target was removed and the trial target was displayed at one of the six positions. A single wavelength of a sine curve with an amplitude of 2cm was drawn from the start position (the tip of the homing target) to the tip of the trial target. Subjects could not advance to the docking phase if the homing tolerance was not met. The pointer was not required to be in any specific orientation for the homing phase.

During the tracing phase subjects were instructed to move the pointer to the trial target while keeping the tip of the pointer as close to the sine curve (composed of a series of small cylindrical segments) as possible. No feedback regarding proximity to either the curve or trial target was given during the tracing phase. Subjects had to make their own determination of proximity using the visual cues available within the particular experiment condition. Subjects middle clicked with the mouse when they were satisfied with their performance. A score was displayed to give subjects some rough feedback about their performance during the trial. After a short delay the system advanced to the next trial.

The score was intended to give subjects some rough idea of how well they performed on the just completed trial. The score was computed by adding the mean error for the tracing to $1/10$ of the final position error.

When drawing a planar curve in 3D the question of which plane to draw the curve on arises. For the X-axis target positions the best case is to draw the curve on the XY plane. In this situation the curve travels left or right and up and down according to the sine function. Both degrees of freedom for the curve are easily visible. The worst case alternative is to draw the curve on the XZ plane. The left or right movement of the curve is easily visible as before, but the sine function now moves the curve in and out along the Z-axis. Based the earlier experiments the Z-axis was known to be the most difficult to visualize. To balance for this effect, half the trials have the curve drawn in the best case orientation and half the trials have the curve drawn in the worst case orientation. In the case of the Z-axis target positions, two similar orientations are used, one on the YZ plane and one on the XZ plane.

Small random viewing shifts to the left, right, up or down were used for the situations in which the sine curve was drawn on the YZ or XZ planes. Without these view shifts the curve would appear as a straight line.

A vection background identical to the one used in the line tracing experiment was also used. The goal was to increase the perceived motion parallax in the head-tracked viewing conditions.

7.5 Design

The experimental design consisted of four within-subject independent variables: display mode, head-tracking mode, target position and session block. For compatibility between the stereoscopic and monoscopic conditions all target positions in the virtual environment were behind the screen surface and were displayed at a size appropriate to a perspective projection based on a nominal viewing position of 40cm in front of the screen surface. To mitigate the effects of the apparatus, subjects wore the stereo glasses and head-tracker in all conditions. The receiver of the Polhemus was encased in modeling material that was shaped and coloured to resemble the pointer in the virtual environment. The physical pointer is shown in Figure 3.2.

An experiment session comprised one of the four combinations of display and head-tracking modes. After each session, subjects completed a short questionnaire dealing only with that session. Each subject performed four sessions, thereby covering all four combinations of display and head-tracking modes. The combinations of display mode and head-tracking mode were counterbalanced across subjects according to a latin square where each condition was performed first, second, third or fourth an equal number of times. After the final session, each subject completed a longer questionnaire about all the conditions.

Each subject participated in four sessions, one for each viewing condition, split across two separate days. Two sessions were conducted the first day and two sessions were conducted on the second day. Each session was divided into five blocks, a training block followed by four experiment blocks. Each experiment block contained 36 trials consisting of six repetitions of each target position – three in the best case orientation and three in the worst case orientation – presented in random order. The training block was an abbreviation of an experiment block consisting of 20 trials. During the first 10 trials of the training block subjects were given feedback telling them when they were within 1cm of the line or trial target. No feedback was given for the remaining 10 trials of the training block. To minimize fatigue subjects were given a one minute break after each block.

Each subject performed a total of 576 experiment trials. Each subject spent between 120 and 150 minutes performing the experiment including time for breaks and time to complete the questionnaires, but not including the time spent on the screening tasks.

7.6 Training

Subjects were trained using the following script. The script was a basis for the experimenter, but was not read directly.

> "Your task is to move your pointer from the starting point to the ending point as quickly and as accurately as possible while attempting to keep the tip of your pointer as close as possible to the curve. The starting point is located at the tip of the stationary object on the screen. Move your pointer to the starting point now. Notice that when the tip of your pointer is close to the tip of the starting target a red box appears around the tip of the target."
>
> [*Experimenter waits while subject moves pointer to start target.*]
>
> "The red box must be present for you to be able to start a trial. To start a trial click the middle mouse button with your left hand. "
>
> [*Experimenter waits for subject to click with the middle button.*]
>
> "The starting target has disappeared and another target has appeared somewhere in the virtual environment along with a curve

that goes from where the starting target was to the tip of the target now visible. Move your pointer to the target while trying to keep the tip of your pointer as close to the curve as you can. The orientation of the pointer with respect to the curve or target does not matter. You can have the pointer in any orientation you choose. All that matters is that you get the tip of the pointer as close as possible to the tip of the target while following the curve as closely as possible."

"If you find that you have moved away from the curve, try to move the pointer closer to the curve. Try not to move the pointer back to the beginning of the curve."

[*Experimenter waits for subject to move pointer to target.*]

"When you are satisfied that you have reached the target with your pointer click the middle mouse button with your left hand. This will end the trial."

[*Experimenter waits for subject to complete trial.*]

"You are now given a score that is a rough indication of how well you did the trial. Lower scores are better. Your goal should be to try to keep your score consistent rather than focusing on getting the lowest possible score."

"Please try to perform the task as quickly and as accurately as possible."

"Do you have any questions?"

The experimenter answered any questions that would not bias the subject's performance in any of the visual modes. Typical questions were, "Does the orientation of the pointer matter?" and "Does it matter what side of the pointer I have facing forward?" The answer to these questions was always no.

No instruction as to what was fast enough, or what was accurate enough was given, but the enforced accuracy requirement for the homing phase may have given subjects some guidance. If subjects asked whether they were to favour speed over accuracy or vice-versa they were told that they would have to decide for themselves how quickly or accurately they should perform the task.

The experimenter observed the subject during the training trials. When the training trials were over the experimenter asked for and answered any questions using the same guideline stated previously. The experimenter then told the subject that he would be on his own during the actual trials. The experimenter then reminded the subject to perform the trials as quickly and as accurately as possible, left the room and started the first experiment block. The experimenter re-entered the room after each block to check on the subject and tell the subject to rest his arm for at least a minute. After waiting for a minute the subject was asked if he was ready to proceed. If the subject responded in the affirmative then the experimenter left the room and initiated the next experiment block.

7.7 Results

The dependent variables stated earlier were gathered for each trial. Across all eight subjects a total of 4608 trials were completed.

Preliminary Screening

The data was screened for potential outliers. There were many trials requiring more than 20 seconds to complete, but only one trial requiring more than 30 seconds. Two trials had path lengths above 100cm. Using the same approach as in the point location and line tracing experiments these trials were considered to be valid and within the range of reasonable variation. Thus, no trials were eliminated and all trials were used in the analysis.

The experiment is a 2 (display mode) $\times$ 2 (head-tracking mode) $\times$ 6 (target position) $\times$ 4 (session block) design with repeated measures on all factors. For the remaining analysis, the repetitions of each target position performed by a subject within a single block are averaged together to yield a single score. The result is a series of 768 measures, one for each subject in each of the conditions.

Residual Analysis and Degree of Freedom Adjustments

Residuals were analysed for homogeneity of variance as described in Chapter 3 and demonstrated in Chapter 4. A log transform was applied to each dependent variable to make the variance more uniform and suitable for analysis. To account for violations of the sphericity assumption, the Huynh-Feldt adjusted degrees of freedom are used to test for significance.

Trial Completion Time

The repeated measures ANOVA for trial completion time yielded significant main effects for target position $F_{HF}(5,35) = 5.87$, $p = 0.0005$ and block $F_{HF}(1.2,8.4) = 10.90$, $p = 0.0086$. No significant interactions were found. Full results of the ANOVA are presented in Table 7.17 located at the end of the chapter.

Table 7.1 gives the trial completion time for each target position. We can see that the shortest trial completion time is for the x-left target position and the longest trial completion time is for the z-near target position. Table 7.2 gives the trial completion time for each block. It is evident that trial completion time improves by 1.47s from the first block to the last block.

Pos.	Time
x-left	8.46
x-right	8.57
y-bottom	8.91
y-top	8.66
z-far	8.92
z-near	9.61

Table 7.1: Mean trial completion time in seconds for each of the target positions.

Block	1	2	3	4
Time	9.73	8.92	8.50	8.26

Table 7.2: Mean trial completion time in seconds in each of the session blocks.

Final Position Error

The repeated measures ANOVA for final position error reveals significant main effects for display mode $F(1,7) = 7.05$, $p = 0.0327$, and target position $F_{HF}(3.4,23.5) = 5.45$, $p = 0.0043$. No significant interactions were found. Full results of the ANOVA appear in Table 7.18 located at the end of the chapter.

Table 7.3 gives the final position error in both display modes. The stereoscopic display mode has a significant reduction upon the final position error. Table 7.4 gives the final position error for all the target positions. The z-far target position has almost twice the error of the other target positions.

Display mode	Pos. error
mono	1.15
stereo	0.82

Table 7.3: Mean of final position error in centimetres for each display mode averaged over the other conditions.

Pos.	Pos. error
x-left	0.68
x-right	0.99
y-bottom	0.81
y-top	0.87
z-far	1.79
z-near	0.80

Table 7.4: Mean of final position error in centimetres for each target position averaged over the other conditions.

Final Position Axis Error

The repeated measures ANOVA for final position axis error reveals significant main effects for target position $F_{HF}(2.8,19.6) = 3.95$, $p = 0.0250$ and axis $F_{HF}(1.3,8.8) = 48.85$, $p < 0.0001$. Significant two-way interactions were found for display mode $\times$ axis $F_{HF}(1.3,9.0) = 15.92$, $p = 0.0022$, and target position $\times$ axis $F_{HF}(3.9,27.3) = 6.98$, $p = 0.0006$. A significant three-way interaction of head-tracking mode $\times$ target position $\times$ axis $F_{HF}(10,70) = 2.15$, $p = 0.0312$ was also found. Because the purpose of this analysis is detect how the error is distributed across axes, only the effect of axis or interactions with axis are examined in detail. Full results of the ANOVA appear in Table 5.22 at the end of the chapter.

Table 7.5 gives the final position error along each of the axes. The Z-axis clearly stands out with almost four times the error of the X- or Y-axes. Table 7.6 gives the final position error along each of the axes for both display modes. The error along the X- and Y-axes is mostly unchanged between the display modes. The stereoscopic display mode significantly reduces the Z-axis final position error.

The target position $\times$ axis and head-tracking mode $\times$ target position $\times$ axis inter-

Axis	X	Y	Z
Error	0.20	0.22	0.86

Table 7.5: Mean of final position error in centimetres along the X, Y and Z axes.

Display Mode	X	Y	Z
mono	0.21	0.22	1.09
stereo	0.20	0.22	0.69

Table 7.6: Mean of final position error in centimetres along the X, Y and Z axes for both display modes.

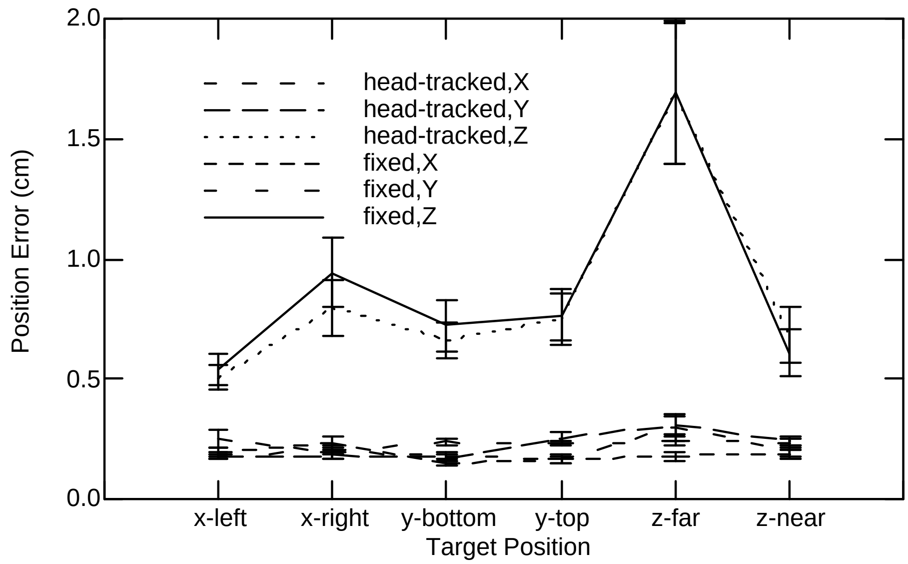


Figure 7.2: Mean of final position error in centimetres across all target positions for all axes in the two head-tracking modes.

actions are considered together. Table 7.7 gives the final position error broken down across head-tracking mode, target position and axis. The relationship between these factors is illustrated in Figure 7.2. The large increase in Z-axis error in the z-far target position stands out in both the head-tracked and fixed viewpoint conditions. The scale compression in Figure 7.2 makes it difficult to observe that the X- and Y-axis error also increase for the z-far target in the head-tracked mode.

Path Length

The repeated measures ANOVA for path length reveals a significant main effect for block $F_{HF}(2.4,17.0) = 23.49$, $p < 0.0001$, a significant display mode $\times$ target position $F_{HF}(3.9,27.0) = 3.39$, $p = 0.0240$, two-way interaction and a significant display mode $\times$ head-tracking mode $\times$ target position $F_{HF}(3.4,23.8) = 3.24$, $p = 0.0352$, three-way interaction. Full results of the ANOVA appear in Table 7.20 at the end of the chapter.

Table 7.8 gives the path length in each of the session blocks. Path length decreases steadily over the blocks, dropping 1.68cm between the first and last blocks.

Table 7.9 gives the path length in each of the target positions for both display modes and both head-tracking modes. Figure 7.3 illustrates the relationship between these variables. The relationship remains relatively stable across the X- and Y-axis target positions. The fixed viewpoint stereoscopic condition had the shortest path lengths and the fixed viewpoint monoscopic condition had the longest path lengths. The path lengths of the two head-tracked conditions generally falls between the two non-head tracked conditions. In the two Z-axis conditions this changes so that the fixed viewpoint stereoscopic condition still has the shortest path lengths, but now the two head-tracked conditions have the longer path lengths. The path length for the fixed viewpoint monoscopic condition has fallen and is just slightly

	Fixed			**Head-tracked**		
Pos.	**X**	**Y**	**Z**	**X**	**Y**	**Z**
x-left	0.20	0.18	0.54	0.25	0.18	0.51
x-right	0.24	0.20	0.94	0.20	0.18	0.80
y-bottom	0.15	0.24	0.72	0.18	0.17	0.66
y-top	0.17	0.23	0.77	0.17	0.26	0.75
z-far	0.18	0.24	1.70	0.30	0.31	1.69
z-near	0.19	0.24	0.61	0.20	0.24	0.68

Table 7.7: Mean of final position error in centimetres along the X, Y and Z axes for each target position in both head-tracking modes.

Block	1	2	3	4
Path Length	19.00	18.20	17.64	17.32

Table 7.8: Mean path length in centimetres in each of the session blocks.

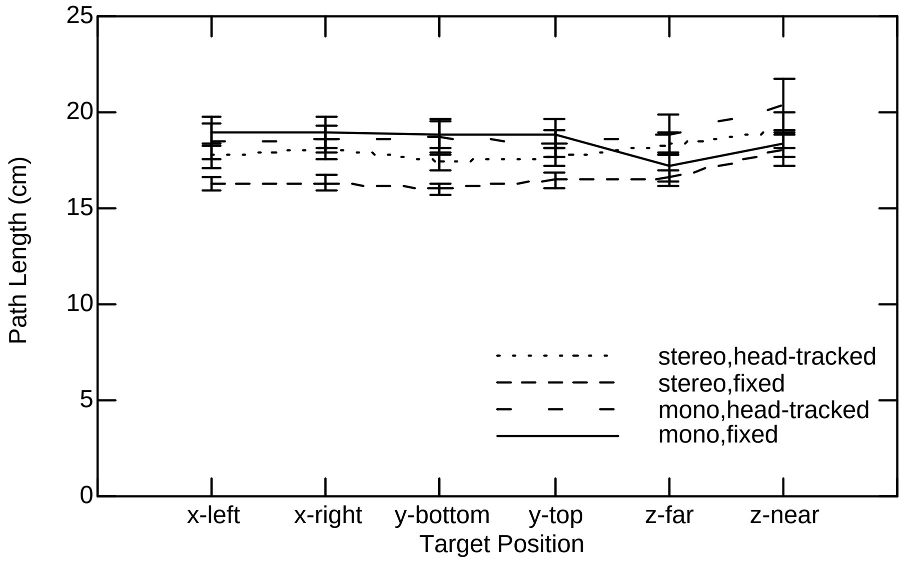


Figure 7.3: Mean of path length in centimetres vs. target position for all combinations of display mode and head-tracking mode.

higher than the corresponding stereoscopic condition.

	Fixed		**Head-tracked**	
Pos.	**mono**	**stereo**	**mono**	**stereo**
x-left	19.02	16.29	18.48	17.76
x-right	18.95	16.31	18.62	18.15
y-bottom	18.86	16.05	18.71	17.47
y-top	18.89	16.50	18.40	17.67
z-far	17.19	16.60	18.85	18.35
z-near	18.35	18.07	20.39	19.07

Table 7.9: Mean path length in centimetres for each target position in both the head-tracked and fixed viewpoint conditions and both display modes.

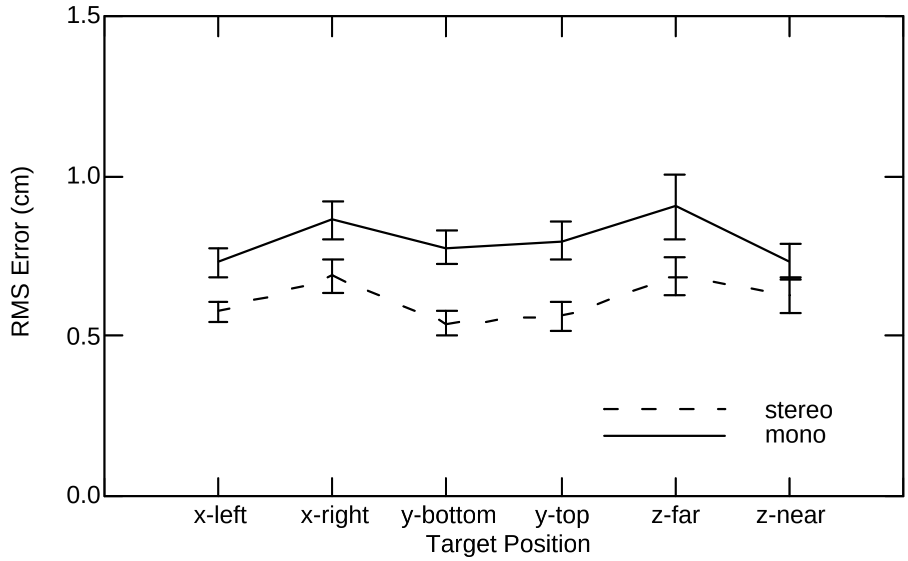


Figure 7.4: Mean of RMS error in centimetres vs. target position for both display modes.

RMS Error

The repeated measures ANOVA for RMS error showed significant main effects for display mode $F(1,7) = 13.33$, $p = 0.0082$ and target position $F_{HF}(4.4,30.5) = 3.50$, $p = 0.0161$, as well as a significant two-way interaction for display mode $\times$ target position $F_{HF}(5,35) = 3.58$, $p = 0.0102$. Full results of the ANOVA appear in Table 7.21 at the end of the chapter.

The display mode and target position factors are examined together because of the significant interaction. Table 7.10 gives the RMS error for each of the target positions in

Pos.	**Mono**	**Stereo**
x-left	0.73	0.58
x-right	0.86	0.69
y-bottom	0.78	0.54
y-top	0.80	0.56
z-far	0.91	0.69
z-near	0.74	0.63

Table 7.10: Mean RMS error in centimetres for each target position in both display modes.

both display modes. Figure 7.4 illustrates the relationship between these factors. The stereoscopic display mode clearly reduces the RMS error across all target positions. As for target position, The z-near target is amongst the targets with the least RMS error in the monoscopic condition and the z-far target has the highest RMS error. In the stereoscopic condition the x-right target and the z-far target have the highest RMS error and the y-bottom target now has the least amount of RMS error.

RMS Axis Error

To gain some additional insight into the source of the RMS error, a repeated measures ANOVA is conducted for axis RMS error. The ANOVA revealed significant main effects for display mode $\mathrm{F}(1,7) = 12.32$, $p = 0.0099$, target position $\mathrm{F}_{\mathrm{HF}}(3.9,27.0) = 6.81$, $P = 0.0007$ and axis $\mathrm{F}_{\mathrm{HF}}(1.4,9.5) = 42.00$, $p < 0.0001$. Significant two-way interactions were found for display mode $\times$ axis $\mathrm{F}_{\mathrm{HF}}(2,14) = 11.45$, $p = 0.0011$, head-tracking mode $\times$ axis $\mathrm{F}_{\mathrm{HF}}(1.6,11.3) = 4.92$, $p = 0.0344$, and target position $\times$ axis $\mathrm{F}_{\mathrm{HF}}(2.8,19.6) = 21.61$, $p < 0.0001$. Two significant three-way interactions were found for display mode $\times$ head-tracking $\times$ target position $\mathrm{F}_{\mathrm{HF}}(4.0,28.0) = 3.55$, $p = 0.0183$, and head-tracking mode $\times$ block $\times$ axis $\mathrm{F}_{\mathrm{HF}}(6,42) = 3.04$, $p = 0.0146$. Full results of the ANOVA appear in Table 7.22 located at the end of the chapter. Because the purpose of this analysis is to examine the effect of axis, only the effect of axis and its interactions are analyzed further.

Table 7.11 gives the RMS error along the axes for both display modes. It is evident that the Z-axis RMS error is larger than the RMS error for the X- and Y-axes. Stereoscopic viewing reduced the RMS error along all the axes, but it had a much larger effect upon the Z-axis error.

Table 7.12 gives the RMS error along the axes for each target position and the relationship amongst these factors is illustrated in Figure 7.5. The Z-axis error was always

Display Mode	**X**	**Y**	**Z**
mono	0.31	0.33	0.62
stereo	0.26	0.29	0.44

Table 7.11: Mean of RMS error in centimetres along the X, Y and Z axes in both display modes.

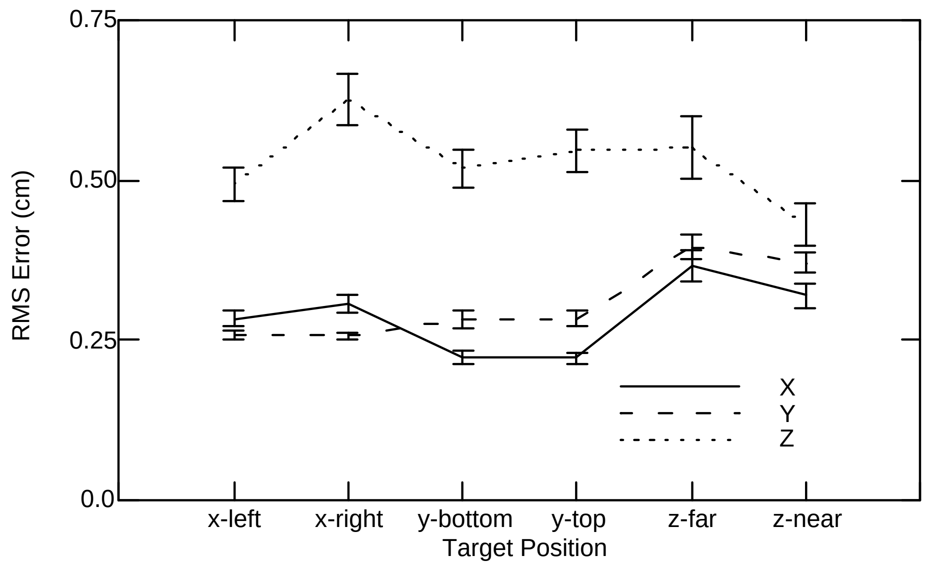


Figure 7.5: Mean of RMS error in centimetres along each of the axes vs. target position.

larger than the error along the X- and Y-axes. However, in the z-near target position the Z-axis RMS error was relatively close to the Y-axis RMS error whereas in the x-right target position the Z-axis RMS error was much larger than the X- and Y-axis RMS error. In fact the x-right target position had the highest Z-axis RMS error of all target positions. The z-far target position exhibited the highest X- and Y-axis RMS error, but the Z-axis RMS error for the z-far target position was close to the Z-axis RMS error of several other target positions. Interestingly, the z-near target position had the second highest RMS error for the X- and Y-axes, but the lowest RMS error for the Z-axis.

Finally, Table 7.13 gives the RMS error along the X, Y and Z axes for both head-tracking modes across the four blocks and Figure 7.6 illustrates the relationship between

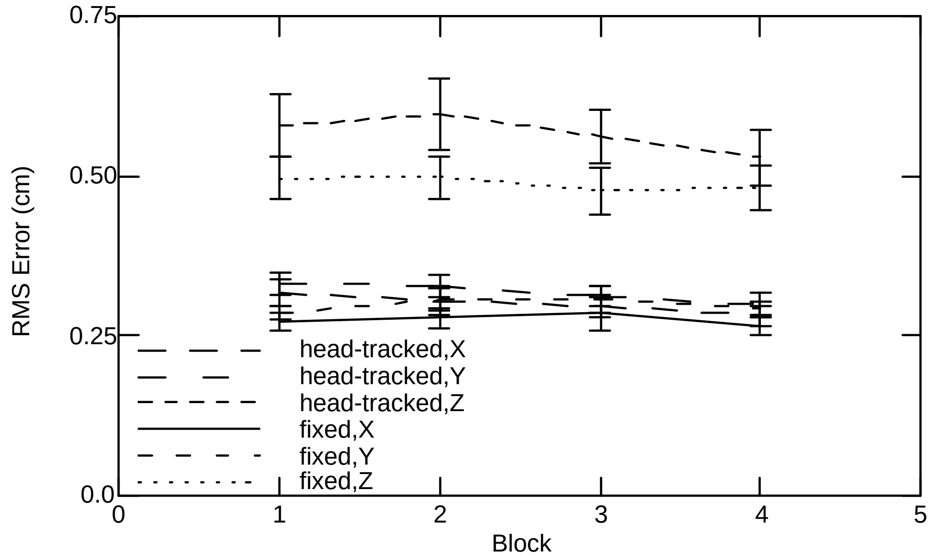


Figure 7.6: Mean of RMS error in centimetres along each of the axes vs. session block for both head-tracking modes.

these factors. It is clear that the RMS error for the head-tracked modes was higher than for the fixed viewpoint modes. However, there is a steady improvement in the RMS error for the head-tracked mode, while in the fixed viewpoint mode the RMS error remained relatively stable across blocks.

Pos.	**X**	**Y**	**Z**
x-left	0.28	0.26	0.49
x-right	0.31	0.26	0.63
y-bottom	0.22	0.28	0.52
y-top	0.22	0.28	0.55
z-far	0.37	0.40	0.55
z-near	0.32	0.37	0.43

Table 7.12: Mean of RMS error in centimetres along the X, Y and Z axes for each of the target positions.

Questionnaires

At the end of each session subjects were required to answer a series of questions about different aspects of the task. Four questions required the subject to rate the difficulty of different components of the task. They were asked to indicate the difficulty of determining the position of the target and the curve. They were also asked to indicate the difficulty of matching the target position and tracing the curve. Each of these questions was answered on a five-point scale where 1 is "easy" and 5 is "hard." Four additional questions asked subjects to indicate if they found any target positions easier or harder to match, or whether they found any curves easier or harder to trace. If a subject indicated that a component was easier or harder a free form answer was used to describe the particular conditions. A copy of the post-session questionnaire is located in Appendix A.6.

At the end of the experiment, subjects were asked to score all the display and head-tracking mode combinations according to task difficulty and ease of use. Additionally, at the end of the experiment subjects were asked to conduct three rankings of the four different viewing conditions. Subjects were asked to rank the viewing modes based on their preference, ease of use and their performance. Low rankings were best and high rankings were worst. A copy of the post-experiment questionnaire can be found in Appendix A.7.

Friedman rank tests failed to find any significant effects for subjects' scores, however, two of the questions achieved almost significant results. Subjects indicated that it was

	Fixed			**Head-tracked**		
Block	**X**	**Y**	**Z**	**X**	**Y**	**Z**
1	0.27	0.29	0.50	0.32	0.33	0.58
2	0.28	0.31	0.50	0.31	0.33	0.60
3	0.29	0.31	0.48	0.30	0.31	0.56
4	0.27	0.29	0.48	0.28	0.30	0.53

Table 7.13: Mean of RMS error in centimetres along the X, Y and Z axes for each target position across the four session blocks.

hardest to determine the position of the curve in the head-tracked monoscopic condition and easiest in the head-tracked stereoscopic condition. In terms of usefulness of a viewing mode, subjects felt that the fixed viewpoint stereoscopic condition was the most useful, while the fixed viewpoint monoscopic condition was the least useful. Determining the position of the curve in the head-tracked monoscopic condition seemed to be one of the hardest elements of the task. Table 7.14 gives the mean scores for the post session and post experiment questions.

In the post experiment questions, subjects indicted that the task was easiest in the two head-tracked conditions. However, they indicated that the fixed viewpoint stereoscopic condition was the most useful.

The post-experiment rankings indicated that subjects had a reasonably strong, yet not significant, preference for the head-tracked stereoscopic viewing condition and they also considered their performance to have been best in this viewing condition. A Friedman rank test for ease of use was just barely significant ($\chi^2_F = 7.80$ with 3 df, p $= 0.05$). Subjects indicated that the fixed viewpoint stereoscopic viewing condition was easiest to use, with the head-tracked stereoscopic condition being somewhat more difficult. The most difficult viewing condition was the fixed viewpoint monoscopic condition. Figure 7.7 provides summary information for subjects' rankings.

Subjects' free form answers to the questions about the type of targets and lines were classified into general categories. Few subjects indicated that any targets were easier or harder. Responses regarding which curves were easier to trace showed that more subjects felt that the curves on the XY plane were easier in the monoscopic modes than in the stereoscopic modes. The descriptions of which curves were harder to trace varied extensively with little apparent clustering of responses into particular categories. To some degree, curves on the XZ and YZ plane were considered to be harder, however, subjects tended to isolate a

	Fixed		Head-tracked			
Question	**Mono**	**Stereo**	**Mono**	**Stereo**	**Friedman**	**p**
det. target pos.	2.9	2.3	3.5	2.2	3.41	0.332
match target pos.	2.6	2.3	2.9	2.6	1.09	0.780
det. curve pos.	3.6	3.1	4.0	2.8	7.69	0.053
trace curve	3.6	3.1	3.6	2.8	3.41	0.332
task difficulty	3.6	2.6	3.1	2.5	4.99	0.173
mode usefulness	3.5	1.6	2.4	2.3	7.39	0.061

Table 7.14: Means of subjects' responses to the subjective questions regarding display mode and head tracking mode conditions. The first four questions were answered immediately after a particular condition, and the last two questions were answered after all conditions had been completed. All questions were answered on a scale of 1 to 5 where 1 indicated easy or useful and 5 indicated hard or not useful.

particular subset of these curves in their descriptions, for example curves on the XZ plane running to the x-right target position. The categorized summaries of subjects' answers can be found in Tables 7.15 and 7.16.

	Easier							
	Target				Curve			
	Fixed		Head-tracked		Fixed		Head-tracked	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**
none	5	6	6	8	2	3	1	2
XY plane	1	2	1		6	2	5	3
X-axis					1			1
Y-axis								1
Z-axis			1		1			
z-near	2					1		
XZ + YZ						1	1	
XY, X						2	1	
XY right							1	
YZ top								1
z-far								1

Table 7.15: Coded summary of subjects' responses to the questions about which targets and curves were easier to match the position of or trace. Responses are categorized by visual feedback conditions, with M indicating the monoscopic display mode and S the stereoscopic display mode. Letter pairs refer to planes and single letters refer to axes. For example, XY, X refers to the curves that lie on the XY plane and run to target positions on the X-axis.

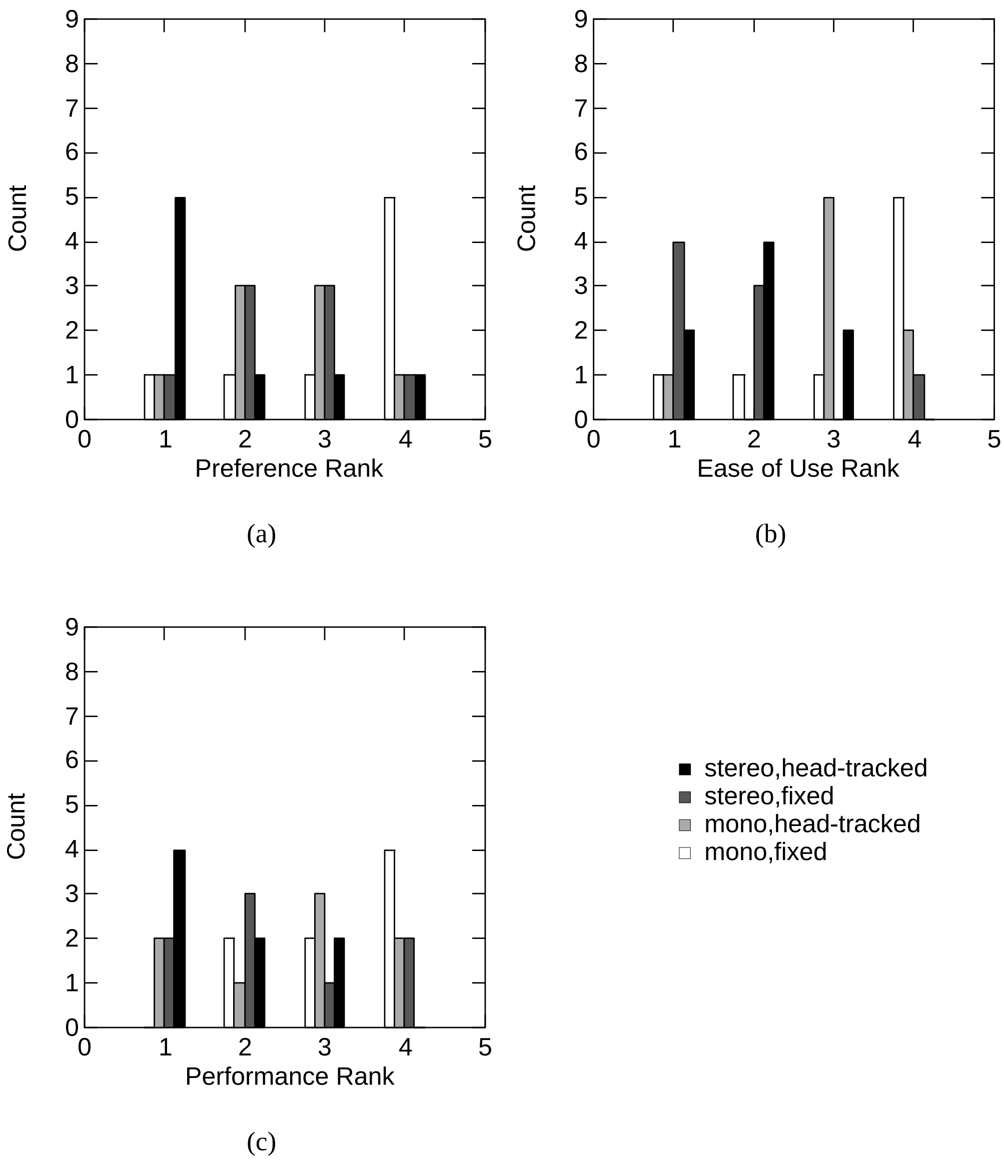


Figure 7.7: Rankings of subjects' preference, perceived ease of use and perceived performance for each of the viewing conditions. The height of a bar indicates the number of times a particular rank was assigned by a subject.

	Harder							
	Target				Curve			
	Fixed		Head-tracked		Fixed		Head-tracked	
Response	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**	**Mono**	**Stereo**
none	6	6	7	8	2	1		3
XY plane							1	
Z-axis		2			2	3	3	2
z-near	1							
z-far	1		1					
XZ + YZ					4	1	1	1
XZ, X						1	1	3
XZ, Z							1	
XZ far					1			
XZ right						1		
YZ, Y						2		
YZ bot.							1	

Table 7.16: Coded summary of subjects' responses to the questions about which targets and curves were harder to match the position of or trace. Responses are categorized by visual feedback conditions, with M indicating the monoscopic display mode and S the stereoscopic display mode. Letter pairs refer to planes and single letters refer to axes. For example, XY, X refers to the curves that lie on the XY plane and run to target positions on the X-axis.

7.8 Discussion

In contrast to earlier experiments where display mode and head-tracking mode had strong main effects, this experiment showed only small effects for these variables in only a few conditions.

7.8.1 Speed

Stereoscopic viewing does not improve trial completion time

The results for trial completion time in this experiment were somewhat surprising, especially in contrast to earlier experiments. There were no significant main effects for either head-tracking mode or display mode. Neither of the two factors that directly influenced viewing had any effect upon the time taken by subjects to complete the task. Earlier experiments and other dependent variables suggest that stereoscopic viewing in particular has a very strong influence. It is puzzling then that no effect was found for this task. It may well be that stereoscopic viewing is not as powerful when occlusion is available throughout the task to provide depth information. However, RMS error showed that stereoscopic viewing was better than monoscopic viewing. Considering trial completion time in conjunction with RMS error suggests that subjects seemed to prefer a particular movement pace and allowed their error level to vary rather than trial completion time.

X-axis positions are fastest, z-near position slowest

In this experiment target position translates into direction of movement as the sine curves were drawn from the starting position to the target positions. Performance seems to be best for targets along the X-axis. The worst performance was for the z-near target position. However, when final position error is also considered we see a much larger final position error for the z-far target position as compared with any of the other target positions. This would suggest that subjects were unable to accurately determine where the target was when it appeared in the z-far position.

Later blocks were faster than earlier ones

Block had a significant effect upon trial completion time and also upon path length. Subjects completed trials faster and with less overall movement in successive blocks. The lack of an effect for block with respect to RMS error indicates that subjects were no less accurate in successive blocks. In combination this suggests that subjects movements were more efficient in later blocks.

7.8.2 Accuracy

Stereoscopic viewing reduces Z-axis error

Stereoscopic viewing did reduce the amount of final position error for this task, however, it is important to keep in mind that final position error is less significant for this task, which was to have subjects trace the entire curve, not just the endpoints. Stereoscopic viewing had a fairly large effect upon the Z-axis final position error, reducing the Z-axis error from 1.04cm in the monoscopic condition to 0.69cm in the stereoscopic condition.

Head-tracking reduced Z-axis error for the x-right target

The head-tracking mode $\times$ target position $\times$ axis interaction draws attention to the reduced Z-axis error for the x-right target position in the head-tracked conditions.

Z-far target had a much higher Z-axis error

The target position $\times$ axis interaction highlights the large increase in Z-axis error for the z-far target position. While the Z-axis error was larger than the X- or Y-axis error across all target positions, the Z-axis error was much higher for the z-far target than for any of the other target positions.

Stereoscopic viewing reduced RMS error

RMS error had significant main effects for display mode and target position as well as a significant interaction between these factors. The RMS error was reduced for the stereoscopic display mode in each target position. However, due to the significant interaction it is evident that the effect of stereoscopic viewing was not uniform. The z-near target position had one of the lower RMS errors in the monoscopic conditions, but in the stereoscopic condition the z-near target position has one of the larger RMS errors. Also of interest is the x-right target position whose RMS error was close to the error for the z-far target in the monoscopic conditions and equal to the z-far target in the stereoscopic conditions.

Higher Z-axis RMS error is reduced over blocks by head-tracking

Across the X, Y, and Z axes, stereoscopic viewing reduced the amount of RMS error along the Z-axis. There was significantly more RMS error along the Z-axis in the head-tracked condition than in the fixed viewpoint condition. Interestingly, there was a reduction in the amount of Z-axis RMS error across blocks in the head-tracked condition, while in the fixed viewpoint condition the amount of RMS error remained relatively stable. In a positive sense this could indicate that subjects were able to make better use of head-tracking with practice. A more negative view for head-tracking might indicate that subjects used head-tracking less and less across blocks and thus their Z-axis RMS error began to converge to the fixed viewpoint level.

7.8.3 Conciseness

The results for path length would at first appear to suggest excellent performance for the z-far target position. However, one must keep in mind that the z-far target position had a final position error that was almost twice that of the other target positions. The path length is low because subjects did not move all the way to the target. It is likely that had subjects moved all the way to the target the path length for the z-far target position would be much higher. The results did indicate a longer path length for the z-near target position. The path lengths for the X- and Y-axis targets do not vary significantly.

Path length is one of the few dependent variables that showed any effect for head-tracking mode. The X- and Y-axis target positions appear to have had slightly longer movement paths in the fixed viewpoint monoscopic condition as compared with the head-tracked monoscopic condition. However this situation reverses itself in the stereoscopic conditions with the fixed viewpoint condition having had shorter movement paths compared to the head-tracked viewpoint condition.

Care must be taken when considering path length for the Z-axis conditions because of the evidence that subjects did not move all the way to the target for the z-far target position. Nonetheless, in both the monoscopic and stereoscopic head-tracked conditions, the movement paths were longer for the Z-axis target positions than in the fixed viewpoint conditions. Recall that the movement paths were shorter for the head-tracked conditions for targets along the X- and Y-axes.

7.8.4 Felicity

As in the line tracing experiment, many of the subjective scores failed to achieve significance, however, in contrast to that experiment there were much stronger trends. The head-tracked stereoscopic condition was generally thought to be easier to use than the head-tracked monoscopic condition that subjects generally considered to be most difficult. Interestingly, in the post-experiment questionnaire where subjects were able to score conditions based upon experience with all of them, the fixed viewpoint stereoscopic condition was considered to be the most useful. It may be that subjects scored the head-tracked stereoscopic condition highly at first because of its novelty, but upon later reflection they considered the fixed viewpoint stereoscopic condition to be most useful.

Subjects' rankings of the visual feedback conditions indicated that they found the head-tracked monoscopic condition easiest to use. However, in ranking their preference and performance, the trend favoured the head-tracked stereoscopic condition. An interesting experiment could be carried out to determine which visual feedback condition subjects would select, if given the choice, both before and after some experience with each of the conditions.

7.9 Conclusions

As in earlier experiments the goal of this experiment was to determine the effects of display mode, head-tracking mode and target position upon the ability to perform a task. The task was to trace a planar sine wave drawn to one of six possible locations.

Stereoscopic viewing did not reduce the amount of time required to perform the task, but it did reduce the amount of final position error and RMS error. Thus stereoscopic viewing did not allow subjects to perform the tracing task any faster, but it did improve the accuracy with which subjects could perform the task.

Head-tracking was generally not beneficial in the monoscopic display mode, often degrading performance. In the stereoscopic condition head-tracking does not appear to have degraded performance and it may have improved performance for certain target positions. Nonetheless, the overall benefit from head-tracking is slim at best as it appears that head-tracking degraded performance in more situations than it improved performance.

Target position continues to be a factor with significant influence on performance. As expected, the z-near and z-far targets generally had poor performance, but the x-right target position also showed poor performance when compared with the x-left target position.

Subjective feedback suggests that subjects may have initialy favoured the head-tracked stereoscopic viewing condition, but after some exposure to all the visual feedback conditions, the fixed viewpoint stereoscopic was considered to be easier to use.

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	2.185	1	2.185	1.26	0.2988		
Error	12.141	7	1.734				
head-track	0.261	1	0.261	0.14	0.7154		
Error	12.672	7	1.810				
d×h	3.106	1	3.106	1.02	0.3452		
Error	21.228	7	3.033				
position	1.201	5	0.240	5.87	0.0005	0.0005	1.00
Error	1.432	35	0.041				
d×p	0.060	5	0.012	1.42	0.2427	0.2427	1.00
Error	0.296	35	0.008				
h×p	0.017	5	0.003	0.33	0.8895	0.7932	0.57
Error	0.354	35	0.010				
d×h×p	0.072	5	0.014	1.50	0.2162	0.2162	1.00
Error	0.339	35	0.010				
block	2.885	3	0.962	10.90	0.0002	0.0086	0.40
Error	1.853	21	0.088				
d×b	0.252	3	0.084	1.92	0.1570	0.1858	0.63
Error	0.920	21	0.044				
h×b	0.027	3	0.009	0.24	0.8651	0.7905	0.68
Error	0.789	21	0.038				
d×h×b	0.401	3	0.134	1.67	0.2044	0.2254	0.64
Error	1.683	21	0.080				
p×b	0.130	15	0.009	1.71	0.0602	0.1378	0.42
Error	0.532	105	0.005				
d×p×b	0.060	15	0.004	0.80	0.6788	0.6788	1.00
Error	0.525	105	0.005				
h×p×b	0.077	15	0.005	1.10	0.3644	0.3644	1.00
Error	0.489	105	0.005				
d×h×p×b	0.059	15	0.004	0.72	0.7586	0.6743	0.54
Error	0.573	105	0.005				
Total	94.399						

Table 7.17: Experiment 4 – Repeated measures ANOVA for log of trial completion time.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	21.782	1	21.782	7.05	0.0327		
Error	21.627	7	3.090				
head-track	0.215	1	0.215	0.39	0.5504		
Error	3.832	7	0.547				
d×h	3.699	1	3.699	4.27	0.0777		
Error	6.070	7	0.867				
position	28.084	5	5.617	5.45	0.0008	0.0043	0.67
Error	36.067	35	1.030				
d×p	1.490	5	0.298	1.11	0.3718	0.3643	0.54
Error	9.376	35	0.268				
h×p	0.530	5	0.106	0.60	0.6993	0.6993	1.00
Error	6.174	35	0.176				
d×h×p	1.620	5	0.324	1.96	0.1095	0.1405	0.69
Error	5.790	35	0.165				
block	0.447	3	0.149	0.90	0.4589	0.4567	0.97
Error	3.484	21	0.166				
d×b	0.184	3	0.061	0.35	0.7929	0.7929	1.00
Error	3.729	21	0.178				
h×b	0.424	3	0.141	1.12	0.3651	0.3651	1.00
Error	2.663	21	0.127				
d×h×b	0.192	3	0.064	0.46	0.7161	0.6781	0.81
Error	2.954	21	0.141				
p×b	1.375	15	0.092	1.02	0.4420	0.4420	1.00
Error	9.447	105	0.090				
d×p×b	0.898	15	0.060	0.66	0.8177	0.7497	0.63
Error	9.526	105	0.091				
h×p×b	0.847	15	0.057	0.76	0.7225	0.7017	0.85
Error	7.850	105	0.075				
d×h×p×b	0.875	15	0.058	0.68	0.7969	0.7455	0.70
Error	8.977	105	0.085				
Total	511.039						

Table 7.18: Experiment 4 – Repeated measures ANOVA for log of final position error.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	14.752	1	14.752	2.81	0.1373		
Error	36.690	7	5.241				
head-track	0.215	1	0.215	0.20	0.6647		
Error	7.344	7	1.049				
d×h	6.541	1	6.541	3.79	0.0925		
Error	12.069	7	1.724				
position	27.460	5	5.492	3.95	0.0060	0.0250	0.56
Error	48.655	35	1.390				
d×p	1.112	5	0.222	0.60	0.7002	0.6366	0.66
Error	12.969	35	0.371				
h×p	3.930	5	0.786	1.87	0.1257	0.1257	1.00
Error	14.748	35	0.421				
d×h×p	2.459	5	0.492	1.52	0.2078	0.2078	1.00
Error	11.300	35	0.323				
block	1.369	3	0.456	1.22	0.3257	0.3257	1.00
Error	7.827	21	0.373				
d×b	0.203	3	0.068	0.16	0.9221	0.9221	1.00
Error	8.901	21	0.424				
h×b	0.816	3	0.272	1.34	0.2890	0.2890	1.00
Error	4.271	21	0.203				
d×h×b	0.528	3	0.176	1.00	0.4106	0.4045	0.87
Error	3.683	21	0.175				
p×b	3.646	15	0.243	1.16	0.3178	0.3250	0.86
Error	22.078	105	0.210				
d×p×b	1.280	15	0.085	0.40	0.9771	0.9736	0.95
Error	22.530	105	0.215				
h×p×b	1.883	15	0.126	0.69	0.7875	0.7569	0.81
Error	19.057	105	0.182				
d×h×p×b	2.444	15	0.163	0.90	0.5683	0.5515	0.79
Error	19.051	105	0.181				

Table 7.19: Experiment 4 – Repeated measures ANOVA for log of final position error along axes (continued on next page).

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
axis	586.025	2	293.013	48.85	$<$ 0.0001	$<$ 0.0001	0.63
Error	83.970	14	5.998				
d×a	29.155	2	14.578	15.92	0.0002	0.0022	0.64
Error	12.818	14	0.916				
h×a	1.662	2	0.831	1.86	0.1916	0.2002	0.83
Error	6.242	14	0.446				
d×h×a	3.454	2	1.727	3.68	0.0519	0.0628	0.85
Error	6.568	14	0.469				
p×a	38.975	10	3.898	6.98	$<$ 0.0001	0.0006	0.39
Error	39.089	70	0.558				
d×p×a	2.829	10	0.283	1.00	0.4522	0.4522	1.00
Error	19.808	70	0.283				
h×p×a	4.357	10	0.436	2.15	0.0312	0.0312	1.00
Error	14.172	70	0.202				
d×h×p×a	1.914	10	0.191	0.93	0.5135	0.5135	1.00
Error	14.443	70	0.206				
b×a	0.374	6	0.062	0.39	0.8787	0.8787	1.00
Error	6.648	42	0.158				
d×b×a	1.086	6	0.181	1.18	0.3331	0.3331	1.00
Error	6.417	42	0.153				
h×b×a	1.454	6	0.242	1.60	0.1699	0.1699	1.00
Error	6.345	42	0.151				
d×h×b×a	0.729	6	0.121	0.89	0.5119	0.5080	0.95
Error	5.740	42	0.137				
p×b×a	2.725	30	0.091	1.04	0.4229	0.4229	1.00
Error	18.425	210	0.088				
d×p×b×a	3.760	30	0.125	1.11	0.3241	0.3241	1.00
Error	23.676	210	0.113				
h×p×b×a	2.888	30	0.096	0.87	0.6659	0.6644	0.99
Error	23.262	210	0.111				
d×h×p×b×a	3.085	30	0.103	1.01	0.4547	0.4535	0.82
Error	21.330	210	0.102				
Total	1895.449						

Table 7.19: Experiment 4 – Repeated measures ANOVA for log of final position error along axes (continued from previous page).

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
display	0.753	1	0.753	2.46	0.1604		
Error	2.137	7	0.305				
head-track	0.382	1	0.382	0.55	0.4807		
Error	4.818	7	0.688				
d×h	0.230	1	0.230	0.39	0.5520		
Error	4.120	7	0.589				
position	0.288	5	0.058	1.21	0.3242	0.3289	0.75
Error	1.662	35	0.047				
d×p	0.174	5	0.035	3.39	0.0134	0.0240	0.77
Error	0.359	35	0.010				
h×p	0.116	5	0.023	1.73	0.1541	0.1960	0.56
Error	0.469	35	0.013				
d×h×p	0.140	5	0.028	3.24	0.0165	0.0352	0.68
Error	0.302	35	0.009				
block	0.894	3	0.298	23.49	$<$ 0.0001	$<$ 0.0001	0.81
Error	0.266	21	0.013				
d×b	0.012	3	0.004	0.28	0.8425	0.7530	0.63
Error	0.301	21	0.014				
h×b	0.046	3	0.015	1.01	0.4083	0.4069	0.97
Error	0.316	21	0.015				
d×h×b	0.134	3	0.045	1.57	0.2256	0.2427	0.65
Error	0.594	21	0.028				
p×b	0.096	15	0.006	0.86	0.6060	0.5381	0.45
Error	0.781	105	0.007				
d×p×b	0.108	15	0.007	1.44	0.1433	0.2321	0.35
Error	0.524	105	0.005				
h×p×b	0.141	15	0.009	1.97	0.0244	0.1114	0.32
Error	0.503	105	0.005				
d×h×p×b	0.098	15	0.007	1.10	0.3647	0.3756	0.27
Error	0.622	105	0.006				
Total	31.124						

Table 7.20: Experiment 4 – Repeated measures ANOVA for log of path length.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	13.369	1	13.369	13.33	0.0082		
Error	7.019	7	1.003				
head-track	1.471	1	1.471	3.78	0.0929		
Error	2.724	7	0.389				
d×h	1.951	1	1.951	2.01	0.1993		
Error	6.799	7	0.971				
position	2.574	5	0.515	3.50	0.0114	0.0161	0.87
Error	5.145	35	0.147				
d×p	0.972	5	0.194	3.58	0.0102	0.0102	1.00
Error	1.901	35	0.054				
h×p	0.249	5	0.050	0.79	0.5667	0.5355	0.74
Error	2.218	35	0.063				
d×h×p	0.769	5	0.154	2.24	0.0717	0.1149	0.59
Error	2.400	35	0.069				
block	0.558	3	0.186	2.42	0.0946	0.1267	0.65
Error	1.614	21	0.077				
d×b	0.032	3	0.011	0.24	0.8700	0.8700	1.00
Error	0.959	21	0.046				
h×b	0.154	3	0.051	0.70	0.5621	0.5262	0.75
Error	1.536	21	0.073				
d×h×b	0.115	3	0.038	0.69	0.5711	0.5396	0.78
Error	1.177	21	0.056				
p×b	0.455	15	0.030	0.89	0.5754	0.5162	0.44
Error	3.569	105	0.034				
d×p×b	0.161	15	0.011	0.47	0.9506	0.8581	0.49
Error	2.393	105	0.023				
h×p×b	0.586	15	0.039	1.29	0.2190	0.2616	0.57
Error	3.169	105	0.030				
d×h×p×b	0.346	15	0.023	0.73	0.7453	0.6316	0.42
Error	3.298	105	0.031				
Total	185.177						

Table 7.21: Experiment 4 – Repeated measures ANOVA for log of RMS error.

Source	**SS**	**df**	**MS**	**F**	**conv. p**	**adj. p**	**H-F ϵ**
display	25.948	1	25.948	12.32	0.0099		
Error	14.741	7	2.106				
head-track	3.785	1	3.785	3.02	0.1260		
Error	8.782	7	1.255				
d×h	4.236	1	4.236	2.36	0.1680		
Error	12.541	7	1.792				
position	14.167	5	2.833	6.81	0.0002	0.0007	0.77
Error	14.561	35	0.416				
d×p	1.355	5	0.271	2.10	0.0890	0.0890	1.00
Error	4.520	35	0.129				
h×p	0.753	5	0.151	1.34	0.2702	0.2702	1.00
Error	3.933	35	0.112				
d×h×p	2.492	5	0.498	3.55	0.0106	0.0183	0.80
Error	4.911	35	0.140				
block	1.808	3	0.603	3.75	0.0266	0.0547	0.62
Error	3.374	21	0.161				
d×b	0.075	3	0.025	0.26	0.8542	0.8542	1.00
Error	2.030	21	0.097				
h×b	0.454	3	0.151	1.06	0.3892	0.3686	0.59
Error	3.011	21	0.143				
d×h×b	0.357	3	0.119	1.13	0.3600	0.3600	1.00
Error	2.216	21	0.106				
p×b	0.852	15	0.057	0.71	0.7684	0.6141	0.32
Error	8.382	105	0.080				
d×p×b	0.416	15	0.028	0.53	0.9163	0.8197	0.51
Error	5.464	105	0.052				
h×p×b	1.217	15	0.081	1.14	0.3289	0.3548	0.41
Error	7.455	105	0.071				
d×h×p×b	0.799	15	0.053	0.73	0.7511	0.6315	0.41
Error	7.680	105	0.073				

Table 7.22: Experiment 4 – Repeated measures ANOVA for log of RMS error along axes (continued on next page).

Source	SS	df	MS	F	conv. p	adj. p	H-F ϵ
axis	124.151	2	62.076	42.00	< 0.0001	< 0.0001	0.68
Error	20.690	14	1.478				
d×a	4.208	2	2.104	11.45	0.0011	0.0011	1.00
Error	2.573	14	0.184				
h×a	0.453	2	0.227	4.92	0.0240	0.0344	0.81
Error	0.644	14	0.046				
d×h×a	0.920	2	0.460	2.66	0.1052	0.1225	0.78
Error	2.423	14	0.173				
p×a	32.578	10	3.258	21.61	< 0.0001	< 0.0001	0.28
Error	10.554	70	0.151				
d×p×a	3.377	10	0.338	9.06	< 0.0001	< 0.0001	0.57
Error	2.610	70	0.037				
h×p×a	0.303	10	0.030	0.53	0.8652	0.6593	0.28
Error	4.027	70	0.058				
d×h×p×a	0.319	10	0.032	0.77	0.6580	0.6260	0.76
Error	2.902	70	0.041				
b×a	0.126	6	0.021	1.00	0.4372	0.4372	1.00
Error	0.883	42	0.021				
d×b×a	0.194	6	0.032	1.56	0.1836	0.2080	0.72
Error	0.873	42	0.021				
h×b×a	0.403	6	0.067	3.04	0.0146	0.0146	1.00
Error	0.928	42	0.022				
d×h×b×a	0.094	6	0.016	0.64	0.6991	0.6389	0.66
Error	1.029	42	0.025				
p×b×a	0.724	30	0.024	1.52	0.0486	0.0940	0.60
Error	3.338	210	0.016				
d×p×b×a	0.520	30	0.017	1.09	0.3510	0.3792	0.41
Error	3.340	210	0.016				
h×p×b×a	0.508	30	0.017	0.97	0.5212	0.5093	0.73
Error	3.682	210	0.018				
d×h×p×b×a	0.351	30	0.012	0.76	0.8160	0.8160	1.00
Error	3.240	210	0.015				
Total	663.577						

Table 7.22: Experiment 4 – Repeated measures ANOVA for log of RMS error along axes (continued from previous page).

Chapter 8

General Conclusions

The goal of this research was to study the influence of visual feedback and spatial organization upon a series of tasks in a three dimensional environment. The tasks were selected to be simple and thus potentially more generalizable than some of the other studies in the literature. Each task was meant to be only slightly different from one of the preceding tasks.

The experiments were conducted using the same hardware and with only minor software differences. A general difficulty with comparing prior work is the variety of different systems used. Because researchers are still uncovering the factors that influence interactive 3D tasks, there is the potential that factors other than those under investigation are cause some of the observed effects. By examining a sequence of experiments conducted using essentially the same hardware and software it is more likely that differences across experiments can be attributed to the conditions under study rather than being artifacts of the experiment itself. In this manner confounding effects can be reduced.

The point location experiment served as the direct basis for the docking and line tracing experiments. The line tracing experiment in turn served as the basis for the curve tracing experiment. There is an almost endless variety of similar experiments that could have been conducted; some are considered in more detail in the chapter on Future Work that follows this one.

In choosing how to alter the task under examination from one experiment to the next, more or less variation was possible. For example, an alternate successor experiment to the point location experiment having less variation would be one in which the orientation of the target changed, but only the position (rather than the position and the orientation) of the target had to be matched. A successor to the line tracing experiment with more variation would be a curve tracing experiment where the curve varied simultaneously in all three dimensions. The group of experiments selected was meant to have reasonably small changes in complexity between them, while still allowing a fair amount of coverage of task complexity.

A variety of different variables were used to measure performance across each of the tasks. It was evident at the outset of this work that considering only one or two dependent variables such as trial completion time or error would not provide a sufficiently complete sense of how the independent variables influenced task performance. Thus, a selection of objective and subjective measurements were used. Nonetheless, there are still many different potential objective and subjective measures that might have been employed. The final set used here was not meant to be exhaustive. These dependent measures were settled upon as being adequate to provide a reasonable understanding of how the independent variables influenced the tasks as well as demonstrating the importance of looking at a broad set of dependent measures.

The subject pool imposes some limitations on the generalizability of the results. All subjects were right-handed male university students drawn from the fields of computer science and engineering. While it was not deliberately screened for, most participants had little experience with 3D computer graphics, and none had experience with computerized 6 DOF interactive techniques. These results cannot be generalized to population groups with different characteristics.

8.1 Stereopsis

Stereoscopic viewing provides significant benefits

Stereoscopic viewing (binocular disparity) proved to be significantly better than monoscopic viewing – as initialy expected – in the first two experiments. Subjects were able to complete the point location task and the docking task in less time when using a stereoscopic display than when using a monoscopic display. This finding is similar to that of most other research that has studied the use of binocular disparity in computer displays.

The point location and docking tasks require a subject to make a determination of the location in space of the target. Ideally this determination should be made before movement is initiated. Binocular disparity improves a subject's ability to make this determination for targets at the same depth as the starting point as well as for targets at different depths. Had the experiments stopped at this point one might be led to conclude that stereoscopic viewing is likely to allow subjects to perform all tasks faster than with monoscopic viewing.

Stereoscopic viewing not always beneficial

The later experiments demonstrated, contrary to expectations, that stereoscopic viewing did not have a consistently powerful effect. Stereoscopic viewing affected most of the dependent variables fairly clearly in the point location experiment. In the docking experiment stereoscopic viewing had a significant effect upon most of the measures associated with the translational component of the task, but not on the rotational measures. In the line tracing and curve tracing experiments stereoscopic viewing had a more limited effect upon overall performance. Table 8.1 provides the trial completion time for the monoscopic and stereoscopic

Experiment	**Stereo**	**Mono**	**% Increase**	**p**	ω^2
1	3.57	5.00	40.1	< 0.0001	0.2357
2	6.81	8.48	24.5	0.0463	0.0792
3	5.45	6.32	16.0	0.0825	0.0192
4	8.29	9.42	13.6	0.2988	0.0047

Table 8.1: Trial completion time in seconds for the monoscopic and stereoscopic conditions for each experiment. The column labeled increase is the increase in trial completion time for the monoscopic condition over the stereoscopic condition. The p column is the probability that the observed difference between overall means for the two conditions is due to chance and comes directly from the ANOVA table for trial completion time for each experiment. The column labeled ω^2 indicates the magnitude of the experimental effect of display mode upon trial completion time in each of the experiments.

display modes for each of the experiments. Even though stereoscopic viewing did improve performance in every experiment, the improvement was only significant in the first two experiments.

It should be pointed out that the power of each experiment is limited by the number of subjects. The use of eight subjects in each experiment was a deliberate choice. These experiments were designed to detect large effects; the kind that would apply to almost all subjects and cause them to change their behaviour as a result. Had more subjects been tested for the tracing experiments, the observed differences might indeed have been significant. Table 8.1 provides the probabilities for the effect of display mode on trial completion time from the respective ANOVA tables for each experiment. The magnitude of the experimental effect (ω^2) is computed from the appropriate sum of squares and mean squares that appear in the respective ANOVA tables.

In the tracing experiments (Experiments 3 and 4), stereoscopic viewing significantly reduced only the amount of RMS error and the amount of position error along the Z-axis (recall that RMS error is the root mean square of the distance from the tip of the pointer to the ideal path at each movement step). Some care must be taken when drawing a conclu-

Experiment	Stereo	Mono	% Increase	p	ω^2
1	1.02	1.38	35.3	0.0046	0.0910
2	1.03	1.59	54.4	0.0032	0.1003
3	0.59	0.66	11.9	0.0886	0.0104
4	0.62	0.80	29.0	0.0082	0.0664

Table 8.2: Mean RMS error in centimetres for the monoscopic and stereoscopic conditions for each experiment. The column labeled % increase is the increase in RMS error for the monoscopic condition over the stereoscopic condition. The p column is the probability that the observed difference between overall means for the two conditions is due to chance and comes directly from the ANOVA table for RMS error for each experiment. The column labeled ω^2 indicates the magnitude of the experimental effect for display mode on RMS error in each of the experiments.

sion based on the variables that were affected. One could argue that subjects were trading speed for accuracy, that is, subjects were keeping their trial completion time fairly constant while allowing their error level to increase as necessary. Had subjects strived to be more accurate in the monoscopic condition it might have taken them longer to perform the task. However, in the point location experiment, stereoscopic viewing reduced trial completion time, RMS error and position error along the Z-axis. In other words, in the point location and docking experiments stereoscopic viewing allowed subjects to increase both their speed and their accuracy. Table 8.2 contrasts RMS error across all experiments for the monoscopic and stereoscopic display modes.

One possible reason for the subdued effect of stereoscopic viewing in the tracing tasks is the continuous availability of an occlusion cue in the display. Studies by Braunstein et al. [20] and Tittle et al. [94] have shown that occlusion is a more powerful depth cue than stereopsis. In the tracing tasks, the ability to make depth judgments via occlusion may have caused subjects to focus less attention upon the disparity-based depth cues. Nonetheless, even when trial completion time was unaffected by stereoscopic viewing, RMS error along the Z-axis was significantly reduced in the stereoscopic display modes. Even though stereoscopic viewing did not reduce trial completion time, it did improve accuracy.

Stereoscopic viewing has little effect upon rotational tasks

A comparison of the trial completion time for the point location and docking experiments (Table 8.3) shows similar increases between the monoscopic and stereoscopic conditions. This allows us to suggest that the added task complexity of the monoscopic and stereoscopic conditions is unchanged, implying that stereoscopic viewing does not have an effect upon rotation performance. To verify this finding it would be best to conduct a experiment that examined only orientation under the influence of display mode.

Experiment	**Mono**	**Stereo**
Point Location	5.00	3.57
Docking	8.48	6.81
Increase	3.48	3.24

Table 8.3: Mean trial completion time in seconds for the monoscopic and stereoscopic display modes in the point location and docking experiments. The row labeled increase gives the change in trial completion time from the point location experiment to the docking experiment.

8.2 Motion Parallax / Head-Tracking

The change in what we see as our head moves in the environment is an essential component of the human perceptual system. Gibson [45] discusses this and other "natural" elements of perception in great detail. Most researchers in the virtual reality field consider the benefit of motion parallax via head-tracking to be a foregone conclusion.

Head-tracking is an essential part of almost any virtual reality system. The definition of virtual reality that I chose focuses explicitly upon head-tracking as the key element of virtual reality. Virtual reality relies upon the ability of a viewer to look anywhere and have the display reflect the direction of their view. In most systems this requires explicit tracking of head position.

Sollenberger and Milgram [89] showed a strong effect for a rotational display in a task where subjects had to determine which of two visual networks contains a specified point. Several other researchers have shown strong benefits for head-tracking. Arthur et al. [6] replicated the results of Sollenberger and Milgram, using head-tracking instead of a rotational display and they found that the head-coupled stereoscopic viewing condition produced the shortest response times and the lowest error rates. Ware and Balakrishnan [101] make use of a head-coupled stereoscopic display in a volume location task. Their primary purpose was to investigate and model the effect of lag and frame rate in a fish-tank VR environment. They do not compare head-tracked interaction against fixed viewpoint interaction.

The overall results for head-tracking have been quite surprising. My initial expectations were that head-tracking should have some effect, either positive or negative. I expected a positive effect would derive from an improvement in the ability of subjects to correctly perceive the stimuli. Negative effects could appear in the form of subjects taking more time to observe the stimuli prior to movement, or in the form of interference between head movement and hand movement. The general lack of effects suggests that subjects may not have made much use of the view point changes offered by head tracking, or that the beneficial and detrimental aspects of changing one's viewpoint somehow cancel each other out.

Head-tracking has almost no effect on trial completion time

Head-tracking often failed to show a significant effect in the experiments where it was present. Table 8.4 compares trial completion time in the fixed viewpoint and head-tracked viewpoint conditions in the three experiments where head-tracking was tested. It is evident from the table that not only did head-tracking not have a significant effect upon trial completion time, head-tracking made almost no difference at all upon trial completion time. The other experiments in the literature suggest that head-tracking should tend to reduce trial completion

Experiment	Fixed	Head-Tracked	% Increase	p	ω^2
1	4.22	4.35	3.1	0.5820	-0.0098
3	5.88	5.90	0.0	0.9116	-0.0130
4	8.84	8.87	0.0	0.7154	-0.0161

Table 8.4: Trial completion time in seconds for the fixed and head-tracked viewpoint conditions in each of the experiments where head-tracking was tested. The column labeled increase is the increase in trial completion time for the head-tracked viewpoint condition over the fixed viewpoint condition. Small differences appear as 0.0 due to rounding. The p column is the probability that the observed difference between overall means for the two conditions is due to chance and comes directly from the ANOVA table for trial completion time for each experiment. The column labeled ω^2 represents the magnitude of the experimental effect for head-tracking mode on trial completion time in each of the experiments.

time.

Head-tracking often detrimental

Over all of the experiments where head-tracking was tested it exhibited a significant main effect in only two situations, RMS error and axis RMS error for the line tracing experiment. Table 8.5 provides a summary of the RMS error in all the experiments where head-tracking was tested. Although the head-tracking mode only had a significant effect on RMS error in the line tracing experiment (Experiment 3), the trend in all experiments was for head-tracking to impair performance.

In the situations where head tracking was an element of a significant interaction effect, the result was generally to impair performance in some manner. In the curve tracing experiment head-tracking appears to have interfered with the block to block reduction in RMS error seen in the fixed viewpoint condition.

Experiment	Fixed	Head-Tracked	% Increase	p	ω^2
1	1.16	1.23	6.0	0.3216	0.0011
3	0.55	0.70	27.3	0.0281	0.0330
4	0.67	0.75	11.9	0.0929	0.0058

Table 8.5: Mean RMS error in centimetres for the fixed and head-tracked viewpoint conditions in each of the experiments where head-tracking was tested. The column labeled % increase is the increase in RMS error for the head-tracked viewpoint condition over the fixed viewpoint condition. The p column is the probability that the observed difference between overall means for the two conditions is due to chance and comes directly from the ANOVA table for trial completion time for each experiment. The column labeled ω^2 indicates the magnitude of the experimental effect for head-tracking mode upon RMS error in each of the experiments.

Benefits of head-tracking appear to be highly selective

In the curve tracing experiment head-tracking appears to have provided a slight benefit in reducing the amount of RMS error along certain axes for specific target locations. Once again in the curve tracing experiment, the head-tracked monoscopic viewing condition appears to have produced shorter movement paths than the other viewing conditions, but this is only significant for a few of the target positions.

In the experiments described here, head-tracking was part of the viewing condition, but the central task was to move a pointer located in the virtual environment using the hand. Most other studies where head-tracking or motion parallax was found to be beneficial employed a task that involved only viewing of a three dimensional stimulus. In the experiments conducted here, head movements large enough to produce an appreciable amount of motion parallax tended to make it difficult for subjects to control the position of their hand and thus the pointer in the virtual environment. This added control difficulty may be what translated into the generally poorer performance for the head-tracked conditions. As subjects learned about the location and structure of the stimuli in the various experiments they may have lessened their use of the head-tracker resulting in performance that converged to the level of the fixed viewpoint conditions.

Additional insight into the role of motion parallax via head-tracking is provided in a study of teleoperation by Ikehara, Cole and Merritt [56]. In their study subjects used a wand controlled by a telemanipulator to touch points identified by lights located within a mass of twisted wire (wire maze). The task itself was structured to allow separate measures of the time required to spot the target versus the time required to touch the target. The structure of the wire maze is dense enough and complex enough to have provided a great deal of occlusion. They found that motion significantly improved performance for the sub-tasks that required perception only. They concluded that “motion’s major benefit occurs because it improves the initial perception of the target. After the initial perception the benefit of motion is less clear …”

The ability to change one’s viewpoint via head-tracking is likely to have had some utility in the line tracing and curve tracing experiments. There are clearly viewing conditions that would have made the tasks extremely difficult to perform without some kind of viewpoint change. In the line and curve tracing experiments, small viewpoint shifts in the manner of [75] were added so that subjects would have some sense of the 3D structure of the stimulus in the fixed viewpoint monoscopic display mode. Without these shifts, a line along the Z-axis would have appeared as a point (small circle) and a curve would have appeared to be a straight line. Thus an a priori decision was made to provide alternate viewpoints in certain conditions. Recall that aside from enhancing the sense of realism, head-tracking is meant to allow for *continuous* change of viewpoint under user control. If all that is needed is an alternate fixed viewpoint, this can be provided without additional hardware, software and computational resources necessary to interface with this hardware.

The general lack of an effect for head-tracking might also be explained by a tradeoff between task performance and viewpoint change. Subjects may have sacrificed movement time in order to obtain a more favourable view of the stimuli in the environment. In other words, subjects may have been able to perform the task better after moving their head to obtain a new view of the stimuli, but this improved performance was at the cost of the time incurred during the head movement. The results so far indicate that even if this was the case, subjects were able to obtain an equivalent, and in most cases better, level of performance without head-tracking.

It remains to be seen what type of interactive task benefits from head-tracking. Earlier work cited above identified the benefit of head-tracking in comprehending complex visual networks. Perhaps head-tracking would be beneficial in an interactive tracing task that used stimuli similar to those of Sollenberger and Milgram. Nonetheless, occlusion is a powerful depth cue and the ability to interactively probe the depth of objects in the environment using the pointer may take precedence over the motion parallax cues. This is similar to the reduction in the strength of the disparity cues in the line tracing and curve tracing experiments.

8.3 Target Position

Target position generally had a significant effect upon many of the measures across all of the experiments that were conducted. Target position has a dual meaning when interpreting the experiments as the position in space indicates both a stimulus location and a movement endpoint. Because all the trials start from a fixed central location in the workspace, each target position also implies a specific movement direction. For example, the near target implies movement of the pointer toward the subject.

Targets along the Z-axis show poorer performance

The overall results of the experiments show that the target positions that differed in depth from the starting point typically had poorer performance than those that were at the same depth, agreeing with initial expectations. However, contrary to expectations, under certain circumstances performance for the target positions along the Z-axis was improved to the level of the other targets. Notably, in the point location experiment the trial completion time for the near target in the stereoscopic display mode was roughly equal to the trial completion time of the targets at the same depth as the start position. While the trial completion time for the far target in the stereoscopic display mode also improved over its monoscopic level, it remained poorer than for the other targets.

Earlier work by Zhai [109] and Massimino et al. [65] showed that the amount of error along the Z-axis would be higher than the error levels on the X and Y axes. Zhai also showed that stereoscopic viewing would reduce the error levels as compared to monoscopic viewing. However, their studies involved a tracking task making it difficult to identify any effects related to the spatial location of the target, or the organization of the task.

Performance not uniform along axis

In broad terms the left and right target positions along the X-axis tended to result in the best overall performance across tasks and dependent measures. One anomaly appears to be the amount of Z-axis error (both final position error and RMS error) for the right target position in the line tracing and curve tracing experiments. The cause of this increased amount of error is not immediately evident. One potential factor is visual interference from the line or curve coupled with some biomechanical factor. Strictly biomechanical factors are ruled out as this increase in error is not apparent in the point location and docking experiments. Strictly visual factors are ruled out because any potential visual interference should have also occurred for the left target position.

The Z-axis target positions tend to have the poorest performance, however, across most tasks there tends to be a distinct advantage for the near target. In many cases performance for the near target tends to match the performance for the targets in the same plane as the start target. Performance for the far target is almost always worse than for any of the other targets.

Most other studies tend to group together all targets in the same direction or along the same axis. The differences in target positions that varied along the same axis have shown that such a grouping is too simplistic and is likely to disguise an important source of variation.

8.4 Target Orientation

Only the docking experiment directly investigated the effects of target orientation. Thus there is little in the way of cross experiment comparison that can be made. Nonetheless, a few of the conclusions from the docking experiment are worth reiterating.

Translation performance may vary as a result of the target orientation

The docking experiment demonstrated that measures associated primarily with translational performance (position error, path length and RMS error) showed interaction effects between the position of the target and the orientation. The lack of any consistent pattern of influence for target position coupled with the large number of position and orientation conditions makes it difficult to reach any broad conclusion regarding the effect of target orientation upon these measures. The large number of conditions also increases the likelihood that essentially random variation is causing the significant interaction. More investigation is needed before any conclusion that can be held with confidence may be reached.

Rotation performance varies as a function of the target orientation

The measures principally associated with rotational performance showed fairly clear and consistent effects for target orientation. The measures for final orientation error and rotation amount showed that the two Y-axis target orientations had poorer performance than the other target orientations. The combination of the poor orientation match coupled with the higher amount of pointer rotation in the Y-axis conditions suggests that subjects had difficulty determining when the orientation of the pointer matched the pointer of the target. More rotation of the pointer seems to indicate that subjects would continue to reorient the pointer until they arrived at a satisfactory visual match.

Studies by Parsons [73] and Pani et al. [72] have indicated that people have a great deal of difficulty visualizing, comprehending and describing rotations that are not aligned with the world axes[1] or with principal axes of the object. In the docking experiment, all the target orientations were a result of 45°rotations about one of the world axes and in principal should not have been any more difficult to comprehend.

The implication of these results is that the visual appearance of the target made it more difficult to match the orientation of the two 45°Y axis target orientation conditions. Interestingly, many researchers of mental rotation in psychology, when describing the selection of stimuli for their studies, indicate that orientations that resulted in self-occlusion or unusual perspective distortion were excluded (e.g. Shepard and Metzler [84], Shepard and Judd [83], and Parsons [73]). No further details of the excluded conditions are provided and thus the exclusion process seems to be subjective.

[1]The Cartesian axes are often not used as a basis for understanding or describing human cognition, being essentially a mathematical construct. Psychologists tend to describe the world coordinate axes as follows. The *transverse horizontal axis* runs left to right, The *vertical axis* runs up and down, and the *line of sight axis* runs toward and away from the viewer. In the experiments described here, these axes correspond directly to the X, Y and Z axes.

8.5 Occlusion

Occlusion is a factor that was not manipulated in any of the experiments reported here. The displays in all the experiments made use of solid shaded objects and thus occlusion was available as a depth cue in all of the experiments. Many studies have demonstrated the importance of occlusion and the over-riding effect that occlusion cues have over disparity cues. I have hypothesized that the strength of the occlusion cue is what resulted in the apparent reduction of the disparity cues in the tracing experiments. However, there is an important distinction between this experiment and many of the other studies that is worth stressing.

These experiments made use a solid shaded virtual pointer under user control, that could occlude or be occluded by other objects in the virtual environment. Many of the other studies have made use of wireframe objects [109] or they did not provide the user with a controllable pointer. This controllable pointer allows a user to interactively probe the depth of objects within the environment using one of the most powerful depth cues available. The availability of occlusion information throughout most of the tracing task may be what eliminated the effect of disparity, an otherwise strong depth cue as well. While the objects were solid shaded, there was no attempt made to mimic other solid properties, especially the normal impenetrability of solid objects.

During the training phase for many of the tasks, I was able to observe the manner in which subjects performed the task. In the point location task some subjects would determine when the position of the pointer matched the position of the target by determining when part of a face of the pointer partially intersected a face of the target. This cue was so strong and valuable that subjects often commented that they felt they were somehow cheating by making use of this cue. In the tracing experiments subjects would position a face of the pointer so that it was partly occluded by the path. By observing changes in the visibility at this “occlusion interface” subjects were able to track the path.

Informally, it seems that subjects made use of depth cues other than occlusion (e.g. perspective, stereoscopic viewing) to determine the relative locations of objects when they were far apart and switched to using the occlusion cue when objects were close together. When objects were far apart occlusion provided a limited amount of information and made interaction harder. When objects were close together occlusion provided a simple, fast and accurate means of adjusting the depth of the pointer to match the depth of a target.

The ability to completely or partially penetrate objects highlights a significant distinction between the manipulation of virtual objects and teleoperation. Because teleoperation makes uses of (real) solid objects, interpenetration is only possible as far as the materials of the objects allow and even then it may not be desirable.

The interactive use of occlusion has serious implications for the design of interactive tasks in virtual reality. A frequent criticism of VR applications is the ability of one solid object to penetrate another solid object. User's of VR systems feel that solid objects such as walls and floors should be impenetrable. The reason for this shortcoming is the high computational cost of collision detection algorithms that determine when one object has intersected some other object in the environment. But if the computational power were available, complete elimination of the ability for some objects to interpenetrate others might still be undesirable as it would eliminate a valuable depth sensing technique, one that is only available in virtual environments.

8.6 Summary Guidelines

One of the goals of this research was to identify guidelines for the use of the depth cues that were studied. This section presents design guidelines based upon the conclusions in the preceding sections.

Stereopsis

Stereopsis is a powerful depth cue that functions well in the absence of any occlusion cue. The presence of occlusion cues may reduce the benefit that may be provided by stereopsis. Under some circumstances stereopsis can eliminate the performance degradation for movements along the depth axis, especially those that are near the viewer.

Motion Parallax

Contrary to initial expectations, motion parallax did not prove to be a strong depth cue for the tasks that were investigated here. Designers would be advised to carefully consider whether their task has a sufficiently strong perceptual requirement that would benefit from the availability of a motion parallax cue. It appears that many of the benefits of head-tracking can be derived through the use of less expensive (both computationally and monetarily) means such as alternative or multiple views.

Target Position

The position of objects within the workspace can have a significant impact upon task performance. Performance when working with objects far from the user is likely to be worse than when working with objects at the same depth. In the absence of stereopsis any change of depth across objects is likely to result in poorer performance. The presence of stereopsis can bring performance for objects nearer to the viewer to the level of objects without any depth change.

Target Orientation

The ability to achieve certain orientations using a 6DOF pointer was not uniform across the target orientations that were tested. Some target orientations were much more difficult to match than others. While it may be practical to restrict the working volume as necessary to achieve a certain performance level, it is unlikely that the full range of orientations can be restricted in a similar fashion. A guideline suggesting that certain orientations be avoided is impractical because it would probably result in an unusable system. These findings are probably most important to the research community. They indicate that orientations require careful advance screening before being used because certain orientations are likely to overstate overall task difficulty if all orientations are grouped together.

Chapter 9

Future Work

The experiments conducted so far have suggested a wide variety of future analyses and experiments that could be performed.

9.1 Additional results from captured data

All tracking data for the head-tracker and pointer were captured in every experiment. At the time the experiments were conducted, only information on trial completion time and final error (position and orientation) had been extracted from the raw data and was readily available. The initial results based on these measures for the point location experiment were presented and published at the ACM Symposium on Virtual Reality Software and Technology (VRST) ‘97 [16]. The initial results for the docking experiment were presented at the IEEE Virtual Reality Annual International Symposium (VRAIS) ‘98 [18].

The full set of descriptive statistics used for the analyses presented in earlier chapters was extracted after all the experiments had been completed. The full availability of a temporal record of the exact actions of the subjects opens the possibility of conducting additional analysis on the already completed experiments. Another dependent measure that might provide additional insight into the subjects' operation during a task could be extracted. For example, in the docking experiment it might be interesting to know whether subjects tended to bring the pointer close to the default orientation at the beginning of the task even though this was not required.

There is an almost endless variety of additional variables that could be examined. Two stand out as providing the most interesting additional analyses. One is to develop a best path analysis and compute error metrics against this best path. The other is to perform an analysis of the captured head motion in search of interesting properties.

Current error metrics are evaluated against the best line or distance from the displayed line or curve. An alternative would be to process the captured data and develop a model of the "ideal" movement path. The error metric for each trial could then be computed against this ideal movement path. In one approach, all movements are "averaged" together to yield the ideal movement path. A somewhat more advanced approach assumes that the best path is still the one displayed. However, rather than computing the error metric as the distance from the pointer to the curve, the curve is parameterized so as to have an ideal position on the curve for each pointer position within the movement.

Head-movement has not been extensively analyzed in the preceding chapters. In the analysis that was done, head-tracking is either available or unavailable to subjects, but no analysis has been done to determine how much head movement subjects actually make. The availability of the original data makes it possible to extract information about the amount of head movement used. It may be possible to determine whether subjects were able to improve their performance while continuing to make use of head-tracking in later blocks of an experiment, or whether they simply discontinued their use of head-tracking.

The virtual environment research community generally feels that head-tracking provides important benefits. This belief is somewhat contrary to the findings reported here where head-tracking was generally detrimental, if it had any effect at all. It may be that the tasks studied here do not require head-tracking.

9.2 Head Movement Analysis

To gain some additional insight into the role of head-tracking in these experiments, a brief analysis was conducted. *Total head movement* is defined as the amount of head movement in centimetres between the start click and end click of a trial and is computed as the sum of the Euclidean distance of consecutive head positions. The total head movement was computed for Experiment 1 and Experiment 3. Recall that these two studies differed only by the addition of a line connecting the start and end positions. Repeated measures ANOVAS were computed using the same independent variables as in those experiments.

For Experiment 1, significant main effects were found for display mode $F(1,7) = 38.40$, p = 0.0004, target position $F_{HF}(1,7) = 7.59$, p = 0.0001, and block $F_{HF}(3,21) = 3.08$, p = 0.0495. A significant two-way interaction of display mode $\times$ target position $F_{HF}(5,35) = 5.57$, p = 0.0007, was also found.

For Experiment 3, a significant main effect for target position $F_{HF}(2.1,14.7) = 25.55$, $p < 0.0001$ was found. A significant two-way interaction was found for head-tracking mode $\times$ target position $F_{HF}(2.1,14.5) = 10.67$, p = 0.0013.

The results for Experiment 1 showed that across all target positions, head movement increased from 4.51cm in the stereoscopic display mode to 6.33cm in the monoscopic display mode. This might indicate that subjects attempted to make use of some motion parallax information in the monoscopic display mode, however, it is not clear that the small amount of head movement in either of these conditions is a result of intentional movement on the part of the subject. The larger amount of head movement may be a result of the longer trial completion time. Across all display mode and target position conditions, the y-top target

Block	Head Movement
1	5.73
2	5.17
3	5.48
4	5.30

Table 9.1: Mean amount of head movement in centimetres for each block of Experiment 1.

Position	Fixed	Head-tracked
x-left	4.79	5.09
x-right	5.27	4.89
y-bottom	4.97	5.58
y-top	5.91	6.46
z-far	8.74	17.34
z-near	7.61	17.34

Table 9.2: Mean head movement in centimetres for each of the target positions in both head-tracking modes for Experiment 3.

position in the monoscopic condition showed the most head movement at 7.55cm and the z-near target in the stereoscopic condition showed the least amount of head movement with 3.70cm.

Table 9.1 shows the amount of head movement across blocks of the experiment. There is a downward trend in the average amount of head movement from 5.73cm in the first block to 5.30cm in the fourth block. This finding is consistent with the hypothesis that subjects reduced their head-movement over time. However, once again, it is not clear that the amount of the reduction in absolute terms is sufficient to be the result of intentional differences on the part of the subject.

The results of Experiment 3 are more illuminating. Table 9.2 shows the amount of head movement for each target position in both head-tracking modes. It is immediately evident that there is a large increase for the two Z-axis target positions over the X- and Y-axis target positions in the head-tracked condition. While there is also an increase in the amount of head movement for the two Z-axis target positions over the X- and Y-axis target positions in the fixed viewpoint condition, it is much smaller than the increase in the head-tracked condition. Head movement of the magnitude obtained for the Z-axis target positions in the head-tracked condition produce large changes in the appearance of the stimuli on the display.

Some amount of head movement occurred regardless of the experimental condition. The amount of head movement in the X- and Y-axis target positions where head-tracking provided little benefit may serve as a baseline from which to draw further conclusions. The average head movement of the X- and Y-axis target positions in both head-tracking modes was 5.73cm. In the fixed viewpoint condition, head movement in the two Z-axis target positions goes above this baseline level. This would seem to indicate that when head-tracking was not enabled, subjects moved their heads in order to obtain a better view, but soon realized that their head movements had no effect upon the display.

The difference in the amount of head movement between Experiments 1 and 3, especially for the Z-axis target positions, demonstrated that subjects made greater use of head movement when the result of that movement was potentially useful. In Experiment 1, the parallax shifts changed the position of the target on the display, but did little to improve perception. In Experiment 3, the line to be traced extended over a large depth and the parallax shifts were able to markedly improve the perception of the Z-axis targets.

The head movement results indicate that subjects made use of the head-tracking mechanism in the conditions where a viewpoint change was most beneficial. However, this head movement failed to translate into improved performance as currently measured in the experiments. If there is some tradeoff between task performance and viewpoint change it may be possible to detect this tradeoff though further analysis of the captured tracking data.

Can hand movement tasks benefit from head-tracking?

A potential challenge for future work would be to find hand movement tasks that benefit from head-tracking. The evidence from the experiments here suggests that the tasks would have to involve a substantial perceptual component in order to benefit. One possibility is a lengthy complex task composed of simpler components where several view changes are required. It may be that each component task gains very little from the availability of motion parallax, but when the components are assembled, head-tracking allows viewing changes that are essential to the performance of the compound task. The question of whether continuous head-tracking is required or just a selection of alternate fixed views is an interesting one with important implications about the kind of hardware necessary for desktop or "fish tank" virtual environments.

9.3 Future Experiments

There are a variety of future experiments that might be conducted as as result of the work reported here. Several opportunities for further study have already been mentioned in earlier chapters during the discussion of the results. This section highlights some of the most interesting potential future studies.

Increased path complexity

There are many variations on the two primary tasks (pointing and tracing) that one might study. One obvious follow-on to the experiments described in earlier chapters would be to test performance in a curve tracing task where all three dimensions are allowed to vary simultaneously. Another alternative is to investigate a curve tracing task where different curves are dynamically generated for each trial. In the curve tracing task described in Chapter 7, the same basic curve, a single sine wavelength with an amplitude of 2cm, was used for all trials. Subjects may have developed muscle memory for the curve, causing them to rely less upon the visual information that was presented.

Investigation of target position and target orientation

The interaction between position and orientation factors in the docking experiment suggested a complex relationship between the position and orientation factors. The results suggest that different pointer orientations might have an effect upon the ability to locate points in space or that different locations in space might have an effect upon the ability to match orientation. The large number of factors involved coupled with the lack of any clear overall pattern make it difficult to make definite statements about how position and orientation factors may influence each other. Further experiments could be conducted to determine whether position and orientation do indeed influence each other or whether there are additional confounding factors that produced the results reported here.

The influence of auxiliary feedback cues

None of the experiments that were conducted provided any feedback to subjects beyond the initial requirement to match the position of the start target. When feedback was present during the training trials subjects relied heavily upon the feedback to guide them through the task. Experiments similar to those described here could be conducted to see how large an effect the presence of feedback has upon task performance.

Point location versus volume location

Some earlier studies that made use of a volume location task did not report any differences in task difficulty based on the position of the target in the environment. Volume location is a useful operation in virtual environments where objects are selected and manipulated by moving the pointer into the volume of the object. However, most graphics applications rely upon the ability to select and manipulate small objects such as the handles of an object or the control vertices of a spline curve. A direct comparison of a point location and volume location task would be useful. Is there a fundamental difference between these two tasks, or is the point location task just more sensitive because of its increased accuracy requirement?

Interactive selection of depth cues

Reinhart [80] [79] conducted several studies of depth cues and found that the availability of a single depth cue resulted in a large performance improvement. Additional depth cues resulted in additional performance gains, but the benefits diminished as each cue was added. His findings highlight the general belief that the best performance will be obtained by adding additional depth cues to the display until all the depth cues available in the real world are also available in the virtual world.

The results of the experiments reported here have shown that in some cases the mechanism for providing a motion parallax depth cue can impair performance. The results also showed that the availability of a occlusion depth cue reduced the effectiveness of the stereoscopic depth cue. The nature of the occlusion cue whereby subjects could make interactive depth probes suggests a potential alternative avenue for further study. The goal of these studies would be to determine what influence the ability to interactively activate and deactivate different depth cues would have on virtual environment interaction. Would all the available depth cues be used? Would some depth cues be used selectively? Would some depth cues not be used because they do not provide sufficient benefit?

Investigation of additional depth cues in an interactive environment

Wanger et al. [100] have studied factors that affect the ability to perceive spatial relationships in static 3D images such as shadows, rendering quality and lighting. There are a wide variety of additional depth cues such as these that could be tested to determine if they have any influence upon performance in an interactive task. The results here have shown that the ability to interact with the environment and probe depth produces different results than tasks involving only perception. The implication is that our understanding of depth cues in non-interactive settings needs to be more thoroughly investigated for interactive environments.

Bibliography

[1] Alias Research Inc. Alias Upfront. Commercial computer software, 1991. Toronto, Ontario, Canada.

[2] Alias Research Inc. Alias Sketch! Commercial computer software for Apple Macintosh, 1992. Toronto, Ontario, Canada.

[3] R. E. Allen, editor. *The Concise Oxford Dictionary of Current English*. Oxford University Press, New York, 8th edition, 1990.

[4] S. L. Altmann. *Rotations, Quaternions and Double Groups*. Oxford: Clarendon Press, 1986.

[5] Marian Annett. A classification of hand preference by association analysis. *British Journal of Psychology*, 61(3):303–321, 1970.

[6] Kevin W. Arthur, Kellogg S. Booth, and Colin Ware. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*, 11(3):239–265, 1993.

[7] Ascension Technology Corporation, Burlington, Vermont, U.S.A. *The Ascension Bird, Position and Orientation Measurement System, Installation and Operation Guide*, June 1992.

[8] David Banks. Interactive manipulation and display of two-dimensional surfaces in four-dimensional space. In *1992 Symposium on Interactive 3D Graphics*, pages 197–207, 1992.

[9] John A. Baro and Stephen Lehmkuhle. InSight 2-InColor. Computer software for Apple Macintosh.

[10] Eric A. Bier. Skitters and jacks: Interactive 3D positioning tools. In *1986 Workshop on Interactive 3D Graphics*, pages 183–196, 1986.

[11] Eric A. Bier. Snap-dragging in three dimensions. In *1990 Symposium on Interactive 3D Graphics*, pages 193–204, 1990.

[12] Eric A. Bier and Maureen C. Stone. Snap-dragging. In *Proceedings of ACM SIGGRAPH '86*, pages 233–240, 1986.

[13] Chuck Blanchard, Scott Burgess, Young Harvill, Jaron Lanier, Ann Lasko, Mark Oberman, and Michael Teitel. Reality built for two: A virtual reality tool. In *1990 Symposium on Interactive 3D Graphics*, pages 35–36, March 1990.

[14] Teresa W. Bleser. *An Input Device Model of Interactive System Design*. Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, The George Washington University, Washington, D.C., 1991.

[15] R.A. Bolt. "Put-That-There": Voice and gesture at the graphics interface. In *Computer Graphics (SIGGRAPH '80 Proceedings)*, pages 262–270. ACM, July 1980.

[16] J. Boritz and K. S. Booth. A study of interactive 3D point location in a computer simulated virtual environment. In *ACM Symposium on Virtual Reality Software and Technology (VRST) '97*, pages 181–187, 1997.

[17] James Boritz. Three dimensional gesture recognition. Course project for CPSC 514 Topics in Computer Graphics, UBC Computer Science, 1994.

[18] James Boritz and Kellogg S. Booth. A study of interactive 6DOF docking in a computerised virtual environment. In *Proceedings of IEEE Virtual Reality Annual Internationl Symposium (VRAIS) '98*, pages 139–146, 1998.

[19] James Boritz, Kellogg S. Booth, and William B. Cowan. Fitts' law studies of directional mouse movement. In *Proceedings of Graphics Interface '91*, pages 216–223, 1991.

[20] M. L. Braunstein, G. J. Anderson, M. W. Rouse, and J. S. Tittle. Recovering viewer-centered depth from disparity, occlusion and velocity gradients. *Perception & Psychophysics*, 40:216–224, 1986.

[21] Fred P. Brooks. Walkthrough – A dynamic graphics system for simulating virtual buildings. In *1986 Workshop on Interactive 3D Graphics*, pages 77–78. ACM, 1986.

[22] William Buxton. Lexical and pragmatic considerations of input structures. *Computer Graphics*, 17(1):31–37, January 1983.

[23] William Buxton. There's more to interaction than meets the eye: Some issues in manual input. In D. A. Norman and S. W. Draper, editors, *User Centered System Design: New Perspectives on Human Computer Interaction*, pages 319–337. Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.

[24] Stuart K. Card, K. W. English, and Betty J. Burr. Evaluation of mouse, rate-controlled isometric joystick, step keys and text keys for text selection on a CRT. *Ergonomics*, 21:601–613, 1978.

[25] Arthur Cayley. On certain results relating to quaternions. *Philosophical Magazine*, xxvi:141–145, February 1845.

[26] M. Chen, S. J. Mountford, and A. Sellen. A study in interactive 3D rotation using 2D control devices. In *Proceedings of ACM SIGGRAPH '88*, pages 121–129, 1988.

[27] Michael Chen. A technique for specifying rotations in three dimensions using a 2D input device. In *Proceedings IEEE Montech '87 – Compint '87*, pages 118–120, 1987.

[28] James C. Chung. A comparision of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams. In *1992 Symposium on Interactive 3D Graphics*, pages 193–196, 1992.

[29] Domenic V. Cicchetti. Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin*, 77(6):405–408, 1972.

[30] James Clark. Designing surfaces in 3D. *Communications of the ACM*, 19(8):454–460, 1976.

[31] D. Brookshire Conner, Scott S. Snibbe, Kenneth P. Herndon, Daniel C. Robbins, Robert C. Zeleznik, and Andries van Dam. Three-dimensional widgets. In *1992 Symposium on Interactive 3D Graphics*, pages 183–188, 1992.

[32] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–72, June 1992.

[33] Michael Deering. High resolution virtual reality. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, pages 195–202, July 1992.

[34] Michael F. Deering. Holosketch: A virtual reality sketching / animation tool. *ACM Transactions on Computer-Human Interaction*, 2(3):220–238, September 1995.

[35] K. Evans, P. Tanner, and M. Wein. Tablet-based valuators that provide one, two or three degrees of freedom. In *Proceedings of ACM SIGGRAPH ‘81*, pages 91–97, 1981.

[36] D. Farnsworth. The Farnsworth dichotomous test for color blindness. The Psychological Corp., 1947.

[37] Steven Feiner, Blair MacIntyre, and Doree Seligmann. Knowledge-based augmented reality. *CACM*, 36(7):52–62, 1993.

[38] S. Fisher, M. McGreevy, J. Humphries, and W. Robinett. Virtual environment display system. In *1986 Workshop on Interactive 3D Graphics*, pages 77–87, 1986.

[39] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–391, June 1954.

[40] James D. Foley and Victor L. Wallace. The art of natural graphic man–machine conversation. *Proceedings of the IEEE*, 62(4):462–470, April 1974. Reprinted in John C. Beatty and Kellogg S. Booth, Tutorial: Computer Graphics.

[41] J.D. Foley, V.L. Wallace, and P. Chan. The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications*, 4(11):13–48, November 1984.

[42] Andrew Forsberg, Kenneth Herndon, and Robert Zeleznik. Aperture based selection for immersive virtual environments. In *Proceedings of UIST ’96*, pages 95–96, 1996.

[43] W. R. Garner. *The Processing of Information and Structure*. Lawrence Erlbaum, Hillsdale, New Jersey, 1974.

[44] W. R. Garner and G. L. Felfoldy. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology 1*, pages 225–241, 1970.

[45] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[46] J. Paul Grice. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics*, volume 3: Speech Acts, pages 41–58. Academic Press, New York, NY, 1975.

[47] Carl Grubin. Derivation of the quaternion scheme via the Euler axis and angle. *Journal of Spacecraft and Rockets*, 7(10):1261–1263, October 1970.

[48] GSPC. Status report of the Graphics Standards Planning Committee of the ACM/SIGGRAPH. *Computer Graphics*, 11(3):19+117, Fall 1977.

[49] GSPC. Status report of the Graphics Standards Planning Committee. *Computer Graphics (SIGGRAPH '79 Proceedings)*, 13(3):274, August 1979.

[50] GSPC. GKS, information processing systems, computer graphics, graphical kernel system (X3H3/83–25r3). *Computer Graphics Special Issue*, page 396, February 1984.

[51] Sir William Rowan Hamilton. On quaternions; or on a new system of imaginaries in algebra. *Philosophical Magazine*, xxv:10–13, July 1844.

[52] Pat Hanrahan and Paul Haeberli. Direct WYSIWYG painting and texturing on 3D shapes. In *Proceedings of ACM SIGGRAPH '90*, pages 215–224, 1990.

[53] Ken Hinckley, Joe Tulio, Randy Pausch, Dennis Proffitt, and Neal Kassell. Usability analysis of 3D rotation techniques. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST-97)*, pages 1–10, New York, October14–17 1997. ACM Press.

[54] David C. Howell. *Statistical Methods for Psychology*. Duxbury Press, Boston MA, 1989.

[55] David C. Howell. *Statistical Methods for Psychology*. Duxbuy Press, 4th edition, 1997.

[56] Curtis Ikehara, Robert E. Cole, and John O. Merritt. Visual-motor correspondence in stereoscopic video displays for teleoperated manipulator tasks. In *Proceedings of SPIE Vol. 1915 Stereoscopic Displays and Applications IV*, pages 167–176, 1993.

[57] Robert J. K. Jacob, Linda E. Sibert, Daniel C. McFarlane, and M. Preston Mullen Jr. Integrability and separability of input devices. *ACM Transactions on Computer-Human Interaction*, 1(1):3–26, 1994.

[58] P. Johnson. Hardware for stereoscopic computer graphics and imaging. *Information Display*, 5(7&8):16–19, 1989.

[59] Timothy E. Johnson. Sketchpad III, a computer program for drawing in three dimensions. In *Proceedings of the Spring Joint Computer Conference*, 1963. Reprinted in *Tutorial and Selected Readings in Interactive Computer Graphics*, Herbert Freeman ed., IEEE Computer Society, Silver Spring MD 1984, pp. 20-26.

[60] Pierre Jolicoeur, Shimon Ullman, and Marilynn Mackay. Visual curve tracing properties. *Journal of Experimental Psychology: Human Perception and Performance*, 17(4):997–1022, 1991.

[61] Jiandong Liang and Mark Green. JDCAD: A highly interactive 3D modeling system. *Computers and Graphics*, 18(4):499–506, July–August 1994.

[62] L. Lipton. Factors affecting ghosting in time-multiplexed planostereoscopic CRT display systems. In *Proceedings of SPIE Vol. 761 True 3D Imaging Techniques and Display Technologies*, pages 75–78, 1987.

[63] Lenny Lipton. Selection devices for field-sequential stereoscopic displays: A brief history. In *Proceedings of SPIE Vol. 1457 Stereoscopic Displays and Applications II*, pages 274–282, 1991.

[64] Jock Mackinlay, Stuart K. Card, and George G. Robertson. A semantic analysis of the design space of input devices. *Human-Computer Interaction*, 5(2-3):145–190, 1990.

[65] M. J. Massimino, T. B. Sheridan, and J. B. Roseborough. One handed tracking in six degrees of freedom. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pages 498–503, 1989.

[66] S. W. McWhorter, L. F. Hodges, and W. E. Rodriguez. Comparison of 3-D display formats for CAD applications. In *Proceedings of SPIE Vol. 1457 Stereoscopic Displays and Applications II*, pages 85–90, 1991.

[67] Mark R. Mine. Virtual environment interaction techniques. Technical Report TR95-018, Department of Computer Science, University of North Carolina - Chapel Hill, May 4 1995.

[68] William M. Newman and Robert F. Sproull. *Principles of Interactive Computer Graphics*. McGraw-Hill, first edition, 1973.

[69] Gregory M. Nielson and Dan R. Olsen, Jr. Direct manipulation techniques for 3D objects using 2D locator devices. In *1986 Workshop on Interactive 3D Graphics*, pages 175–182. ACM, 1986.

[70] Donald A. Norman. *The Psychology of Everyday Things*. Basic Books, New York, 1988.

[71] R. C. Oldfield. The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9:97–113, 1971.

[72] John R. Pani, Colin T. William, and Gordon T. Shippey. Determinants of the perception of rotational motion: Orientation of the motion to the object and to the environment. *Journal of Experimental Psychology: Human Perception and Perfromance*, 21(6):1441–1456, 1995.

[73] L. Parsons. Inability to reason about an object's orientation using an axis and angle of rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6):1259–1277, 1995.

[74] Robert Patterson and Wayne L. Martin. Human stereopsis. *Human Factors*, 34(6):669–692, 1992.

[75] Cary B. Phillips, Norman I. Badler, and John Granieri. Automatic viewing control for 3D direct manipulation. In *1992 Symposium on Interactive 3D Graphics*, March 1992.

[76] Polhemus Incorporated, Colchester, Vermont, U.S.A. *3SPACE FASTRAK User's Manual,*, revision f edition, February 1993.

[77] E. C. Poulton. *Tracking Skill and Manual Control*. Academic Press, New York, 1974.

[78] Frederick H. Raab, Ernest B. Blood, Terry O. Steiner, and Herbert R. Jones. Magnetic position and orientation tracking system. *IEEE Transactions on Aerospace and Electronic Systems*, AES-15(5):709–718, 1979.

[79] William F. Reinhart. Depth cueing for visual search and cursor positioning. In *Proceedings of SPIE Vol. 1457 Stereoscopic Displays and Applications II*, pages 221–232, 1991.

[80] William F. Reinhart, Robert J. Beaton, and Harry L. Snyder. Comparison of depth cues for relative depth judgements. In *Proceedings of SPIE Vol. 1256 Stereoscopic Displays and Applications*, pages 12–21, 1990.

[81] John M. Reising, Kristen K. Liggett, C1C Chris Rate, and David C. Hartsock. 3-D target designation using two control devices and an aiding technique. In *Proceedings of SPIE Vol. 1669, Stereoscopic Displays and Applications III*, pages 146–154, 1992.

[82] Christopher Schmandt. Spatial input/display correspondence in a stereoscopic computer graphic workstation. In *Proceedings of ACM SIGGRAPH '83*, pages 253–259, 1983.

[83] R. N. Shepard and S. A. Judd. Perceptual illusion of rotation of three-dimensional objects. *Science*, 191:952–954, 1976.

[84] R. N. Shepard and J. Metzler. Mental rotation of three-dimensional objects. *Science*, 171:701–703, 1971.

[85] Ken Shoemake. Animating rotations with quaternions. In *Proceedings of ACM SIGGRAPH '85*, pages 245–254, 1985.

[86] Ken Shoemake. Arcball: A user interface for specifying three-dimensinal orientation using a mouse. In *Proceedings of Graphics Interface '92*, pages 151–156, May 1992.

[87] Shooting Star Technology, Rosedale, British Columbia, Canada. *ADL-1 6 Degree of Freedom Tracking System User's Manual*, version 2.1 firmware edition.

[88] J. F. Soechting, F. Lacquanti, and C. A. Tersuolo. Coordination of arm movements in three-dimensional space. Sensorimotor mapping during drawing movement. *Neuroscience*, 17(2):295–311, 1986.

[89] R. L. Sollenberger and P. Milgram. Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors*, 35(3):483–499, 1993.

[90] David Sturman. *Whole–hand Input*. Ph.D. dissertation, Media Arts and Sciences Section, School of Architecture and Planning, Massachusetts Institute of Technology, 1992.

[91] Ivan E. Sutherland. Sketchpad, a man-machine graphical communication system. In *Proceedings of the Spring Joint Computer Conference*, 1963. Appears in *Tutorial and Selected Readings in Interactive Computer Graphics*, Herbert Freeman ed., IEEE Computer Society, Silver Spring MD 1984 pp. 2–19.

[92] Ivan E. Sutherland. Head mounted three dimensional display. In *Proceedings of the Fall Joint Computer Conference*, pages 757–764, 1968.

[93] Haruo Takemura, Akira Tomono, and Yukio Kobayashi. An evaluation of 3-D object pointing using a field sequential stereoscopic display. In *Proceedings of Graphics Interface '88*, pages 112–118, 1988.

[94] J. S. Tittle, M. W. Rouse, and M. L. Braunstein. Relationship of static stereoscopic depth perception to performance with dynamic stereoscopic displays. In *Proceedings of the 32nd Annual Meeting of the Human Factors Society*, pages 1439–1442, Santa Monica, CA, 1988. Human Factors Society.

[95] James S. Tittle and Myron L. Braunstein. Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics*, 54(2):157–169, 1993.

[96] Anita Trip. Toward the virtual hand lab in the summer of 1994. Practicum Report for SFU Kinesiology 352, September 1994.

[97] D. Vickers. *Sorceror's Apprentice: Head Mounted Display and Wand*. Ph.D. dissertation, Department of Computer Science, University of Utah, Salt Lake City, 1974.

[98] John Vince. *The Language of Computer Graphics*. Van Nostrand Reinhold, 1990.

[99] Yanqing Wang, Christine L. MacKenzie, Valerie A. Summers, and Kellogg S. Booth. The structure of object transportation and orientation in human-computer interaction. In *Proceedings of ACM CHI '98 Conference on Human Factors in Computing Systems*, pages 312–319, 1998.

[100] Leonard R. Wanger, James A. Ferwerda, and Donald P. Greenberg. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, 12(3):44–58, May 1992.

[101] Colin Ware and Ravin Balakrishnan. Reaching for objects in VR displays: Lag and frame rate. *ACM Transactions on Computer-Human Interaction*, 1(4):331–356, 1994.

[102] Colin Ware and Ravin Balakrishnan. Target acquisition in fish tank VR: The effects of lag and frame rate. In *Proceedings of Graphics Interface '94*, 1994.

[103] Colin Ware and S. Osborne. Exploration and virtual camera control in virtual three dimensional environments. In *1990 Symposium on Interactive 3D Graphics*, pages 175–183, 1990.

[104] Colin Ware and Leonard Slipp. Using velocity control to navigate 3D graphical environments: A comparison of three interfaces. In *Proceedings of the Human Factors Society 35th Annual Meeting*, volume 1 of *Computer Systems: Input Techniques*, pages 300–304, 1991.

[105] *Webster's 7th Collegiate Dictionary*. Merriam-Webster Inc., 1963.

[106] C. D. Wickens, S. Todd, and K. Seidler. Three-dimensional displays: Perception, implementation and applications. Technical Report 89-001, CSERIAC, Wright Patterson Air Force Base, Ohio, 1989.

[107] Robert C. Zeleznik, Kenneth P. Herndon, Daniel C. Robbins, Nate Huang, Tom Meyer, Noah Parker, and John F. Hughes. An interactive 3D toolkit for constructing 3D widgets. In James T. Kajiya, editor, *Computer Graphics (SIGGRAPH '93 Proceedings)*, pages 81–84, August 1993.

[108] S. Zhai and P. Milgram. Human performance evaluation of manipulation schemes in virtual environments. In *Proceedings of VRAIS '93: IEEE Virtual Reality Annual International Symposium*, pages 155–161, 1993.

[109] S. Zhai and P. Milgram. Asymmetrical spatial accuracy in 3D tracking. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*, 1994.

[110] Shumin Zhai. *Human Performance in Six Degree of Freedom Input Control*. Ph.D. dissertation, University of Toronto, Department of Industrial Engineering, 1995.

[111] Shumin Zhai, William Buxton, and Paul Milgram. The “silk cursor”: Investigating transparency for 3D target acquisition. In *Proceedings of ACM CHI ‘94 Conference on Human Factors in Computing Systems*, pages 459–464, 1994.

[112] Shumin Zhai and Paul Milgram. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proceedings of ACM CHI ‘98 Conference on Human Factors in Computing Systems*, pages 320–327, 1998.

[113] Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson, and Young Harvill. A hand gesture interface device. In J.M. Carroll and P.P. Tanner, editors, *Proceedings of Human Factors in Computing Systems and Graphics Interface ’87*, pages 189–192, April 1987.

Appendix A

Study Forms

A.1 Consent Form

A.2 Edinburgh Inventory

A.3 Post-Session Questionnaire Experiment 1

A.4 Post-Session Questionnaire Experiment 2

A.5 Post-Session Questionnaire Experiment 3

A.6 Post-Session Questionnaire Experiment 4

A.7 Post-Experiment Questionnaire

Appendix B

Ethics Submission

B.1 Initial Submission

B.2 Amendment One