Improving Depth Perception in 3D Interfaces with Sound

by

Stephen Wesley Mereu

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 1995

I hearby declare that I am the sole author of this thesis.

I authorize the University of Waterloo to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the University of Waterloo to reporduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

The University of Waterloo requires the signatures of all persons using or photocoping this thesis. Please sign below, and give address and date.

Abstract

This thesis investigates the use of sound to improve user depth perception in 3D computer interfaces. Two experiments were performed to determine a) the effect that sound has on depth perception and b) how this effect varies over an extended period of use. When compared to a “no sound” environment with limited visual depth cues, sound feedback helped to reduce task errors while at the same time increased user appreciation. Task completion under the use of sound did, however, take longer than the no sound environment. “Sound only” trials, which have no visual feedback, were also tested and in some cases reached the same performance level of the no sound case suggesting a possible use in applications not requiring a visual display. Three types of audible sound environments were studied each using different sound feedback: a simple tone, a musical piece, and a more involved musical piece using a fixed orchestra arrangement. Overall, the tonal environment proved to have the lowest error levels, yet was more annoying than the other two. An extended period study was performed, that required subjects to repeat the experiment on a Monday, Wednesday, and Friday, to measure the effect that experience had on performance. The results showed little improvement in task accuracy but a significant improvement in task time. Extra trials were also performed to determine if sounds at random locations could be learned and later located by memory. Although performance for these memory trials did decrease compared to the regular trials, there was still a notable improvement over the no sound condition. In conclusion, sound was found to be an effective mechanism to improve depth perception.

Acknowledgements

This research was funded by a NSERC research grant, and ICR and UW Graduate scholarships. I would like to thank my supervisor, Rick Kazman, for support and guidance throughout the project and my readers William Cowan and Graham Strong.

My parents and house mates were also very encouraging of which I am very thankful. I would like to give special thanks to my dear friend Melanie Lawson for her prayers and encouraging words that kept me going in the final stages.

Most of all I would like to thank my Lord and Saviour Jesus Christ. Not only has He saved my soul but He continues to bless and make intercession for me saying, “I will never leave thee nor forsake thee” and “Him that cometh to me I shall in no wise cast out”.

Trademarks

AutoCAD is a registered trademark of Autodesk Inc.
Alias is a registered trademark of Alias Research Inc.
Pentium is a registered trademark of Intel Inc.
Prisms 3D Animation is a registered trademark of Side Effects Software Inc.
SAS is a registered trademark of SAS Institute Inc.
UltraSound MAX is a registered trademark of Advanced Gravis Computer Technology Ltd.

Any other trademarks referred to in this thesis but have not been acknowledged are omissions and should be given their proper acknowledgment.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Over the years, computers have become faster and more powerful. With such computational power, users are able to do more tasks using computers than ever before. One recent area of advancement has been in the performance of interactive 3-dimensional (3D) graphics, thus opening the way for numerous 3D applications. Some of these include computer aided design (CAD) packages like AutoCAD, or modeling and animation applications like Alias and Prisms 3D Animation.

People rely on many different stimuli in their natural environment to help them to perceive depth. A 3D application must therefore simulate these depth messages so that the user can perceive three dimensions even though it is displayed on a 2D screen. A lack of depth information provided by a computer interface will hinder the user's depth perception and thus degrade performance. This is the greatest problem with most commercial 3D packages available today.

Many applications, like AutoCAD, attempt to overcome this difficulty by providing four views of the modeled scene; three orthogonal views: top, front, side, and one perspective view. These views help, but they place a greater workload on the user who has to reconstruct the 3D model in their mind rather than simply viewing it.

Other alternatives to having multiple model views include improving the input device from a standard 2D mouse to one that offers six degrees of freedom: left/right, up/down, front/back, and then the rotations pitch, yaw, and roll. [Ware 91] Some input devices, like a mouse, are used only for input and give no output information. Others provide *haptic feedback* that resists the user's motion. This extra resistance varies as the user moves the cursor around the 3D world. When an object is contacted it is felt directly through the input device thus giving the user a greater sense of depth perception. [Brooks, Ouh-Young, Batter & Kilpatrick 90], [Das, Zak, Kim, Bejczy & Schenker 92].

These enhanced input devices help depth perception through the user's sense of touch. Other methods, however, involve their sense of sight. The CrystalEyes shutter glasses, for example, gives the wearer of the glasses an enhanced view of the screen. Each lens is continually turning on and off to either block all the light to the eye or to let it all pass through. The two lenses are switching asynchronous to each other and alternate at such a high rate that switching is unnoticeable to the user. While they are alternating, the video screen is synchronized to each lens so that it displays the correct image for the eye that is allowed to see. The user then perceives a stereo image that helps their depth perception.

Besides the stereo effect, the shutter glasses also provide a head tracking device that allows the application to monitor the user's head position and orientation. The image displayed can then be adjusted to conform to the user's current head position so that by moving their head, the user can "look around" the model and get a greater sense of depth.

Some methods improve depth perception through the sense of touch, others through the sense of sight, but what about the sense of hearing? Is it possible to improve depth perception using audio feedback? This thesis explores this area by testing three different sound environments each using a different type of sound feedback. Chapter 2 reviews background material covering visual perception, audio perception, the interaction between the two, and discusses examples that use audio feedback in computer interfaces.

Chapter 3 describes the experimental design, including design decisions and pretest results. Two experiments were undertaken: one studying the use of sound as an aid to depth perception, the second investigating the effect that learning has on the use of sound feedback.

The results from both experiments are organized in tables and graphs and are shown in Chapter 4 with a few brief comments. The main discussion, however, is found in Chapter 5 where the use of each sound environment is considered, conclusions drawn and the possibility of future research noted.

Overall, sound feedback proved to be an effective medium to help depth perception. Task accuracy was improved while using sound but the time to complete the task suffered. The extended study, which looked at the learning effect on performance over a period of time, showed that task accuracy remained consistent over the period while the time required to complete the task improved. In conclusion, depth perception was aided by sound and even the degradation of task time while using sound was reduced as the users became more familiar with the sound environments.

Chapter 2 Background

2.1 Visual Depth Perception

2.1.1 Depth Cues

In a natural 3D environment, people rely on many aspects of the environment, or *depth cues*, that help them to perceive depth. Adding a sense of depth to a computer interface then involves simulating these depth cues on the display terminal. Since users rely extensively on visual cues to perceive depth, an interface that provides few cues is not only difficult to work with but can hinder user performance. Understanding the depth cues that people naturally rely upon is therefore essential when designing a 3D computer application and is therefore important for research in this area.

Coren and Ward identified twelve different depth cues that people rely on in their natural environment. [Coren & Ward 89]. These visual cues can be categorized into four groups: colour, size, position, and physiological. Table 2.1 lists these categories with their depth cues and identifies which are used frequently in commercial 3D computer applications.

Category	Depth Cue	Use in 3D Applications
Colour	Aerial Perspective	
	Object Shading	✓✓
	Texture Gradient	✓
Size	Retinal	✓✓
	Familiar	✓
Position	Interposition	✓✓
	Linear Perspective	✓✓
	Height in the Plane	✓
	Stereopsis	✓
	Motion Parallax	✓✓
Physiological	Accommodation	
	Eye Convergence	

Table 2.1 Depth Cues

Many of the cues that help people to perceive depth in their natural environment are simulated in 3D applications to make the image to appear to have depth. The check marks above indicate the use of each cue in commercial 3D applications, where the most significant cues in these applications are shown with more check marks.

2.1.1.1 Colour Cues

Colouring depth cues help people to perceive depth based on differences in object colours. There are three main types of colour depth cues: aerial perspective, shading, and texture.

Aerial Perspective

When a person looks at an object in the distance, like a mountain, the light has to travel a long way to reach the person's eye. Along the journey to the eye, the light collides with many dust and water particles that will scatter different light frequencies in different ways. Generally, distant objects will look bluer as well blurrier. This type of depth cue is called *aerial perspective*.

Object Shading

All objects have a natural colour which is often consistent across the entire surface. This consistency, however, changes when a light is shone upon the surface causing the shade of the colour to vary as the shape of the surface and distance to the light varies. By observing this variation in colour shading, viewers can perceive a change in depth by comparing the colours of two points on an object.

Texture Gradient

A *texture* is a pattern, like a wood grain, that appears on the surface of an object. How this texture varies with depth, or the *texture gradient,* helps viewers to perceive depth since points closer to the viewer have a texture that is spaced farther apart than distant points. Distant points appear closer together and thus the texture appears to have a finer grain.

2.1.1.2 Size Cues

Along with colouring cues, viewers rely heavily on size cues. The apparent size of the object indicates its distance from the viewer. Size cues include: retinal size and familiar size.

Retinal Size

Objects that are closer to the viewer appear larger than distant objects. This is because of a difference in the image size that the object produces on the retina. Figure 2.1 shows how near objects produce a large image, and distant objects a small image. Viewers use this cue to determine which object is closer than the other.

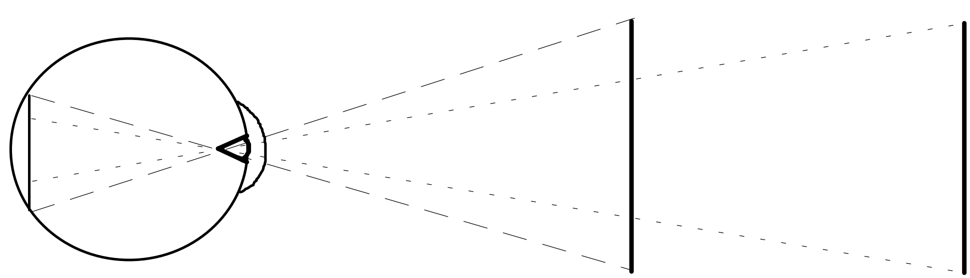


Figure 2.1 Retinal Size
The two bars of equal length appear different to the eye: near objects produce a large image on the retina, whereas distant objects produce a smaller image.

Familiar Size

The viewer's knowledge of relative object size, known as *familiar size*, is another helpful depth cue. In this case, the viewer bases the distance between two objects on the difference in retinal size and previous knowledge of the actual size of the objects. For example, Figure 2.2 shows two cases: one where the book is larger than the car, and one where it is smaller than the car. Since a book is normally smaller than a car, the book appears to be closer in the top image, but at about the same distance in the lower image.

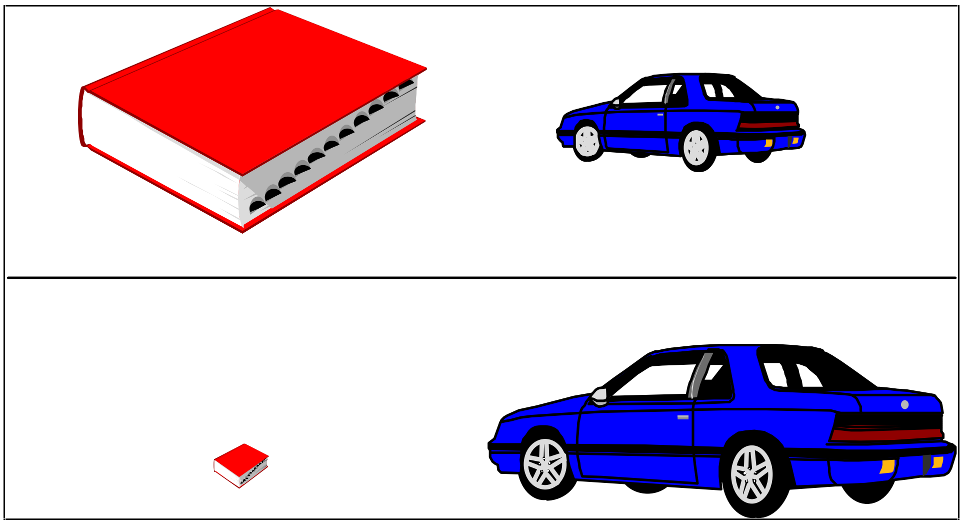


Figure 2.2 Familiar Size
Being familiar with object size helps to determine which object is closer. In the top image the book appears to be closer since it is usually smaller than a car. The lower image shows a smaller book and it appears to be at about the same depth as the car.

2.1.1.3 Position Cues

Colour depth cues help viewers' depth perception based on the object's shade, size cues help based on object size, whereas position cues give depth information based on where the object is located. There are five types of positioning depth cues: interposition, linear perspective, height in the plane, stereopsis, and motion parallax.

Interposition

When one object is partially occluded by another, it is assumed that the second is the closer object. For example, in Figure 2.3 there are two diskettes and a computer monitor. The diskettes appear to be in front of the monitor since they occlude it. Even considering the diskettes, one appears to be closer than the other since the second is not fully visible. This type of depth cue is called *interposition*: closer objects block the line of sight of farther ones.

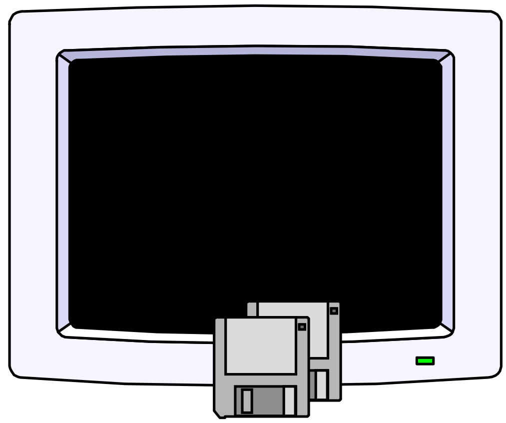


Figure 2.3 Interposition
An object that covers another will appear to be in front of the covered object. This is known as interposition. In the figure, the diskettes appear to be closer since they cover part of the monitor.

Linear Perspective

Linear perspective is the type of depth cue observed when looking down a set of parallel railroad tracks. The tracks are physically parallel, but they appear to “join” at the horizon. In Figure 2.4 the seven lines appear to be converging indicating a depth change. Linear perspective is most apparent with “hard” edged objects like a cube, as opposed to objects with less edge definition like a sphere.

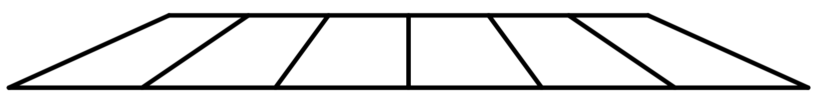


Figure 2.4 Linear Perspective
With linear perspective, parallel lines appear to converge in the distance.

Height in the Plane

The vertical position of an object is a depth cue called *height in the plane*; the farther an object is from the viewer, the closer that object is to the horizon. For instance, Figure 2.5 shows two boats on the water; one closer to the horizon than the other. The boat closest to the horizon appears to be farther from the viewer than the other boat. This not only applies with objects below the horizon, but also to those above, as seen with the two birds.

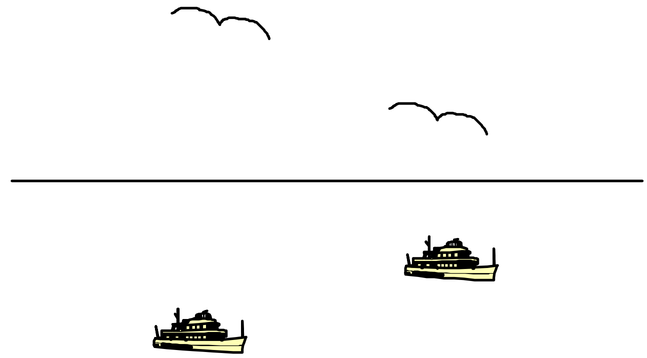


Figure 2.5 Height in the Plane
Object vertical position with respect to the horizon helps to determine the object's depth. In the figure above, the boat and bird nearest the horizon line both appear to be more distant than the other boat and bird.

Stereopsis

Since people's eyes are separated by a horizontal distance, each eye receives a different image of the scene. This stereo view, or *Stereopsis,* consisting of two images is then blended into one. The viewer perceives these two retinal images as one three dimensional image and can determine depth by comparing them and noting the differences. In Figure 2.6, a left and right image are shown of a van as if the viewer was looking through a window. By interposition, the viewer knows that the van is behind the window since part of the van is occluded. Further depth information can be found in comparing the images. The left image shows more of the driver's mirror than the right image. The more this differs, the closer the object is to the viewing window. Very little difference would indicate that the van is far from the window.

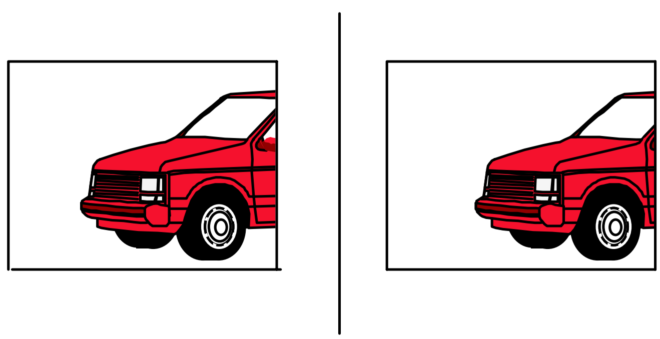


Figure 2.6 Stereopsis
Since each eye receives a slightly different image, a comparison can be done to help indicate depth. The figure above shows a sample left and right eye image of a van through a window. The greater the difference, the closer the object is to the viewer. In this case, the left image is showing more of the driver's mirror than the right image.

Motion Parallax

All the preceding depth cues have been dealing with stationary objects. *Motion parallax* is the cue that deals with object movement. Looking out a window while stepping to the side will cause the objects in the view to move. Close objects will appear to move in the opposite direction while distant objects move in the same direction. In the middle there will be a fixation point that will not appear to move at all. The viewer can determine distance by noting the speed at which the objects move. Objects further from the fixation point move faster than objects close to it. For instance, a distant mountain will appear to be keeping up to the viewer who is watching it from a moving a car, while a nearby tree will seem to speed right past.

2.1.1.4 Physiological Cues

The last three categories of depth cues have been based on the appearance of the object to the viewer. The final category, however, uses physiological changes of the viewer to determine depth. These changes include: accommodation and convergence of the eyes.

Accommodation

Accommodation is the change in focus of the eye's lens when focusing on an object. For example, by staring out a window, the eyes might be focused on something outside. Then they can be refocused to view an insect on the window pane. This physical change in the eyes' focus gives the user a sense of depth.

Eye Convergence

Another physiological effect is that of the angle between the viewer's two eyes. When the viewer is looking at a close object the eyes converge and the angle between them is small. Looking at distant objects cause the eyes to diverge and the angle between them increases. Both accommodation and the convergence angle between the eyes offer a small effect on depth perception with objects at great distance.

Of all of these natural environment depth cues that viewers rely on, many are not relevant to all 3D computer applications. Most applications, for example, wouldn't simulate stereopsis since it may require more hardware and software than it is worth. Physiological cues are also hard to simulate and wouldn't be provided by a computer application. This is due to a fixed screen distance that requires no change in eye focus and convergence while the user works with the application.

Excluding depth cues from an interface not only reduces the amount of cues to help the user's depth perception, but it may also be a hindrance. For example, based on the retinal size depth cue, the user would expect that a cursor moving in depth should get smaller with distance. An application that does not have this ability to change cursor size with depth may therefore confuse the user. In this case the inability to provide an appropriate depth cue has caused a "negative" depth cue to be perceived since a contrary response was received from what was expected.

2.1.2 Fitt's Law Extended to 3D

Object selection, or *picking*, with a mouse is one of the most common tasks performed in computer applications. How picking and time are related is therefore important to an application designer who wants to reduce the time a user requires to do a task.

Fitt's law states that one dimensional movement time to a target object is a function of the distance to the target and its size. [Fitts 54]. Specifically,

$$MT = a + b\log_2\left(\tfrac{2A}{W}\right)$$

where *MT* is the movement time, *a* and *b* are constants, *A* is the amplitude (distance) and *W* is the width of the target. The relationship is therefore as follows: as target size decreases or distance increases, movement time to the target increases.

Mackenzie and Buxton extended this law for two dimensional tasks and found a similar relationship by assuming that the 2D target was a rectangular object. [Mackenzie & Buxton 92]. They found that setting *W* to be the shortest side of the target rectangle and *A* to be the distance to its centre, the movement time function was defined as:

$$MT = 230 + 166\log_2\left(\tfrac{A}{W} + 1\right)$$

As with the one dimensional case, movement time in the 2D case was also a function of target distance and size. Zhai, Buxton and Milgram applied a similar approach to their 3D cursor but defined *A* to be the distance between two targets and *W* was the width of the 3D cursor. [Zhai *et al. 94*]. The movement time was then the time required to move the cursor between the two targets. Again as distance increased, movement time also increased. Picking in the experiments of this thesis will be in 3D so it is expected that a similar result is seen.

2.2 Audio Perception

2.2.1 The Ear and Hearing

Truly an amazing organ, the ear "is so sensitive that it can almost hear the random rain of air molecules bouncing against the eardrum. Yet in spite of its extraordinary sensitivity the ear can withstand the pounding of sound waves strong enough to set the body vibrating." [von Bekesy 57].

The ear itself, shown in Figure 2.7, is made up of three parts: the outer ear, the middle ear and the inner ear. The outer ear has the responsibility of capturing the sound wave on the pinna and channeling it to the eardrum. Vibrations to the eardrum move the small bones called the ossicles of the middle ear that are connected to it. These bones then transmit the message through another membrane, called the oval window, to the fluid in the inner ear, which amplifies the signal approximately 22 times.

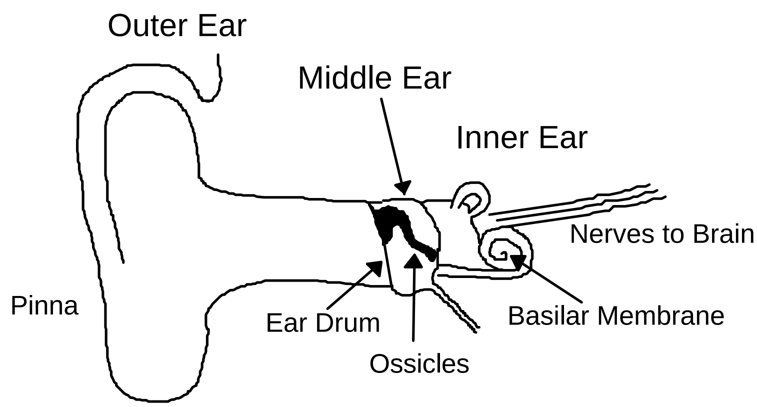


Figure 2.7 Anatomy of the Ear
The pinna focuses the sound wave through the ear canal to the ear drum. Sound waves then move the ear drum, which vibrate the connected ossicles bones. These bones transmit the sound energy to the fluid in the inner ear where the basilar membrane resides. As the membrane moves in this fluid it picks out the individual wave frequencies and transmits them to the brain.

In this fluid is the basilar membrane that acts like a bank of filters. Various locations along the membrane filter out specific wave frequencies. For example, the base of the membrane connected near the oval window captures the low frequencies whereas the opposite free end retrieves the high frequencies.

The ear's threshold of hearing sensitivity peaks at about 4kHz, and has much less sensitivity at both the high and low ends. For example, sensitivity of a 100Hz tone is 1000 times lower than a tone at 1000Hz. The upper limits, which can be as high as 20kHz, decrease as the age of the hearer increases. In particular, von Bekesy noted that the upper limit drops by 80Hz every 6 months after a person reaches 40 years of age. [von Bekesy 57], [Rosen & Howell 91]. These thresholds were measured on subjects performing the thesis experiments to determine how well they could hear, as described in section 3.3.1.4.

2.2.2   Perception

Before designing a *sound interface,* which uses sound as a medium to relay information to the user, it is important to understand how people differentiate sound messages, or *sound cues*. A particular sound that may seem reasonable to the designer could be completely inappropriate for the situation in which it is used simply because the user misinterprets its purpose.

Much study has been made in the area of sound perception. Bregman and Steiger found that when subjects were presented with a sequence of two tones in an ascending or descending order of pitch, they perceived a spatial change in a corresponding manner. [Bregman & Steiger 80]. They went on to show that two concurrent sounds can be segregated by subjects into individual *sound streams* by introducing one sound before both together and that this too can effect spatial localization. It was noted that a difference of 10 ms in sound onset times can produce the desired stream segregation.

In a natural setting, sounds are generated from several sources but are all combined and channeled through the ears for analysis. This process, termed *auditory scene analysis*, segregates the incoming sound into separate source streams. [Bregman 90]. Multiple stream segregation can be encouraged by varying different *sound dimensions*, like volume or pitch. An interface can then vary the sound dimensions of its auditory cues to pass information onto the hearer. How these dimensions are perceived is thus of great interest to the sound cue designer. It has been found that humans can differentiate up to 20 sound dimensions. [Yeung 80]. Some of these dimensions include: pitch, volume, balance, duration, timing, timbre, attack, decay, oscillation, and spatial location.

By using the sound dimensions, segregation of two sound cues increases as the difference in at least one sound dimension increases. For example, consider two sound cues that differ only in their pitch dimension. As one cue increases in pitch a difference will be perceived and the hearer will separate the two incoming streams. Differentiation improves further as the number of dimensions between two sound messages differs. [Buxton, Bly, Frysinger, Lunney, Mansur, Mezrich & Morrison 85].

Which sound dimension is most important depends on how it is to be used. Balance, for example, is naturally perceived to be a side to side action, whereas pitch, as noted above, has an up-down correlation. On their own, without considering how they might be used, Wickens noted that there is little evidence as to which sound dimension is the most important dimension. [Wickens 84].

Other methods of stream segregation, that are not involved in the experiments described in this thesis are: temporal orders, rhythmic patterns, and spatial separation of sound sources. See [Curso 80], [Gerth 92], [Clark 89], [Schulze 89], and [Good & Gilkey 92] for more information.

In addition to considering sound segregation, the sound designer must also be aware of *sound bonding*. People associate different sounds with different objects and actions. Any deviation from this association, like a sound delay after an action is performed, causes confusion. Gates and Bradshaw studied the effect that sound delay has on user performance when the sound feedback from the user’s action was delayed. [Gates & Bradshaw 74]. The results showed a significant decline in performance when the sound was delayed by 0.18 seconds. They compared several combinations of delayed feedback and noted that immediate feedback had the best performance, followed by no feedback, then immediate feedback mixed with delayed feedback, extraneous feedback, and finally delayed feedback.

Often users are required to remember a cue, and in a sound interface this may be a sound cue. Factors that affect user’s recollection of sound and how long of a delay between hearing and recalling a sound are therefore an important issue in design. Botte, Baruch and Mönikheim experimented to see how volume and time delay between hearing and recalling effected memory and found that performance increased as volume increased and it decreased as delay increased. [Botte, Baruch & Mönikheim 92].

2.2.3 Individual Differences

2.2.3.1 Gender

One might expect to find there to be individual differences in audio perception. On a whole these differences, however, have not been significant. Prior and Troup noted that there was no significant difference in rhythm change detection due to gender. [Prior & Troup 88]. The delay experiment mentioned earlier by Gates also found no differences due to gender. Kroeger, however, noted large individual differences in audio spatial location. [Kroeger 94].

Pishkin and Blanchard studied this area in detail and tested the effects of gender on several sound dimensions: duration, laterality (balance), frequency, amplitude (volume), and number of sound repetitions. [Pishkin & Blanchard 64]. Their initial test modified only one test sound dimension at a time and found no significant difference due to gender for any of the dimensions. However, once extra dimensions were also modified in addition to the test dimension, males performed significantly worse in their perception of balance than females.

Noting that this difference was found using adult subjects, Pishkin and Rosenbluh sought to determine if performance was affected by age. They performed a second experiment, this time using adolescent subjects. [Pishkin & Rosenbluh 66]. They found that male adolescent subjects made fewer errors than the female subjects, however, there was no significant difference in determining balance between the two genders. Upon comparing the results of the original experiment, they concluded that the ability to determine balance deteriorates in males with age.

2.2.3.2 Musical Ability

Since some people are more musically inclined than others, one would expect a difference in performance in auditory tasks based on musical ability. Prior and Troup, however, studied this area and found that, although musicians performed faster than non-musicians, there was no significant difference in instrument identification between the two groups. [Prior & Troup 88]. They also found that both musicians and non-musicians chose similar strategies when analyzing musical rhythms, though in this case musician performance was better.

Many musicians have what is called *perfect pitch*, in that they can identify a musical note as soon as it is heard. Even though they have this ability, it has been found that they often locate the note in the wrong octave. [Deutsh 86].

2.3 Visual Perception vs. Audio Perception

People continually process different signals that are being received by their senses. How these signals are attended to is therefore an issue when designing a multi-sensory interface. Is there any interaction between visual and audio perception or is their processing completely separate?

Psychologists define two types of concentration of mental activity, or *attention*: divided attention and selective attention. *Divided attention* refers to the responding to several stimuli at once, whereas in *selective attention* the individual tries to focus on only one. [Matlin 94]. During attention of specific stimuli or tasks, unattended tasks can sometimes interrupt and be attended to as well. An example of this occurs when a person hears their name mentioned in another conversation even though they were not attending to that conversation. [Kahneman 34].

For simple or familiar tasks, *preattentive processing* takes place where several tasks are accomplished in parallel. Difficult or unfamiliar tasks, however, take much longer since the attention is focused and processing is done serially. [Treisman & Gelade 80]. Treisman and Gormican suggest that there is not a fixed division between preattentive processing and focused attention, but a continuum since with practice tasks become more familiar. [Treisman & Gormican 88].

How is attention governed by the different types of stimuli received? Colivata found that subjects acknowledged visual stimuli over audio ones when both were presented at the same time. [Colivata 74]. Moreover, several of the subjects were not even aware that the audio signal had occurred. Even when the audio stimuli had twice the subjective intensity as the visual one, the majority of the subjects still responded to the visual cue. She concluded that between the human visual and auditory senses, the visual sense is more dominant.

Auditory processing is not completely separate, however, and in fact it has been shown that it shares a common processing space with visual processing. Auerbach and Sperling studied this phenomenon to test whether subjects use a disjunct space, that is, one for the auditory direction and one for the visual, but found that they actually used one combined space. [Auerbach & Sperling 74].

In comparing the effectiveness of sound and visual cues to task accuracy, Brown, Newsome and Glinert found that subjects responded equally well when only one cue was presented, though audio cues tended to take longer to process. [Brown, Newsome, Glinert 89].

Although visual cues dominate, sound cues are just as helpful as Brown's experiment proved. By becoming more familiar with them, users of a sound interface should be able to process the sound cues more rapidly in parallel with the visual cues. Many examples of such interfaces are discussed in the following section.

2.4 Audio in Computer Interfaces

Research on the use of audio in human-computer interfaces has only blossomed recently; the first International Conference on Auditory Display (ICAD) being held in 1992. Although it is a new area of study much research has already been completed.

Auditory Display deals with the ways that audio information is presented, or "displayed", to convey information to the user. Since it is a new area many terms have not yet been settled. *Audification, sonification,* and a*uralization* all refer to a similar idea where the sound heard is based on some type of data supplied. [Kramer 94].

Kramer lists several pros and cons for using auditory displays, many of which have been known to exist long before computers became widely available. One advantage, for example, is that the user can simply monitor a task audibly without actually viewing it. Mereu and Kovach used this technique to listen to seismic data when setting up seismic instruments in the field. At the time, paper output was the usual method of quickly verifying that an instrument was set up correctly, but the addition of an audio response allowed instruments to be easily positioned and monitored at night or in poor weather. [Mereu & Kovach 70].

Other advantages with auditory displays include: processing volumes of data quickly by turning it into an audio signal, being able to "background" sound until a significant event occurs, and the ability to listen to several events in parallel. Disadvantages include: low resolution in sound dimensions, a tendency to become annoying after an extended time, the difficulty of using it while speaking with others, and that it can easily disturb neighbouring workers.

Of the recent research, two specific types of sound messages have developed: the earcon, and the auditory icon. Both are similar in their purpose, that is to convey auditory information, yet they differ in their composition.

The *auditory icon*, largely due to Gaver's work, uses a familiar natural sound to represent an action or object. The sound will vary in its sound dimensions to indicate something is taking place. Gaver's SonicFinder, for example, was built on top of the standard Macintosh file Finder to enhance its feedback. When a user copies a file, a pouring sound is heard that rises in pitch as it was nearing completion, just like the sound heard when pouring a glass of water. [Gaver 86, 89].

Earcons, on the other hand, are abstract synthetic tones structured in small sequences called *motives* with a different motive representing a different event or object. A sound message is then created using several of these earcons to indicate the different aspects of an event. If, for example, the earcons A and B represent a file and deletion action respectively, then some combination of those two earcons, AB would indicate that the file was deleted. [Blattner, Papp, & Glinert 94], [Brewster, Wright, & Edwards 94a].

When applied to real applications, earcons and auditory icons open the door to many possibilities. Gaver, Smith and O'Shea, for instance, used auditory icons in a simulation of a cola factory in which the user could monitor several processes occurring at once. The bottle capper machine, for example, would make a continuous rhythmic sound under normal working conditions. A change in this pattern would immediately alert the user even if they were monitoring a different area of the plant. [Gaver, Smith, & O'Shea 91].

Besides monitoring background tasks, sound can help the user to process information in foreground tasks that are often difficult to comprehend visually. Di Giano and Baecker, for example, added sound to a programming environment. By listening to the sound of code running, semantic errors like infinite loops became easily detected. [Di Giano & Baecker 92]. Parallel programs are also very difficult to debug. Jackson and Francioni found that adding sound to a monitoring tool improved the programmers understanding of program execution. [Jackson & Francioni 94].

Other uses of sound in human computer interfaces include listening to some form of the data itself. Hayward applied this approach to listen to seismic signals that were converted to audio signals. [Hayward 94]. Once familiar with the different sounds, users can become proficient at identifying an earthquake over other earth noise such as caused by a train or an explosion. Fitch and Kramer on the other hand used several sound streams to play back physiological responses, like blood pressure, heart and respiratory rate of a patient. In this case auditory icons were used to make it seem more natural using, for example, a breathing noise for the respiratory rate. [Fitch & Kramer 94].

Probably the most useful applications of sound to an interface have been for the visually impaired. Until now the applications described here have been enhancements to normal visual applications making them more powerful and easier to use. Several studies have been done in this area for the blind including Mansur, Blattner and Joy's Sound-Graphs, which mapped xy data to the audio domain so that a blind person could understand the relationship. This was done by moving along the x direction and varying a tone's pitch according to the y value. [Mansur, Blattner, & Joy 85]. Edwards on the other hand showed how to enhance a regular windowing interface so that it could be used by visually disabled users. Each window, including the edge of the screen, was assigned a different tone. As the user moved the mouse around the screen, the entered window would play its tone. A currently active window was designated by playing a pair of tones instead of its regular tone. [Edwards 88].

Although many sound interfaces have been studied, little has been done to extend a 3D interface to use sound feedback. Many of the techniques and results, however, found in the works mentioned above can be applied to the 3D setting.

Chapter 3 Experiment Design

3.1 3D Interface

The hypothesis in this thesis is that sound, as a depth cue, can aid user depth perception in a 3D user interface, and thus improve user performance. To test this, an interface was designed in which the user was required to perform a 3D task thus relying on their depth perception. Various sound cues were then provided to see which, if any, aided the user's depth perception and improved task performance the most.

Colivata's results, discussed previously in section 2.3, showed that visual cues dominate over sound cues even when the subjective intensity of the sound cue is twice that of the visual cue. To isolate the effects of the sound cues, the visual depth cues were minimized so that the user was more dependent on the sound. Certainly, in most applications, the visual depth cues will be provided and therefore will dominate. The emphasis here, however, is not to determine whether sound cues are better than visual cues, but whether they can aid in depth perception.

3.1.1 Interface Task Requiring Depth Perception

One of the most common tasks in 3D computer interfaces, discussed in section 2.1.2, is picking. Picking is, however, more difficult in a 3D application than in a real world setting due to the lack of depth information. The user never has the same depth perception confidence performing an action on a computer generated image than in real life. [Jaubert 95]. To help users, some systems, like Alias, provide four simultaneous views of the scene: top, side, front, and perspective. This gives the user greater confidence in the 3D location of the cursor. The multiple views, however, become cumbersome since it requires the user to constantly change their focus as well as mentally fusing the views.

Ideally, only one view is needed which displays a perspective image enhanced with extra depth cues. This was the approach taken in designing the experiment interface where the extra depth cues consist of sound cues. Figure 3.1 shows the single perspective view of an object on the interface screen. Section 3.1.2 below describes how the "blobby" object shown was designed.

From the figure, a target location on the object surface can be seen. The interface task required the user to pick this target located in three-space with the cursor. This task was chosen since picking is common to almost all 3D applications. Target accuracy and the time required to complete the task were recorded to measure user performance. To counteract learning effects, both the rotation of the object and the surface position of the target were randomly determined for each user trial.

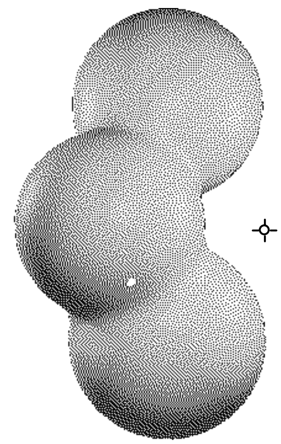


Figure 3.1 "Blobby" Object with Target Location
The "blobby" object's irregular shape and lack of hard edges reduce the amount of visual depth cues provided to the user. Both the object's orientation and the location of the white target on the object's surface are randomly determined to reduce learning effects for each user trial. The crosshair cursor grows smaller as depth "into" the screen increases.

Movement of the cursor in a 3D environment can involve up to six degrees of freedom: x, y, z, pitch, yaw, and roll. Since most picking tasks require only absolute positioning, the rotational degrees of freedom pitch, yaw and roll were not considered. This left the three coordinates - x, y, and z - to be controlled by the user's input.

Input was accomplished using a regular 2D mouse to give the x and y values, and the up and down arrow keys to give the z values that moved the cursor "into" and "out of" the screen respectively. Since the representation of the mouse cursor is located on a 2D screen plane in the 3D environment, its x and y values were not used directly. These values instead represent the perspective values of the actual 3D cursor's x and y values. The actual x and y values were therefore computed based on these original x and y values and the current value of z. Figure 3.2 shows the difference between the mouse cursor's y value and the actual 3D cursor's y value.

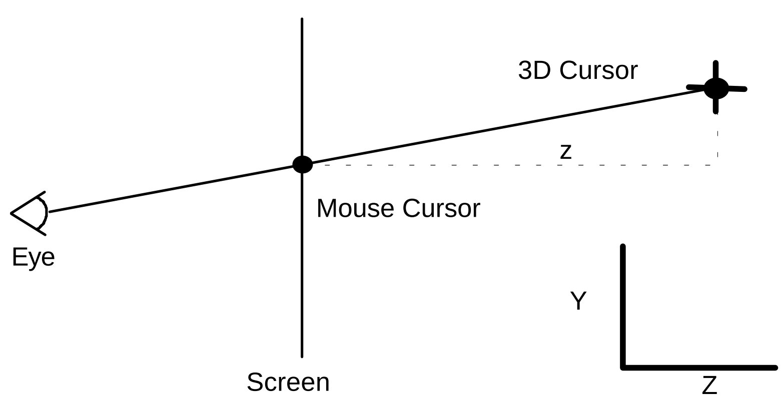


Figure 3.2 3D Cursor Location
The actual 3D cursor's x and y values are computed by projecting the mouse cursor's location back to the current z value.

3.1.2 Minimizing Visual Depth Cues

Section 2.1.1 described all the visual depth cues that people rely on in everyday life. It was also noted that some of these are more important than others in computer interfaces. By minimizing these visual depth cues the user should become more dependent on the sound cues so that the effect of the sound cues on user depth perception can be isolated. The following paragraphs describe how these visual depth cues were minimized.

Colouring depth cues are reduced by not implementing any aerial perspective colouring or texture gradients. Shading of the object is maintained, however, so that it still appears to have some depth.

Size depth cues, which include familiar and retinal size cues, are all removed by not using any objects of familiar shape nor having multiple object from which depth can be compared. Instead, a single irregular shaped “blobby” object was used as shown in Figure 3.1.

In addition to reducing size depth cues, this single object also reduces positioning depth cues. It has no hard edges that might have otherwise provided a linear perspective cue. Spheres could have been used but since they are uniform, the depth across them would have been easier to determine than with a non-uniform blobby shape. Using only one object with no background horizon also removes any interposition depth cues and height in the plane cues that would otherwise have been possible.

Other positioning cues like stereopsis and motion parallax are also removed, by only displaying one fixed image of the object and not allowing either the object or the viewing position to move. One consequence from this restriction is that the target location must always appear on the visible side of the object.

The last of the visual depth cues mentioned in section 2.1.1 are the physiological cues: accommodation and eye convergence. Both of these are not normally supported by computer interfaces since the user is always focusing on a fixed distance. The experiment interface therefore does not support either of these visual depth cues either.

The only other object seen besides the blobby object is the cursor, which is simply a 2D crosshair that grows smaller as depth increases. This size change with depth is introduced as mentioned in section 2.1.1, not as an accurate measure of depth, but as a way to eliminate any “negative” depth cues that would arise if the cursor’s size didn’t change with depth. With no size change, the user would be receiving negative feedback as the cursor was moved in depth since visually it would not appear to be moving. To make this depth cue present, yet unreliable as an accurate measure of depth, the cursor size is mapped to one of five fixed sizes.

3.1.3 Adding Sound Depth Cues

There are many ways that sound can be added to return depth information. Navy ships, for instance, use sonar to send out sound waves that reflect off the sea bottom. The depth is then determined based on the echo received. A similar approach could be done where the user can "listen" to echoes at specific points on the screen to determine their depth. Other possibilities include sound emitting devices such as those used in Gaver's Arkola simulation plant described in section 2.4. Objects farther away would sound quieter than objects near the user's position.

For this interface, however, picking is the main user task and thus the cursor and target locations are the key points of interest for which sound is mapped. Just as there are many ways to use sound, there are also many ways to map a specific sound to a spatial location. These are discussed more fully in section 3.2.

During an experiment task, the sound mapped to the cursor location is modified as the cursor is moved around the 3D environment. This allows the user to continually monitor the location of the cursor audibly. Along with this feedback, the user can, at any time, depress the left mouse button to hear the target location. While the button is depressed, the cursor sound is silent so that the target can be clearly heard. By pressing and releasing the button consecutively a comparison can be made to determine how far away the cursor is from the target. Once the user believes that the cursor is at the target, the space bar is used to signal that the task is complete.

3.2 Sound As a Positional Cue

Using sound to indicate a 3D spatial location can be done in a variety of ways introducing many possible design decisions. Not only is the type of sound, like a simple tone or noise, important but how that sound is mapped to the 3D location is also an issue. For example, is the location's sound mapping a relative or an absolute mapping? Is the sound a general earcon that, as described in section 2.4, changes in its sound dimensions or is it an auditory icon whose sound resembles that of a real life object or action? How is the sound played: tonally, chordally, musically, as an actual 3D spatial sound, or by some other sound mapping? These design issues will be discussed in the following sections.

3.2.1 Mapping Sound to Position: Relative vs. Absolute

Mapping sound to a spatial location can be done by either a relative mapping or an absolute mapping. A relative mapping requires the system to know two locations and the resulting sound is based on some combination to those two positions. For example, the linear distance between two points is a scalar that could be mapped to a sound dimension like volume. For the task described in section 3.1.1, there are two known locations: the cursor and the target, so initially a relative sound mapping might seem to be an ideal choice. Knowing the target location is, however, only a product of the experiment task. In most real 3D applications such as a CAD program, the computer system does not know what the user's desired target is and thus a relative mapping is impossible.

The alternative to a relative mapping is an absolute mapping. Here the sound played is specific to a spatial location, like the cursor location. As the cursor moves its three location dimensions x, y, z, the attached sound also changes in three of its dimensions. Some possible sound dimensions, described in section 2.2.2, include volume, balance and pitch. For the experiment task, there are two locations that require an absolute sound mapping: the cursor and the target. The issue then becomes, how many sounds are playing at a time: one or two? Do they play only when requested or are they constantly playing?

Since the user's task is to locate the target with the cursor, it is of interest to know whether they relied on the sound cues. For this reason, the cursor's location sound cue is always playing. The target location, however, does not move so it is not as critical that it is always heard. The user can request to hear the target location, as described in section 3.1.3, but when this occurs the cursor sound is turned off so that a clear comparison can be made. Another reason for playing only one sound at a time is that it would be difficult to distinguish which sound cue was attached to each location. Not only would there be this identification problem, but the sounds themselves might conflict and cause confusion. For example, one of the possible sound mappings, discussed later in section 3.2.2, maps a location dimension, like y, to music tempo. Playing sounds from two different locations would certainly be confusing since the user would simultaneously hear music playing at two different tempos.

As already noted, a real application would not normally know the user's desired target location. Since no target sound would then be played, no comparison can be made as described here in the experiment task. If, however, the sound cues do help the user to perceive depth, then with experience, the user should be able to associate different sounds with their spatial positions and thus be able to locate them with the cursor.

3.2.2 Possible Sound Mappings

Probably the greatest drawback with sound feedback in most applications is that over time the sound becomes annoying. This means that no matter how useful it is to the user's productivity, if its annoying, it won't be used. Annoyance is then the most important characteristic to consider when picking a sound mapping.

Jones and Furner showed that even though auditory icons use more natural every day sound events and that it is easier to identify the associated object or action, users preferred earcons. [Jones & Furner 89] Certainly there are an infinite number of possible earcons mappings to consider, but something that is pleasant to listen to and offers great variety and flexibility is music. An application that can pass information to the user by making small changes to how the music played has great potential. Not only does it provide an extra mechanism for message passing, but it is enjoyable to listen to and can also be configured it to play different songs throughout the day according to the individual tastes of the user.

To determine the music's effectiveness as a 3D positional cue, four sound environments, added to the experiment interface of section 3.1, were selected and tested: a no sound environment, a simple tonal sound, a musical piece and a more complex orchestra arrangement. The following sections describe each of these sound environments in detail.

3.2.2.1 No Sound

The no sound environment generates no sound cues to help the user. It is used as a control test block from which the performance of the other sound environments can be compared.

3.2.2.2 Tonal

Some systems use a simple pure tone that varies its sound dimensions to pass on a message to the user. Brewster, Wright and Edwards, for instance, added this type of sound feedback to an auditory enhanced scrollbar. [Brewster, Wright, and Edwards 94b]. A similar type of sound is used here but it maps the location's x, y, and z values to three separate sound dimensions of the tone. Possible sound dimensions that were considered included: volume, balance, tone oscillation, pitch bend and noise. These are described more fully in section 3.2.4.

3.2.2.3 Musical

Since simple tones tend to become annoying after extensive use, replacing them with music is a reasonable alternative to consider. As with the tonal environment, the location's x, y, and z values are mapped to three of the sound's dimensions. Along with the possible dimensions mentioned for the tonal sound, music also offers two other possible dimensions: tempo and key. These too are discussed more fully in section 3.2.4.

3.2.2.4 Orchestral

Can the music mapping be extended in a way that takes advantage of the distinctive natures of the instruments that are playing? The orchestral environment tries to accomplish this by setting up eight instrument sections in a four by two grid. This grid is then set on the x-z plane with the greatest grid granularity given to the z dimension as shown in Figure 3.3. Each section is assigned a different instrument in a manner similar to a real orchestra layout: strings in front, winds in the middle and percussion instruments at the back.

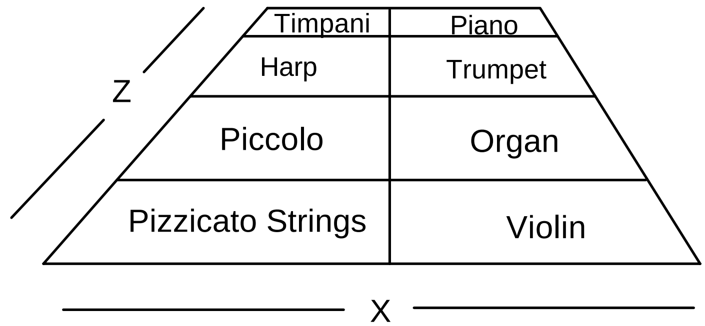


Figure 3.3 Orchestra Arrangement
The orchestral environment uses an orchestra consisting of eight instrument sections laid out on its x-z horizontal plane.

Since the orchestra is arranged on the horizontal plane, the x and z components of a location will locate a point somewhere in the orchestra. To indicate to the user where this location is, the instrument section or sections at that location will play the melody the loudest. Other instruments from the orchestra may also be playing, but they will all be playing a background part more quietly than the foreground melody. As the user moves through the orchestra, each newly entered instrument section will begin to play the melody, while previously entered sections will continue to play the background part. These instrument boundaries aren’t rigid, but gradually increase and decrease the volumes of the new and old instrument sections respectively.

Instrument identification is essential for this orchestral sound environment to be useful. If an average user cannot distinguish one instrument from the next, they will not be able to determine what location corresponds to the combination of playing instruments. It is therefore necessary to find the top eight most identifiable instruments and place those instruments in the orchestra. As mentioned in section 2.2.3.2, Prior and Troup found that instrument timbre was equally identified between musicians and non-musicians, so user musical ability should not be an issue. [Prior & Troup 88].

3.2.3 Auditioning Orchestral Instruments

Selecting the top eight recognizable instruments to use in the orchestra was done through an auditioning process where the results of this process ranked each instrument contender according to its ease of identification. The standard General MIDI patch set[1] consists of 128 musical instruments. From this group, 32 instruments were selected that could play at least two octaves of notes and did not make any obvious conflicts with other selected instruments. For example, where several instruments of a particular style were available, like the six possible organs, only one was selected.

---

[1] The General MIDI (Musical Instrument Digital Interface) patch set is a standard that defines which instrument patches are assigned to each patch number. See Appendix G of the UltraSound User's Guide for a full listing of this set.

Instrument auditioning consisted of playing a musical piece using a Gravis UltraSound MAX card installed in a Pentium computer. Since some instruments may be more identifiable on certain pieces than on others, eight musical pieces were selected from which a random piece could be drawn. To make each piece as consistent as possible, each piece had approximately the same musical note range, from G4 to A5, and lasted for exactly eight bars. Appendix A lists all the songs used with their note ranges.

The range that a piece is played in varies with each instrument. Consider, for example, a tuba verses a piccolo. A tuba generally plays music with very low notes, whereas a piccolo very high notes. The selected piece was therefore translated up or down from its original range into one that was more familiar for the instrument that was playing it. Although some instruments have a large note range, like a piano, only one range was selected in which it would most likely play the experiment pieces. The 32 instruments tested and their ranges are listed in Appendix B.

Since the average person may not have a fresh recollection of what each instrument sounds like, a learning period was given. Subjects were presented with three lists of instruments: winds, strings and percussion, from which they could pick an instrument to play any of the possible musical pieces. The instruments were organized into the three groups to help the subject to learn the instruments and make comparisons with similar instruments in the same grouping.

Once the subject was fairly confident with instrument identification, a test was conducted in which 30 instruments and pieces were randomly selected and played. Subjects were instructed to stop the music as soon as they recognized the instrument being played and were to select the instrument's name from an alphabetical list. The single list was used rather than the three categorized lists to counteract any learning biases that would be associated by learning the instrument's location on the previous lists. Since instruments were played randomly, multiple or even no occurrences of a given instrument were possible. This removed any possibility of identification by elimination.

The instrument identification experiment was run with six volunteer subjects, consisting of 4 males and 2 females, having a mix of musical and nonmusical backgrounds. When analyzing the results, the goal was to find the greatest number of identifiable instruments; that is, those that had the fewest number of conflicts with other instruments. It soon became apparent that certain instruments were always being confused with other specific instruments. By grouping these together into one instrument *equivalence class*, exactly one instrument could be drawn from this class of instruments for the orchestra.

Of the initial 32 possible instruments, six were removed for having a recognition rate of less than 30%: they were identified correctly fewer than 3 times in 10. These were: the honky tonk piano, oboe, tuba, steel drums, bassoon, and music box. Next, three more instruments: clarinet, dulcimer and harpsichord were removed for having the most number of conflicts (four to five) with other instruments. This left

23 instruments from which to group instrument equivalence classes. To make the greatest number of equivalence classes, marimba, English horn and synth voice were removed as their conflicts spanned several of the instrument classes.

Once those 12 instruments were removed, 20 instruments were left to form 13 instrument equivalence classes that did not conflict with any other class. Table 3.1 lists these 13 classes. The first totals column lists the number of times that instruments in that equivalence class were tested. Some of these tests caused conflicts with the instruments previously mentioned. Removing those conflicting instruments, left the number of times the instruments in the equivalence classes were tested conflict free as shown in the second column. This column shows that these 13 instrument equivalence classes are very identifiable since there were no conflicts made in the 94 tests.

	Instrument Equivalence Classes	**Total including conflicts**	**Tested conflicts removed**	**Eight Final Instruments Selected**
1	Accordian	4	4	
2	Banjo, Koto, Pizzicato Str.	19	17	Pizzicato Strings
3	Bass, Cello, Violin	19	14	Violin
4	Bells	4	4	
5	Calliope, Shakuachi	6	5	
6	Guitar (Distorted)	6	6	
7	Harp	5	5	Harp
8	Organ, String Ensemble	8	7	Organ
9	Piano	7	6	Piano
10	Piccolo	10	9	Piccolo
11	Sitar	5	3	
12	Timpani	7	7	Timpani
13	Trombone, Trumpet	10	7	Trumpet
		110	94	

Table 3.1 Instrument Equivalence Classes
Instruments that were often confused with one another were grouped into an equivalence class from which only one instrument would be selected for the orchestra.

The final eight instruments for the orchestra were selected from these thirteen equivalence classes to make up the most "normal" orchestra possible as shown in Figure 3.3. Originally, the string ensemble was chosen over organ since it appears in a regular orchestra, but it was later replaced with the organ to ensure that it would not be mistaken by the other string sections. It was found that there was no significant difference in user performance between these two instrument choices.

3.2.4 Location Dimension to Sound Mapping Pretest

Each of the three audible sound environments has several possible sound mappings, as noted in section 3.2.2, that can be ascribed to a location's x, y, or z dimension. To be able to properly compare the sound environments, it is therefore desirable to have the best sound mappings per location dimension for each environment.

Although listeners can differentiate between many sound dimensions, Wickins, discussed in section 2.2.2, noted that there was little evidence regarding the degree to which one dimension was better than another. Initially it may seem that any sound dimension can be used, but since a 3D location requires three such mappings, one for x, y, and z, pretests were performed to determine which of the sound dimensions would map the best to each of these location dimensions. Table 3.2 lists the sound dimensions that were tested with each of the sound environments.

Sound Dimension	Tonal	Musical	Orchestral	Mapping
Volume	✓	✓	X	Sound level rises.
Balance	✓	✓	X	Sound pans from the left ear to the right.
Oscillation (Virbrato)	✓	✓	✓	Sound begins to take on a wavering aspect to its pitch.
Pitch Bend (Tuning)	✓	✓	✓	Notes played raise in pitch up to two semitones.
Noise	✓	✓	X	Additional noise to the sound increases.
Tempo	X	✓	✓	Rate at which the music plays increases.
Key Shift	X	✓	✓	Notes played raise in key well beyond two semitones.

Table 3.2 Sound Dimensions Tested
Ensuring the best sound mapping for each sound environment required testing several sound dimensions to determine which maps to a location's x, y, and z dimension the best under that sound environment. The check marks in the table above indicate which sound dimensions were tested for each sound environment. The last column describes how each sound dimension changes as its mapped location dimension increases from a low to high value.

3.2.4.1 Location Dimensions

Each sound environment has three location dimensions which can have sound mapped to them: x, y, and z. Since the orchestral environment uses an orchestra metaphor on its x-z plane, both x and z already have sound mapped to them. Seven pretests were therefore required to find the best sound mappings for tonal x, y, z, musical x, y, z, and orchestral y.

The best sound mapping for each location dimension was found by testing that dimension with each possible sound mapping. The 3D interface described in section 3.1 was used to test the sound environment and location dimension with each of these sound dimensions. Sound dimensions were judged based on the dependent vairaibles of their subjective appeal, the target accuracy attained while used, and the time required to attain that accuracy. Since an annoying sound mapping that gives the best accuracy and time will not likely be used, subjective ratings were treated with greater importance than both target accuracy and time.

3.2.4.2 Pretest Experiment Details

The seven pretest experiments were performed using the interface and its user task of section 3.1, a Gravis UltraSound MAX card and a Pentium computer. Subjects entered into a small experiment room containing this equipment and were given similar information sheets to the ones in Appendix D. Once they were ready, they put the headphones on and began the pretest experiments. The ages of the subjects for the seven pretests ranged from 19 to 31. All the data means for these seven pretests are shown in Appendix C.

Tonal

Eleven paid subjects, consisting of 9 males and 2 females, performed the tonal pretests. Subject musical ability and 3D interface experience were measured on a scale of 0 to 3 to indicate subjects with no experience to extensive experience respectively. The average subject musical and 3D interface experience was found to lie in the middle at 1.3 for both measures. Each subject tested the tonal mappings for x, y and z. The order of these tests was randomized between subjects to counterbalance any order biasing.

Musical

The musical pretests were performed by a total of 9 male paid subjects; eight tested x, nine tested y and seven tested z. The reason for the difference was due to one person unable to finish the z test, and an extra person was tested on y to try to improve the level of significance. Again the musical ability and 3D interface experience were medial with averages at 1.2 and 1.4 respectively. As with the tonal pretests, each subject tested the musical mappings for x, y and z with the order of these adjusted to prevent order biasing.

Orchestral

Eight male paid subjects performed the orchestral pretest, having an average musical ability rating of 1.3 and an average 3D experience rating of 1.3.

3.2.4.3 Tonal Pretest Results

Tonal X

An initial four subjects were used to test the five types of sound mappings of the tonal environment: volume, balance, oscillation, pitch bend and noise[2]. A one-way ANOVA test with repeated measures showed a significant subjective difference of B, P and V over N; so N was removed, as seen in Table 3.3. More subjects were tested on the remaining mappings, bringing the total number of subjects up to nine. Performing another one-way ANOVA test on subjective preference showed a significant difference between B, P, and V over O; thus O was

[2] These five tonal sound mappings will be abbreviated as follows volume (V), balance (B), oscillation (O), pitch (P) and noise (N) respectively.

removed. Testing eleven subjects on B, P and V, showed no significant differences between them regarding subjective preference, target error or time to reach the target, ($F(2,20)=0.10$, $p=0.90$; $F(2,20)=0.90$, $p=0.42$; $F(2,20)=0.94$, $p=0.41$ respectively). Possible sound mappings for tonal x were therefore B, P, and V.

Map1	**Map2**	**Measure**	**df**	**F** **Pr > F**
Balance	Noise	Subjective Preference	(1, 3)	**25.00** **0.0154**
Pitch Bend	Noise	Subjective Preference	(1, 3)	**12.00** **0.0405**
Volume	Noise	Subjective Preference	(1, 3)	**24.00** **0.0163**
Balance	Oscillation	Subjective Preference	(1, 8)	**25.00** **0.0011**
Pitch Bend	Oscillation	Subjective Preference	(1, 8)	**78.40** **0.0001**
Volume	Oscillation	Subjective Preference	(1, 8)	**64.00** **0.0001**

Table 3.3 Significant Differences in Sound Mappings for Tonal X
One-way ANOVA tests with repeated measures[3] were performed on subjective preference, target accuracy and time to reach the target for the five sound mappings of tonal x. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results both the noise and oscillation sound mappings were eliminated.

Tonal Y

The sound mappings for tonal y were tested in a similar way, using an initial four subjects. A one-way ANOVA test with repeated measures showed a significant subjective difference of V over N; so as with tonal x, N was removed as seen in Table 3.4. More subjects were tested on the remaining mappings, bringing the total number of subjects up to nine. Another one-way ANOVA test was performed on subjective preference showing a significant difference between B, P, and V over O; thus O was removed. Testing 11 subjects on B, P and V, showed V to be significantly better subjectively than P, but took a significantly longer time. Target error showed no significant difference for all three mappings, ($F(2,20)=1.57$, $p=0.23$). Although P was not preferred, it was not immediately eliminated since it might be the only choice once x and z mappings have been selected. Possible choices for tonal y were the same as tonal x: B, P and V.

---

3 The F values and probability, p, that it occured by chance are shown in the table, where the probability is shown in a smaller typeface. Any significant result, that is one having an F value with $p<0.05$, are shown in bold face print and found in a shaded box.

Map1	Map2	Measure	df	F Pr > F
Volume	Noise	Subjective Preference	(1, 3)	**22.09** **0.0182**
Balance	Oscillation	Subjective Preference	(1, 8)	**12.90** **0.0071**
Pitch Bend	Oscillation	Subjective Preference	(1, 8)	**12.25** **0.0081**
Volume	Oscillation	Subjective Preference	(1, 8)	**146.29** **0.0001**
Volume	Pitch Bend	Subjective Preference	(1, 10)	**13.06** **0.0047**
Pitch Bend	Volume	Time	(1, 10)	**6.72** **0.0269**

Table 3.4 Significant Differences in Sound Mappings for Tonal Y

One-way ANOVA tests with repeated measures were performed on subjective preference, target accuracy, and time to reach the target for the five sound mappings of tonal y. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results both the noise and oscillation mappings were eliminated.

Map1	Map2	Measure	df	F Pr > F
Balance	Noise	Subjective Preference	(1, 3)	**24.00** **0.0163**
Pitch Bend	Noise	Subjective Preference	(1, 3)	**12.00** **0.0405**
Volume	Noise	Subjective Preference	(1, 3)	**22.09** **0.0182**
Balance	Oscillation	Subjective Preference	(1, 8)	**25.00** **0.0011**
Pitch Bend	Oscillation	Subjective Preference	(1, 8)	**78.40** **0.0001**
Volume	Oscillation	Subjective Preference	(1, 8)	**64.00** **0.0001**

Table 3.5 Significant Differences in Sound Mappings for Tonal Z

One-way ANOVA tests with repeated measures were performed on subjective preference, target accuracy and time to reach the target for the five sound mappings of tonal z. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results both the noise and oscillation mappings were eliminated.

Tonal Z

The tonal z results resemble those of tonal x, with B, P, and V all being significantly better subjectively than N and O, as seen in Table 3.5. A total of 11 subjects were tested on B, V, and P but no significant differences were found between them concerning subjective preference, target error or time to reach the target, ($F(2,20)=0.31$, $p=0.74$; $F(2,20)=0.05$, $p=0.95$; $F(2,20)=0.46$, $p=0.64$ respectively). As with tonal x and y, the possible sound mappings for tonal z were B, P and V.

All three location dimensions, x, y, and z, found similar results by eliminating N and O due to the subjective preference of B, P and V. Since x has a natural side to side motion, balance was selected to be its sound mapping. This leaves P and V for y and z. Of these two location dimensions, z was the most important and was given priority since it holds the depth information. Although no significant difference for z was noted, V was rated subjectively better than P and task time was smaller. Assigning volume to z left pitch bend for y. This assignment is reasonable since, as noted in section 2.2.2, pitch change is perceived to have a vertical spatial relationship which the y dimension also has.

3.2.4.4 Musical Pretest Results

Musical X

An initial seven subjects were used to test the seven types of sound mappings of the musical environment: volume, balance, oscillation, pitch bend, noise, tempo, and key shift[4]. A one-way ANOVA test with repeated measures showed a significant subjective difference of B, K, T and V over N; thus N was removed as seen in Table 3.6. Eight subjects in total were tested on the remaining mappings. Performing another one-way ANOVA test on subjective preference showed a significant difference between B over K, O, P; and V over K, O, P and T. No significant differences were found between B and V regarding subjective preference, target error or time to reach the target, ($F(1,7)=0.13$, $p=0.73$; $F(1,7)=0.80$, $p=0.40$; $F(1,7)=0.41$, $p=0.54$ respectively). Possible sound mappings for musical x were therefore B and V.

[4] These seven musical sound mappings will be abbreviated as follows: volume (V), balance (B), oscillation (O), pitch (P), noise(N), tempo (T) and key (K) respectively.

Map1	Map2	Measure	df	F Pr > F
Balance	Noise	Subjective Preference	(1, 6)	**16.62** 0.0065
Key Shift	Noise	Subjective Preference	(1, 6)	**6.25** 0.0465
Tempo	Noise	Subjective Preference	(1, 6)	**10.80** 0.0167
Volume	Noise	Subjective Preference	(1, 6)	**16.62** 0.0065
Balance	Key Shift	Subjective Preference	(1, 7)	**14.54** 0.0066
Balance	Oscillation	Subjective Preference	(1, 7)	**40.38** 0.0004
Balance	Pitch Bend	Subjective Preference	(1, 7)	**24.65** 0.0016
Volume	Key Shift	Subjective Preference	(1, 7)	**28.00** 0.0011
Volume	Oscillation	Subjective Preference	(1, 7)	**14.91** 0.0062
Volume	Pitch Bend	Subjective Preference	(1, 7)	**7.00** 0.0331
Volume	Tempo	Subjective Preference	(1, 7)	**5.60** 0.0499

Table 3.6 Significant Differences in Sound Mappings for Musical X
One-way ANOVA tests with repeated measures were performed on subjective preference, target accuracy and time to reach the target for the five sound mappings of musical x. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results the noise, key shift, oscillation, pitch bend and tempo sound mappings were eliminated.

Musical Y

The sound mappings for tonal y were also tested with an initial seven subjects. A one-way ANOVA test with repeated measures showed a significant subjective difference of B, K, T and V over N, as seen in Table 3.7. As for musical x, N was also removed for musical y. Eight subjects in total were tested on the remaining mappings. Another one-way ANOVA test was performed on subjective preference showing a significant difference between B, P, T, and V over O, thus O was removed. Nine subjects in total were tested on B, P, T, K, and V; no significant differences were, however, found regarding subjective preference, target accuracy and time to reach the target, ($F(1,8)=0.98$, $p=0.43$; $F(1,8)=0.02$, $p=0.99$; $F(1,8)=1.28$, $p=0.30$ respectively). Possible choices for musical y were therefore B, P, T, K, and V.

Map1	Map2	Measure	df	F Pr > F
Balance	Noise	Subjective Preference	(1, 6)	**36.00** **0.0010**
Key Shift	Noise	Subjective Preference	(1, 6)	**6.35** **0.0453**
Tempo	Noise	Subjective Preference	(1, 6)	**11.29** **0.0152**
Volume	Noise	Subjective Preference	(1, 6)	**15.00** **0.0082**
Balance	Oscillation	Subjective Preference	(1, 7)	**10.72** **0.0136**
Pitch Bend	Oscillation	Subjective Preference	(1, 7)	**5.73** **0.0479**
Tempo	Oscillation	Subjective Preference	(1, 7)	**9.33** **0.0185**
Volume	Oscillation	Subjective Preference	(1, 7)	**7.63** **0.0280**

Table 3.7 Significant Differences in Sound Mappings for Musical Y

One-way ANOVA tests with repeated measures were performed on subjective preference, target accuracy and time to reach the target for the seven sound mappings of musical y. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results both the noise and oscillation sound mappings were eliminated.

Map1	Map2	Measure	df	F Pr > F
Balance	Noise	Subjective Preference	(1, 5)	**8.93** **0.0305**
Tempo	Noise	Subjective Preference	(1, 5)	**7.50** **0.0409**
Volume	Noise	Subjective Preference	(1, 5)	**29.14** **0.0029**
Volume	Balance	Subjective Preference	(1, 6)	**6.25** **0.0465**
Volume	Key Shift	Subjective Preference	(1, 6)	**84.00** **0.0001**
Volume	Pitch Bend	Subjective Preference	(1, 6)	**12.79** **0.0117**
Volume	Oscillation	Subjective Preference	(1, 6)	**66.69** **0.0002**
Volume	Tempo	Subjective Preference	(1, 6)	**11.29** **0.0152**

Table 3.8 Significant Differences in Sound Mappings for Musical Z

One-way ANOVA tests with repeated measures were performed on subjective preference, target accuracy and time to reach the target for the seven sound mappings of musical z. In this table, sound mappings in column Map1 were significantly better than those in column Map2, with the number of degrees of freedom shown in the df column. From these results all sound mappings except for volume were eliminated.

Musical Z

As with musical x and y, N was also eliminated. A total of seven subjects tested the remaining mappings B, K, P, O, T, and V. Performing a one-way ANOVA test with repeated measures showed that V had a significantly better subjective rating than B, K, P, O and T as seen in Table 3.8.

Volume was preferred over all the other sound mappings for musical z and was thus selected for this dimension. The mappings for x had been reduced to either balance or volume, but since z was assigned to volume, x was assigned to balance. This choice is also favourable since x already has an inherent side to side metaphor as does balance. This left either key shift, tempo or pitch bend to become the y sound mapping. No significant differences were found, but pitch bend was selected since, as mentioned for tonal y, pitch maps naturally to a vertical dimension. Assigning musical y to pitch also maintains a sound mapping consistent with the tonal environment.

3.2.4.5 Orchestral Pretest Results

Orchestral Y

Eight subjects tested the four sound mappings for orchestral mappings: oscillation, pitch bend, tempo, and key shift. One-way ANOVA tests with repeated measures were performed on the subjective preference, target accuracy, and time to reach the target data, with no significant differences found, ($F(3,21)=1.54$, $p=0.23$; $F(3,21)=0.87$, $p=0.47$; $F(3,21)=2.75$, $p=0.07$ respectively). In comparison to the others, key shift and pitch bend required the most time to reach the target so removing them left two choices: oscillation and tempo. It was decided that oscillation would be the better of the two since music that is constantly changing in tempo is disturbing.

3.3 Experiment Procedure

3.3.1 Experiment 1 - Sound as a Position Cue

The purpose of the *Sound Cue* experiment, was to determine if sound could be used as a general position cue but more specifically as a depth cue. Of the three audible sound environments selected for testing, do any provide enough information to the user so that their depth perception is improved? Brown, discussed in section 2.3, noted that processing audio cues for screen location mapping took longer than visual cues. It is necessary to determine if this result also holds for these three sound environments.

Did the subjects really rely on the audio cues? As also noted in section 2.3, Colivata found that a high percentage of subjects acknowledged only the visual cues. It has also been noted that information that is not useful becomes annoying and likely ignored. [Brewster, Wright, & Edwards 94b]. What would be the effect on performance if no visual aids were presented but the user had to rely solely on sound cues? Would performance increase or decrease? Answers to these questions were sought in the Sound Cue experiment.

3.3.1.1 Design

As in the pretest experiments of section 3.2.4.2, a Gravis UltraSound MAX sound card and a Pentium computer were used to run the experiment interface and subject tasks described in section 3.1. Subjects used a standard mouse and keyboard for input and heard all sound responses over a pair of headphones

The experiment consisted of four test blocks each using a different sound environment: no sound, tonal, musical and orchestral. Testing these four blocks was done using a within-subject design where each subject tested each of the blocks. Within-subject testing was chosen over between-subject testing since it requires fewer subjects, less overall instruction and training time, and there is less variability between treatments due to individual differences. [Martin 77].

The disadvantage with within-subject testing is that subjects learn. Later test blocks could therefore obtain better results since the user has become more familiar with the interface. To eliminate these learning effects, counterbalancing was used following a Latin Square design. Each subject was randomly assigned one of four block orderings as seen in Table 3.9. With this kind of ordering, each block appears in each position exactly once and comes before and after every other block twice.

Order	Block A	Block B	Block C	Block D
1	No Sound	Tonal	Orchestral	Musical
2	Tonal	No Sound	Musical	Orchestral
3	Musical	Orchestral	Tonal	No Sound
4	Orchestral	Musical	No Sound	Tonal

Table 3.9 Latin Squares

The Latin squares block ordering ensures that each block appears in each position exactly once as seen in the table. The second ordering is based on the first by swapping the order of block A and B, as well as block C and D. Ordering three is found by swapping blocks AB of order two with blocks CD. The last block ordering is similar to the second, by swapping block A with B and block C with D of the third block ordering.

3.3.1.2 Subjects

Twenty paid subjects with no severe visual or hearing impairments, consisting of 14 males and 6 females, agreed to participate in the experiment. Subject ages ranged from 20 to 25. Eleven subjects came from a computer science background, 6 from science, and the remaining 3 from other disciplines. All subjects were familiar with using a mouse, though very few had had prior 3D interface experience. Musical experience of the subjects was fairly balanced with an average musical ability of 1.2 on a range of 0 to 3, representing no experience to extensive experience respectively.

Twenty were chosen to assign an equal number of subjects to each of the four block orderings.

3.3.1.3 Measurements Taken

Several measures were recorded for each subject including: subject biographics, and the subject's performance and subjective rating on each sound environment. The subject biographics form, shown in Appendix D, recorded the participant's age, gender, area of study, amount of musical experience, and 3D interface experience.

Two measures were recorded for the subject's performance of the experiment task described in section 3.1.1: target accuracy and the time required to reach that accuracy. Subjects were informed before the experiment that time and accuracy were equally important. Accuracy was measured by computing linear distance in units from the final cursor position to the target position, while time was measured in milliseconds.

Once the experiment was complete, subjects filled out the subjective rating form shown in Appendix D. This form requested a subjective rating for each of the sound environments in four different areas: subjective performance, preference, usability, and marketability. Rankings were on a scale of 1 to 4 where 1 was very poor, and 4 was very good.

3.3.1.4 Ear Test

It was desirable that the results from this thesis would be true for the general user and not specific to any specialized group. For this reason, individual hearing ability was not used for screening subjects but simply to account for any discrepancies in the data of subjects with poor hearing. This was accomplished by performing a two minute ear test.

The ear test was designed after Bekesy's ear test, which tested hearing ability over the frequency range of 100Hz to 10000Hz [Bekesy 60]. Rather than use this large range, only the musical frequencies from the playable notes A0 (27.5Hz) to A7 (3520Hz) were tested. Other frequency ranges were not considered since only the playable notes A0 to A7 would be used in the experiments. Later it was noted that this is a weakness since instruments use a whole spectrum of harmonics. This isn't, however, of great concern since, as noted above, the ear test was not used to screen subjects, but simply to note if any couldn't even hear these main frequencies.

Subjects were presented with an initial tone of 27.5Hz that constantly rose to the final frequency of 3520Hz. As the frequency was rising, the subject adjusted the volume level to be at the threshold of hearing. This was done using a timer that would decrease the volume by 2% every 80ms when there was no user input. To counteract this decrease, the subject could increase the volume by holding down the left mouse button. The subject was instructed to increase the volume level until it was just audible, then to release the mouse button and allow the volume to decrease until it was just inaudible. The subject would then depress the button again and repeat the process until the final frequency of 3520Hz was reached. The results of this volume fluctuation versus frequency, graphed on an *audiogram*, maps out a subject's threshold of hearing over the entire frequency range.

Figure 3.4 shows two sample audiograms from the ear test. Graph A shows the hearing threshold of a subject with good hearing, whereas graph B shows of one with poorer hearing. For both graphs, the very low frequencies were hard to hear and the volume level was at its maximum. Once the tone was detected they both showed a decrease in volume level required. The subject of graph A did not require the sound level to be as high as the subject of graph B for the remaining frequencies. This indicates that subject B has a weaker hearing ability for the higher frequencies.

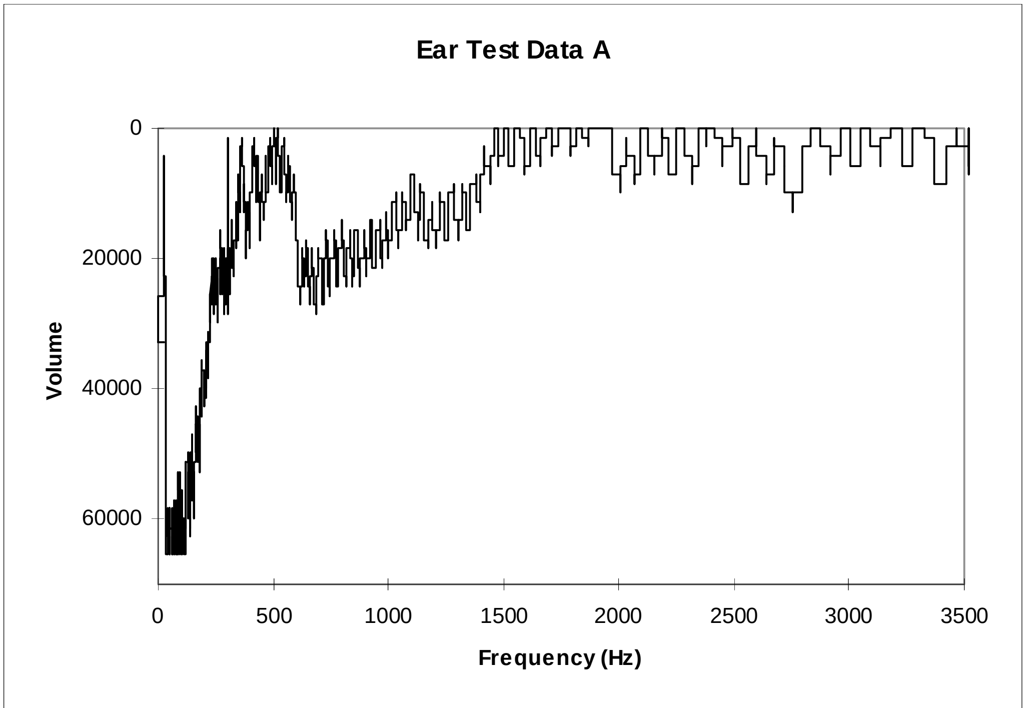
Before using the ear test on subjects, it was tested for its reliability using two separate steady test tones: A3 (220Hz) and A7 (3520Hz). By playing one of these fixed tones in the ear test rather than a frequency range, it was expected that the audiogram would be fairly constant with minor fluctuations. This was found to be the case at both the low and high test frequencies showing that the ear test could produce consistent results over the test period.

3.3.1.5 Procedure

Testing was carried out in a room in which no exterior noise could be heard by the subject because of sound baffling from the walls and the headphones that were worn. Subjects were instructed on the experiment procedure using the forms given in Appendix D and assigned to one of the four block orderings mentioned in section 3.3.1.1. Once they understood the experiment procedure, they performed the two minute ear test followed by the sequence of four test blocks. The entire experiment lasted for about one hour.

Each of the four blocks consisted of 25 task trials as described in section 3.1.1. The trials were broken down into three groups: 5 learning trials, 15 visual trials, and 5 blind trials. During the learning trials, subjects were to take as long as they wanted to become familiar with the current sound mapping. They were notified that during this time no data collected would be used. The remaining 20 trials, however, recorded both time and accuracy information.

The only difference between the visual and blind trials, was a completely blank screen for the blind trials. The purpose of these blind trials was to see how well the subjects performed using only sound feedback, after they had completed 5 learning and 15 visual trials.



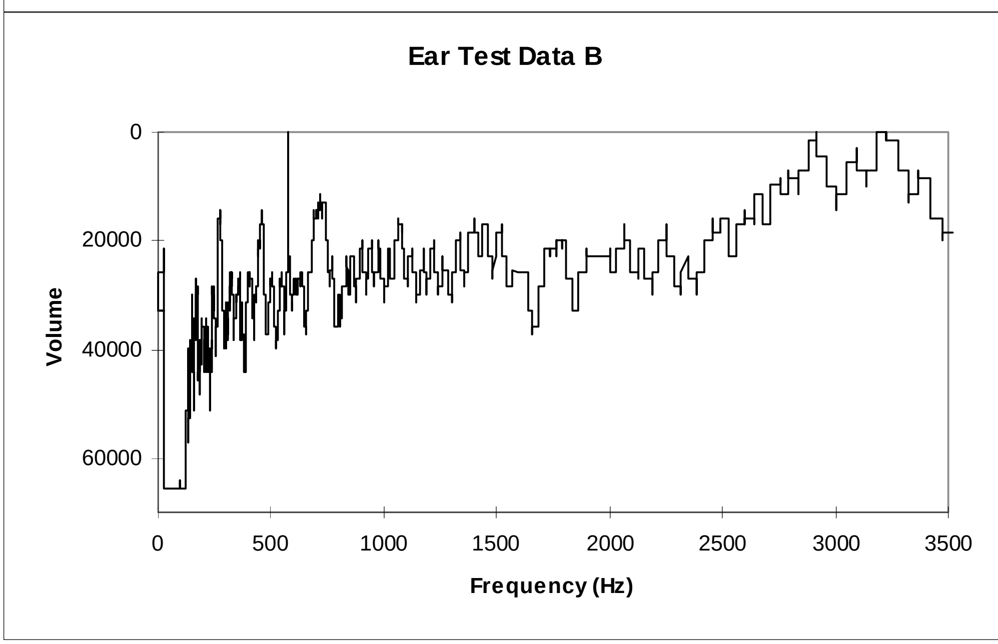


Figure 3.4 Ear Test Results

The above two audiograms show the auditory responses of two subjects with the given ear test. Subject A's had a good threshold of hearing and was able to hear most frequencies at a low volume level. The dip around 1000Hz was observed in several subject's audiograms. Subject B, began to hear the low frequency at about the same place on the frequency scale, however, the volume level was much higher.

3.3.2 Experiment 2 - Extended Period Trials

Skills are often improve as they are excercised over a period of time. The purpose of experiment 2, or the *Extended Period* experiment, was to determine how the results found in the Sound Cue experiment would change if subjects were to perform the same experiment on three separate days: Monday, Wednesday and Friday.

Also of interest was how learning affects the various sound environments' usability. In section 3.2.1 it was mentioned that a real 3D application could not normally play the target sound since it did not know the user's intentions. If a user spent time learning a sound mapping using target sounds, could they develop the ability to continue to use the sound cues even if the target sounds were unavailable? That is, could an experienced user find a location by recalling the sound of that location from a previous use? Do they create any associations between the visual cues and the sound cues or do they treat these cues as two independent sources? These questions were tested on the third day when subject experience was the greatest.

As with the Sound Cue experiment, this experiment used the Latin Squares block ordering and recorded the same measurements. Subjects never received the same block ordering that they had had on any previous day.

3.3.2.1 Subjects

Seven paid subjects, consisting of 6 males and 1 female, agreed to participate in the Extended Period Experiment. Subject ages ranged from 20 to 25 with 1 subject having a computer science background, 4 from science, and the remaining 2 from other disciplines. Musical experience was again at an average level at 1.3 on a range of 0 to 3, and few had had prior 3D interface experience.

3.3.2.2 Procedure

The first and second days proceeded exactly as in the Sound Cue Experiment, where subjects performed 25 trials: 5 learning, 15 experiment, and 5 blind trials. For the third day, however, 20 extra trials were added to each audible sound environment block. After they performed the original 15 visual and 5 blind trials, they were tested on 10 *prehear* and 10 *quiet* target trials.

Prehear Trials

It is thought that frequent users of a 3D application using one of these sound environments would develop a mental audio map of the sound environment and get to know what sounds occur at different locations. Locating the target then becomes a simple task of recalling the sound at that location and moving the cursor to a position that sounds like it.

To test this, recalling a sound was simulated by having subjects prehear the target sound before performing the task. Once they initiated the task, the target sound ceased and they could only hear the cursor's current location. Matlin noted that sound echoing in the auditory system, *echoic storage*, lasted from 2 to 3 seconds on average. [Matlin 94]. Since it was found in the Sound Cue Experiment that the task took much longer to do than 3 seconds, subjects would have to remember what the target sounded like rather than rely on echoic storage. The results of the trials would then show whether a subject was able to locate a target based on a previously heard sound.

Quiet Trials

The quiet target trials were similar to the prehear trials, yet in this case the target sound was never heard. As mentioned earlier, the purpose was to determine whether subjects make any associations between the visual cues and the audio ones. By not playing any target sound, subjects would have to rely solely on the cursor sound and the visual cues. Since the cursor size was made to be an unreliable depth cue, the main visual cues will come from the "blobby" object itself, described in section 3.1.2.

The only difference between these trials and the no sound environment trials is that there is a sound attached to the cursor. An improvement with the quiet trials over the no sound trials in task performance would therefore indicate that subjects had successfully made an association between the audio and the visual cues.

Chapter 4 Results

4.1 Experiment 1 - Sound as a Position Cue

4.1.1 Preparing the Data for Analysis

As described in section 3.3.1, four sound environments were tested in this experiment: no sound, tonal, musical, and orchestral. Since the last three present the user with an audible sound cue they were all tested under visual and no-visual cases. This makes seven different sound environment cases: no-sound, tonal-blind, tonal-visual, musical-blind, musical-visual, orchestral-blind, orchestral-visual.[5] The experiment results consisted of two measurements per environment (target error, and time to reach the target), and four subjective ratings (performance, preference, usability, and marketability). Measurement data was recorded for each trial according to subject and sound environment as shown in Table 4.1. Trials that were more than three standard deviations from the mean were considered outliers and were removed. The remaining subject trials were then averaged together to give one response value per measure for each subject and sound environment. The resulting data is found in Appendix F.

Measure	Sound Environments
Target Error, Time	N, TB, TV, MB, MV, OB, OV
Subjective Performance, Preference, Usability, Marketability	N, T, M, O

Table 4.1 Recorded Measures per Sound Environment
The two data measures, target error and time, were recorded for the seven possible sound environments. The subjective measures, however, were recorded for each general sound environment and made no distinction between the presences or absence of the visual cue.

Data among each subjects were grouped together by each measure and were analyzed using analysis of variance (ANOVA)[6] tests. Since ANOVA tests assume homogeneity of variance, any set of data that did not pass this test was adjusted using a log based function. The function was applied to all data values in the group so that the adjustment will not affect its statistical validity. Table 4.2 lists all the adjusting functions used to conform to the homogeneity of variance requirement.

[5]Hereafter these sound environments will be abbreviated N, TB, TV, MB, MV, OB, and OV respectively. Subjective ratings are listed as N, T, M and O since no distinction was made for the presence or absence of visual cues.
[6] All data was analyzed using SAS. See Appendix E for a sample SAS program.

Measure	Adjusting Function
Target Error	log (x)
Time	log (x)
Subjective Preference	
Subjective Performance	log (x)
Subjective Usability	
Subjective Marketability	

Table 4.2 Data Adjusting Functions for Homogeneity of Variance
Adjusting functions were applied to some of the measures to reduce the differences in variances which is required for ANOVA tests.

4.1.2 Subject and Block Order Effects

Subject differences (gender, area of study, musical experience, and 3D interface experience), and subject test block order were analyzed to determine if they had any effect on the measured results. These factors, discussed in section 3.3.1.2 and the actual subject data listed in Appendix F, are shown in Table 4.3 with the number of measured levels. Each of these factors was tested against the sound environments using a two-way ANOVA test with repeated measures on the sound environments.

Factor	Number of Levels	Levels
Gender	2	M, F
Area of Study	4	CS, SCI, ARTS, OTHER
Musical Experience	4	None, Little, Familiar, Proficient
3D Interface Experience	2	None, Little
Block Order	4	NTMO, TNMO, MONT, OMTN

Table 4.3 Measured Levels for Subject Factors and Block Order
The above table summarizes all the recoreded levels for each of the subject factors and block ordering.

4.1.2.1 Interaction of Subject Factors & Block Order with Sound Environment

Table 4.4 shows the interaction effect of these factors with sound environments on the six measures: target error, task time, subjective performance, preference, usability and marketability. The F values and probability, p, that it occured by chance are shown in the table with p being listed in a smaller typeface. Any significant result, that is one having an F value with $p<0.05$, are showin in bold face print and found in a shaded box.

The table shows that only 3D interface experience with subjective performance showed a significant interaction effect. This interaction is graphed in Figure 4.1. Subjects with no previous 3D interface experience found both the M and O sound environments to subjectively perform equally well, whereas users with some 3D experience gave a higher subjective performance rating to M, but a lower one to O. Comparing these two responses for each sound environment, however, showed no significant difference. ($F(1, 18) = 0.07$, $p = 0.7952$). This shows that although an interaction was detected, the difference in responses between the subject groups per sound environment was insignificant.

Factor	df	Error	Time	df	Prefer-ence	Perform-ance	Usabil-ity	Market-ability
		F Pr > F	F Pr > F		F Pr > F	F Pr > F	F Pr > F	F Pr > F
Gender	(6, 108)	1.09 0.3718	1.37 0.2328	(3, 54)	0.36 0.7841	0.17 0.9130	1.51 0.2220	0.54 0.6597
Area of Study	(18, 96)	0.83 0.6669	1.48 0.1153	(9, 48)	1.01 0.4487	0.79 0.6307	1.10 0.3787	2.00 0.0593
Music Exp	(18, 96)	1.31 0.1976	0.90 0.5818	(9, 48)	1.98 0.0632	1.00 0.4531	0.70 0.7020	1.33 0.2489
3D Exp	(6,108)	1.00 0.4272	1.30 0.2647	(3, 54)	1.45 0.2394	**2.91 0.0428**	2.11 0.1100	2.01 0.1237
Block Order	(18, 96)	0.78 0.7166	1.00 0.4617	(9, 48)	1.19 0.3248	0.45 0.8999	1.27 0.2775	0.92 0.5134

Table 4.4 Interaction of Subject Factors and Block Order with Sound Environment

Two-Way ANOVA with Repeated Measures on One Factor (Sound Environment)
Univariate Test of Hypothesis for Within Subject Effects
The table lists all the F values and probability, p, that this value occured by chance. All significant F values, whose $p<0.05$, are shown in boldface. The number of degrees of freedom, shown in the df column, varies depending on the number of levels in each factor and the number of sound environments per measure. The table here shows that there was no interaction between subject factors and the sound environments except for 3D interface experience and subjective performance.

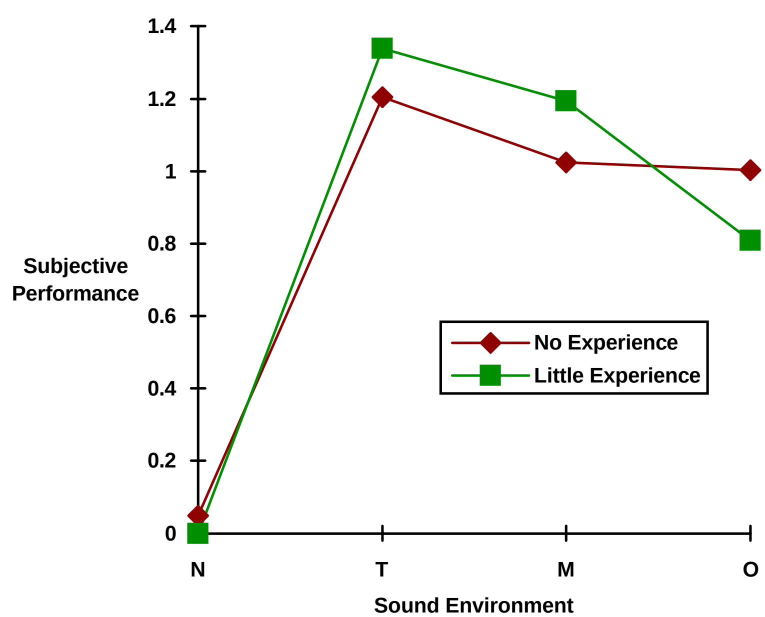


Figure 4.1 Interaction of 3D Graphics Experience and Sound Environment with Respect to Subjective Performance
Two types of 3D graphics experience were reported by the subjects: none and little. The graph above shows that these two groups responded differently depending on the sound environment. This suggests an interaction between 3D graphics experience and sound environment. Comparing each pair of responses per sound environment, however, showed no significant difference.

4.1.2.2 Effect of Subject Factors and Block Order on Measured Data

Table 4.5 shows the effect of each of the subject factors on the six recorded measures. Only subject gender caused a significant difference in marketability ratings. Females rated all sound environments with a higher marketable rating than males with a mean of 2.75 over 2.37 on a subjective scale of 1 to 4. Other than this difference, subject factors and block ordering made no significant effect on any of the six measured results.

Factor	df	Error	Time	Preference	Performance	Usability	Marketability
		F Pr > F	F Pr > F	F Pr > F	F Pr > F	F Pr > F	F Pr > F
Gender	(1,18)	0.11 0.7485	1.36 0.2584	0.44 0.5166	0.93 0.3464	0.15 0.7040	**6.19** **0.0229**
Area of Study	(3, 16)	0.94 0.4446	0.57 0.6432	2.43 0.1033	2.93 0.0656	0.29 0.8337	0.24 0.8643
Music Exp	(3, 16)	0.43 0.7348	1.18 0.3491	0.06 0.9783	1.61 0.2269	0.85 0.4851	0.08 0.9676
3D Exp	(1, 18)	0.54 0.4708	0.30 0.5910	1.44 0.2449	0.07 0.7952	0.15 0.7040	2.99 0.1010
Block Order	(3, 16)	0.66 0.5882	1.13 0.3670	0.77 0.5250	0.89 0.4669	0.71 0.5610	1.24 0.3295

Table 4.5 Effect of Subject Differences and Block Order on Measured Data
Two-Way ANOVA with Repeated Measures on One Factor (Sound Environment)
Test of Hypothesis for Between Subject Effects
How subject differences and test block ordering affected the data was tested and found no significant effects except for gender and subjective marketability. In general females rated sound environments as being more marketable than did males.

4.1.3 Interaction Effects of Visual Cue with Sound Environment

The interaction effect of visual cue (visual and blind)[7] with sound environment on target error and time was tested using a two-way ANOVA test with repeated measures on both factors. Table 4.6 shows that interaction was significant, for both target error and time and graphed Figure 4.2, and 4.3 respectively.

Interaction Factors	df	Error	Time
		F Pr > F	F Pr > F
Visual Cue & Sound Environment	(2, 38)	**3.66** **0.0353**	**7.76** **0.0015**

Table 4.6 Interaction of Visual Cue with Sound Environment
Two-Way ANOVA Repeated Measures on Two Factors (Visual Cue & Sound Environment)
Univariate Test of Hypothesis for Within Subject Effects
The interaction between visual cue and sound environment is significant for both target error and time.

[7] Visual cues are abbreviated as V, B respectively.

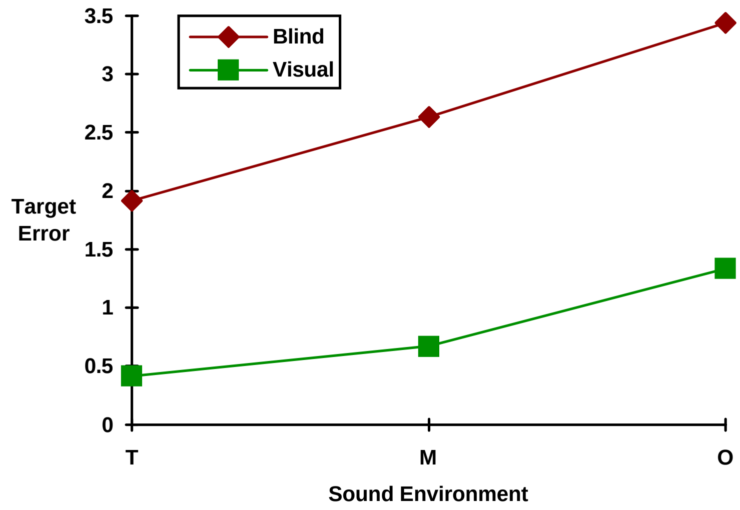


Figure 4.2 Interaction of Visual Cue and Sound Environment with Respect to Target Error

Parallel lines would indicate that no interaction was occurring as sound environment changes. In the graph above the visual cue lines are not parallel indicating an interaction between visual cue and sound environment. It also shows that the blind cues caused more errors that the visual ones.

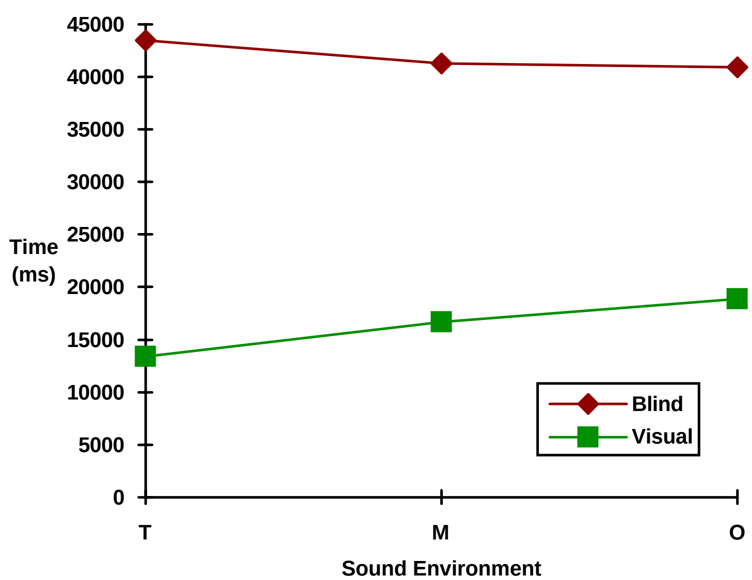


Figure 4.3 Interaction of Visual Cue and Sound Environment with Respect to Time

As in Figure 4.2, interaction occurred between the visual cue and the sound environment, shown here by two non-parallel lines. In this case, blind cue time decreased from sound environment T to O, but visual cue time increased.

4.1.4 Correlation of Target Error and Time

The Pearson correlation coefficient of target error and time was computed for each sound environment and found to have no statistical significance, as seen in Table 4.7. This is somewhat surprising since it was thought that a correlation would have been found where the more time a subject spent on the task, the less the target error. This result, however, shows that no such correlation exists but that subject task time and target error are essentially independent from one another.

N	TB	TV	MB	MV	OB	OV
R **Pr > R**	**R** **Pr > R**	**R** **Pr > R**	**R** **Pr > R**	**R** **Pr > R**	**R** **Pr > R**	**R** **Pr > R**
-0.31624 0.1743	-0.20910 0.3763	-0.40959 0.0729	-0.23914 0.3099	-0.15905 0.5030	-0.06434 0.7876	0.03773 0.8745

Table 4.7 Correlation of Target Error and Time for Each Sound Environment
The Pearson correlation coefficients showed no significant correlation between target error and time.

4.1.4 Sound Environment Comparison

The sound environments (N, TB, TV, MB, MV, OB, and OV) were all compared using one-way ANOVA tests with repeated measures on the sound environment. Target error and time both had significant differences among the sound environments with F(6,114)=62.47 (p=0.0001) and 90.80 (p=0.0001) respectively. The subjective measures (performance, preference, usability, and marketability) also reached significance with F(3,57)=8.56 (p=0.0001), 120.01 (p=0.0001), 5.31 (p=0.0027) and 12.61 (p=0.0001) respectively. Table 4.8 and 4.9 make three type of comparisons: no sound vs. sound, blind vs. visual cue, and a comparison for the best blind and visual sound environments. The means are then graphed in Figure 4.4, 4.5, and 4.6 for target error, time and subjective ratings respectively.

4.1.4.1 No Sound vs. Sound

In general, the no sound environment proved to have medial target errors and small task times. Visual sound environments (TV, MV, OV) all showed significantly less target errors than the no sound environment, however, it took longer for them to reach that level of accuracy. The blind sound environments (TB, MB, OB) took significantly longer than the no sound environment. It interesting to note, however, that the TB and MB sound environments showed no significant difference in target error over the no sound environment.

Subjectively, the audible sound environments (T, N, and O) were ranked significantly better than the no sound environment in subjective performance, preference, usability and marketability.

Comparison df (1, 19)	S1	S2	Error F (Pr > F)	Time F (Pr > F)
No Sound vs. Sound	N	TB	0.89 (0.3571)	**145.10 (0.0001)**
	N	MB	3.17 (0.0908)	**258.39 (0.0001)**
	N	OB	**23.37 (0.0001)**	**186.64 (0.0001)**
	N	TV	**100.56 (0.0001)**	**47.92 (0.0001)**
	N	MV	**65.46 (0.0001)**	**99.11 (0.0001)**
	N	OV	**9.83 (0.0054)**	**79.82 (0.0001)**
Blind vs. Visual	TV	TB	**112.27 (0.0001)**	**133.51 (0.0001)**
	MV	MB	**92.13 (0.0001)**	**100.41 (0.0001)**
	OV	OB	**120.11 (0.0001)**	**84.34 (0.0001)**
Audible Sound Environments	TB	MB	**6.65 (0.0184)**	0.04 (0.8526)
	TB	OB	**23.59 (0.0001)**	0.03 (0.8683)
	MB	OB	**8.29 (0.0096)**	0.20 (0.6619)
	TV	MV	**27.08 (0.0001)**	**11.84 (0.0027)**
	TV	OV	**102.19 (0.0001)**	**15.97 (0.0008)**
	MV	OV	**87.39 (0.0001)**	2.47 (0.1323)

Table 4.8 Comparing Sound Environments Using Target Error and Time
One-Way ANOVA with Repeated Measures on Sound Environment
Analysis of Variance of Contrast Variables
The above table compares the target error and time for some of the pairs of sound environments. A significant difference between the task error or time of the sound environment in column S1 and the sound environment in column S2 is shown in boldface. This simply shows that the difference was significant, compare with the graphs of Figure 4.4 and 4.5 to determine which were better.

4.1.4.2 Blind vs. Visual

Comparing the two possible visual cues (blind and visual) for each sound environment showed that visual cues gave significantly less errors and were significantly faster as was expected.

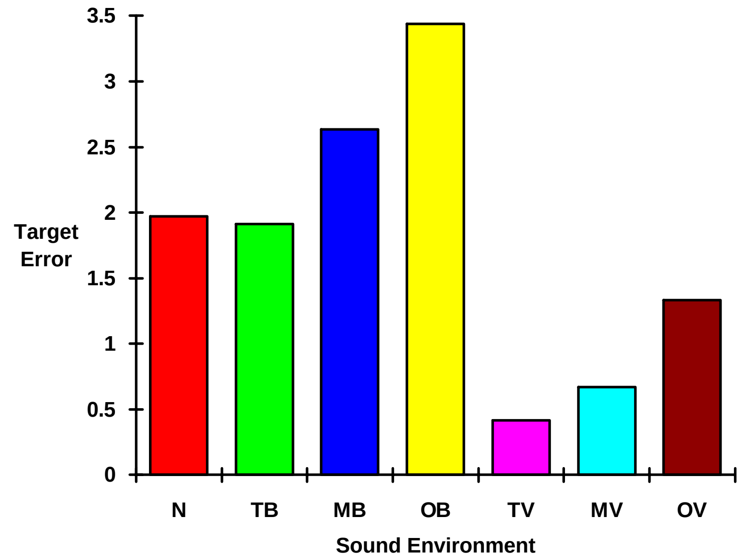


Figure 4.4 Sound Environment vs. Target Error

The above graph shows that the visual sound environments (TV, MV, and OV) had smaller target errors than the no sound environment. As expected, the blind sound environments showed an increase in error over their visual equivalent. It is interesting to note that TB and MB showed no significant difference over N, suggesting that either of these sound-only environments could be used in place of the visual-only case. Within the audible sound environments, T was better than M followed by O in both the blind and visual cases.

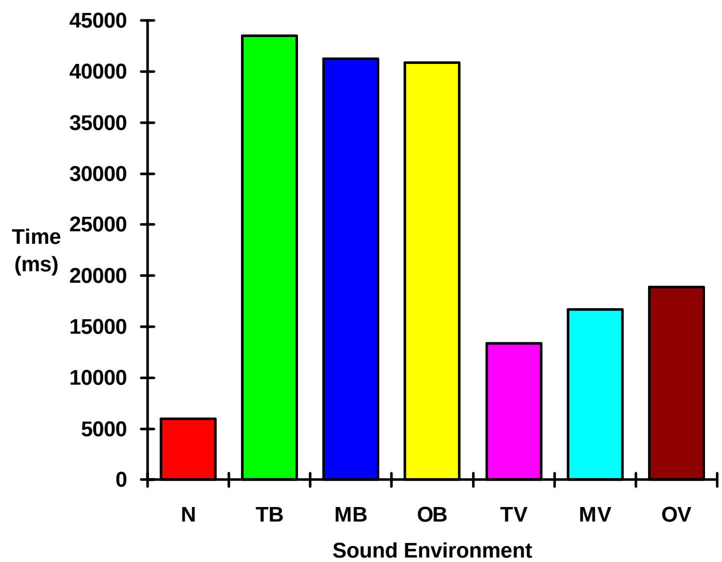


Figure 4.5 Sound Environment vs. Time

This figure shows the great difference between the time of the no sound environment and the other audible sound environments. As with target error the visual environments had significantly better time performance than the blind environments. It is interesting that TV was significantly better than MV and OV, yet TB showed no significant difference over MB and OB.

4.1.4.3 Audible Sound Environments

Of the audible sound environments (T, M, and O), which performed the best? In terms of target error T was significantly better than M, followed by O in both the blind and visual cases. Time to complete the task followed the same ordering in the visual case with TV being significantly better than MV and OV. There was, however, no significant difference between MV and OV. In the blind case, no significant difference in time was noted across any of these sound environments.

Subjectively, there was no significant difference in subject preference for the T, M and O sound environments. T was, however, ranked as performing better and was found to be easier to use than the M and O sound environments. M also showed to be easier to use and more marketable than the O sound environment.

Comparison	S1	S2	Prefer-ence	Perform-ance	Usabil-ity	Market-ability
df (1, 19)			F Pr > F	F Pr > F	F Pr > F	F Pr > F
No Sound vs.	N	T	**8.43** **0.0091**	**450.28** **0.0001**	**9.21** **0.0068**	**22.77** **0.0001**
Sound	N	M	**37.70** **0.0001**	**293.52** **0.0001**	**2.50** **0.1306**	**37.33** **0.0001**
	N	O	**13.04** **0.0019**	**222.35** **0.0001**	0.11 0.7481	**10.33** **0.0046**
Audible	T	M	3.62 0.0724	**5.92** **0.0250**	**12.84** **0.0020**	0.05 0.8336
Sound	T	O	0.77 0.3900	**17.83** **0.0005**	**23.75** **0.0001**	2.67 0.1189
Environments	M	O	0.16 0.6939	2.48 0.1319	**5.15** **0.0351**	**3.67** **0.0705**

Table 4.9 Comparing Sound Environments Using Subjective Ratings
One-Way ANOVA with Repeated Measures on Sound Environment
Analysis of Variance of Contrast Variables
The above table compares the subjective ratings for some of the pairs of sound environments. A significant difference between these ratings of the sound environment in column S1 and the sound environment in column S2 is shown in boldface. This simply shows that the difference was significant, compare with the graph in Figure 4.6 to determine which was better.

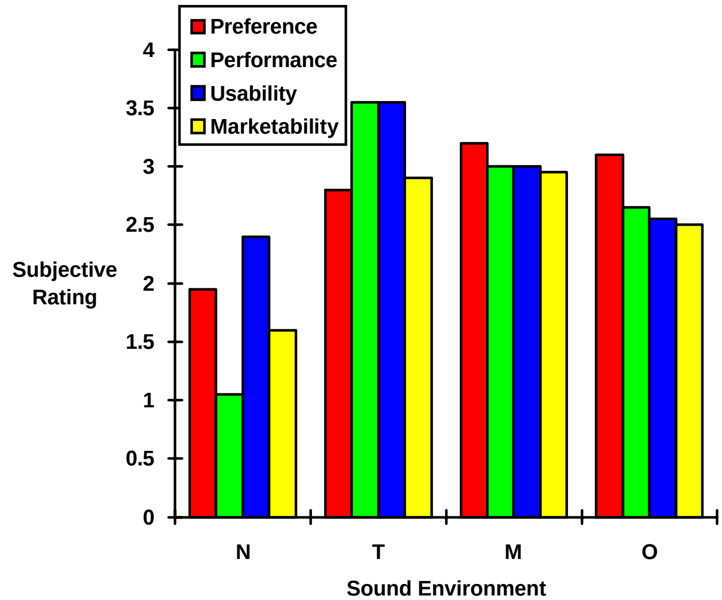


Figure 4.6   Sound Environment vs. Subjective Ratings
In this figure, larger subjective results are better. Overall, subjects rated all the audible sound environments over the no sound environment with respect to subjective performance, preference, usability and marketability. Tonal was perceived to perform the best and was found to be the easiest to use. The environments which played music (M and O) were preferred over tonal, however, the difference was not significant. Both T and M were rated as more marketable over O.

4.2   Experiment 2 - Extended Period Trials

4.2.1   Preparing the Data for Analysis

The Extended Period experiment recorded data for three separate days: Monday, Wednesday and Friday of a single week. As with the Sound Cue experiment, the data measures target accuracy and time were preprocessed to remove any outlying trials and then averages were computed. Data failing to approach homogeneity of variance were adjusted using a log based function. These functions are listed in Table 4.10.

Measure	Adjusting Function Day 1	Adjusting Function Day 2	Adjusting Function Day 3
Target Error	log (*x*)	log (*x* ^ *0.5*)	log (*x*)
Time	log (*x*)	log (*x*)	log (*x* ^ *0.5*)
Subjective Preference			
Subjective Performance	log (*x*)	log (*x*)	log (*x*)
Subjective Usability			
Subjective Marketability			

Table 4.10   Data Adjusting Functions for Homogeneity of Variance
Adjusting functions were applied to some of the measures to reduce the differences in variances which is required for ANOVA tests.

4.2.2 Interaction Effects of Visual Cue, Sound Environment and Day

The interaction effect of visual cue with sound environment and day on target error and time was tested using a three-way ANOVA test with repeated measures on three factors. Table 4.11 shows that there was no significant interaction between these factors except for visual cue and sound environment. This does not suggest that there was no learning effect, but that the type of sound environment and visual cue did not alter the learning effect. The interaction between visual cue and sound environment with respect to target error and time are shown in Figures 4.7 and 4.8 respectively.

Interaction Factors	df	Error F Pr > F	Time F Pr > F
Visual Cue, Sound Env. & Day	(4, 24)	0.71 0.5914	0.73 0.5831
Visual Cue, Day	(2, 12)	0.20 0.8175	0.15 0.8624
Sound Env. & Day	(4, 24)	0.65 0.6356	0.57 0.6862
Visual Cue & Sound Env.	(2, 12)	6.29 0.0136	5.25 0.0230

Table 4.11 Interaction of Visual Cue, Sound Environment, and Day on Target Error and Time

Three-Way ANOVA Repeated Measures on Three Factors

Univariate Test of Hypothesis for Within Subject Effects

No significant interaction was found between the experiment day and either visual cue or sound environment. This does not suggest that day did not have an effect on the data, but simply that it did not interact with the type of sound environment or visual cue. As in Table 4.6, visual cue and sound environment interacted in both target error and time.

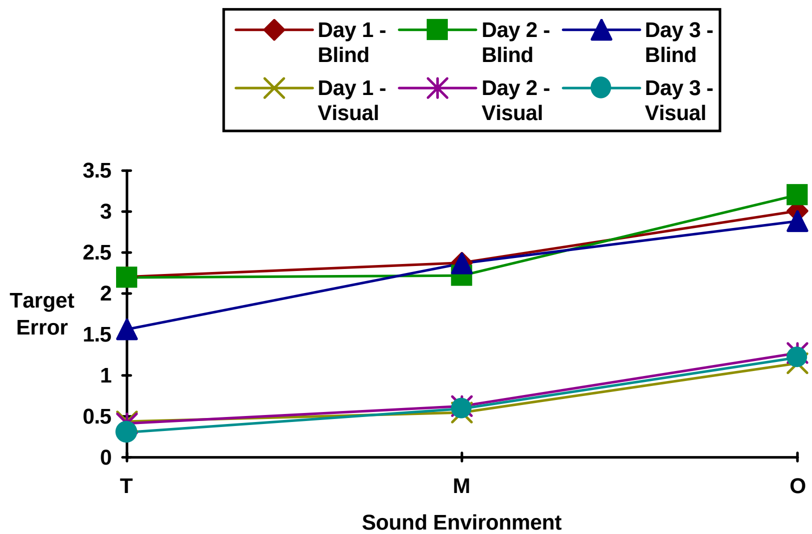


Figure 4.7 Interaction of Visual Cue and Sound Environment on Target Error

There was no significant interaction found between day and visual cue or sound environment. Visual cue did however interact significantly with sound environment as seen in the non-parallel blind and visual cue lines.

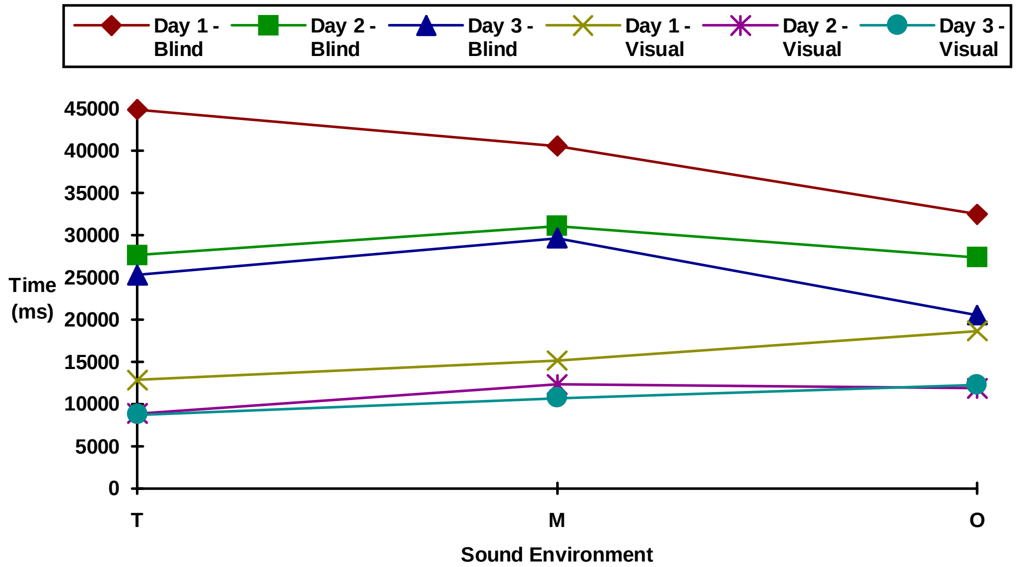


Figure 4.8 Interaction of Visual Cue and Sound Environment on Time
Again, no significant interaction existed between the day and the visual cue or sound environment. Even though the three days' blind lines are not parallel, the difference did not prove to be significant. As with target error, however, visual cue did interact significantly with sound environment.

The interaction effect of day and sound environment on the subjective ratings was tested using a two-way ANOVA test with repeated measures on the two factors. Table 4.12 shows that there was no significant interaction between these factors except for the subjective preference measure as seen in Figure 4.9. The graph shows that the sound environments that played music (musical and orchestral) both became more preferable by subjects as the experiment period progressed.

Interaction Factors	df	Preference	Performance	Usability	Marketability
		F Pr > F	F Pr > F	F Pr > F	F Pr > F
Day & Sound Environment	(6, 36)	2.64 0.0319	0.94 0.4816	0.62 0.7101	0.62 0.7127

Table 4.12 Interaction of Day and Sound Environment on Subjective Ratings
Two-Way ANOVA Repeated Measures on Two Factors (Day & Sound Environment)
Univariate Test of Hypothesis for Within Subject Effects
Only subjective preference was affected by a significant interaction between experiment day and sound environment.

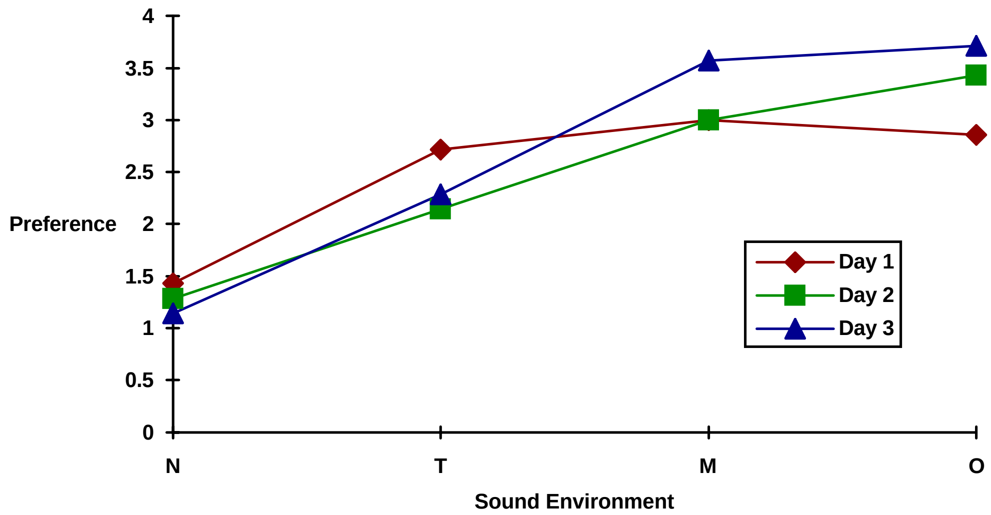


Figure 4.9 Interaction of Day and Sound Environment on Subjective Preference

The day lines here show that there is a very significant interaction between day and sound environment with respect to subjective preference, where a larger score is better. As the experiment period progressed, the musical and orchestral sound environments both attained higher preference over the no sound and tonal environments.

4.2.3 Sound Environment Comparison

The sound environments (N, TB, TV, MB, MV, OB, and OV) were all compared each day using one-way ANOVA tests with repeated measures on the sound environment. The day of the week showed to have no significant effect on target error ($F(2,12)=1.20$, $p=0.3343$). It did, however, have a significant effect on the time to reach the target ($F(2,12)=23.52$, $p=0.0001$). Sound environment caused a significant effect of target error, task time, and the subjective ratings recorded for the three days, as seen in Table 4.13 and Table 4.14. Only subjective usability showed no significant effect due to sound environment. Tables 4.15 and 4.16 make three type of comparisons: no sound vs. sound, blind vs. visual cue, and a comparison for the best blind and best visual sound environment. The means of the target error, time, and subjective ratings are graphed in Figures 4.10 to 4.12 respectively.

Day df (6, 36)	Error F Pr > F	Time F Pr > F
1	20.34 0.0001	24.53 0.0001
2	32.20 0.0001	36.80 0.0001
3	22.01 0.0001	43.48 0.0001

Table 4.13 Effect of Sound Environment on Target Error and Time

As seen in the table, the type of sound environment created a significant effect on target error and task time for each of the three days.

Day df (3, 18)	Preference F Pr > F	Performance F Pr > F	Usability F Pr > F	Marketability F Pr > F
1	**4.11** **0.0219**	**44.85** **0.0001**	3.02 0.0566	**4.89** **0.0117**
2	**9.65** **0.0005**	**86.75** **0.0001**	1.22 0.3297	**4.67** **0.0139**
3	**26.03** **0.0001**	**58.90** **0.0001**	1.29 0.3090	**7.95** **0.0014**

Table 4.14 Effect of Sound Environment on the Subjective Ratings

The type of sound environment caused a significant effect each day for subjective performance, preference, and marketability. There was, however, no significant effect on usability, suggesting that all sound environments were rated at approximately the same level of usability.

Comparison df (1, 6)	S1	S2	Day 1 Error F Pr > F	Day 1 Time F Pr > F	Day 2 Error F Pr > F	Day 2 Time F Pr > F	Day 3 Error F Pr > F	Day 3 Time F Pr > F
No Sound vs Sound	N	TB	0.02 0.8861	**25.58** **0.0023**	3.32 0.1183	**47.65** **0.0005**	1.24 0.3083	**66.64** **0.0002**
	N	MB	1.37 0.2861	**49.68** **0.0004**	3.96 0.0939	**80.83** **0.0001**	**22.01** **0.0034**	**191.41** **0.0001**
	N	OB	4.15 0.0878	**58.49** **0.0003**	**14.04** **0.0095**	**45.47** **0.0005**	**53.84** **0.0003**	**146.16** **0.0001**
	N	TV	**26.19** **0.0022**	**27.46** **0.0019**	**36.74** **0.0009**	**42.00** **0.0006**	**65.87** **0.0002**	**44.77** **0.0005**
	N	MV	**32.84** **0.0012**	**32.11** **0.0013**	**20.31** **0.0041**	**69.96** **0.0002**	**15.18** **0.0080**	**235.69** **0.0001**
	N	OV	5.89 0.0514	**17.10** **0.0061**	0.58 0.4738	**33.61** **0.0012**	0.59 0.4728	**45.78** **0.0005**
Blind vs Visual	TV	TB	**30.63** **0.0015**	**19.21** **0.0047**	**95.19** **0.0001**	**39.61** **0.0007**	**89.03** **0.0001**	**56.95** **0.0003**
	MV	MB	**42.93** **0.0006**	**39.11** **0.0008**	**28.07** **0.0018**	**52.79** **0.0003**	**26.74** **0.0021**	**133.73** **0.0001**
	OV	OB	**29.10** **0.0017**	**23.82** **0.0028**	**264.51** **0.0001**	**45.82** **0.0005**	**47.13** **0.0005**	**20.26** **0.0041**
Sound Comparison	TB	MB	0.47 0.5189	0.06 0.8199	0.08 0.7829	0.83 0.3987	**13.69** **0.0101**	2.91 0.1391
	TB	OB	4.15 0.0877	0.68 0.4417	**6.20** **0.0471**	0.03 0.8763	**14.22** **0.0093**	0.52 0.4989
	MB	OB	1.85 0.2223	8.18 0.0288	**6.06** **0.0490**	1.65 0.2464	3.37 0.1159	**15.92** **0.0072**
	TV	MV	4.96 0.0675	8.19 0.0287	**6.20** **0.0472**	**14.60** **0.0087**	**12.60** **0.0121**	5.53 0.0569
	TV	OV	**26.55** **0.0021**	4.35 0.0822	**40.23** **0.0007**	**8.11** **0.0292**	**65.89** **0.0002**	4.79 0.0713
	MV	OV	**19.76** **0.0044**	1.36 0.2871	**23.89** **0.0027**	0.62 0.4597	**12.15** **0.0131**	0.41 0.5474

Table 4.15 Comparing Sound Environments Using Target Error and Time
One-Way ANOVA with Repeated Measures
Analysis of Variance of Contrast Variables

The above table compares the target error and time for some of the pairs of sound environments. A significant difference between the task error or time of the sound environment in column S1 and the sound environment in column S2 is shown in boldface. This simply shows that the difference was significant, compare with the graphs of Figure 4.10 and 4.11 to determine which environment was better.

4.2.3.1 No Sound vs Sound

Figure 4.10 shows that over the three days the target error remained about the same for MB, MV, and OV; but improved for N, TB, TV, and OB. Time to reach the target, however, improved as days progressed for all sound environments as seen in Figure 4.11. The tables and graphs show that by day three N had significantly better error performance than MB and OB but still no significant difference with TB. When compared to the visual sound environments, N was still significantly poorer than TV and MV, but its error difference with OV had become less significant. N still proved to have the best time showing significant differences over all the other sound environments even by the third day.

Subjectively, N’s preference decreased over the week, while M and O’s increased causing an even greater significant difference, as seen in Figure 4.12. N’s subjective performance made no change and was significantly poorer than the audible sound environments. Although N’s usability rating was lower than the other environments, it improved over the week and showed no significant difference to the other environment’s ratings. N’s marketability rating showed no change over the week but its difference to the marketability rating of M and O became more significant by the end of the week. The loss of marketability significance between N and O on the second day was due to a larger variance of the original data that day.

4.2.3.2 Blind vs. Visual

As found in the Sound Cue experiment, the visual sound environments out-performed their blind equivalents by a significant amount in both target error and task time.

4.2.3.3 Audible Sound Environments

Comparing the audible sound environments showed similar results to the Sound Cue experiment. T generally performed better than M and O both visually and in the blind case. Of interest, TB’s target error improved the most over MB and OB so that a significant difference was detected. TV also improved in this area and also caused a significant difference over MV and OV. With respect to time, all sound environments improved over the experiment period, but few significant differences between sound environments were detected. On last day, however, OB’s task time reduced a significant amount over MB’s time. It is interesting to note that OB’s time was consistently better than TB and MB.

M and O were almost always improving in all four subjective areas over the week. T on the other hand showed a decrease in subjective preference, usability, and marketability, as seen in Figure 4.12. In comparison, M and O showed a significant difference over T with respect to subjective preference. This too was seen in the marketability rating with both the musical environments (M and O) being more marketable than T. For subjective performance and usability, T was found to help the best and was the easier to use, however, the differences to M and O were not significant.

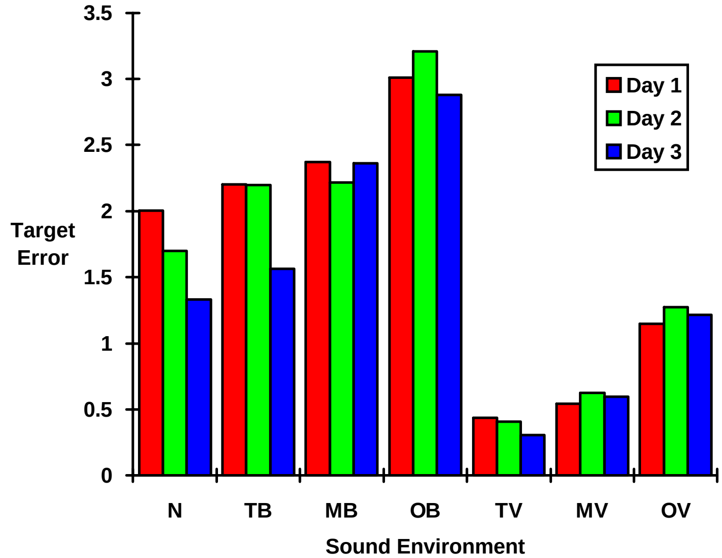


Figure 4.10 Sound Environment vs. Target Error

In general, target error improved over the experiment period. The tonal cases TB and TV still proved to be the best audible sound environment. Also of interest was the fact N's error rate also improved. This shows that the subjects were becoming more aquainted with the 3D interface.

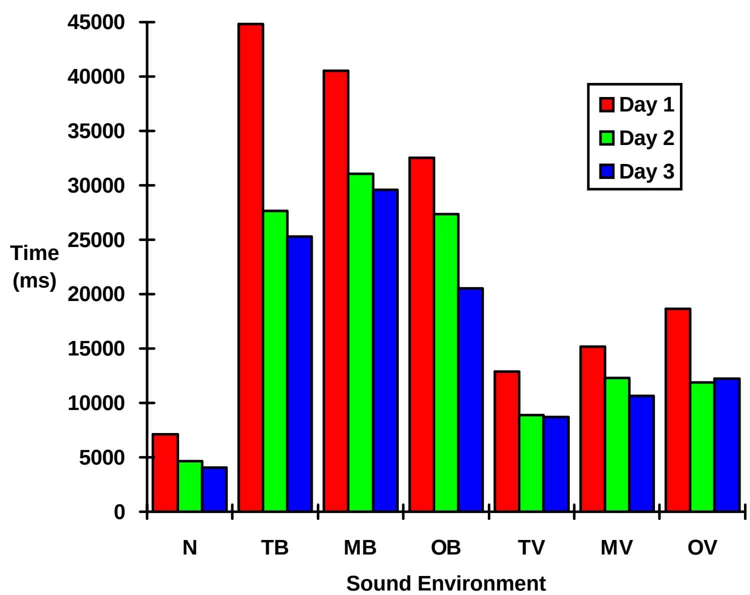


Figure 4.11 Sound Environment vs. Time

As with target error, task time also decreased. This indicates that learning improves both time and accuracy. It is interesting to see OB's time being consistently lower than the other blind environments over the three days.

Day 1 df = (1, 6)	S1	S2	Preference F Pr > F	Performance F Pr > F	Usability F Pr > F	Marketability F Pr > F
No Sound vs	N	T	1.17 0.3208	**604.15** **0.0001**	5.26 0.0616	3.98 0.0930
Sound	N	M	**11.11** **0.0157**	**89.10** **0.0001**	2.10 0.1975	**10.68** **0.0171**
	N	O	**7.36** **0.0349**	**118.21** **0.0001**	1.00 0.3559	**11.11** **0.0157**
Sound	T	M	**6.35** **0.0453**	1.62 0.2503	2.40 0.1723	0.36 0.5686
Comparison	T	O	1.60 0.2534	4.44 0.0796	**36.00** **0.0010**	0.18 0.6891
	M	O	0.13 0.7358	0.28 0.6150	1.00 0.3559	0.18 0.6891

Day 2						
No Sound vs	N	T	2.40 0.1723	**604.15** **0.0001**	3.69 0.1030	2.40 0.1723
Sound	N	M	**10.80** **0.0167**	**495.31** **0.0001**	1.07 0.3410	**6.25** **0.0465**
	N	O	**28.12** **0.0018**	**118.21** **0.0001**	0.46 0.5222	5.45 0.0582
Sound	T	M	**10.80** **0.0167**	1.35 0.2894	0.30 0.6036	4.50 0.0781
Comparison	T	O	**9.35** **0.0223**	5.72 0.0539	1.41 0.2797	**6.35** **0.0453**
	M	O	1.35 0.2894	2.56 0.1610	1.00 0.3559	0.00 1.0000

Day 3						
No Sound vs	N	T	5.05 0.0656	**1071.38** **0.0001**	5.25 0.0618	3.48 0.1112
Sound	N	M	**66.69** **0.0002**	**162.54** **0.0001**	1.83 0.2248	**16.62** **0.0065**
	N	O	**74.77** **0.0001**	**89.10** **0.0001**	0.43 0.5352	**13.44** **0.0105**
Sound	T	M	**20.25** **0.0041**	3.79 0.0996	0.18 0.6891	**6.35** **0.0453**
Comparison	T	O	**15.00** **0.0082**	2.97 0.1356	0.79 0.4072	2.88 0.1403
	M	O	1.00 0.3559	0.01 0.9074	0.63 0.4571	0.30 0.6036

Table 4.16 Comparing Sound Environments Using Subjective Rating
One-Way ANOVA with Repeated Measures
Analysis of Variance of Contrast Variables

This table is similar to Table 4.15, but compares the different sound environments with respect to the subjective ratings. A significant difference between a rating of the sound environment in column S1 and the sound environment in column S2 is shown in boldface. This simply shows that the difference was significant, compare with the graph in Figure 4.12 to determine which environment was better.

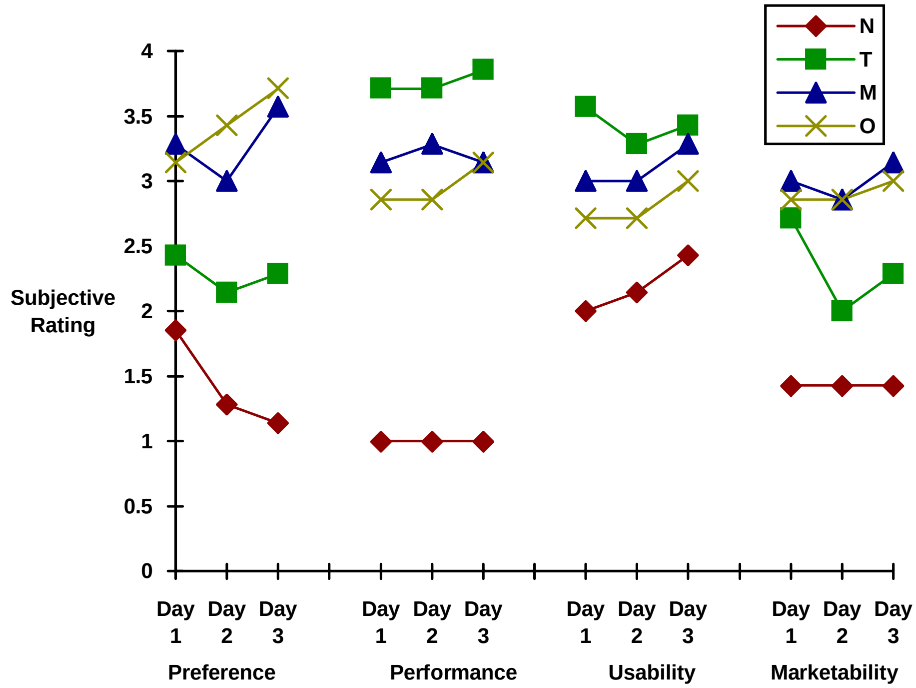


Figure 4.12   Day vs. Subjective Ratings

The graph shows how the subjective ratings changed as the experiment period progressed, where a larger rating is better. The two musical environments (M and O) both showed general improved in subjective performance, preference, usability and marketability. The tonal environment, on the other hand decreased in subjective preference, usability and marketability. It did, however, improve in its subjective performance rating. The no sound environment was found to become easier to use through the experiment period, but it suffered in its subjective preference.

4.2.4 Prehear and Quiet Target Trials

On the third day of second experiment, the two extra sound trails, prehear target and quiet target[8], were performed, as described in section 3.3.2.2. Both of these tests attempted to determine if an experienced subject had developed a mental audio map of the sound environment and was able to perform the task without hearing the target sound during the trial. The prehear tests allowed the subject to "prehear" the target sound before beginning, whereas the quiet tests played no target sound. These were both tested on the three types of audible sound environments: T, M, and O. The resulting thirteen sound environments (N, TB, TV, TP, TQ, MB, MV, MP, MQ, OB, OV, OP, and OQ) were all compared using one-way ANOVA tests with repeated measures on the sound environment. The effect of the sound environment showed a significance with respect to target error and task time with $F(12,72)=22.01$, $p<0.0001$ and $F(12,72)=43.48$, $p<0.0001$. Table 4.17 makes a number of statistical comparisons with these two new types of trials. The target error and time means are graphed in Figures 4.13 and 4.14 respectively.

---

[8] These will be abbreviated as P and Q respectively.

df (1, 6)	S1	S2	Error F (Pr > F)	Time F (Pr > F)
No Sound vs Prehear Target	N	TP	**10.42** **0.0179**	**42.31** **0.0006**
	N	MP	5.06 0.0656	**152.81** **0.0001**
	N	OP	3.61 0.1061	**87.89** **0.0001**
Sound & Visual vs Prehear Target	TV	TP	**79.56** **0.0001**	0.24 0.6401
	MV	MP	4.99 0.0670	**12.60** **0.0121**
	OV	OP	3.49 0.1111	1.00 0..3569
Prehear Target Comparison	TP	MP	1.09 0.3366	**22.02** **0.0034**
	TP	OP	2.98 0.1352	4.63 0.0749
	MP	OP	0.20 0.6704	0.06 0.8196
No Sound vs Quiet Target	N	TQ	0.05 0.8351	3.24 0.1219
	N	MQ	5.01 0.0666	**206.11** **0.0001**
	N	OQ	5.24 0.0620	**23.47** **0.0029**
Sound & Visual vs Quiet Target	TV	TQ	**75.33** **0.0001**	**16.52** **0.0066**
	MV	MQ	**6.90** **0.0392**	**112.22** **0.0001**
	OV	OQ	1.11 0.3322	**44.83** **0.0005**
Quiet Target Comparison	TQ	MQ	2.50 0.1649	5.27 0.0614
	TQ	OQ	2.61 0.1576	3.41 0.1142
	MQ	OQ	0.00 0.9849	0.17 0.6911

Table 4.17 Comparing Prehear and Quiet Target Sound Environments Using Subjective Rating
One-Way ANOVA with Repeated Measures
Analysis of Variance of Contrast Variables
This table makes some comparisons between the prehear and quiet sound environments with other sound environments. As in other tables of this nature, a significant difference between a rating of the sound environment in column S1 and the sound environment in column S2 is shown in boldface. This simply shows that the difference was significant, compare with the graphs in Figure 4.13 and 4.14 to determine which environment was better.

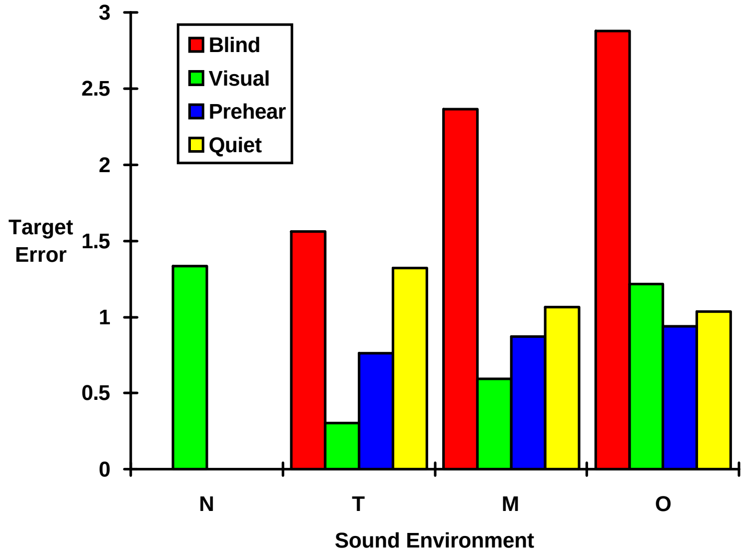


Figure 4.13 Prehear and Quiet Target Trials vs. Target Error

Although the prehear and quiet target environments generally had more target errors than their standard visual equivalent environment, they both showed an improvement over the no sound condition. The orchestral environment was a special case which actually improved under the prehear and quiet target environments over its standard visual case.

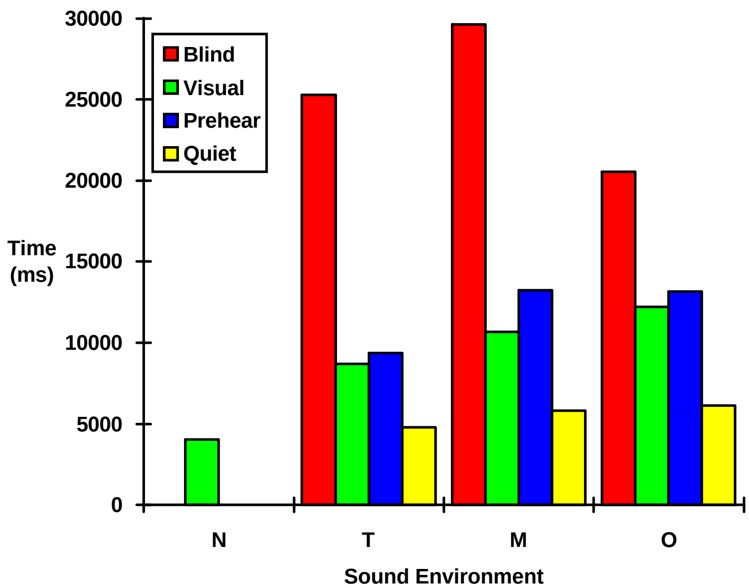


Figure 4.14 Prehear and Quiet Target Trials vs. Time

As with all other audible sound environments, the prehear and the quiet target environments required the subject more time to complete the task. It is interesting to note, however, that the quiet target environment required less time than the standard visual case. The tonal case in particular showed no significant difference in time to the no sound environment.

4.2.4.1 Prehear Trial Results

Target error showed a significant reduction using TP over N. MP and OP sound environments, however, made no significant improvement. N's time was significantly smaller than all of the prehear environments.

Comparing the prehear trials to the visual trials of the same sound type showed a significant increase in target error for T, but no significant difference with M nor O. In fact O had a reduction in target error. With respect to time, M showed a significant increase under the prehear condition compared to the standard visual condition. The other two environments T and O showed no significant difference in time.

Of the prehear trials (TP, MP, and OP), there was no significant difference between them with respect to error, however, TP's time was significantly smaller than that of MP.

4.2.4.2 Quiet Trial Results

Although the quiet target trials did not yield a significant difference in target error over N, there was reduction in errors using these sound environments. In particular MQ and OQ's error difference to N was approaching significance. N's time, however, was significantly smaller than both MQ and OQ.

Comparing the quiet target trials to the standard visual trials showed a significant increase in target errors for both TQ and MQ. OQ, however, had a decrease in errors, though this result was not significant. In all three cases the quiet trials took significantly less time to reach the target than the standard visual trials.

No significant difference was found in target error or time amongst TQ, MQ, and OQ.

Chapter 5 Discussion

5.1 Experiment 1 - Sound as a Position Cue

The Sound Cue experiment showed that subject traits had little effect on the recorded results. Two effects were, however, noted. First, an interaction was detected between subject 3D interface experience and sound environment with respect to subjective performance. This first effect can be virtually ignored since the difference in the responses per sound environment from those with different 3D backgrounds was insignificant.

The second subject trait that caused a significant difference in the data was due to gender. Females rated all the tested environments as more marketable than did males. Since this effect was uniform across all sound environments, no significant interaction between gender and sound environment was detected.

Other than these two cases, no additional measured result including: target error, time to complete the task, subjective performance, preference, usability and marketability, was affected by the subject's gender, area of study, musical experience, 3D interface experience, or the order in which the subjects tested each sound environment.

Comparing target error with time, it was originally thought that a correlation would be found. For example, as time to target increases, target error would decrease, and vice-versa. Surprisingly, no correlation was found to be significant. This is not due to outliers since they were removed and an average response was found for each subject. It suggests that subject response is not as predictable; some are quick yet make small errors, while others take longer to make large errors.

5.1.1 Single Cue Environments: Visual-Only and Sound-Only

The no sound environment (visual-only) was significantly faster than the blind (sound-only) environments. With respect to target error, however, the tonal and musical environments under blind condition caused no significant increase in target error over the no sound case. This suggests that either of these two sound-only environments could be used in applications where the user's visual attention is required elsewhere and still produce equivalent results to the visual-only environment.

5.1.2 Combination Cue Environments: Visual and Sound

Combination environments having both sound and visual cues reduced the target error significantly. Error reduction was measured at 78.9%, 66.1%, and 32.4% for the tonal, musical and orchestral sound environments respectively. The subjective performance, preference, usability, and marketability of all of these audible sound environments were better than the no sound environment. There was a consequence with this improved error rates and subjective appreciation, however, and that was found in task time. The combination sound with visual environments caused the task to take 123.4%, 178.4%, and 215.1% longer than with the no sound environment for tonal, musical, and orchestral respectively. The net result is that a sound enhanced visual environment reduces error and increases subjective appreciation but requires more time.

Comparing the three audible sound environments shows the tonal environment to be best for the single usage session that subjects did in the experiment. It had the lowest target error for both the visual and blind conditions and had the smallest time in the visual case. Subjectively it was found to have the highest subjective performance, easy to use and marketable. Preference was the only weak point as both the musical and orchestral environments were preferred, though this difference was not significant. Many commented about this, saying that they found the tonal environment to be very easy and helped the best, yet they thought it would likely become annoying. An annoying sound environment is not likely to be used no matter how much improvement it yields.

Following tonal, the next best audible sound environment was the musical environment. Although its measures were not as good as the tonal environment, it still had a reasonable target error and time with good subjective ratings. The loss in accuracy when compared to the tonal environment is likely due to the variability in the music pieces themselves. Song variation probably distracted the subjects somewhat thus causing the increase in task time. However, because of its ability to play any song, it may prove to be the sound environment of choice. After all, which environment is better? A really efficient annoying environment that will never be used, or a less efficient yet enjoyable sound environment that still improves user performance over the regular no sound environment? It may also improve as familiarity of a particular musical piece grows, but this hypothesis was not tested.

The orchestral environment proved to be the least effective sound environment, falling behind the others in target error, time and subjective ratings. Although it has the flexibility of playing any song of the user's choice, its added complexity seemed to outweigh those benefits. Many subjects commented on the difficulty using the environment and this resulted in long task time. The increase in error rates of this environment over the other two audible environments may have been due to its lack of depth granularity. The sound card provided sixteen output channels that allowed for eight different instrument sections, each having a foreground and a background channel. This limited the orchestra to be 2 by 4 in size, thus only allowing 4 different instrument sections to give the depth information. Perhaps this method would improve with the ability to improve the granularity of the z dimension by having more instrument sections.

5.2 Experiment 2 - Extended Period Trials

Learning effects are always of interest when designing an interface. Does the tonal sound environment still have the least target errors as time progresses or do the differences to the other environments become less apparent?

The results showed that by the third day all audible sound environments maintained consistent target error, except for tonal. Tonal's error level decreased, thus increasing the difference with the other audible environments. Another interesting result found that the no sound target error also decreased over the test period.

Although the other environments showed no major improvement in error, all environments including the no sound environment improved significantly with respect to task time. This indicates that as the subjects became more familiar with the task they were able to perform it equally well with respect to error yet taking less time.

The degree that the no sound environment was different from the combination environments, which used both sound and visual cues, changed in interesting ways over the trial period. On the first day the tonal, musical and orchestral environments showed a reduction in target error over the no sound environment by 78%, 73%, and 43% respectively. By the third day, all of these environments had improved but the reduction using the tonal, musical, and orchestral over the no sound environment became 77%, 55%, and 9% respectively. With respect to time, the tonal, musical and orchestral environments took longer than the no sound environment by 81%, 113% and 163% respectively on the first day and 116%, 164%, and 203% on the third day. This is showing the significant improvement in both target error and time of the no sound environment as discussed in 4.2.3.1.

Subjectively, all the audible environments maintained a higher subjective performance level over the no sound environment. Both the musical and orchestral environments showed either improvements or no significant change over the test period. Orchestral, for example, had the highest preference rating by the third day. Subjects were perhaps getting more accustomed to it and did not find it as complicated, but more enjoyable. Tonal and no sound both showed a decrease in preference as was expected. Over time, tonal sound became more annoying and less preferred.

5.2.1 Use of Sound Environments In Real Applications

In practice, most applications will not be able to play a target sound since the target sound location is not known. The prehear and quiet target environments, in which the target sound cannot be heard during the trial, were tested to validate the use of the auditory enhancements in a more realistic setting. As expected, target error levels increased in both the tonal and musical environments over the standard visual condition though it did not reach significance in the musical case. The orchestral environment, however, showed a surprising result in that the target error actually decreased when the target sound could not be heard during the trial. Subjects commented that they found it easy to locate themselves in depth since they knew which instrument section they were in.

The prehear target environment trials indicated that if a user becomes familiar enough with an environment and knows the sounds at different locations, that these locations can be located more accurately than with no sound. The quiet target environments show that even if subjects do not specifically know a target sound from previous experience, they will become accustomed to using both the sound and visual cues together and target errors will improve. The quiet target environments did not reduce the target error significantly compared to the no sound environment. The time required to complete the tasks was, however, significantly lower than all other audible sound environments.

Both of these results show that the sound environments are viable in standard 3D interface tasks where a user will not be able to hear their target. A modelling application, like Alias, would therefore likely benefit since users have a greater sense of the depth of their cursor.

5.3 Future Work

The emphasis of the Sound Cue experiment was to determine if sound could in fact help users with depth perception. The Extended Period experiment went on to see how prolonged use over three days would effect the performance of the three audible sound environments and investigated more realistic environments where no target sound is played. Only three sound mappings were, however, tested: tonal, musical, and orchestral. The results found here should not be generalized to all sound mappings since there are many other possible mappings available that may prove to give better or worse results than those tested here.

Future research in this area could therefore involve testing other possible mappings. With sound there are an unlimited number to choose from. Those chosen here were primarily music based since music has the advantage of being less annoying to users than arbitrary sounds as has been shown. Other mappings might include nature sounds like a waterfall, or birds singing. These could be replace the tonal sound with all the directional mappings remaining the same. Like the musical mappings, these "natural" sounds may be less annoying as well.

Another important use of sound in a computer interface is the effect it has with visually impaired users. The no visual (blind) sound environments tested here, where no visual aid was given, do not fairly test this area since the subjects were all sighted users. A real test would involve sight impaired subjects. They are known to have a greater dependence on their hearing ability so it wouldn't be surprising if they benefited more from a sound based application than sighted users. This would open up the doors for the visually impaired, since they would be able to perceive depth and perform 3D computer tasks from which they are currently excluded.

5.4 Conclusions

In conclusion, sound environments were found to improve user depth perception. Tonal sound offered the best results but was found to be the most annoying to use. Musical performed well and also supported the flexibility of playing a variety of songs chosen by the user. Orchestral was more complicated and less accurate but with experience, users preferred it over the other environments.

Even if the target sound was not played during the trial, experienced sound environment users were able to perform better than if no sound was presented. Orchestral sound reduced the target error under these conditions, and users found they could locate themselves very easily in one of the instrument sections.

When designing an application, the main question then becomes: how tolerant will the users be with a tonal sound environment? Certainly, tonal offers the best performance, but its degree of annoyance could cause it to be abandoned quickly. If this is a problem, the musical environment would be next best choice. It offers reasonable improvements over the no sound environment's error levels, it has high subjective ratings indicating its general user approval, and it has great flexibility since it is user configurable.

The results found in this study have shown that depth can be perceived through the use of sound cues. Often they showed that performance was improved over the no sound environment. These improvements, however, were done using an interface in which the visual depth cues were minimized. In a real application, visual cues will certainly be more helpful and thus the added benefit of sound would perhaps be less significant. For low vision users or users whose focus is away from the screen, however, a sound enhanced 3D interface would definitely be an asset.

References

Advanced Gravis Computer Technology Ltd. (1994). *UltraSound User's Guide*, Burnaby, B.C.: Advanced Gravis Computer Technology Ltd., 97-99.

Auerbach, C., & Sperling, P. (1974). A common auditory-visual space: Evidence for its reality, *Perception & Psychophysics*, 16 (1), 129-135.

Blattner, M.M, Papp, A.L.III, Glinert, E.P. (1994). Sonic Enhancement of Two-Dimensional Graphics Displays, *Auditory Display, Sonification, Audification and Auditory Interfaces*, Reading, MA: Addison-Wesley Publishing Co., 447-470.

Botte, M.C., Baruch, C., & Mönikheim, S. (1992). Memory for Loudness: the Role of Loudness Contour. *Advances in the Biosciences: Auditory Physiology and Perception*. Pergamon Press: Oxford, Vol 83, 305-311.

Bregman, A. S. & Steiger, H. (1980). Auditory streaming and vertical localization: Interdependence of "what" and "where" decisions in audition. *Perception & Psychophysics*, 28 (6), 539-546.

Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Brewster, S.A, Wright, P.C, & Edwards, A.D.N. (1994a). A Detailed Investigation into the Effectiveness of Earcons, *Auditory Display, Sonification, Audification and Auditory Interfaces*, Reading, MA: Addison-Wesley Publishing Co., 471-498.

Brewster, S.A, Wright, P.C, & Edwards, A.D.N. (1994b). The Design and Evaluation of an Auditory-Enhanced ScrollBar. *Proceedings of CHI'94*, New York: ACM, 173-179.

Brooks, A.F., Ouh-Young, M., Batter & J., Kilpatrick, P. (1990). Project GROPE - Haptic Displays for Scientific Visualization, *Computer Graphics*, 24 (4), 177-185.

Brown, M.L, Newsome, S.L., & Glinert E.P. (1989). An Experiment into the Use of Auditory Cues to Reduce Visual Workload. *Proceedings of CHI'89*, New York: ACM, 339-346.

Buxton, W., Bly, S., Frysinger, S., Lunney, D., Mansur, D., Mezrich, J. & Morrison, R. (1985). Communicating with sound. *Proceedings of CHI'85*, New York: ACM, 115-119.

Clarke, E.F. (1989). The perception of expressive timing in music. *Psychological Research*, 51, 2-9.

Cody, R.P & Smith, J. (1991). *Applied Statistics and the SAS Programming Language, 3rd Ed.*, Amsterdam: Elsevier Science Publishing Co.

Colivata, F.B. (1974). Human sensory dominance. *Perception and Psychophysics*. 16 (2), 409-412.

Coren, S., Ward, L. (1989). *Sensation & Perception*, 273-294. Orlando: Harcourt Brace Jovanovich, Publishers.

Curso, G.M. (1980). Auditory temporal order and perceived fusion-nonfusion. *Perception & Phychophysics*, 28 (5), 465-470.

Das, H., Zak, H., Kim, W.S., Bejczy & A.K., Schenker, P.S. (1992). Operator Performance with Alternative Manual Control Modes in Teleoperation, *Presence*, 1 (2), 201-217.

Deutsch, D. (1986). Auditory Pattern Recognition, *Handbook of Perception and Human Performance*. Toronto: Wiley. Vol 2, chap 32.

Di Giano, C.J. & Baecker R.M. (1992). Program Auralization: Sound Enhancements to the Programming Environment. *Proceedings of Graphics Interface '92*. Toronto: CIPS. 44-52.

Edwards A.N. (1988). The Design of Auditory Interfaces for Visually Disabled Users. *Proceedings of CHI*, New York: ACM, 83-88.

Fitch, W.T. & Kramer, G. (1994). Sonifying the Body Electric: Superiority of an Auditory over a Visual Display in a Complex, Multivariate System, *Auditory Display, Sonification, Audification and Auditory Interfaces*, Reading, MA: Addison-Wesley Publishing Co., 307-326

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.

Gates, A. & Bradshaw, J. (1974). Effects of auditory feedback on musical performance task. *Perception & Psychophysics*. 16 (1), 105-109.

Gaver, W.W, Smith, R.B, & O'Shea, T. (1991). Effective Sounds in Complex Systems: The Arkola Simulation. *Proceedings of CHI*, New York: ACM, 85-90.

Gaver, W.W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*. 2, 167-177.

Gaver, W.W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction*. 4 (1), 67-94.

Gerth, J.M. (1992). Systematic Errors Occur in the Discrimination of Complex Audio Signals. *Proceedings of the Human Factors Society*, 36, 263-267.

Good, M.D., & Gilkey, R.H. (1992). Masking Between Spatially Separated Sounds. *Proceedings of the Human Factors Society*, 36, 253-257.

Hayward, C. (1994). Listening to the Earth Sing, *Auditory Display, Sonification, Audification and Auditory Interfaces*, Reading, MA: Addison-Wesley Publishing Co., 369-416.

Howell, D.C. (1992). *Statistical Methods for Psychology, 3rd Ed.* Belmont, CA: Duxbury Press.

Jaccard, J. & Becker, M. (1983). *Statistics for the Behavioral Sciences*, 2nd Ed. Belmont, CA: Wadsworth Publishing Company.

Jackson, J.A. & Francioni J.M. (1994). Synchronization of Visual and Aural Prallel Program Performance Data, *Auditory Display, Sonification, Audification and Auditory Interfaces,* Reading, MA: Addison-Wesley Publishing Co., 291-306.

Jaubert, F. (1995). *A Study of Delays and De-Synchronization in a Multiple-View Direct Manipulation Task,* Master's Thesis: University of Waterloo.

Jones, S.D. & Furner, S.M. (1989). The construction of audio icons and information cues for human computer dialogues. *Contemporary Ergonomics, Proc. of the Ergonomics Society's 1989 Annual Conference*, Reading, England: Taylor & Francis. 436-441.

Kahneman, D. (1934). *Attention and Effort*, Englewood Cliffs, NJ: Prentice Hall.

Kramer, G. (1994). An Introduction to Auditory Display, *Auditory Display, Sonification, Audification and Auditory Interfaces*, Reading, MA: Addison-Wesley Publishing Co., 1-78.

Mackenzie, I. S., Buxton, W. (1992). Extending Fitt's Law To Two-Dimensional Tasks, *Proceedings of CHI '92 Conference on Human Factors in Computing Systems,* New York: ACM, 219-226.

Mansur, D.L, Blattner M.M, Joy, K.I. (1985). Sound-Graphs: A Numerical Data Analysis Method for The Blind. *Proc. of the 18th Annual Hawaii International Conference on System Sciences*. 198-203.

Martin, D.W. (1977). *Doing Psychology Experiments*, Brooks/Cole Publishing Company: California, 51-85.

Matlin, M.W. (1994). *Cognition*, 3rd Ed., Orlando, FL: Harcourt Brace Jovanovich Publishers. 43-55.

Mereu, R.F. & Kovach, R.J. (1970). A Portable Seismic System for Crustal Studies, *Bulletin of the Seismic Society of America,* 60 (5), 1607-1613.

Pishkin, V., & Blanchard, R. J. (1964). Auditory concept identification as a function of subject sex and stimulus dimensions. *Psychon. Sci.,* 1, 177-178.

Pishkin, V., & Rosenbluh, E. S. (1966). Concept identification of auditory dimensions as a function of age and sex. *Psychon. Sci.* Vol 4, 165-166.

Prior, M. & Troup, G.A. (1988). Processing of Timbre and Rhythm in Musicians and Non-Musicians. *Cortex*, 24, 451-456.

Rosen, A., Howell, P. (1991). *Signals and Systems for Speech and Hearing*. 236-257.

Schulze, H.H. (1989). Categorical perception of rhythmic patterns. *Psychological Research*, 51, 10-15

Treisman, A.M. & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psycology,* 12, 97-136.

Treisman, A.M. & Gormican, S. (1988) Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review,* 95, 15-48.

von Bekesy, G. (1957). The Ear. *Scientific American*, 197 (2), 66-78.

Ware, C. & Slipp, L. (1991). Using Velocity Control To Navigate 3D Graphical Environments: A Comparison of Three Interfaces, *Proc. of the Human Factors Society*, 300-304.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Merrill.

Yeung, E.S. (1980). Pattern recognition by audio representation of multivariate analytical data. *Analytical Chemistry*. 52 (7), 1120-1123.

Zhai, S., Buxton, W., Milgram, P. (1994). The "Silk Cursor": Investigating Transparency for 3D Target Acquisition. *Proceedings of CHI '94 Conference on Human Factors in Computing Systems,* New York: ACM, 459-464.

Appendix A Pretest Music Piece Ranges

Musical Piece	Time	Key	Low Note	High Note
Auld Lang Syne	4/4	C	G4	A5
My Bonnie	3/4	F	A4	A5
Frere Jacques	4/4	C	G4	A5
Jesus Loves Me	4/4	A	A4	A5
He's Got The Whole World	4/4	C	B4	A5
Yankee Doodle	4/4	E	B4	A5
Brahm's Lullaby	3/4	A	B4	A5
Eine kleine Nachtmusik	4/4	D	A4	A5

Appendix B Pretest Instruments and Ranges

Range	Notes
Very Low	C2 - C4
Low	C3 - C5
Mid	C4 - C6
High	C5 - C7
Very High	C6 - C8

	Instrument	Classification	Range
1	Accordian	Wind	Mid
2	Banjo	String	Mid
3	Bass	String	Very Low
4	Bassoon	Wind	Mid
5	Bells	Percussion	Mid
6	Calliope	Wind	Mid
7	Cello	String	Low
8	Clarinet	Wind	High
9	Dulcimer	Percussion	Mid
10	English Horn	Wind	Low
11	Guitar (Distorted)	String	Mid
12	Harp	String	Mid
13	Harpsichord	String	Mid
14	Honky Tonk Piano	Percussion	Mid
15	Koto	Percussion	Mid
16	Marimba	Percussion	Mid
17	Music Box	Percussion	Mid
18	Oboe	Wind	Mid
19	Organ	Wind	Mid
20	Piano	Percussion	Mid
21	Piccolo	Wind	High
22	Pizzicato Strings	String	Mid
23	Shakuachi	Wind	High
24	Sitar	String	Mid
25	Steel Drums	Percussion	Mid
26	String Ensemble	String	Mid
27	Synth Voice	Wind	Mid
28	Timpani	Percussion	Very Low
29	Trombone	Wind	Low
30	Trumpet	Wind	Mid
31	Tuba	Wind	Very Low
32	Violin	String	High

Appendix C Sound Mapping Pretest Data

Tonal X

Subjective Data

OBS	B	N	O	P	V
1	3	2	2	3	3
2	4	1	4	4	3
3	4	1	1	2	3
4	4	1	2	4	4
5	4	-	1	4	4
6	3	-	1	3	3
7	2	-	1	4	3
8	2	-	1	2	2
9	3	-	1	3	4
10	2	-	-	3	3
11	2	-	-	2	2

Variable	N	Mean	Std Dev	Variance
B	11	3.000	0.894	0.800
N	4	1.250	0.500	0.250
O	9	1.556	1.014	1.028
P	11	3.091	0.831	0.691
V	11	3.091	0.701	0.491

Target Error Data

OBS	B	P	V
1	1.847	0.424	1.621
2	1.307	1.609	1.896
3	1.638	1.954	2.246
4	1.295	1.547	1.310
5	1.372	1.111	2.421
6	4.075	3.570	3.947
7	3.297	5.513	3.852
8	4.104	4.831	4.275
9	2.081	1.220	1.166
10	1.194	2.427	2.099
11	2.127	3.541	2.675

Variable	N	Mean	Std Dev	Variance
B	11	2.212	1.101	1.212
P	11	2.522	1.631	2.662
V	11	2.501	1.080	1.167

Time Data

OBS	B	P	V
1	4375.80	6978.83	3591.67
2	6969.71	9440.38	11282.83
3	9629.00	7040.86	10078.75
4	8590.00	9589.17	6683.88
5	4014.00	3601.67	3051.43
6	11392.25	8733.29	9116.86
7	2687.60	2796.00	3802.20
8	2888.83	2089.33	1745.33
9	19299.50	13747.17	13698.00
10	16241.00	13623.00	13456.60
11	11444.00	12340.83	10933.00

Variable	N	Mean	Std Dev	Variance
B	11	8866.518	5460.163	29813376.490
P	11	8180.047	4142.437	17159781.331
V	11	7949.140	4353.821	18955761.038

Tonal Y

Subjective Data

OBS	B	N	O	P	V
1	3	3	2	2	4
2	3	1	2	2	4
3	4	1	1	2	3
4	3	1	2	4	4
5	2	-	1	1	2
6	1	-	1	2	3
7	2	-	1	2	3
8	1	-	1	2	2
9	3	-	1	2	3
10	3	-	-	3	3
11	2	-	-	2	2

Variable	N	Mean	Std Dev	Variance
B	11	2.455	0.934	0.873
N	4	1.500	1.000	1.000
O	9	1.333	0.500	0.250
P	11	2.182	0.751	0.564
V	11	3.000	0.775	0.600

Target Error Data

OBS	B	P	V
1	4.398	1.105	0.707
2	1.816	1.064	2.119
3	1.718	1.734	2.258
4	2.235	2.611	2.547
5	1.987	1.737	1.375
6	3.482	3.950	3.510
7	4.267	3.382	3.567
8	4.350	4.372	3.569
9	2.100	1.642	1.165
10	1.736	1.499	1.575
11	3.040	4.035	3.063

Variable	N	Mean	Std Dev	Variance
B	11	2.830	1.112	1.236
P	11	2.466	1.251	1.566
V	11	2.314	1.028	1.056

Time Data

OBS	B	P	V
1	2425.75	5042.60	3508.50
2	9220.83	6166.33	7422.43
3	7451.00	9517.71	13451.57
4	7655.14	5588.29	8716.75
5	7470.86	4665.14	7401.43
6	12981.86	7902.00	11992.29
7	8312.00	9162.20	5958.83
8	2188.25	4575.60	4744.67
9	28861.33	13022.33	19993.40
10	14585.50	10225.60	12957.20
11	14413.17	16064.00	21206.40

Variable	N	Mean	Std Dev	Variance
B	11	10505.972	7365.932	54256959.064
P	11	8357.437	3718.014	13823631.444
V	11	10668.497	5884.688	34629555.626

Tonal Z

Subjective Data

OBS	B	N	O	P	V
1	3	2	1	3	3
2	3	1	3	4	4
3	4	1	1	2	3
4	3	1	2	4	4
5	2	-	1	2	2
6	4	-	1	3	3
7	3	-	1	3	4
8	3	-	1	3	2
9	3	-	2	3	4
10	3	-	-	3	3
11	3	-	-	3	3

Variable	N	Mean	Std Dev	Variance
B	11	3.091	0.539	0.291
N	4	1.250	0.500	0.250
O	9	1.444	0.726	0.528
P	11	3.000	0.632	0.400
V	11	3.182	0.751	0.564

Target Error Data

OBS	B	P	V
1	0.663	0.394	0.404
2	0.241	0.197	0.245
3	0.154	0.217	0.197
4	0.319	0.257	0.216
5	0.448	0.517	0.319
6	0.205	0.255	0.154
7	1.921	1.545	3.509
8	1.736	1.793	0.966
9	0.153	0.339	0.305
10	0.216	0.300	0.184
11	1.858	2.334	1.988

Variable	N	Mean	Std Dev	Variance
B	11	0.719	0.735	0.540
P	11	0.741	0.765	0.586
V	11	0.772	1.059	1.121

Time Data

OBS	B	P	V
1	4466.50	6986.60	5339.20
2	8997.86	10388.50	14015.29
3	4856.67	10242.83	7485.71
4	10090.29	8977.00	9117.43
5	6841.33	7849.38	6902.33
6	11128.14	11054.14	11009.75
7	6223.00	10770.80	9141.33
8	8236.17	7431.83	7645.80
9	11980.00	7915.60	10698.60
10	13519.00	10817.67	15764.50
11	13237.83	14875.20	9757.25

Variable	N	Mean	Std Dev	Variance
B	11	9052.435	3219.057	10362331.016
P	11	9755.414	2266.845	5138585.924
V	11	9716.109	3077.836	9473073.893

Musical X

Subjective Data

OBS	V	B	T	K	O	P	N
1	4	4	3	3	2	3	1
2	3	3	2	2	2	3	2
3	4	3	3	3	1	2	2
4	1	3	2	1	2	2	1
5	4	4	1	2	1	2	1
6	4	3	4	3	2	2	3
7	3	3	2	2	1	2	1
8	3	4	1	2	1	2	-

Variable	N	Mean	Std Dev	Variance
V	8	3.250	1.035	1.071
B	8	3.375	0.518	0.268
T	8	2.250	1.035	1.071
K	8	2.250	0.707	0.500
O	8	1.500	0.535	0.286
P	8	2.250	0.463	0.214
N	7	1.571	0.787	0.619

Target Error Data

OBS	B	K	P	T	O	V
1	1.010	1.716	2.034	1.453	2.090	0.641
2	3.145	4.122	2.853	4.056	2.279	3.617
3	2.773	2.448	3.737	2.626	2.780	2.756
4	4.948	4.443	5.147	4.100	4.496	4.241
5	4.765	3.911	4.510	3.906	2.915	4.221
6	3.820	5.052	3.909	3.620	3.927	4.463
7	4.650	3.573	4.336	4.705	4.767	5.064
8	3.024	6.135	7.392	8.151	5.365	6.475

Variable	N	Mean	Std Dev	Variance
B	8	3.517	1.320	1.743
K	8	3.925	1.396	1.950
P	8	4.240	1.603	2.570
T	8	4.077	1.935	3.746
O	8	3.577	1.228	1.507
V	8	3.935	1.712	2.932

Time Data

OBS	B	K	P	T	O	V
1	6936.71	12049.29	6491.88	8047.13	8798.13	8482.00
2	21173.29	42281.29	22819.75	36896.25	47122.86	15947.67
3	13985.14	24025.71	24217.50	17528.00	22903.13	21997.00
4	4190.40	6029.60	11861.40	3257.00	8278.83	2991.40
5	4569.17	8305.80	4006.33	7172.40	7209.60	4505.20
6	2622.50	2708.40	1626.00	1478.00	1544.80	1109.00
7	2268.67	2208.40	1871.83	2883.00	1958.00	3383.00
8	1259.00	4483.33	4318.83	4783.80	3194.83	5876.40

Variable	N	Mean	Std Dev	Variance
B	8	7125.610	6947.373	48265991.729
K	8	12761.477	13861.252	192134294.21
P	8	9651.691	9147.015	83667881.228
T	8	10255.697	11876.255	141045429.70
O	8	12626.272	15510.763	240583777.98
V	8	8036.458	7269.457	52845008.285

Musical Y

Subjective Data

OBS	V	B	T	K	O	P	N
1	3	3	3	3	2	3	1
2	2	2	3	2	3	4	1
3	3	3	3	2	1	2	2
4	1	3	1	1	1	1	1
5	4	4	1	2	1	2	1
6	4	3	4	2	2	1	2
7	2	3	2	3	1	3	1
8	4	2	3	3	1	2	-
9	1	1	1	1	-	1	-

Variable	N	Mean	Std Dev	Variance
V	9	2.667	1.225	1.500
B	9	2.667	0.866	0.750
T	9	2.333	1.118	1.250
K	9	2.111	0.782	0.611
O	8	1.500	0.756	0.571
P	9	2.111	1.054	1.111
N	7	1.286	0.488	0.238

Target Error Data

OBS	B	K	P	T	V
1	1.606	1.118	1.879	2.097	1.272
2	2.301	4.861	3.456	3.855	3.206
3	2.674	4.008	2.582	3.754	2.979
4	2.838	2.209	2.760	2.120	2.971
5	4.312	3.443	4.086	3.046	4.024
6	3.654	3.620	3.159	4.048	4.316
7	5.462	4.387	4.946	4.787	4.323
8	4.465	3.769	4.116	4.281	4.174
9	3.797	4.166	4.799	3.524	4.255

Variable	N	Mean	Std Dev	Variance
B	9	3.457	1.208	1.460
K	9	3.509	1.160	1.346
P	9	3.531	1.039	1.079
T	9	3.501	0.924	0.855
V	9	3.502	1.015	1.030

Time Data

OBS	B	K	P	T	V
1	6265.71	17340.43	8957.86	8826.71	8754.71
2	23391.25	14035.71	18088.29	16052.33	26517.50
3	5056.00	7583.57	8557.33	11959.38	6915.57
4	42083.25	44672.38	29168.88	28742.14	23260.29
5	11418.50	11538.67	11067.40	16615.80	12552.17
6	10475.67	37065.00	12717.60	9242.50	6761.00
7	1366.60	1530.67	1455.33	3191.00	1291.50
8	3118.00	3337.00	4593.00	2947.50	2646.67
9	2793.60	1370.80	1468.00	876.80	959.60

Variable	N	Mean	Std Dev	Variance
B	9	11774.287	13204.089	174347976.85
K	9	15386.025	15582.716	242821027.59
P	9	10674.854	8771.412	76937667.545
T	9	10939.352	8723.880	76106075.850
V	9	9962.112	9276.028	86044687.352

Musical Z

Subjective Data

OBS	V	B	T	K	O	P	N
1	4	3	3	1	2	4	1
2	4	3	3	2	1	2	3
3	3	3	2	1	1	1	1
4	4	4	1	2	1	2	1
5	4	3	4	3	3	3	3
6	4	4	3	2	1	4	1
7	4	2	3	2	1	2	-
8	4	-	1	-	-	3	-

Variable	N	Mean	Std Dev	Variance
V	8	3.875	0.354	0.125
B	7	3.143	0.690	0.476
T	8	2.500	1.069	1.143
K	7	1.857	0.690	0.476
O	7	1.429	0.787	0.619
P	8	2.625	1.061	1.125
N	6	1.667	1.033	1.067

Target Error Data

OBS	B	K	P	T	O	V
1	0.485	0.852	3.487	0.462	1.673	0.715
2	0.404	3.526	2.736	1.904	1.763	0.706
3	0.806	1.484	2.843	1.270	0.947	0.765
4	4.731	2.207	3.423	2.627	5.681	3.310
5	3.425	3.085	4.563	3.400	2.801	1.602
6	2.888	1.767	4.069	2.754	3.250	3.320
7	3.739	7.130	5.563	6.463	2.795	4.244

Variable	N	Mean	Std Dev	Variance
B	7	2.354	1.765	3.115
K	7	2.864	2.093	4.379
P	7	3.812	1.004	1.008
T	7	2.697	1.930	3.724
O	7	2.701	1.539	2.369
V	7	2.095	1.497	2.240

Time Data

OBS	B	K	P	T	O	V
1	5390.14	15179.13	7888.83	8116.17	9982.50	5058.50
2	10052.83	10936.86	8208.17	17166.75	14117.38	10159.71
3	19224.88	16172.14	14602.00	15250.00	14272.14	12812.57
4	5716.80	29146.50	7619.40	7043.67	8991.60	6586.00
5	9361.50	11527.00	3611.00	8291.40	12186.00	6271.20
6	6530.67	7431.40	4314.40	8540.50	6890.50	7254.33
7	3392.80	8473.00	2670.80	3490.33	12153.17	4714.83

Variable	N	Mean	Std Dev	Variance
B	7	8524.231	5252.801	27591913.901
K	7	14123.718	7355.391	54101778.245
P	7	6987.800	4038.615	16310408.514
T	7	9699.831	4795.971	23001340.967
O	7	11227.612	2731.167	7459275.480
V	7	7551.022	2926.973	8567173.662

Orchestral Y

Subjective Data

OBS	K	O	P	T
1	4	3	3	2
2	1	1	1	1
3	3	3	3	3
4	2	1	1	1
5	2	1	3	3
6	2	2	3	2
7	2	2	2	2
8	4	2	2	3

Variable	N	Mean	Std Dev	Variance
K	8	2.500	1.069	1.143
O	8	1.875	0.835	0.696
P	8	2.250	0.886	0.786
T	8	2.125	0.835	0.696

Target Error Data

OBS	K	O	P	T
1	1.342	0.680	0.423	1.431
2	2.361	3.179	2.611	2.081
3	0.841	0.771	0.867	0.712
4	0.857	0.866	0.488	1.328
5	0.646	1.067	0.663	1.170
6	2.288	1.272	2.513	2.037
7	1.045	0.618	1.099	1.361
8	4.380	3.648	4.270	4.325

Variable	N	Mean	Std Dev	Variance
K	8	1.720	1.258	1.583
O	8	1.513	1.199	1.436
P	8	1.617	1.377	1.896
T	8	1.806	1.111	1.235

Time Data

OBS	K	O	P	T
1	8532.43	8481.00	9876.43	7301.88
2	17123.75	28995.71	22648.00	22639.43
3	19840.00	14742.00	10688.57	12723.86
4	20417.57	18704.43	14282.43	14740.00
5	27378.33	12370.83	25458.17	15133.60
6	60194.17	21468.00	34268.17	25535.67
7	13812.33	10890.00	9248.17	9643.83
8	10568.50	493.83	50.80	169.60

Variable	N	Mean	Std Dev	Variance
K	8	22233.385	16463.488	271046450.89
O	8	14518.226	8671.615	75196915.028
P	8	15815.091	10926.566	119389843.15
T	8	13485.983	8143.182	66311408.459

Appendix D   Experiment Forms

University of Waterloo

March 1, 1995

Dear Participant:

I am a graduate student in the Department of Computer Science at the University of Waterloo. Under the supervision of Professor Rick Kazman, I am conducting research on the use of sound in 3D computer interfaces to improve depth perception. Lack of depth perception is generally one of the greatest difficulties that users experience with 3D interfaces. The results of this study should help determine whether sound can be used to improve the depth cues that interfaces provide.

I appreciate your participation in the study and your answers to the brief questionnaire at the end. The experiment consists of two stages: an ear test followed by several 3D interface tasks. Ear tests take approximately two minutes and simply map out your hearing ability over a range of frequencies. After this, you will be given several 3D tasks to perform with the help of depth perception sound cues. Both your speed and accuracy will be measured to see which sound cues improve your depth perception the most. During both of these stages, you will have full control of the volume of the sound, so there will be no risk of ear damage. The experiment is not fatiguing and should be enjoyable.

The experiment should take about an hour, and you will be paid $10. Although results from the entire experiment are desired you may end your participation at any time. When you are ready to begin, a numerical code will be given to you that you will use throughout the experiment. This code is used only for data identification between the various experiment components and is never associated with your name. Since I am interested only in the general results of several participants and not those of individuals, your name will not appear in any reports produced. Once finished, the numerically identified data collected will be stored for later analysis.

This project has been reviewed and approved for ethics through the Office of Human Research & Animal Care at the University of Waterloo. If you have any questions or concerns resulting from your participation in this study, please contact this office at 885-1211 Ext. 6005.

Thank you in advance for your assistance with this project. If you have any questions about this project, please feel free to ask before beginning or contact me later at 725-1865 or swmereu@watcgl.

Sincerely,
Stephen Mereu

Ear Test Instructions

The purpose of the ear test is to determine your hearing ability over the musical range A0 (27.5 Hz) to A7 (3250 Hz). Results from this test will only be used to determine the reliability of the results for the remaining sound experiments.

The experiment begins by playing a tone at 27.5 Hz and slowly increases its frequency until it reaches 3250 Hz. As the tone is increasing in frequency, the volume level will be changing based on your input. When you hold down the left mouse button, the volume will increase, and when you release it, the volume will decrease.

Your goal is to increase the volume with the mouse button until you can just hear the tone. Once heard, release the mouse button to decrease the volume until the tone is no longer heard. Continue in this manner over the entire frequency range at which point the experiment will end. The results will determine your threshold of hearing over the frequency range played.

Since the tone being played will always be at the lower limit of your hearing, try not to be distracted by any hiss or noise that may be present, especially at low frequencies.

3D Interface Task

The purpose of this experiment is to determine whether sound can improve your ability to perceive depth into the screen. To measure this, you are to perform a specific 3D task aided by a limited number of visual depth cues, as well as the additional sound cues.

When you begin, you will see a randomly rotated 3D "blobby" shape. On the surface of this shape is a randomly placed spot that will always be on the visible side of the shape. Your task is to move the cursor in three dimensions to the spot and indicate when you've reached it.

The mouse moves the cursor left/right and up/down on the screen. To increase and decrease the cursor's depth in the 3D environment, use the up and down arrow keys respectively. The cursor will change size according to its depth but this size is only a clue that the depth is changing. It should not be regarded as an accurate depth measure.

Upon arriving at the target spot, press the spacebar to indicate that the task is finished. Speed and accuracy are equally important and are both recorded.

During the experiment, the following sound environments will be used: tonal, musical, and orchestral. Tonal sound is simply a constant tone that will vary as you move the cursor around the 3D world. Musical sound also varies under cursor movement, but rather than using a tone it will play a randomly selected musical piece. Then finally, for orchestral sound, an orchestra layout will be used which allows the music to change depending on which section of the orchestra your cursor is currently positioned.

To perceive the 3D positioning, the cursor's three coordinates (x,y,z) are mapped to three of the sound's dimensions. Some of the possible sound dimensions include: volume, balance, note oscillation, pitch bending of notes, and which instrument is playing the loudest. As you move the cursor around in the 3D environment, the sound will change according to the position and the current sound mapping.

The position of the target can also be heard by holding down the left mouse button. This allows you to compare the difference in the sound feedback to determine your distance to the target. Releasing the mouse button reverts back to playing the sound for the cursor position.

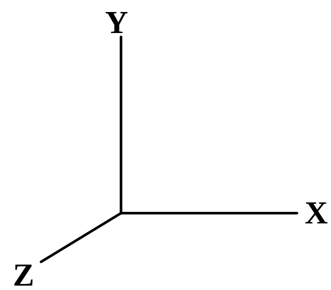
There are four test blocks: no sound, tonal, musical and orchestral. Each block consists of 25 trials where the first 5 are learning trails, followed by 15 visual experiment trials, followed by 5 blind experiment trials. The data recorded during the learn trials is not examined to allow time to become familiar with the sound mappings. The remaining trials are all examined with the blind trials being identical to the previous visual ones except that no visual aid is provided. The order that you will do the blocks in will be made known to you prior to commencing the experiment. At any time during the experiment you are free to discontinue with no obligation to finish.

No Sound

This sound environment plays no sound to help with depth perception.

The current sound mappings are

X	⇨	**No Sound**
Y	⇨	**No Sound**
Z	⇨	**No Sound**

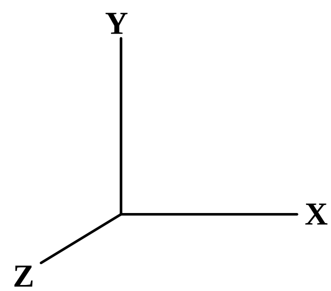


Tonal Sound

This sound environment plays a single constant tone for the duration of the experiment. As described earlier the sound will either indicate the position of the cursor or, while the left mouse button is held down, the target's position. To relate the sound to the position, three of its sound dimensions have been mapped to the position's (x,y,z) coordinates. The sound will then be modified as the position is moved.

The current sound mappings are

X ⇨ **Balance**
Y ⇨ **Pitch Bend**
Z ⇨ **Volume**

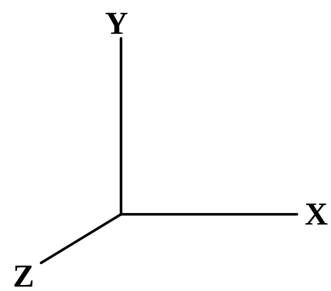


Musical Sound

The musical sound environment plays a musical piece for the duration of the experiment. As described earlier the sound will either indicate the position of the cursor or, while the left mouse button is held down, the target's position. To relate the sound to the position, three of its sound dimensions have been mapped to the position's (x,y,z) coordinates. The sound will then be modified as the position is moved.

The current sound mappings are

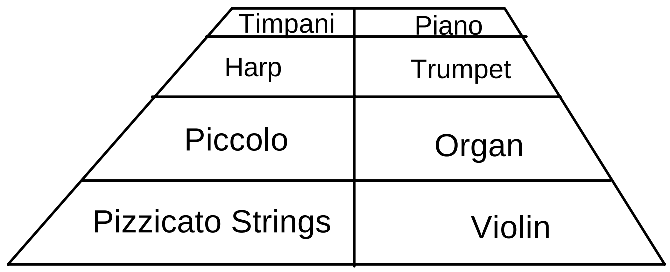
X ⇨ **Balance**
Y ⇨ **Pitch Bend**
Z ⇨ **Volume**



Orchestral Sound

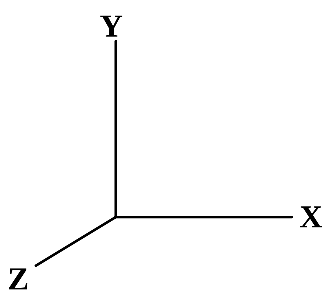
The orchestral sound environment uses a orchestra to help you perceive depth. As described earlier the sound will either indicate the position of the cursor or, while the left mous button is held down, the target's position. To relate the sound to the position, the sound dimensions have been mapped to the position's (x,y,z) coordinates. The sound will then be modified as the position is moved.

The orchestra has been assembled in the horizontal xz-plane. On this plane there are eight instrument sections as shown below. As you move your cursor around through these sections you will hear the various instruments as they play.



The current sound mappings are

XZ ⇨ **Instrument playing the melody the loudest**
Y ⇨ **Oscillation**



Department of Computer Science

Faculty of Mathematics

University of Waterloo

Consent to Participate

This study on the use of sound to improve depth perception in 3D graphical environments is being conducted by Stephen Mereu of the Department of Computer Science at the University of Waterloo. I have read the information letter describing the purposes and the tasks involved in participation in the study and understand that should the information that I provide be used in publications my identity will be protected. I acknowledge that I may withdraw my consent to participate at any time.

This study has been reviewed by, and has received ethics clearance through, the Office of Human Research at the University of Waterloo. This Office will receive any complaints or concerns with regard to my involvement in this study.

Participant's Name: (please print) ____

Participant's Signature: ____

Witness's Signature: ____

Date: ____/____/____
D M Y

Department of Computer Science

Faculty of Mathematics

University of Waterloo

Please answer the following questions about yourself:

1) Numerical code used for identification during the experiment: ____________________

2) Age: ____________

3) Gender: M / F

4) Area of Study: ________________________

5) Musical experience (specify instruments played, number of years, courses taken, or other related musical experience):
____
____
____
____
____
____
____

6) 3D Interface experience (specify the amount of experience with other 3D applications):
____
____
____
____
____

Please answer the following questions about the sound environments that you experienced in the experiment. 1 - very poor, 2 - poor, 3 - good, 4 - very good

Participant Preference How did you like the:	**1**	**2**	**3**	**4**
no sound environment?	☐	☐	☐	☐
tonal sound environment?	☐	☐	☐	☐
musical sound environment?	☐	☐	☐	☐
orchestral sound environment?	☐	☐	☐	☐

Sound Environment Performance
How was the improvement to your depth perception using the:

no sound environment?	☐	☐	☐	☐
tonal sound environment?	☐	☐	☐	☐
musical sound environment?	☐	☐	☐	☐
orchestral sound environment?	☐	☐	☐	☐

Sound Environment Usability
How easy to use and understand was the:

no sound environment?	☐	☐	☐	☐
tonal sound environment?	☐	☐	☐	☐
musical sound environment?	☐	☐	☐	☐
orchestral sound environment?	☐	☐	☐	☐

Please answer the following question about the sound environments.
1 - never, 2 - not likely, 3 - likely, 4 - absolutely

Sound Environment Marketability If you had a 3D application that used sound to improve depth perception, how often would you use the sound if the sound environment was the:	**1**	**2**	**3**	**4**
no sound environment?	☐	☐	☐	☐
tonal sound environment?	☐	☐	☐	☐
musical sound environment?	☐	☐	☐	☐
orchestral sound environment?	☐	☐	☐	☐

Other Comments About the Experiment

____
____
____

Appendix E Sample SAS Program

```
/* Experiment 1 - Target Error

   Twenty subjects ran through each of the blocks in one of four
   random orders:  N T M O
                   T N O M
                   O M T N
                   M O N T

      Blocks:
              N - No Sound
              TB - Tonal Sound, Blind
              TV - Tonal Sound, Visual
              MB - Musical Sound, Blind
              MV - Musical Sound, Visual
              OB - Orchestral Sound, Blind
              OV - Orchestral Sound, Visual

   This data represents their average error distance to the target.

   Testing the homogeneity of variance found a large spread in the variances.
   The data was then adjusted to reduce this spread by computing its log.
   This formula makes the variance in the adjusted data almost the same.


   This SAS program prints

       1) Subject information
       2) Target error data
       3) Target error means, std and variance
       4) Adjusted target error means, std and variance
       5) Effect of gender on target error
       6) Effect of area of study on target error
       7) Effect of musical experience on target error
       8) Effect of 3D interface experience on target error
       9) Effect of block order on target error
      10) The interaction of sound type and visual cues
      11) An analysis of sound types
*/

DATA Exp_1;
INPUT Gender $ AreaStdy $ Mus_Exp Grph_Exp Order $ MB MV N OB OV TB TV;

Adj_N = LOG(N);
Adj_TB = LOG(TB);
Adj_TV = LOG(TV);
Adj_MB = LOG(MB);
Adj_MV = LOG(MV);
Adj_OB = LOG(OB);
Adj_OV = LOG(OV);

CARDS;
F CS     0 0 TNOM 3.452   0.538   0.553   3.424   0.875   1.284   0.476
M ARTS   3 0 OMTN 2.602   1.174   1.120   3.825   2.564   1.149   0.322
M CS     0 1 MONT 3.476   0.661   1.208   5.270   1.266   2.611   0.370
F CS     3 0 OMTN 5.698   0.915   2.090   3.303   1.552   3.591   1.009
M OTHER  2 1 MONT 0.634   0.721   2.929   3.893   1.344   0.669   0.323
F CS     0 0 NTMO 1.897   0.868   1.681   4.315   1.401   0.969   0.280
M CS     0 0 TNOM 2.047   0.336   2.092   2.717   0.723   1.905   0.326
M OTHER  3 1 TNOM 2.071   0.545   1.174   3.863   0.835   0.459   0.194
M CS     0 0 OMTN 4.737   0.901   3.834   5.110   2.933   4.314   0.392
F OTHER  0 0 NTMO 3.473   0.610   1.715   3.557   0.931   0.969   0.247
M CS     2 0 MONT 1.416   0.916   2.070   2.917   1.509   2.438   0.782
F CS     1 0 OMTN 2.573   0.529   2.380   2.994   1.415   0.439   0.242
M CS     1 1 TNOM 1.991   0.865   2.558   2.534   1.260   2.036   0.270
M SCI    2 0 NTMO 2.749   0.537   4.002   2.887   0.880   1.360   0.263
F SCI    3 0 NTMO 1.760   1.055   1.300   2.924   1.503   2.549   0.862
M SCI    1 0 MONT 2.191   0.342   0.956   3.422   0.876   0.642   0.278
M SCI    0 1 TNOM 3.427   0.385   2.265   3.423   1.873   3.633   0.299
M ARTS   2 0 OMTN 2.698   0.449   2.162   1.532   1.180   0.516   0.267
M ARTS   1 0 MONT 2.734   0.561   1.716   3.785   1.113   3.424   0.794
```

```
M CS     0 1 NTMO 1.047   0.463   1.622   3.089   0.614   3.306   0.303

PROC PRINT;
   TITLE "Subject Info";
   VAR Gender AreaStdy Mus_Exp Grph_Exp Order;

PROC PRINT;
   TITLE "Target Error Data";
   VAR N TB TV MB MV OB OV;

PROC MEANS N MEAN STD VAR MAXDEC=3;
   TITLE "Means of Target Error Data";
   VAR N TB TV MB MV OB OV;

PROC MEANS N MEAN STD VAR MAXDEC=3;
   TITLE "Means of Adjusted Target Error Data - Log(x)";
   VAR Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV;

PROC ANOVA;
   CLASSES Gender;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
One Factor";
   TITLE3 "The Effect of Gender on Target Error";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = Gender /nouni;
   REPEATED Sound 7 (1 2 3 4 5 6 7) / NOM;
   MEANS Gender;

PROC ANOVA;
   CLASSES AreaStdy;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
One Factor";
   TITLE3 "The Effect of Area of Study on Target Error";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = AreaStdy /nouni;
   REPEATED Sound 7 (1 2 3 4 5 6 7) / NOM;
   MEANS AreaStdy;
   RUN;

PROC ANOVA;
   CLASSES Mus_Exp;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
One Factor";
   TITLE3 "The Effect of Musical Experience on Target Error";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = Mus_Exp /nouni;
   REPEATED Sound 7 (1 2 3 4 5 6 7) / NOM;
   MEANS Mus_Exp;
   RUN;

PROC ANOVA;
   CLASSES Grph_Exp;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
One Factor";
   TITLE3 "The Effect of 3D Graphical Experience on Target Error";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = Grph_Exp/nouni;
   REPEATED Sound 7 (1 2 3 4 5 6 7) / NOM;
   MEANS Grph_Exp;
   RUN;

PROC ANOVA;
   CLASSES Order;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
One Factor";
   TITLE3 "The Effect of Block Order on Target Error";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = Order /nouni;
   REPEATED Sound 7 (1 2 3 4 5 6 7) / NOM;
   MEANS Order;
   RUN;

PROC ANOVA;
   TITLE "Two Factor ANOVA on Adjusted Target Error with Repeated Measures on
Two Factors";
   TITLE3 "The Interaction of Sound Type and Visual Cue";
   MODEL Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = /nouni;
   REPEATED Sound 3, Cue 2 / NOM;
   RUN;
```

```
PROC ANOVA;
   TITLE "One Factor ANOVA on Adjusted Target Error with Repeated Measures";
   TITLE3 "Analysis of Sound Type";
   MODEL Adj_N Adj_TB Adj_TV Adj_MB Adj_MV Adj_OB Adj_OV = /nouni;
   REPEATED Sound 7 CONTRAST(1) / NOM SUMMARY;
   REPEATED Sound 7 CONTRAST(2) / NOM SUMMARY;
   REPEATED Sound 7 CONTRAST(3) / NOM SUMMARY;
   REPEATED Sound 7 CONTRAST(4) / NOM SUMMARY;
   REPEATED Sound 7 CONTRAST(5) / NOM SUMMARY;
   REPEATED Sound 7 CONTRAST(6) / NOM SUMMARY;
   RUN;
```

Appendix F Sound Cue Experiment Data

Subject Info

OBS	GENDER	AREASTDY	MUS_EXP	GRPH_EXP	ORDER
1	F	CS	0	0	TNOM
2	M	ARTS	3	0	OMTN
3	M	CS	0	1	MONT
4	F	CS	3	0	OMTN
5	M	OTHER	2	1	MONT
6	F	CS	0	0	NTMO
7	M	CS	0	0	TNOM
8	M	OTHER	3	1	TNOM
9	M	CS	0	0	OMTN
10	F	OTHER	0	0	NTMO
11	M	CS	2	0	MONT
12	F	CS	1	0	OMTN
13	M	CS	1	1	TNOM
14	M	SCI	2	0	NTMO
15	F	SCI	3	0	NTMO
16	M	SCI	1	0	MONT
17	M	SCI	0	1	TNOM
18	M	ARTS	2	0	OMTN
19	M	ARTS	1	0	MONT
20	M	CS	0	1	NTMO

Target Error Data

OBS	N	TB	TV	MB	MV	OB	OV
1	0.553	1.284	0.476	3.452	0.538	3.424	0.875
2	1.120	1.149	0.322	2.602	1.174	3.825	2.564
3	1.208	2.611	0.370	3.476	0.661	5.270	1.266
4	2.090	3.591	1.009	5.698	0.915	3.303	1.552
5	2.929	0.669	0.323	0.634	0.721	3.893	1.344
6	1.681	0.969	0.280	1.897	0.868	4.315	1.401
7	2.092	1.905	0.326	2.047	0.336	2.717	0.723
8	1.174	0.459	0.194	2.071	0.545	3.863	0.835
9	3.834	4.314	0.392	4.737	0.901	5.110	2.933
10	1.715	0.969	0.247	3.473	0.610	3.557	0.931
11	2.070	2.438	0.782	1.416	0.916	2.917	1.509
12	2.380	0.439	0.242	2.573	0.529	2.994	1.415
13	2.558	2.036	0.270	1.991	0.865	2.534	1.260
14	4.002	1.360	0.263	2.749	0.537	2.887	0.880
15	1.300	2.549	0.862	1.760	1.055	2.924	1.503
16	0.956	0.642	0.278	2.191	0.342	3.422	0.876
17	2.265	3.633	0.299	3.427	0.385	3.423	1.873
18	2.162	0.516	0.267	2.698	0.449	1.532	1.180
19	1.716	3.424	0.794	2.734	0.561	3.785	1.113
20	1.622	3.306	0.303	1.047	0.463	3.089	0.614

Variable	N	Mean	Std Dev	Variance
N	20	1.971	0.887	0.787
TB	20	1.913	1.244	1.548
TV	20	0.415	0.240	0.058
MB	20	2.634	1.195	1.428
MV	20	0.669	0.244	0.060
OB	20	3.439	0.850	0.723
OV	20	1.332	0.584	0.342

Time Data

OBS	N	TB	TV	MB	MV	OB	OV
1	4987.05	24113.20	9064.33	46596.60	13649.40	35044.4	18443.80
2	8726.60	25217.80	8821.79	27985.60	14326.53	25870.4	14352.40
3	3272.65	14451.60	6319.00	22916.80	16660.40	24995.6	14910.13
4	4609.26	22359.40	10547.80	28930.40	15747.45	69165.6	31798.58
5	5377.55	32719.40	8012.11	50763.00	11952.50	66936.6	12221.15
6	4887.73	49952.80	20433.20	24170.75	13869.87	22360.2	12959.47
7	4985.07	47857.60	10876.67	42277.40	16906.67	39820.2	19575.00
8	4719.58	60962.40	18952.83	49622.80	11707.25	70512.0	16363.90
9	1883.50	36235.60	10267.33	22860.20	11147.33	25741.4	16858.40
10	11054.47	45759.00	14567.53	49179.00	20480.40	38017.4	20156.67
11	5183.47	36246.60	8923.07	28845.80	13234.20	27843.4	12013.93
12	5038.14	91401.40	38214.33	55364.60	43885.60	108073.0	42721.07
13	5292.93	69081.80	12583.64	91784.00	23697.20	35924.2	14333.60
14	3614.60	37225.75	9859.40	46050.75	14623.07	33269.0	25491.53
15	12169.40	17584.80	15554.80	20894.00	16488.07	22751.0	12206.73
16	6009.84	40189.80	14868.00	49613.20	19169.53	37277.6	14918.87
17	9284.15	62680.80	16548.27	37578.60	19431.85	37818.4	34862.87
18	7651.55	26085.00	13548.73	38982.40	13677.93	27790.4	15225.50
19	6482.15	21511.40	8131.64	46733.60	11354.79	30980.4	12318.53
20	4508.26	108570.20	11400.00	43981.00	11362.93	37668.0	15580.60

Variable	N	Mean	Std Dev	Variance
N	20	5986.898	2579.591	6654289.093
TB	20	43510.318	24764.999	613305173.77
TV	20	13374.724	6967.511	48546204.010
MB	20	41256.525	16249.854	264057744.96
MV	20	16668.648	7256.971	52663633.029
OB	20	40892.960	21579.980	465695548.01
OV	20	18865.637	8454.676	71481546.614

Subjective Preference Data

OBS	N	T	M	O
1	2	2	3	4
2	2	4	3	3
3	2	3	3	3
4	2	3	3	2
5	1	4	3	2
6	2	3	4	3
7	1	2	3	4
8	4	3	4	1
9	3	4	3	3
10	2	2	3	4
11	3	2	3	4
12	1	4	3	3
13	1	3	3	4
14	2	2	3	4
15	1	3	4	2
16	2	1	3	4
17	2	2	3	3
18	2	4	3	3
19	1	3	4	4
20	3	2	3	2

Variable	N	Mean	Std Dev	Variance
N	20	1.950	0.826	0.682
T	20	2.800	0.894	0.800
M	20	3.200	0.410	0.168
O	20	3.100	0.912	0.832

Subjective Performance Data

OBS	N	T	M	O
1	1	4	2	3
2	2	4	3	3
3	1	4	4	2
4	1	3	3	2
5	1	3	3	2
6	1	2	3	2
7	1	2	2	3
8	1	4	3	2
9	1	4	3	3
10	1	4	3	2
11	1	3	3	2
12	1	4	3	3
13	1	4	3	4
14	1	4	2	3
15	1	3	4	3
16	1	4	2	3
17	1	4	3	2
18	1	4	3	3
19	1	3	4	4
20	1	4	4	2

Variable	N	Mean	Std Dev	Variance
N	20	1.050	0.224	0.050
T	20	3.550	0.686	0.471
M	20	3.000	0.649	0.421
O	20	2.650	0.671	0.450

Subjective Usability Data

OBS	N	T	M	O
1	4	4	3	3
2	3	4	3	3
3	4	3	3	2
4	2	4	3	1
5	4	3	3	1
6	1	3	3	3
7	1	4	3	4
8	4	4	3	1
9	4	3	3	3
10	1	4	3	2
11	4	3	2	2
12	1	4	4	4
13	1	3	3	3
14	1	4	2	3
15	2	4	4	3
16	4	3	2	2
17	1	4	3	3
18	1	4	3	3
19	1	3	4	3
20	4	3	3	2

Variable	N	Mean	Std Dev	Variance
N	20	2.400	1.429	2.042
T	20	3.550	0.510	0.261
M	20	3.000	0.562	0.316
O	20	2.550	0.887	0.787

Subjective Marketability Data

OBS	N	T	M	O
1	3	4	3	3
2	1	4	3	2
3	1	3	3	1
4	1	3	3	2
5	2	3	3	1
6	2	2	4	3
7	1	3	3	4
8	3	3	3	1
9	1	3	2	2
10	2	2	3	2
11	2	3	2	3
12	2	3	4	3
13	1	3	2	3
14	2	3	2	3
15	1	4	4	3
16	1	2	2	3
17	1	3	3	3
18	1	4	3	3
19	1	2	4	3
20	3	1	3	2

Variable	N	Mean	Std Dev	Variance
N	20	1.600	0.754	0.568
T	20	2.900	0.788	0.621
M	20	2.950	0.686	0.471
O	20	2.500	0.827	0.684

Appendix G Extended Period Experiment

Subject Info

OBS	GENDER	AREASTDY	MUS_EXP	GRPH_EXP	ORDER
1	M	SCI	2	0	NTMO
2	F	SCI	3	0	NTMO
3	M	SCI	1	0	MONT
4	M	SCI	0	1	TNOM
5	M	ARTS	2	0	OMTN
6	M	ARTS	1	0	MONT
7	M	CS	0	1	NTMO

Day 1

Target Error Data

OBS	N	TB	TV	MB	MV	OB	OV
1	4.002	1.360	0.263	2.749	0.537	2.887	0.880
2	1.300	2.549	0.862	1.760	1.055	2.924	1.503
3	0.956	0.642	0.278	2.191	0.342	3.422	0.876
4	2.265	3.633	0.299	3.427	0.385	3.423	1.873
5	2.162	0.516	0.267	2.698	0.449	1.532	1.180
6	1.716	3.424	0.794	2.734	0.561	3.785	1.113
7	1.622	3.306	0.303	1.047	0.463	3.089	0.614

Variable	N	Mean	Std Dev	Variance
N	7	2.003	0.992	0.985
TB	7	2.204	1.346	1.811
TV	7	0.438	0.268	0.072
MB	7	2.372	0.781	0.609
MV	7	0.542	0.239	0.057
OB	7	3.009	0.725	0.526
OV	7	1.148	0.425	0.181

Time Data

OBS	N	TB	TV	MB	MV	OB	OV
1	3614.60	37225.75	9859.40	46050.75	14623.07	33269.0	25491.53
2	12169.40	17584.80	15554.80	20894.00	16488.07	22751.0	12206.73
3	6009.84	40189.80	14868.00	49613.20	19169.53	37277.6	14918.87
4	9284.15	62680.80	16548.27	37578.60	19431.85	37818.4	34862.87
5	7651.55	26085.00	13548.73	38982.40	13677.93	27790.4	15225.50
6	6482.15	21511.40	8131.64	46733.60	11354.79	30980.4	12318.53
7	4508.26	108570.20	11400.00	43981.00	11362.93	37668.0	15580.60

Variable	N	Mean	Std Dev	Variance
N	7	7102.851	2921.437	8534795.837
TB	7	44835.393	31880.645	1016375513.5
TV	7	12844.406	3132.159	9810422.625
MB	7	40547.650	9661.378	93342233.304
MV	7	15158.309	3353.380	11245159.350
OB	7	32507.829	5743.729	32990418.939
OV	7	18657.805	8429.473	71056022.972

Subjective Preference Data

OBS	N	T	M	O
1	2	2	3	4
2	1	3	4	2
3	2	1	3	4
4	2	2	3	3
5	2	4	3	3
6	1	3	4	4
7	3	2	3	2

Variable	N	Mean	Std Dev	Variance
N	7	1.857	0.690	0.476
T	7	2.429	0.976	0.952
M	7	3.286	0.488	0.238
O	7	3.143	0.900	0.810

Subjective Performance Data

OBS	N	T	M	O
1	1	4	2	3
2	1	3	4	3
3	1	4	2	3
4	1	4	3	2
5	1	4	3	3
6	1	3	4	4
7	1	4	4	2

Variable	N	Mean	Std Dev	Variance
N	7	1.000	0.000	0.000
T	7	3.714	0.488	0.238
M	7	3.143	0.900	0.810
O	7	2.857	0.690	0.476

Subjective Usability Data

OBS	N	T	M	O
1	1	4	2	3
2	2	4	4	3
3	4	3	2	2
4	1	4	3	3
5	1	4	3	3
6	1	3	4	3
7	4	3	3	2

Variable	N	Mean	Std Dev	Variance
N	7	2.000	1.414	2.000
T	7	3.571	0.535	0.286
M	7	3.000	0.816	0.667
O	7	2.714	0.488	0.238

Subjective Marketability Data

OBS	N	T	M	O
1	2	3	2	3
2	1	4	4	3
3	1	2	2	3
4	1	3	3	3
5	1	4	3	3
6	1	2	4	3
7	3	1	3	2

Variable	N	Mean	Std Dev	Variance
N	7	1.429	0.787	0.619
T	7	2.714	1.113	1.238
M	7	3.000	0.816	0.667
O	7	2.857	0.378	0.143

Day 2

Target Error Data

OBS	N	TB	TV	MB	MV	OB	OV
1	1.530	1.513	0.250	2.666	0.452	2.960	1.154
2	0.894	2.407	0.626	2.297	0.534	3.249	1.588
3	0.806	0.744	0.294	1.153	0.708	2.568	1.036
4	2.498	3.204	0.316	3.680	0.600	4.023	1.631
5	2.940	2.818	0.421	2.309	0.979	3.169	0.982
6	2.383	3.592	0.794	2.604	0.581	3.003	1.328
7	0.829	1.102	0.158	0.813	0.538	3.467	1.190

Variable	N	Mean	Std Dev	Variance
N	7	1.697	0.902	0.813
TB	7	2.197	1.093	1.195
TV	7	0.408	0.226	0.051
MB	7	2.217	0.967	0.935
MV	7	0.627	0.173	0.030
OB	7	3.206	0.456	0.208
OV	7	1.273	0.256	0.065

Time Data

OBS	N	TB	TV	MB	MV	OB	OV
1	3535.05	43910.8	7390.07	42728.50	16437.20	42367.2	13795.80
2	7782.58	17522.8	10165.13	20774.40	13213.40	16906.8	10738.60
3	3779.00	34325.4	10907.87	51463.60	12266.57	40604.0	13892.40
4	5163.84	27780.4	9711.27	30637.80	12933.13	35957.0	16572.27
5	3791.65	24253.4	7353.67	23736.20	9552.93	9790.4	7058.50
6	4110.35	9676.6	6052.67	26571.20	9325.07	17453.0	6919.53
7	4505.05	35933.2	10497.92	21522.75	12324.07	28307.2	14059.00

Variable	N	Mean	Std Dev	Variance
N	7	4666.790	1479.068	2187642.642
TB	7	27628.943	11652.897	135790010.56
TV	7	8868.370	1898.023	3602491.491
MB	7	31062.064	11713.092	137196514.17
MV	7	12293.196	2406.573	5791592.857
OB	7	27340.800	12855.164	165255247.93
OV	7	11862.300	3733.612	13939858.340

Subjective Preference Data

OBS	N	T	M	O
1	1	1	2	4
2	1	3	4	4
3	1	2	3	4
4	1	3	3	3
5	1	3	3	3
6	1	2	4	3
7	3	1	2	3

Variable	N	Mean	Std Dev	Variance
N	7	1.286	0.756	0.571
T	7	2.143	0.900	0.810
M	7	3.000	0.816	0.667
O	7	3.429	0.535	0.286

Subjective Performance Data

OBS	N	T	M	O
1	1	4	3	4
2	1	3	4	3
3	1	4	3	3
4	1	4	3	2
5	1	4	3	3
6	1	3	4	3
7	1	4	3	2

Variable	N	Mean	Std Dev	Variance
N	7	1.000	0.000	0.000
T	7	3.714	0.488	0.238
M	7	3.286	0.488	0.238
O	7	2.857	0.690	0.476

Subjective Usability Data

OBS	N	T	M	O
1	2	4	2	3
2	1	3	4	4
3	4	3	2	2
4	1	4	3	2
5	2	4	4	4
6	1	2	4	3
7	4	3	2	1

Variable	N	Mean	Std Dev	Variance
N	7	2.143	1.345	1.810
T	7	3.286	0.756	0.571
M	7	3.000	1.000	1.000
O	7	2.714	1.113	1.238

Subjective Marketability Data

OBS	N	T	M	O
1	1	2	1	3
2	1	3	4	4
3	1	2	3	3
4	3	2	2	1
5	2	2	3	3
6	1	2	4	3
7	1	1	3	3

Variable	N	Mean	Std Dev	Variance
N	7	1.429	0.787	0.619
T	7	2.000	0.577	0.333
M	7	2.857	1.069	1.143
O	7	2.857	0.900	0.810

DAY 3

OBS	N	TB	TV	TP	TQ	MB	MV	MP	MQ	OB	OV	OP	OQ
1	0.886	0.659	0.223	0.846	1.575	1.383	0.415	1.331	0.530	3.170	0.845	0.674	0.949
2	0.866	1.489	0.395	0.766	1.552	2.443	0.601	0.664	1.043	2.511	1.364	1.486	0.499
3	0.828	1.008	0.289	0.508	0.879	1.746	0.323	0.735	0.549	1.677	0.650	0.731	0.749
4	1.561	1.898	0.172	0.378	1.468	3.838	0.172	0.661	1.568	3.557	1.468	0.610	1.059
5	1.704	2.294	0.335	0.775	1.145	1.918	0.806	1.256	1.470	2.714	1.947	1.196	1.888
6	2.075	2.488	0.505	1.428	1.304	3.166	1.487	0.872	1.625	3.440	1.368	1.066	1.044
7	1.412	1.108	0.214	0.638	1.337	2.054	0.362	0.572	0.679	3.095	0.878	0.809	1.071

Variable	N	Mean	Std Dev	Variance
N	7	1.333	0.486	0.237
TB	7	1.563	0.688	0.473
TV	7	0.305	0.117	0.014
TP	7	0.763	0.336	0.113
TQ	7	1.323	0.247	0.061
MB	7	2.364	0.862	0.744
MV	7	0.595	0.443	0.197
MP	7	0.870	0.304	0.092
MQ	7	1.066	0.489	0.239
OB	7	2.881	0.647	0.419
OV	7	1.217	0.450	0.203
OP	7	0.939	0.321	0.103
OQ	7	1.037	0.429	0.184

Time Data

OBS	N	TB	TV	TP	TQ	MB	MV
1	3775.70	36476.5	9452.56	9692.89	5947.70	38607.4	10964.50
2	5856.60	19408.0	11037.10	9507.11	6587.67	22407.0	11307.00
3	3329.36	48866.2	11203.50	10996.56	3852.89	41172.0	11306.91
4	4441.36	19552.6	7630.00	6775.75	3712.33	29799.8	12286.10
5	2765.17	11765.6	6538.27	8112.88	3651.00	19663.8	7698.09
6	4096.30	13882.6	5449.50	13456.33	6140.10	31956.6	10681.50
7	3976.36	27088.4	9665.10	6921.88	3638.88	23633.8	10435.60

OBS	MP	MQ	OB	OV	OP	OQ
1	11782.67	5517.38	20757.4	13893.55	11602.13	7306.60
2	12673.00	7958.56	14122.8	9769.36	10341.00	7476.88
3	14395.75	4837.44	24434.4	13260.27	15327.11	6568.90
4	14563.78	6483.00	22638.2	21727.75	15341.14	7249.89
5	12243.75	4603.11	15826.0	7384.73	9329.00	3572.22
6	15909.00	5559.00	20492.4	6949.20	11354.20	4475.50
7	10992.56	5647.89	25457.4	12600.20	18661.50	6118.75

Variable	N	Mean	Std Dev	Variance
N	7	4034.408	971.956	944697.697
TB	7	25291.414	13321.431	177460511.69
TV	7	8710.861	2221.493	4935032.550
TP	7	9351.913	2372.354	5628063.473
TQ	7	4790.081	1357.485	1842766.475
MB	7	29605.771	8233.069	67783425.179
MV	7	10668.529	1437.430	2066205.849
MP	7	13222.929	1765.337	3116415.720
MQ	7	5800.911	1128.864	1274333.363
OB	7	20532.657	4226.918	17866836.716
OV	7	12226.437	5026.452	25265217.878
OP	7	13136.583	3367.185	11337938.116
OQ	7	6109.819	1523.271	2320353.122

Subjective Preference Data

OBS	N	T	M	O
1	1	1	3	4
2	1	4	4	4
3	1	2	3	3
4	1	3	4	4
5	1	3	4	4
6	1	2	4	4
7	2	1	3	3

Variable	N	Mean	Std Dev	Variance
N	7	1.143	0.378	0.143
T	7	2.286	1.113	1.238
M	7	3.571	0.535	0.286
O	7	3.714	0.488	0.238

Subjective Performance Data

OBS	N	T	M	O
1	1	4	2	4
2	1	4	4	4
3	1	4	3	3
4	1	4	3	2
5	1	4	3	3
6	1	3	4	4
7	1	4	3	2

Variable	N	Mean	Std Dev	Variance
N	7	1.000	0.000	0.000
T	7	3.857	0.378	0.143
M	7	3.143	0.690	0.476
O	7	3.143	0.900	0.810

Subjective Usability Data

OBS	N	T	M	O
1	3	4	3	4
2	1	3	4	4
3	4	4	3	2
4	2	4	3	3
5	2	3	3	3
6	1	3	4	4
7	4	3	3	1

Variable	N	Mean	Std Dev	Variance
N	7	2.429	1.272	1.619
T	7	3.429	0.535	0.286
M	7	3.286	0.488	0.238
O	7	3.000	1.155	1.333

Subjective Marketability Data

OBS	N	T	M	O
1	2	2	2	3
2	1	4	4	4
3	1	3	3	2
4	2	2	3	2
5	2	2	3	3
6	1	2	4	4
7	1	1	3	3

Variable	N	Mean	Std Dev	Variance
N	7	1.429	0.535	0.286
T	7	2.286	0.951	0.905
M	7	3.143	0.690	0.476
O	7	3.000	0.816	0.667

Where is the wise? where is the scribe? where is the disputer of this world? hath not God made foolish the wisdom of this world?

For after that in the wisdom of God the world by wisdom knew not God, it pleased God by the foolishness of preaching to save them that believe.
1 Corinthians 1:20-21

A wise man will hear, and will increase learning; and a man
of understanding shall attain unto wise counsels:
... The fear of the LORD is the beginning of knowledge
Proverbs 1:5,7

So then faith cometh by hearing, and hearing by the word of God.
But I say, Have they not heard?
Yes verily, their sound went into all the earth,
and their words unto the ends of the world.
Romans 10:17-18

Then whosoever heareth the sound of the trumpet,
and taketh not warning; if the sword come, and take him away,
his blood shall be upon his own head.
Ezekiel 33:4

The wind bloweth where it listeth,
and thou hearest the sound thereof,
but canst not tell whence it cometh, and whither it goeth:
so is every one that is born of the Spirit.
John 3:8

Blessed is the people that know the joyful sound:
they shall walk, O LORD, in the light of thy countenance.
Psalms 89:15