The Use of 3D Sound as a Navigational Aid in Virtual Environments

by

Ryan T. C. Gunther

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, 1997

I hereby 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 reproduce 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 photocopying this thesis. Please sign below, and give address and date.

Abstract

Because virtual environments are less visually rich than real-world environments, careful consideration must be given to their design, otherwise much of the potential for virtual reality will be lost. One important design criteria is to make certain that adequate navigational cues are incorporated into a complex virtual world. The usability of virtual environments will be decreased if users often become lost or disoriented, as is frequently the case now.

This thesis investigates the use of spatialized sound as a navigational aid for people using virtual reality systems. An experiment was performed to determine if incorporating spatialized sound into a virtual environment a) helps people find specific objects in the environment, and b) influences the rate at which people acquire spatial knowledge.

The results show that the addition of spatialized sound to a virtual environment improved the ability of people to find objects, although the addition of sound did not increase the amount of configurational knowledge a person was able to acquire. In fact, it appears that the addition of sound may have suppressed the development of configurational knowledge.

Acknowledgements

Funding for this research was provided by the Natural Sciences and Engineering Research Council of Canada, the University of Waterloo Institute for Computer Research, and the Computer Graphics Lab at the University of Waterloo. I would also like to thank QSound Labs Inc. for generously donating QMixer95, their real time 3D audio mixer.

In addition, I would like to thank my two supervisors, Rick Kazman and Carolyn MacGregor, for their support and guidance throughout this project, and my readers Steve Mann and Darrell Raymond. I also extend thanks to all of the members of the Computer Graphics Lab for their encouragement and friendship during my stay in Waterloo.

Finally, I would like to thank my parents for all of their love and support.

Trademarks

DataDesk is a registered trademark of Data Description, Inc.

Pentium is a registered trademark of Intel Inc.

QSound is a registered trademark of QSound Labs Inc.

Sapphire 2SX is a trademark of Fujitsu Corporation.

UltraSound MAX is a registered trademark of Advanced Gravis Computer Technology Ltd.

Virtual i-glasses is a registered trademark of Virtual i.O.

Windows NT is a registered trademark of Microsoft Corporation.

World Up is a registered trademark of Sense8 Corporation.

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

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The term Virtual Reality (VR) is a widely misunderstood term and many people only have a vague notion of what constitutes a virtual reality system. Rheingold has suggested that all VR systems are composed of two fundamental elements, immersion and navigation (Rheingold 1991). Immersion is the illusion of being inside a computer-generated scene, and navigation is the ability of a user to move around, as if inside, the computer-generated environment. It is not necessary for either of these two elements to be implemented with one specific kind of technology. Rheingold has written the following:

> "Images can be created with optics or electronics or both. Gestural input can come from gloves and keyboards and steering wheels. Head-mounted displays, gloves that sense finger position and movement, and magnetic head-position trackers can create the feeling of immersion and grant the power of navigation through simulated environments."

Over the last several years VR technology has shown significant improvement and this has allowed VR to become more feasible and popular. However, careful consideration must be given to the design and usability of virtual environments, otherwise much of the potential for VR will be lost. One important design criteria is to make certain that adequate navigational cues are incorporated into a virtual environment. The usability of virtual environments will be severely crippled if users often become lost or disoriented, as is frequently the case now. Unfortunately, this is not an easy criteria to meet since it is not yet well understood how to design real world environments that are easy to navigate.

One reason research on navigation in real world environments has progressed slowly is the difficulty in controlling environmental variables in the real world. It is usually time consuming and expensive, if not impossible, to alter real world environments. Virtual environments can provide the control and flexibility needed to study navigation if it can be shown that exposure to a virtual environment causes navigators to exhibit similar behaviors as they would if exposed to a real world environment.

Virtual environments also have the potential to be used as training environments for real world navigation. A real world environment can be replicated as a virtual world and enhanced with additional navigational cues. These enhanced virtual environments will be powerful training tools if it is possible for a navigator to retain their navigational knowledge when the additional cues are absent and they are navigating in the real environment.

A significant amount of work (Beaumont, Gray, Moore, and Robinson 1984; Butler, Acquino, Hissong, and Scott 1993; Holding and Holding 1989; Moeser 1988; Passini 1984; Peponis, Zimring, and Choi 1990; Thorndyke and Goldin 1983) has been done studying how people navigate in real environments, although this work has progressed slowly and there remains much to be learned. Only during the last five years have people begun to extend this work to include virtual environments.

The amount of time people spend immersed in virtual environments is short, so often people navigating in virtual environments only want to be able to find the specific location they are interested in. Usually, they are not trying to acquire any type of configurational knowledge about the environment. However, people navigating in real world environments generally spend a substantial amount of time in an environment, so it is important for them to be able to acquire configurational knowledge of the environment.

Navigational aids designed for assisting people in finding specific locations are not always appropriate for helping people acquire configurational knowledge. For example, research has shown that although signs make it easy to find specific locations, they sometimes reduce the cognitive load of a person to such an extent that the acquisition of configurational knowledge is suppressed (Moeser 1988). It is important to determine the navigational aids which are suitable for specific tasks.

A wide range of people are beginning to use virtual reality and the hardware being used by these people is quite diverse. In order for a virtual environment to be accessible to a large number of people the visual complexity of the environment must be kept low, otherwise only people with access to a powerful machine will be able to explore the environment. This makes it necessary to construct navigational aids that are appropriate for visually sparse environments.

Almost all of the previous research on navigation in both real and virtual environments has focussed on either the physical layout of the environment or the addition of visual navigational aids. There has been almost no work exploring the use of auditory cues as navigational aids, even though auditory cues may be ideal for environments with low visual complexity. As well, auditory cues might aid navigation because even without any head movement auditory information can be perceived in all directions. This is unlike visual information, which without head movement is restricted to a 150 degree field of view. Auditory cues can also provide redundancy by complementing visual cues, and might be useful navigational aids for visually impaired users.

This thesis will attempt to contribute to the research by exploring the effect that the addition of spatialized auditory cues to a virtual environment has on navigation and spatial knowledge acquisition. Chapter 2 reviews background material covering spatial knowledge, navigational tools, navigation of real and virtual environments, and spatialized sound.

Chapter 3 describes the experimental design, including design decisions and the pilot study results. An experiment was conducted to determine if spatialized auditory cues can enhance navigational ability, and also to examine their affect on the acquisition of spatial knowledge. The experiment required subjects to explore a visually sparse virtual environment and then perform several tasks. These tasks included several object search trials, a landmark recollection and placement task, and a test of spatial ability.

The results from the spatialized sound experiment are analyzed and organized into tables and graphs in Chapter 4. Chapter 5 contains a detailed discussion of these results, where conclusions are drawn and the possibility of future research is discussed.

The results of this study show that the addition of spatialized sounds to a virtual environment can improve a person's ability to locate objects in a virtual environment. However, the results also show that spatialized sounds do not appear to enhance the acquisition of survey knowledge, and may even suppress its development.

Chapter 2

Background

2.1 Spatial Ability

Spatial ability is comprised of both the capability to perceive an environment through the senses, and the cognitive process of learning the environment. Spatial ability appears to be composed of several basic cognitive skills: visual memory, visualization, and spatial orientation (Thorndyke and Goldin 1983). Visual memory is the ability to encode and retain purely visual information. Visualization is the ability to manipulate and transform a visual image in order to solve spatial problems. Spatial orientation is the ability to maintain a consistent frame of reference in a transformed or rotated visual array. Spatial ability influences how well a person is able to acquire spatial knowledge as well as the different strategies people use when navigating. Spatial knowledge is information about the structural configuration of an environment. The effect of spatial ability on spatial knowledge acquisition and navigational strategies will be discussed further in Sections 2.4.1 and 2.5.1.

2.2 Spatial Knowledge

Spatial knowledge is acquired from many different sources: maps, photographs or videos, written or verbal directions, navigation, and so on. Thorndyke and Goldin (Thorndyke and Goldin 1983) have classified spatial knowledge according to three different categories: landmark knowledge, procedural knowledge, and survey knowledge. Each type of knowledge represents different aspects of an environment, is acquired from different sources, and is useful for different tasks.

2.2.1 Landmark Knowledge

Landmark knowledge represents the visual details of specific locations in an environment. Objects that become landmarks are those that are visually distinctive or have personal meaning (Lynch 1960). Some attributes of an object which may distinguish it are architectural style, size, and color. As well, any object that provides directional information is likely to become a prominent landmark. Research performed by Thorndyke and Goldin (Thorndyke and Goldin 1983) has shown that landmark knowledge can be acquired from either direct or indirect exposure to an environment. In the study conducted by Thorndyke, viewing a film of the environment was considered indirect exposure and navigating in the environment was considered direct exposure.

2.2.2 Procedural (Route) Knowledge

Procedural knowledge, also known as Route knowledge, has an egocentric frame of reference and represents information about the sequence of actions that are needed to follow specific routes. This information includes knowledge of critical points along a route where turns occur, and the action required at each point. Procedural knowledge implies knowing the distance between route segments, the direction of turns, and the ordering of landmarks. This kind of knowledge can be acquired from navigation of the actual route.

2.2.3 Survey Knowledge

Survey knowledge has an exocentric viewpoint and represents object locations and inter-object distances with respect to a fixed, global coordinate system. Survey knowledge allows people to estimate distances between landmarks and infer alternate routes that have never been traveled. Darken and Sibert believe that survey knowledge is the most important information a person requires to successfully navigate throughout an environment (Darken and Sibert 1996).

Thorndyke and Goldin have discovered that survey knowledge can be acquired either from studying a map or from extensive navigation of an environment. Survey knowledge learned from a map is called secondary knowledge and survey knowledge acquired from direct navigation is referred to as primary knowledge. Secondary knowledge is inferior to primary knowledge because it is orientation specific. This means that it can most easily be used when the environment is viewed from the same orientation (i.e., forward in the environment corresponds to upward on the map) as it was represented on the map (Levine, Marchon, and Hanley 1984; Warren, Scott, and Medley 1992; Tlauka and Wilson 1996). For example, if the map was oriented with north in the up direction then the information gained from studying the map could be used most easily when facing north in the environment. In the study by Tlauka and Wilson they found that participants who had only seen a map of a virtual environment took 40% longer on contraligned (mean of 14 seconds) landmark pointing trials as opposed to aligned (mean of 10 seconds) trials. In aligned trials a participant's orientation was aligned with the orientation in which the map had been seen and in contraligned trails a participant's orientation was misaligned from the aligned perspective by 180°. Participants who had navigated freely through the virtual environment were equally fast on both contraligned and aligned pointing trials. Survey knowledge acquired from direct navigation is orientation independent, which means it is not affected by viewpoint (Thorndyke and Hayes-Roth 1982). The number of viewpoints from which a space is viewed may be one factor that determines orientation specificity (Evans and Pezdek 1980). During navigation an environment is usually seen from many different viewpoints, whereas a map is customarily viewed from only one viewpoint. This may explain why survey knowledge acquired from a map is orientation specific and survey knowledge gained from navigation is orientation independent.

2.2.4 Acquisition of Spatial Knowledge

To achieve complete knowledge of an environment a person must go through a dynamic process that incorporates all three types of spatial knowledge. A model to describe this, consisting of three steps, was introduced by Siegel and White (Siegel and White 1975). The first thing a person learns to do is recognize prominent landmarks. Then, after traveling between landmarks a person begins to learn the relative distances between them. Landmark sequencing is ascertained which allows routes and links to be formed. Finally, after sufficient navigation a person achieves survey knowledge and is able to infer alternate routes and estimate straight line distances between landmarks.

2.3 Environmental Design

When designing an environment it is important to make sure a suitable configuration is chosen. The configuration of an environment is the way in which spaces are related to one another, not only pairwise but also with respect to the overall pattern that they constitute (Peponis, Zimring, and Choi 1990). The configuration of an environment is important because it is difficult for people to orient themselves in an environment that has a poor configuration. As well, the type of spatial knowledge a person develops is influenced by the configuration of an environment. For example, wayfinding and the acquisition of route knowledge are easy in an environment where the paths have been constrained by limiting the number of decision points along each path. However, even after substantial navigation a person may have failed to develop survey knowledge. Survey knowledge is much easier to acquire in an environment with many wide open spaces, but there are many unconstrained paths in this kind of environment so wayfinding and the development of route knowledge are difficult.

It is possible to incorporate navigational cues into the structure of an environment. Paths, landmarks, and districts should be used to divide an environment into smaller and clearly connected sections (Lynch 1960). Designing different sections of an environment to be easily distinguishable is one cue that can be built into the structure. Sections can be made distinguishable by varying size, form, color, and architectural style. Creating sections of an environment that have high visual access to other parts of the environment is another way navigational cues can be built into the structure of an environment. An environment should also have a basic organizational principle underlying it, and any map of the environment should clearly illustrate this organizational principle (Passini 1984). For example, it is a good idea to organize streets as a grid.

Navigational cues can also be added to an environment after it has been constructed. Assigning names to buildings and labeling rooms with numbers are examples of explicit cues that can be added on top of the structure of an environment. Frequent directional cues should also be added so that a navigator is able to maintain their orientation (Lynch 1960).

2.3.1 Classification of Environments

Darken and Sibert (Darken and Sibert 1993) have classified virtual environments according to three attributes: size, density, and activity. Their classification of virtual environments can be applied equally well to real world environments.

Size

An environment can be one of three sizes, small, large, or infinite. A small environment is one where all of the world can be seen from a single viewpoint. An environment is classified as large if there is no vantage point from which the entire environment may be seen. Infinite environments are those in which a person is able to travel along some dimension forever without encountering an edge of the environment.

Density

Density refers to both the number of objects that are distributed throughout an environment as well as the way in which these objects are distributed. Objects can be spread out uniformly or clustered around specific locations. Both the distribution and number of objects in an environment determine if it is classified as sparse, dense or cluttered. A sparse environment has many large open spaces that have very few objects or navigational cues. An environment with a larger number of objects and navigational cues is categorized as dense, and a cluttered environment is one with so many objects that important landmarks and navigational cues are obscured.

Activity

An environment can be defined as either static or dynamic. A static environment is one in which objects do not change position or appearance over time. An environment is dynamic if objects move around and change appearance. Dynamic environments are more complex to navigate in than are static environments because in a dynamic environment routes can be changed when the position and appearance of landmark objects is altered.

2.4 Wayfinding

Wayfinding is defined by Peponis et al., (Peponis, Zimring, and Choi 1990) as "how well people are able to find their way to a particular destination without delay or undue anxiety." Successful wayfinding requires the use of all three types of spatial knowledge. Landmarks are used by a navigator to acquire and maintain orientation, as well as to recognize destinations. Route knowledge is needed to follow a route, and survey knowledge is required to choose the most appropriate route. Darken and Sibert (Darken and Sibert 1996) have concluded that survey knowledge is the most important type of knowledge needed for successful wayfinding in any environment. However, it is possible to be a successful navigator with just route knowledge, provided there are sufficient navigational cues built into the environment.

Wayfinding tasks can be broken down into three distinct categories: naive search, primed search, and exploration. A naive search is any task where a navigator is looking for a specific target, but has no knowledge of where the target is located. In a primed search a navigator knows the location of the target. During exploration a navigator is not looking for any target. The goal of a navigator during exploration is to acquire survey knowledge. It is not unusual while wayfinding to alternate between several of these tasks. For example, suppose a navigator is unaware of the exact location of a target, but knows its general proximity. In this case the navigator will begin with a primed search and finish with a naive search.

Often a navigator has no knowledge of where a target is located so an exhaustive search of the environment must be undertaken. If an environment is unstructured or lacking sufficient navigational cues a navigator will be unable to perform an efficient search. Some areas may be missed while others are seen several times. Since naive searches are more common in virtual environments than in real world environments it is especially important to construct virtual environments that allow searches to be executed efficiently.

2.4.1 Wayfinding Strategies

Darken and Sibert found that path following is a natural spatial behavior, and that people will create paths using structure from the environment. In one experiment they observed people exploring a large, unstructured ocean environment and noticed that people would follow coastlines (Darken and Sibert 1996). When a radial grid was superimposed on the environment people would follow the gridlines. Dead reckoning, inferring one's present position from a known past position and a constant velocity, was another strategy several people used to help them navigate.

Peponis et al., (Peponis, Zimring, and Choi 1990) observed that people avoid backtracking and Satalich (Satalich 1995) found that even when lost people are reluctant to backtrack. Peponis also concluded that if all else is equal people will continue along the same line they are traveling on. People usually only divert from their current line of movement if there is a new view that allows them to see more space and activity. The areas people often look for are those that have high visual access to other areas.

A study conducted by Butler et al., (Butler, Acquino, Hissong, and Scott 1993) attempted to determine if people prefer simple routes rather than efficient ones. Simple routes have few decision points and efficient routes have short distances to travel. The evidence from the Butler study gives some indication that people prefer routes that conserve energy, regardless of their complexity.

2.5 Individual Differences in Spatial Ability and Wayfinding

People vary a lot in their ability to navigate, and they usually know if they are good navigators (Kozlowski and Bryant 1977). As well, people who are good navigators have a tendency to like, own, and use maps. The opposite is true of poor navigators. People who are highly spatial like to draw maps for others, and also want maps drawn for themselves (Streeter and Vitello 1985).

2.5.1 Low level Spatial skills

Thorndyke and Goldin (Thorndyke and Goldin 1983) ran a study comparing the wayfinding strategies used by two individuals: a 21-year old female and a 23-year old male. The study ran for five days, and each day both subjects were driven over the same 20-mile route in West Los Angles. An analysis of the verbalizations made by each of the subjects showed extensive differences in the strategies adopted by each of them. The strategy adopted by the 21-year old female was characterized as visual/perceptual and the strategy used by the 23-year old male was described as verbal/analytic. The visual/perceptual subject focused on perceptual features of the environment such as the appearance of buildings and characteristics of the landscape. Little effort was made to organize knowledge into a survey representation of the environment. The focus of her attention seemed to be gathering route knowledge. The verbal/analytic subject immediately adopted a grid-like framework and used the primary compass directions to organize his knowledge. He attempted to learn street names and use the relations among streets to establish global, survey knowledge.

The strategy used by each subject seemed to relate to their basic spatial abilities. The visual/perceptual subject had a much higher score on spatial ability, especially visual memory, than on verbal memory. The verbal/analytic subject had high scores on both spatial and verbal abilities. This indicates that people adopt strategies that focus on their strong capabilities and limit demands on their weaker abilities.

Both subjects were eventually able to obtain a fairly accurate cognitive map of their environment, but their rate and pattern of knowledge acquisition was quite different. The visual/perceptual subject gained landmark knowledge more rapidly than the verbal/analytic subject. Over the course of the experiment her ability to recognize critical intersections and non-critical intersections improved by 43% and 34% respectively. This was unlike the performance of the verbal/analytic subject. His ability to recognize critical intersections only improved by 30% and his ability to recognize non-critical intersections actually decreased by 3%. However, the verbal/analytic subject was able to achieve survey knowledge almost immediately. After the first day he only had a mean angular error of 9.4° on the map drawing task, unlike the visual/perceptual subject who had a mean error of 53.8°. By the end of the experiment both subjects performed approximately the same on the map drawing task. The results from the Thorndyke and Goldin study indicate that the information people acquire from an environment is influenced by the strengths and weaknesses of their base level abilities.

2.5.2 Psychometric Tests

There are numerous psychometric tests that attempt to measure spatial ability. Most of these tests are unable to predict navigational ability because the test problems are presented in small-scale spaces where the entire object or environment can be seen from at least one viewpoint. However, the Guilford-Zimmerman Spatial Orientation Test has been shown to predict performance in large-scale spaces (Infield 1991). Recently, it has also been found to predict performance in virtual environments (Satalich 1995). All of the questions on the Guilford-Zimmerman Spatial Orientation test show two pictures of a boat heading in a specific direction. The goal of each question is to indicate the change in both direction and orientation of the boat between the first and second picture.

2.6 Gender Differences

Many studies have attempted to determine gender differences in wayfinding performance and spatial understanding. Unfortunately, these studies have provided conflicting results so it has been difficult to draw any solid conclusions. Beaumont et al., (Beaumont, Gray, Moore, and Robinson 1984) found no gender differences in the wayfinding performance of new visitors to an office building. Two other studies discovered no differences between men and women in their ability to point to landmarks along a route (Montello and Pick 1993; Sadalla and Montello 1989). As well, a recent study by Lawton et al., (Lawton, Charleston, and Zieles 1996) found no gender differences in either wayfinding patterns or overall time to complete a wayfinding task.

However, several other studies have found significant differences between men and women in wayfinding and spatial ability. The study conducted by Lawton et al., found that men were 46% more accurate than women in locating the direction of landmarks. Various other studies have also provided evidence that men are more accurate than women in locating the direction of landmarks (Bryant 1982; Galea and Kimura 1993; Holding and Holding 1989). There is also evidence that men are more accurate than women in keeping track of position relative to distant reference points (Lawton 1994). Studies have also shown that men are more likely to give directions in terms of cardinal reference points (Ward, Newcombe, and Overton 1986), while women are more likely to refer to landmarks (Miller and Santoni 1986) and are generally better at recalling landmarks (Galea and Kimura 1993). As well, women are more likely than men to report using routes when wayfinding (Lawton 1994), and often express anxiety and uncertainty during a wayfinding task even when their actual performance is similar to that of men (Lawton, Charleston, and Zieles 1996; Lawton 1994). One explanation for the higher levels of anxiety and uncertainty expressed by women may be that wayfinding has been stereotyped as a masculine activity (Harris 1981) and women often underestimate their ability on these types of tasks (Beyer 1990).

2.7 Navigational Tools

A lot of effort has gone into designing tools that can help people navigate in different environments. The majority of these tools have been visual, although a small amount of work has studied verbal auditory tools. Signs and maps have received a lot of attention.

2.7.1 Signs

When signs are properly designed they are effective navigational aids. Signs can provide a large amount of navigational information with only a small cost in cognitive resources. However, it is possible for signs to relieve cognitive load to a point where the acquisition of survey knowledge is inhibited (Moeser 1988). There are several guidelines that should be followed when designing signs (Passini 1984). Signs should attract attention and be graphically dissimilar to other signs in the environment. They should also be placed in consistent locations and contain only 3 to 4 bits of information.

2.7.2 Maps

Maps can be powerful navigational aids because they allow people to extract spatial information very quickly. A lot of work has been done studying how to design maps, as well as when and how to display them. Boff and Lincoln (Boff and Lincoln 1988) have studied 'you-are-here' maps. These are maps that are placed throughout an environment and indicate a navigator's current orientation and position. They have determined three fundamental design principles for these types of maps:

Two Point Theorem A map reader must be able to relate two points on the map to the corresponding two points in the environment. This will orient the space properly and allow the map to be used for navigation.

Alignment Principle A map should be aligned with the terrain. A line between any two points in the space should be parallel to the line between those two points on the map.

Forward-Up Equivalence Principle A map should always be oriented so that the up direction (assuming the map is viewed perpendicular to the floor) always shows what is in front of the map reader.

As well, it is generally a good idea for a map to show all of the organizational elements from the environment that it represents. For example, features such as paths, landmarks, and districts should be easily identifiable on the map. Ideally, a map should also show the navigator's current position in the environment. Unfortunately, this is hard to achieve when a map is being used in a real world environment. The forward-up equivalence principle is also difficult to include in maps used in real environments. However, both of these design principles are easy to incorporate into maps used in virtual environments.

As mentioned previously, there is a problem with a map when a person only views it from one perspective. Typically, a person is best able to use the information provided by a map only when they have the same orientation as the map had when it was viewed (Levine, Marchon, and Hanley 1984; Warren, Scott, and Medley 1992). People have tried to solve this problem by creating maps that rotate so that the forward-up equivalence principle is maintained and the map is viewed from many different perspectives. Although rotating maps are able to provide an orientation free representation of the environment they are not without limitations. North-up maps are better for allowing people to learn the locations of landmarks on a map, as well as the relative locations of these landmarks in the environment (Wickens 1992).

2.7.3 Comparison of Navigational Aids

Numerous experiments have been run that attempted to compare the effectiveness of different navigational aids. There is some evidence that both signs and verbal directions are effective navigational aids for a person who is attempting to find a goal, whereas a map can be disadvantageous. Streeter et al., (Streeter, Vitello, and Wonsiewicz 1985) have found that while driving it is more difficult to interpret a customized route map (a static map showing only information relevant to the particular route) than it is to receive taped verbal directions. Participants who used a map made 67% more errors and drove 11% more miles than participants who received taped verbal directions. This may be because viewing a map and driving both compete for the same type of working memory. They also found that taped verbal instructions alone are more effective than a combination of a map and taped instructions. Participants who used a map and received verbal directions made 46% more errors and drove 5% more miles than participants who received only the verbal directions. This suggests that providing more navigational tools does not always mean performance will improve.

Butler et al., (Butler, Acquino, Hissong, and Scott 1993) performed an experiment comparing signs with 'you are here' maps. In their experiment they compared the effect that signs and 'you are here' maps have on wayfinding performance. Subjects who used the signs took the shortest amount of time to find a room, less than 1 minute, followed by subjects who had no aids provided. The subjects who used the 'you are here' maps took the longest amount of time, approximately 6 minutes. Some of the extra time taken by the map group subjects was spent studying and encoding the map.

Satalich (Satalich 1995) found that participants who were allowed to use a map during exploration and wayfinding in a virtual environment performed slightly worse than participants who did not have access to a map. In her study, points were given to participants according to how efficient a route they took during a wayfinding task. It was possible to obtain a maximum of 4 points, and participants who used a map scored an average of 1.3 points. Participants who did not have access to a map performed slightly better, scoring an average of 2.0 points. Satalich believes that the map may have interfered with subjects by drawing some of their visual attention away from the environment and to the map.

It appears that whenever a task involves high cognitive spatial loads the use of a map can be disadvantageous. In these cases signs and verbal directions appear to be better navigational aids. However, as mentioned previously, signs can relieve the cognitive load to a point where survey knowledge is inhibited.

2.7.4 Navigational aids for virtual environments

A significant amount of work has been performed by Darken and Sibert (Darken and Sibert 1993) attempting to apply real world navigational techniques and tools to virtual worlds. Table 2.1 (Darken and Sibert 1993) shows a list of real world tools and techniques that they have attempted to incorporate into large virtual environments:

The acoustic landmark was a spatial audio cue that was set to a steady positional tone. It was not audible throughout the entire environment, but when it became

Technique	**Real World Analog**
Flying	Avian navigation
Spatial audio	Avian landmarking
Breadcrumb markers	Trailblazing
Coordinate feedback	Global position indicator
Districting	Urban environmental cues
Landmarks	Urban environmental cues
Grid navigation	Contour map orientation
Mapview	Map organization & presentation methodologies

Table 2.1: Navigation Techniques

The above table lists all of the navigation techniques studied by Darken & Sibert(Darken and Sibert 1993)

audible people used it for rough direction finding. It had the effect of enlarging the target area. The breadcrumb markers could be dropped either manually or automatically and landmarks were made distinct and placed randomly. A textual readout in either Cartesian or polar coordinates of a participant's current position was used as the coordinate feedback system.

After an informal study of these tools and techniques they came to several conclusions. They noticed that people usually take advantage of environmental cues in predictable ways, such as using them to partition spaces. As well, participants appeared to be able to combine visual and auditory cues to make targets easier to locate. The navigational tools also appeared to influence a person's behavior. Participants exhibited different behaviors depending on the tools being used. The results from this study indicate that it may be possible to incorporate real world navigational tools and techniques into virtual worlds.

2.8 Comparison of Navigation in Virtual and Real Environments

It is unclear whether navigation in real environments is the same as navigation in virtual environments. It appears that there may be several differences. For example, Satalich (Satalich 1995) found that in a virtual environment that represented a real environment people underestimated the true distances of objects, although she did not provide any quantitative measures so it is unclear how extreme this effect was.

One important question yet to be answered is whether navigation of virtual environments provides a person with primary or secondary survey knowledge. A study by Tlauka and Wilson (Tlauka and Wilson 1996) gave evidence that people can acquire primary knowledge from navigating in a virtual environment. However, a recent study performed by Satalich (Satalich 1995) found that people who navigated in a virtual environment either did the same or slightly worse than people who had only seen a map of the environment when performing navigational tasks. The navigational tasks that people had to perform included a directional pointing task, a route distance estimation task, a Euclidean distance estimation task, and several wayfinding tasks.

The virtual environment in her study was a U shaped building and was displayed to subjects with a head mounted display (HMD) that had a 100 degree horizontal by a 30 degree vertical field of view. Subjects moved in the direction their head was pointed and a 3-D joystick was used to change position and orientation. The low fidelity of the virtual environment used in the Satalich study may have been one factor that limited the ability of people to acquire navigational knowledge. The intrusiveness of the VR interface may have caused people to behave unnaturally and served as a distraction that reduced the fidelity of the VR experience. Problems that current VR systems suffer include sluggish response, high weight, low resolution, and low field of view of an HMD, and the artificiality of using hand motion to simulate walking.

Enhancing a virtual environment with auditory cues may be one way to increase the amount of navigational information that is present. Since it will probably be several years before it is possible to generate high-fidelity environments, it is important to determine what kind of navigational aids can be incorporated into low fidelity environments.

2.9 Spatialized Sound

2.9.1 Definition

Begault (Begault 1994) defines spatialized sound as a technique where:

> "...the outer ears are either directly implemented or modeled as digital filters. By filtering a digitized sound source with these filters, one can potentially place sounds anywhere in the virtual space about a headphone listener."

The ability to perceive sounds in space is called spatial perception. Generally, an egocentric frame of reference is used to describe the position of a sound, which means that the distance and orientation of a sound are given with respect to the listener's position. The reference point is located in the middle of the listener's head, between the ears and at approximately eye level. Figure 2.1 illustrates the terms used to describe the position of a sound in space (Begault 1994).

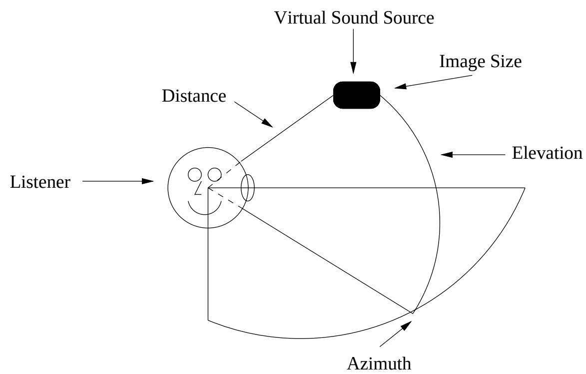

Figure 2.1: Spatial Hearing

The above figure illustrates the terms used to describe the position of a sound in space. Azimuth and elevation describe the perceived position of a sound in terms of its location on the surface of a sphere surrounding a listener's head.

The orientation of a sound source is usually given in degrees and described using two terms, azimuth and elevation. As can be seen in Figure 2.1, azimuth and elevation describe a sound's perceived position in terms of its location on the surface of a sphere surrounding the listener's head. Zero degrees azimuth and elevation describe the point directly in front of the listener. Azimuth increases in a counterclockwise direction around the listener's head, and elevation increases upwards or downwards ninety degrees to a point directly above or below the listener. Azimuth and elevation do not give any indication of the perceived distance for a sound, so a measure of distance must also be included when describing the position of a sound source. Since the human ears are located at opposite positions on either side of the head it is easier for humans to determine the orientation of a sound in terms of azimuth rather than elevation.

2.9.2 Localization Cues

People rely on eight different cues when attempting to localize a sound in space (Burgess 1992).

Interaural Time Difference

Interaural Time Difference (ITD) is the delay between the time when a sound reaches the closer ear compared to the time when it reaches the farther one. It is the primary cue used to determine the lateral position of a sound source. Figure 2.2 provides an example of how ITD works (Begault 1994). When the sound source is at position A the path lengths to both ears are equal, but when the sound source is at position B the path to the left ear is longer than the path to the right ear. This causes the sound wavefront to arrive at the left ear after it has already reached the right one, and this difference in path lengths is the basis of the interaural time difference cue (Begault 1994).

Interaural Intensity Difference

For waveform components smaller than the diameter of the head (i.e., frequencies greater than 1.5 kHz) there will be a significant interaural intensity difference (IID) in the sound arriving at the two ears. This is because the head is an obstacle that sounds have to pass around to reach the far ear, and the effect increases as the size of a sound wavelength becomes smaller. However, this cue does not work well for sounds with frequencies below approximately 1.5 kHz because these sounds have long wavelengths that diffract around the surface of the head, thereby minimizing the intensity differences (Begault 1994).

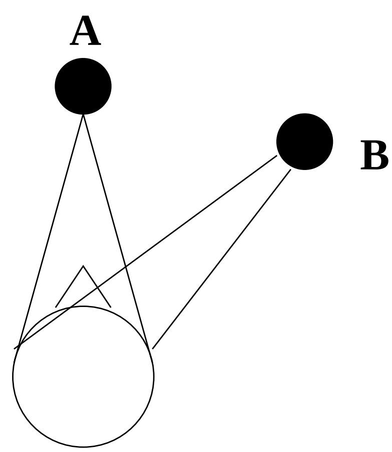

Figure 2.2: Lateralization Cues

The above figure illustrates how ITD works. The path lengths to both ears are equal for a sound source at position A, but for a sound source at position B the path to the left ear is longer than the path to the right ear.

Pinna Response

The pinna is the name given to the outer ear. The folds of the pinna cause small time delays (0-300 μ sec) that alter the spectral content of the sound source (Batteau 1968), and the asymmetry of the pinna's shape causes this spectral modification to change as a function of sound source position (Begault 1994). The pinnae only have a significant effect on sounds with a frequency greater than 4kHz and they are effective for determining both the azimuth and elevation of these sound sources (Burgess 1992).

Shoulder Echos

Although less important than the previously mentioned cues, the shoulders can provide both elevation and azimuth information for certain frequencies (approximately

1-3kHz) (Searle, Braida, Davis, and Colburn 1976). These frequencies reflect off the shoulders and upper body (Gardner 1973) and reach the ears with a delay dependent on the elevation of the sound source. As well, the effects of the reflection on the spectrum of the sound source are direction-dependent (Burgess 1992).

Head Motion

Studies have shown that people tend to move their heads to get a better sense of a sound's direction. Head movement can improve localization and reduce the number of front-back reversals (Thurlow and Runge 1967; Wallach 1940; Thurlow, Mangels, and Runge 1967). The problem of front-back reversal is discussed in Section 2.9.4.

Because of this localization strategy, any virtual reality system incorporating spatialized sounds should be equipped with headtracking to allow a person to move their head when attempting to localize a sound.

Vision

People use visual cues to help them determine where a sound is located and will often ignore auditory directional cues if they disagree with the visual cues (Burgess 1992). An example of this occurs when we watch a movie and perceive an actor's voice as coming from his mouth rather than the actual theater speakers.

Early Echo Response & Reverberation

The sounds that we hear are usually accompanied by reflections from surfaces in the environment. Early echo response (Moore 1990) is the name given to the clear echos that we hear, but do not consciously perceive, in the first 50 to 100ms after a sound starts. Reverberation (Gardner 1973) is the term used to describe the dense echos that follow the early echos. Both of these cues are thought to be important for determining the distance and direction of a sound (Burgess 1992).

2.9.3 Generation of Spatialized Sound

The spectral filtering of a sound source before it reaches the ear drum that is caused primarily by the outer ear is called the Head Related Transfer Function (HRTF) (Begault 1994). By empirically measuring the HRTF from the ears of a specific person (Wightman and Kistler 1989a; Wightman and Kistler 1989b) it is possible to artificially spatialize a sound. This is done by placing microphones near the eardrums of a person. These microphones are used to monitor a sound as it reaches the eardrums, and take into account all the effects of the HRTF for that person. Speakers are placed at various locations around the person and then bursts of pseudorandom noise are played through the speakers. By comparing the original and recorded waveforms, the subject's HRTF can be computed for all of the different positions of the sound source. This is done by using digital filter algorithms to transform the spectrum of the sound source in the same way it would be transformed under normal spatial hearing.

2.9.4 Problems with Spatialized Sound

When sounds are artificially generated, if there are a lack of externalization cues then the sounds will all appear to come from points inside a person's head and the ability to localize the sounds is lost. Loomis (Loomis, Hebert, and Cicinelli 1990) ran a study to determine whether virtual sounds produced without detailed pinna modeling could be perceived as external to the head. In his study Loomis had participants attempt to walk to both real and virtual sound sources. Performance in the virtual condition compared favorably with that in the real condition. Participants in the virtual condition were only 6% slower and were actually 13% more accurate than participants in the real condition. Loomis concluded that a simple virtual display that approximates most of the primary cues of range and azimuth can be effective in creating an impression of external sounds that people can readily find.

Front-Back Confusion

Front-Back confusion occurs when a person is unable to determine whether a sound source is located in front of them or behind them. This problem occurs for both real and artificially generated sounds. It is caused by the roughly spherical shape of the head and the fact that both ITD and IID play an important role in localizing sounds. The cause of the confusion is that identical values of ITD and IID can be calculated for sound sources that are located in opposite positions on a conical surface extending out from the ear. This conical surface is called the cone of confusion and is illustrated in Figure 2.3; identical values of ITD and IID would be calculated for sounds located at positions A and B.

The primary cues that are used to distinguish front from back are pinna response and head movement (Burgess 1992).

NonIndividualized Head Related Transfer Functions

One of the problems with artificially spatialized sounds is that the HRTFs used are measured from the sounds arriving at one particular person's eardrums. Thus, anyone else listening to the generated sounds is hearing them through another per-

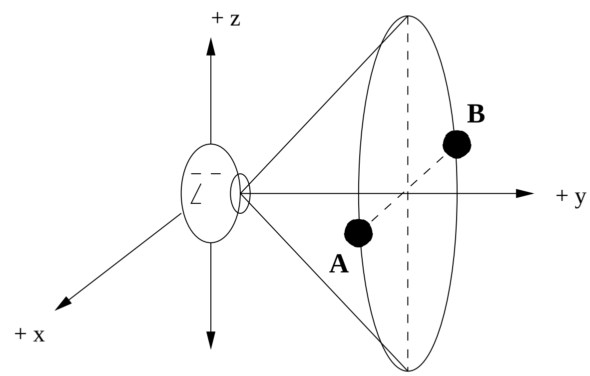

Figure 2.3: Cone of Confusion

The above figure illustrates why front-back reversal of a sound source occurs. The sound sources located at positions A and B have identical values of ITD and IID.

son's ears. Fortunately, Wenzel et al., (Wenzel, Arruda, Kistler, and Wightman 1993) have found that one set of HRTF filters may be suitable for a large portion of the population. They ran a study looking at individualized versus nonindividualized HRTFs and found that the localization of stimuli based on nonindividualized HRTFs only became slightly worse for a good localizer as long as the HRTF was derived from another good localizer. Large errors were made in judging elevation when the stimuli from a poor localizer's HRTF were used. The errors in elevation appeared to be correlated with the presence or absence of elevation-dependent, acoustical features in the 5 to 10 kHz region of a subject's HRTFs. Even when listening through a good localizer's ears, people with poor localization ability were unable to improve their elevation accuracy.

The studies performed by Wenzel et. al, suggest that it is possible for listeners to obtain useful directional information from an auditory display, particularly for the horizontal dimension, without requiring individually tailored HRTFs.

2.9.5 Advantages of Spatialized Sound

There are several reasons why it is important to incorporate sound into computer applications. Virtual reality developer Brenda Laurel gave the following statement regarding the use of audio in video games (Tierney 1993):

> "...really high-quality audio will actually make people tell you that games have better pictures, but really good pictures will not make audio sound better; in fact, they make audio sound worse. So in the (virtual) world we're building, one of the features is a rich auditory environment."

A study conducted by Hendrix (Hendrix and Barfield 1995) found that when spatialized sound was added to a stereoscopic display it increased a user's sense of presence. Presence is the degree to which a person feels as if they are actually in the environment. Hendrix found that spatialized sound increased a user's sense of presence more than non-spatialized sound. However, he also found that including spatialized sound in a virtual environment did not increase a person's sense of realism. Realism is a term used to describe how accurately a virtual world reflects the real world.

Sound can also provide information that is redundant to that provided by the visual cues. This makes it less likely that a person will misinterpret or lose important information. As well, a person's focus of attention for sound can be switched at will unlike that for vision, which often requires eye or head movement. This is useful when many things need to be monitored or attended to at one time. Spatialized sound can also provide information about things that are happening in a 360 degree circle around a person, unlike vision which can only provide information for roughly 150 degrees.

It has only recently been possible to generate 3D sound in real time so almost no work has explored the use of 3D sound in practical applications. However, now that 3D sound systems are becoming less expensive, people are beginning to consider using 3D sound to enhance real world applications. A recent study conducted by Begault (Begault 1993) found that when commercial airline crew members used a 3D auditory display to acquire spatial information about surrounding aircraft they were approximately 2.2 seconds, or 47%, faster in locating aircraft than crew members who used only one-earpiece headsets.

The study performed by Darken (Darken and Sibert 1993) appears to be the only work so far that has attempted to examine the usefulness of 3D sound as a navigational aid. However, 3D sound was not the main focus of the Darken study. The intent of this thesis is to examine more closely how the addition of 3D sound to a virtual environment affects wayfinding and the acquisition of spatial knowledge.

Chapter 3

Experiment Design

3.1 Research Objectives

The main objective of this thesis is that assigning spatialized sounds to specific objects in a virtual environment will decrease the amount of time it takes people to locate these objects. It is also the goal of this research to determine if the use of these auditory cues to locate objects affects the number of objects from the environment a person is able to remember, as well as the amount of survey knowledge they are able to acquire.

To answer these questions a visually sparse virtual environment was constructed, and then spatialized sounds were placed in specific locations throughout the environment. Participants were required to explore the virtual environment, and then locate specific objects from the environment as quickly as possible. After locating all of the required objects, participants were asked to recall objects from the environment, and then place each object in its correct location on a map of the environment. There were three different groups of participants, each of which experienced the virtual environment under different sound conditions. One group of participants experienced the virtual environment with spatialized sounds present during the entire experiment, another group never heard any sounds, and the third group only heard spatialized sounds during part of the experiment.

The emphasis of this study was not to compare the performance of visual and auditory cues as navigational aids, but to examine the effect that spatialized auditory cues have on wayfinding and the acquisition of the three different types of spatial knowledge. The environment was intentionally designed to be visually sparse so that the effects of the spatialized sounds would be easier to isolate.

3.2 Equipment

Two Pentium^{TM} PCs were used in this study. A machine with a 133 MHz processor and 64 megabytes of RAM was used to generate graphics and control movement through the virtual environment. This computer was running Windows NT^{TM} 3.51 and had a Fujitsu Sapphire 2SX^{TM} graphics card installed in it[1]. The Sense8^{TM} 3D authoring package, WorldUp^{TM}, was used to construct the virtual environment, and the WorldUp Player was used to allow people to navigate throughout the environment[2].

The second PC was used to produce the spatialized sound. This machine had a 60 MHz processor, 32 megabytes of RAM, and was running Windows95^{TM}. The QMixer95 realtime 3D audio mixer[3], donated by QSound, was used to generate the

[1]See Appendix A for technical specifications.

[2]See Appendix A for technical specifications.

[3]See Appendix A for technical specifications.

spatialized sounds. A Gravis UltraSound MAX^{TM} soundcard was used to play the spatialized sounds.

The two PCs were connected with an ethernet cable, and all information was passed between the two machines via this cable. During initialization of the virtual environment, the PC generating the graphics sent the location of the virtual sound sources to the PC producing the spatialized sounds. The graphics computer also continually sent a person's position and orientation to the sound processing PC.

A head-mounted display, the Virtual i-glasses^{TM}, was used to display the virtual environment to participants during the experiment[4]. The Virtual i.O HMD had a 30 degree field of view, and a resolution of 180,000 pixels per LCD panel. It was also equipped with a pair of stereo headphones, through which spatialized sounds from the environment could be heard. The virtual environment was also displayed on a computer monitor, which was watched by the experimenter so that the movement of participants through the virtual environment could be studied.

The head-mounted display was equipped with a headtracker containing three degrees of freedom, yaw, pitch, and roll. These are illustrated in Figure 3.2. However, for this study pitch and roll were disabled. The view of the virtual environment was only changed by a left or right head rotation. Pitch and roll were disabled because during the experiment there was no reason for participants to look up or down, and it was felt that limiting the amount of head motion would prevent people from developing motion sickness.

A 2-button mouse was also used to allow participants to move through the virtual environment. Movement during the experiment was constrained to the horizontal dimension; no movement in the vertical direction was possible. Pushing

[4]See Appendix A for technical specifications.

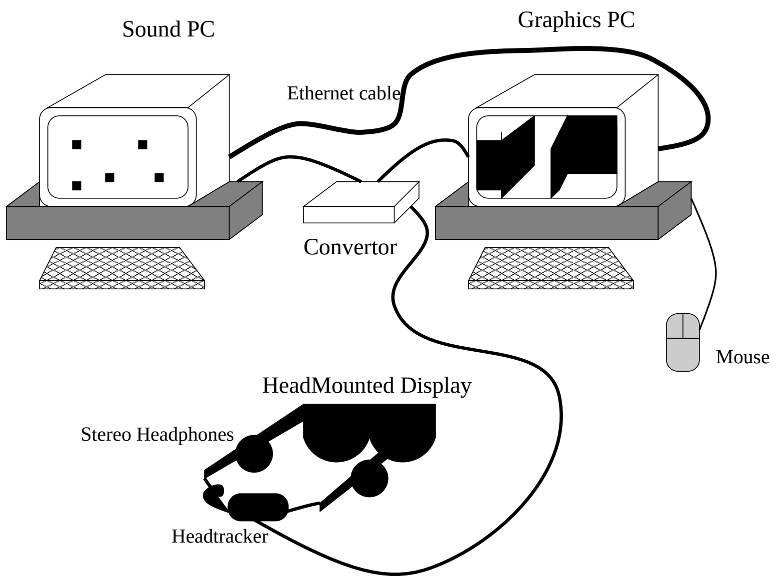

Figure 3.1: Experiment Room Setup

This is a diagram of how the equipment was setup for the spatialized sound experiment. The sound and graphics PC were connected with an ethernet cable, and the HMD had an auditory connection to the sound PC and a video connection to the graphics PC.

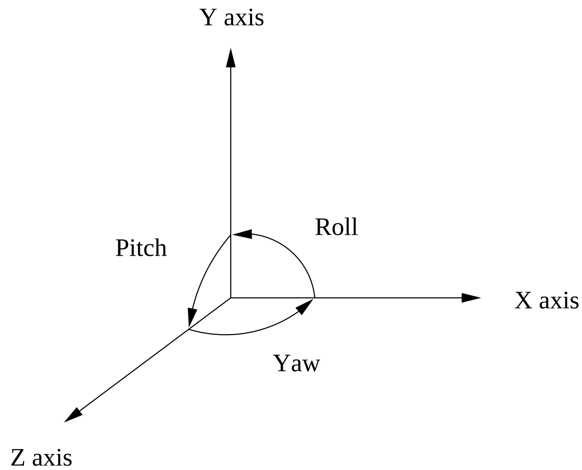

Figure 3.2: Headtracker Orientation

This diagram illustrates the three degrees of freedom for the head-mounted display, yaw, pitch, and roll.

the left mouse button had the same effect as a left head rotation, and pushing the right mouse button was equivalent to a right head rotation. Simultaneously holding down both buttons allowed a participant to move forward; no sideways or backwards motion was possible. All possible movement could be accomplished with the mouse; it was never necessary for people to rotate their heads, although they were free to do so at any time. A small blue sphere was used to indicate the current direction of forward motion and was always visible in the middle of the screen.

An audio cassette recorder was used to capture what was spoken by participants during tasks where they were asked to verbalize their thinking.

3.3 Environmental Design

The virtual environment was designed to be visually sparse so that the effects of the spatialized sounds would be easy to isolate. An effort was made to design an environment that was sufficiently complex, given the amount of time participants were immersed within the environment, to make navigation challenging.

Participants were not provided with a stereoscopic view of the virtual environment because the Sense8 software did not support this. This was not considered a problem for our study because stereopsis is not crucial in architectural modeling. Fred Brooks (Brooks 1992) stated the following when reporting on stereopsis:

> "Stereo noticeably adds to the illusion of presence, but it makes much less difference in architectural modeling than in most applications. Architectural models have lots of parallel lines and planes, known to be perpendicular... Obscuration, perspective, and the kinetic depth effect provide so strong a set of cues that stereo is almost frosting."

Wickens et al., (Wickens, Todd, and Seidler 1989) also found that when the cue of relative motion is available the enhancement of depth perception offered by stereoscopic displays is greatly diminished.

3.3.1 Configurational Layout

The virtual environment used for this experiment was a maze that consisted of twelve distinct rooms, as well as an easily identifiable starting location. Figure 3.3 shows a map of the virtual environment. The starting location, situated in the lower-middle of the environment and indicated by an arrow, provided a large amount of visual access to other sections of the environment. There were four exits from the starting location, situated in the middle of the north, south, west, and east walls. The rooms of the environment varied in size, and ten of them had only one entrance. The other two rooms each had a pair of entrances.

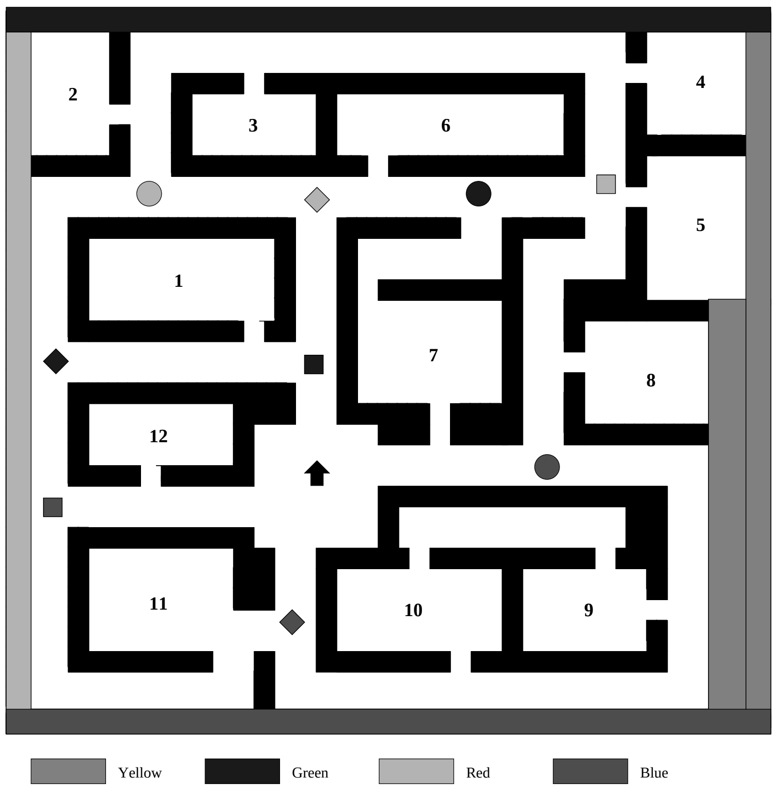

Figure 3.3: Map of Environment

This is a map of the environment that participants explored during the experiment. The initial starting location is indicated by a black arrow and the room numbers indicate the order of the rooms during the tour. The tour is described in Section 3.5.3.

The hallways in the environment had grey floors, and the floor in each of the rooms was light yellow. The ceiling was black everywhere, and all of the interior walls were brown. Each of the four exterior walls was a different color: north was green, west was red, south was blue, and east was yellow. The purpose of color coding the exterior walls was to provide participants with some visual directional information, as well as visual information informing them if they were in the interior or exterior of the environment. Colored geometric shapes were placed on walls in the environment at most of the intersection points. These shapes were squares, diamonds, and circles, and were one of three colors, red, green, or blue. There were a total of nine different geometric landmarks and there were also several objects placed in each of the 12 rooms.

3.3.2 Objects and Sounds

There was exactly one spatialized sound placed in each of the twelve rooms of the environment, and inside each of these rooms was at least one object that was clearly associated with the room sound, for example a dog in the room with the barking sound. As well, placed in each room was at least one object that had no obvious association with any of the sounds from the environment. Henceforth in this thesis the 12 objects that have sounds associated with them will be called Auditory objects and the objects that have no association with any sounds will be called NonAuditory objects. Table 3.1 is a list of all the objects located in the environment.

Sounds were chosen to be as distinctive as possible, and to have a strong association with specific objects in the environment. Each sound could be classified into one of three categories: musical, animal, or object. The category of each sound, as well as the room it was located in, is listed in table 3.2.

All of the sounds were spatially localized in the environment, and the volume of each sound decreased with distance. None of the sounds could be heard throughout

Room Number (refer to figure 3.3)	**Object Description**	**Object Classification**
1	Violin	Auditory
	Hammer;Bar Table;Bar Stools*(2)*	NonAuditory
2	Telephones*(2)*	Auditory
	Desks*(2)*;Bookshelf;Sunglasses	NonAuditory
3	Helicopter	Auditory
	Baseball Bats*(3)*	NonAuditory
4	Bowling Alley	Auditory
	Garbage Cans*(2)*	NonAuditory
5	Birds*(4)*	Auditory
	Balloon;LifeGuard Chair	NonAuditory
6	Ping-Pong Table	Auditory
	Dart Board	NonAuditory
7	Horses*(3)*	Auditory
	Picnic Tables*(2)*;Mailbox	NonAuditory
8	Keyboard	Auditory
	Chairs*(2)*;Present	NonAuditory
9	Frogs*(4)*	Auditory
	Plants*(2)*	NonAuditory
10	Swingset;Monkey Bars; Teeter-Toter	Auditory
	Bench	NonAuditory
11	Dogs*(4)*	Auditory
	Cage;Barbell	NonAuditory
12	Clocks*(3)*	Auditory
	Coffee Table;Couches*(2)*;Teapot	NonAuditory

Table 3.1: Objects in the Environment

The above table lists all of the objects in the environment and classifies each object as either a Auditory or NonAuditory object. The numbers in brackets indicate how many of each particular object exist in the environment.

Sound	**Category**	**Room (refer to figure 3.3)**
Violin Music	Musical	1
Telephone Ringing	Object	2
Helicopter	Object	3
Bowling Alley	Object	4
Seagulls	Animal	5
Ping-Pong Game	Object	6
Horses	Animal	7
Keyboard	Musical	8
Frogs	Animal	9
Laughing Children	Animal *(Human)*	10
Barking Dogs	Animal	11
Ticking Clock	Object	12

Table 3.2: Spatialized Sounds

The above table lists all of the spatialized sounds in the environment and also lists which room each sound is located in.

the entire environment, but sounds located near the middle of the environment could be heard throughout more of the environment than sounds located near one of the exterior walls of the environment. The walls in the virtual environment were "transparent" with respect to the sounds, that is they had no effect on the sounds at all. However, when a subject entered a room the volume of the sound source for that room was increased so that it was obvious which sound source was located in that room.

Objects in the environment were either 3D models, or images texture mapped onto rectangular polygons. Texture mapped images were used instead of complex polygonal models to increase performance and achieve a faster rendering frame rate. Figure 3.4 shows two screen shots from the virtual environment. In the left image there are polygonal models of couches, a coffee table, and a teapot. On the back wall are three images of clocks that have been texture mapped onto rectangular polygons. The image on the right shows a room with a picnic table, mailbox, and two texture mapped images of horses.

3.3.3 Collision detection

Collision detection was implemented so that participants were unable to walk through walls or room objects. In the original implementation of collision detection a person's movement stopped when they collided with an object. However, this made movement awkward so the collision detection was redesigned so that when a person collided with an object or wall they would not come to an abrupt halt, but would slide along the object or wall.

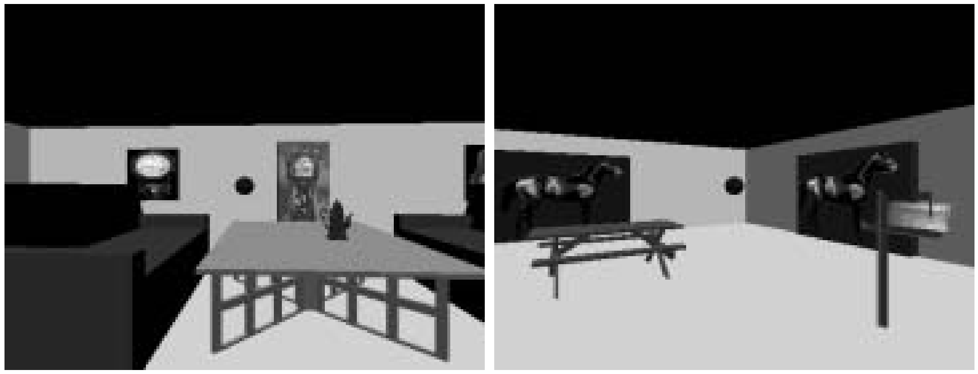

Figure 3.4: Screen Shots from the Virtual Environment

Above are screen shots of two rooms in the environment. The picture on the left is a screen shot of the clock room and the picture on the right is a screen shot of the horse room.

3.4 Pilot Study

This experiment received approval from the Office of Human Research and Animal Care at the University of Waterloo. Before running the actual experiment, a pilot study was performed to test several things: equipment usability, recognizability of objects and sounds, environmental complexity, usability of visual and auditory cues, and appropriateness of paper and pencil tasks. After observing participants during the pilot study and listening to their comments it was decided that several changes would be made to the experimental design. A detailed description of these changes can be found in Appendix D.

Except for several minor changes the experimental design appeared to be adequate. Participants of the pilot study had little trouble recognizing objects or sounds from the environment, and they were able to move around the environment without difficulty. As well, the environmental design seemed to be appropriate. Within the time frame of the experiment most of the participants were able to acquire some spatial knowledge, but the configuration of the environment was complex enough that participants found most of the tasks that tested spatial knowledge challenging.

3.5 Spatialized Sound Experiment

The experiment took place in a small, quiet room situated in the Computer Graphics Lab at the University of Waterloo. During navigation of the virtual environment the lights were turned off, and participants sat in a chair in front of a desk and wore a head mounted display. As previously described, a computer mouse was used to move through the environment.

3.5.1 Participants

Participants were primarily[5] graduate and undergraduate students from the University of Waterloo, ranging from 16 to 35 years of age. A total of 50 people participated in the experiment, 36 men and 14 women. 12 participants were members of the Computer Graphics Lab at the University of Waterloo and were not paid since members of the Graphics Lab are required to volunteer for experiments. The remaining 38 participants were each paid fifteen dollars at the end of the experiment.

[5]Two participants were not students from the University of Waterloo

3.5.2 Conditions

Although it was expected that there would be large individual differences among participants with respect to their spatial and navigational abilities it was decided to run a *between-subject* experiment instead of a *within-subject* one. There were two reasons why a between-subject design was chosen rather than a within-subject design: First, it was important to limit the amount of time that participants spent immersed in the virtual environment so that they would not become fatigued and motion sick. Running a within-subject experiment would either have increased the total time participants spent immersed in the virtual world or limited the amount of data that could be collected within each condition. Second, it was felt that participants would learn their environment very quickly and also that the route they took on one object search task would effect how well they were able to perform on a subsequent one. It was felt that these learning effects would cause a problem in a within-subject experiment so a between-subject experiment was chosen even though it was expected that there would be a large variance within each condition.

Participants were randomly assigned to one of three conditions: NoSound, FullSound, or PartialSound. An effort was made to balance the conditions according to gender since previous research has found evidence of gender differences in both wayfinding and the acquisition of spatial knowledge (Lawton 1994; Galea and Kimura 1993; Ward, Newcombe, and Overton 1986; Bryant 1982).

The only difference among the three conditions was the amount of time that spatialized sounds were heard during the experiment. Participants placed in the NoSound condition were only presented with the visual environment and never heard spatialized sounds. Participants in the FullSound condition were presented with both the visual and auditory environment and were able to hear sounds during the entire experiment. In the PartialSound condition participants were able to hear sounds during most of the experiment, except when searching for the second set of objects during the Object Search phase of the experiment.

As well, the NoSound and PartialSound conditions were each composed of two groups of participants, an A and a B group. These groups are labeled PartialSoundA, PartialSoundB, NoSoundA, and NoSoundB throughout the remainder of this thesis. Within each condition, the only difference between the A and the B participant groups was the order of the objects that participants had to find during the Object Search phase of the experiment. Participants were required to locate a total of eight objects during that phase of the experiment. The eight objects were split into two sets of four, ObjectSetA and ObjectSetB. Participants in an A group had to locate the objects in ObjectSetA and then the objects in ObjectSetB. Participants in a B group had to locate the objects in ObjectSetB and then the objects in ObjectSetA. These two sets of objects will be described in more detail later in this chapter.

Reversing the order of object search for the two groups helped control for two different kinds of effects that may be present: learning effect, and the possibility that one group of objects was easier to locate than the other.

Participants in the FullSound condition were not split into an A and a B group because the main purpose of including the FullSound condition was not to examine search times. The FullSound condition was included primarily to see if the addition of spatialized sounds to a virtual environment has an effect on the acquisition of survey knowledge.

3.5.3 Procedure

The experiment took approximately 1 1/2 hours and each participant was asked to perform several tasks during the course of the experiment. Below is the order of the tasks, as well as the approximate time for each task.

1. Participant Background Questionnaire (2 minutes).
2. Equipment Training (5 minutes).
3. Map and Tour (5 minutes).
4. Unguided Exploration (8 minutes).
5. Object Search - Phase One (8 minutes).
6. Object Search - Phase Two (8 minutes).
7. Object Recall & Placement (10 minutes).
8. Guliford-Zimmerman Orientation Test (10 minutes).
9. Final Questionnaire (5 minutes).

The instructions for each task were read to participants from a script[6]. This was to ensure that all participants received identical instructions, although participants were allowed to ask questions if they were confused about something.

[6]Shown in Appendix B.

Background Questionnaire

At the beginning of the session each participant was read a brief introduction and description of the experiment. Then a letter of information, along with two copies of a consent form were given to each participant[7]. Once the consent form had been signed, participants were given a short background questionnaire[8] that recorded the following information about a participant:

1. Age.
2. Gender.
3. Dominant hand.
4. Any known hearing Impairments.
5. Any known visual Impairments.
6. Any known problems distinguishing between colors.
7. Hours per week spent playing DOOM©.
8. Hours per week spent using a computer.
9. Hours per week spent listening to music.
10. Hours per week spent playing a musical instrument.
11. Familiarity with virtual reality.
12. Assessment of navigational ability compared to other people.

[7] The letter of information and consent form are shown in Appendix B.

[8] Shown in Appendix B

Clearly, any kind of auditory or visual impairment might influence a participant's performance on certain tasks of this experiment. It was possible that the amount of time participants spent listening to music or playing a musical instrument would have an effect on how useful the spatialized sounds were. Experience with computers and familiarity with virtual reality were two other factors that might affect the performance of participants. As well, since the virtual environment that was designed was similar to the environment used in the computer game DOOM, it was felt that the amount of time participants had spent playing DOOM could affect their performance. None of this information was used as a screening tool; the purpose of gathering this information was to provide some indication of how balanced the three experimental conditions were with respect to relevant participant demographics.

Equipment Training

After participants were finished filling out the background questionnaire they were given a training session to become comfortable using the headmounted display and mouse for moving around in the virtual environment. The headmounted display and mouse were used to move through the environment as described in Section 3.2. The purpose of the training session was to ensure that all participants had a similar level of competence using the equipment and moving around in the environment. A small virtual environment, consisting of five rooms, was used for the training environment. Participants were permitted to practice using the equipment for as long as they wanted. They were told that once they felt comfortable moving around the environment they should remove the head mounted display, at which point the instructions for the next task would be read to them.

Map and Tour

Once participants finished the training session they were shown a map of the environment that they would be immersed in during the rest of the navigational tasks. This map was shown in Figure 3.3[9]. As described in Section 3.3.1, each of the four exterior walls in the environment was a different color. As well, there were nine colored geometric shapes on several walls of the environment. The geometric shapes and exterior walls were represented in color on the map that each participant received. The location and orientation of participants at the beginning of the tour, exploration, and object search tasks was indicated on the map with an arrow. Each participant was allowed to view the map while the environment was loaded into the computer, which took approximately 90 seconds. The map was then taken away from all participants.

The purpose of providing participants with this map was to reduce the amount of time that they would need in the virtual environment before they were able to acquire some survey knowledge. It was important to limit the amount of time that participants were immersed in the environment due to the concern that some participants might develop motion sickness.

Once the virtual environment was loaded into the computer participants were taken on a tour of the environment, which brought them to each room of the environment. During the tour participants were unable to control movement, although they could change their view of the environment by turning their head. The tour consisted of participants being transported to each room entrance, where they were given 14 seconds to examine all of the objects and sounds in the room.

The intent of the tour was not to provide participants with any configurational

[9] The rooms on the map that participants saw were not labeled

information about the environment, only to familiarize them with the objects and sounds in each room. During the tour participants were asked to verbalize their thoughts and these were recorded on an audio cassette. If participants were unable to recognize a particular object or sound they were asked to verbalize this. The order of the rooms was the same for each participant and participants were told prior to the tour that the order of the rooms was not significant. The starting location, indicated on the map, was both the first and last location on the tour. The numbering of the rooms in Figure 3.3 represents the order of the rooms during the tour.

Unguided Exploration

After the tour was completed participants began the unguided exploration phase of the experiment. There were two reasons why this phase was included in the experiment. First, to determine if the addition of spatialized sound to a virtual environment has an effect on peoples' exploration strategies or patterns of movement. Second, to provide participants with some time to become familiar with the configurational layout of the environment.

Movement during this phase of the experiment was exactly the same as during the practice session. Each participant was given exactly eight minutes to explore the environment. Once again, participants were told to verbalize their thinking process and this was recorded on an audio cassette.

The position of each participant was continually recorded by the computer throughout the exploration phase. As well, the environment was divided into sections and whenever a person entered a section the computer wrote the section name and the current time to a file. A map showing how the environment was divided into sections is shown in Figure 3.5.

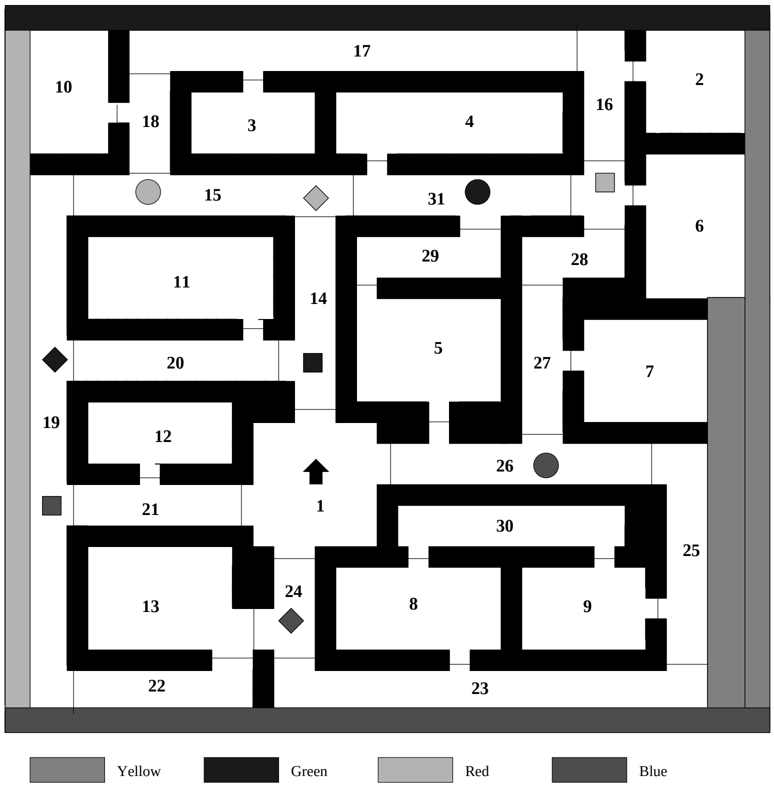

Figure 3.5: Map of the Environment Divided Up into Sections

This shows a map of the environment divided up into the sections that we used to track the movement of subjects. There were a total of 31 sections and these sections are shown numbered on the map.

The purpose of recording participants' movement and having participants voice their thoughts was so that a qualitative analysis could later be made of their exploration strategy.

After exploring the environment for six minutes participants heard a warning bell that signaled that exactly two minutes of exploration time were left. After eight minutes were up a message appeared on the screen informing participants that their time was up.

Before beginning the unguided exploration all participants were told that they should try to become as familiar with the environment as possible. They were also told that after the exploration phase was over they would be asked to find objects from the environment and then recall as many objects from the environment as possible, placing each object in the correct location on a map.

Object Search

Once participants had finished exploring the environment they were told that their next task was to locate eight objects from the environment [10]. This task was divided into two phases. Participants looked for four objects during phase1, and then took a short break before attempting to locate the last four objects during phase2. Before beginning to locate each object participants were placed at the start location, facing the north wall. A keyboard was placed on the desk in front of each subject, and pressing the <space bar> on the keyboard was the mechanism used by participants to indicate that they were ready to begin looking for the next object. Once the <space bar> was pressed, an image of the object to look for was displayed on the screen for two seconds. As well, the computer recorded the current time so that the amount of time it took a participant to locate the object could be determined. Pressing the <space bar> at any other time brought up an image of the object a participant was looking for, and this image was displayed for approximately two seconds.

Participants were told to indicate they had found an object by centering the

[10] There were a total of 60 objects in the environment

blue sphere on the object and moving into it. The computer signaled that they had found the object by displaying a message on the screen. After finding the object, participants were transported back to the start location and pressing the <space bar> indicated that they were ready to begin looking for the next object.

Participants had four minutes to locate each object. After four minutes had passed a message appeared on the screen indicating that time had expired. After this message appeared participants were placed back at the start location and pressing the <space bar> indicated that they were ready to begin looking for the next object. Participants were asked to verbalize their thinking process and this was recorded on an audio cassette. As well, the time and path taken to locate each object was recorded for all of the participants.

As stated earlier, two sets of objects were created (ObjectSetA and ObjectSetB) and each of these sets consisted of four objects. Each of these objects was clearly associated with a distinct sound from the environment. It was thought that objects located in corner rooms might be easier to locate than other objects because participants could associate a corner room with two colored walls. Therefore, included in each set of objects were two objects that were situated in a corner room. As well, an equal number of objects located in perimeter and interior rooms were placed in both sets, and two objects which could be heard from the start location were placed in each set. Table 3.3 lists the objects in each set and Figure 3.6 shows the location of these objects in the environment.

Participants in the NoSoundA, PartialSoundA, and FullSound groups looked for the objects in ObjectSetA during phase1 and the objects in ObjectSetB during phase2. Participants in the NoSoundB and PartialSoundB groups searched for the objects in ObjectSetB during phase1 and the objects in ObjectSetA during phase2. The order of the objects within each object set remained the same for all

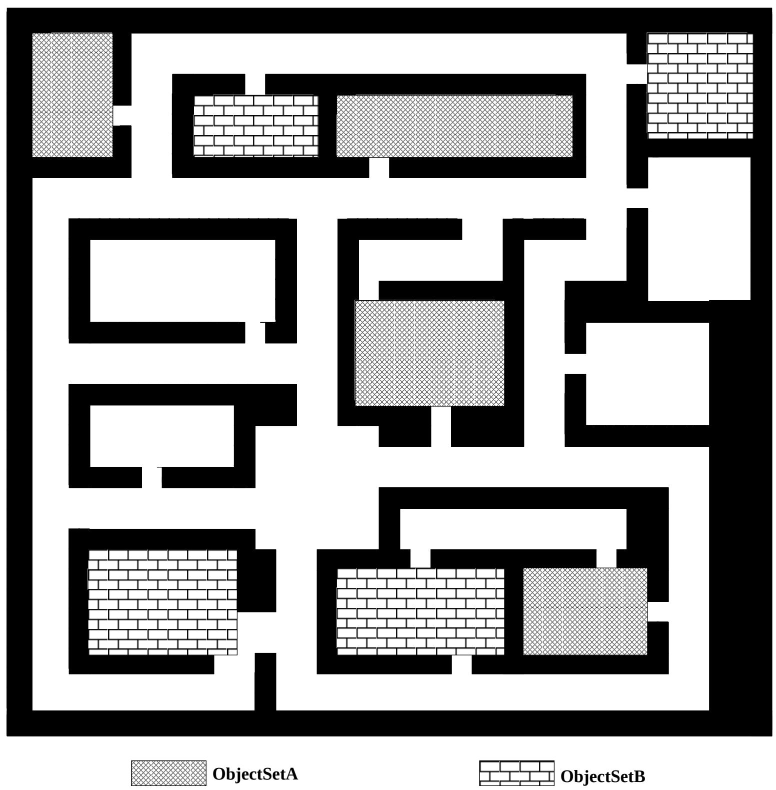

Figure 3.6: Location of Rooms for Each Object Set

This diagram illustrates which rooms in the environment contained objects from one of the two object sets.

ObjectSetA	**ObjectSetB**
Red Telephone	Dog (North wall of room)
Horse (Green Background)	Helicopter
Ping-Pong Table	Swing-Set
Frog (West wall of room)	Bowling Alley

Table 3.3: Two Sets of Objects

Some of the object types exist more than once in the environment. A description is given in brackets beside each of these objects to clarify exactly which object belonged to the set.

participants.

There were two reasons why the order of the objects was not random. The first reason was so that it would be possible to compare the routes taken by participants. Randomizing the order of the objects would have made this difficult. The second reason was to ensure that all of the objects a participant was required to search for were well spread out in the environment.

The procedure to locate the final four objects was exactly the same as the one to find the first four objects. However, all sound was removed from the environment for participants who were in the PartialSound condition. The purpose of this was to compare their time to find the final four objects with the time of participants in the NoSound condition who had never experienced sound.

Object Recall & Placement

After participants located the final object they were given the map of the virtual environment they had seen prior to the tour phase and a list of numbers[11]. Partic-

[11]Shown in Appendix B.

ipants were asked to recall as many objects from the environment as possible and write down the name of each object beside one of the numbers. For each object a participant wrote down they were asked to attempt to place it in the correct room on the map by writing down the number assigned to that object somewhere in the correct room. Once again, participants were asked to verbalize their thinking process and this was recorded on an audio cassette. Participants were told that their total time would be recorded, but that it was much more important to be accurate and write down as many objects as possible than to be fast. Participants were also told to place objects on the map even if they had to guess where an object was located. However, they were also told that if they had no idea where an object was located it was not necessary to place it on the map, although they should still include it in the list of objects.

The purpose of this task was to measure how much landmark and survey knowledge participants had been able to acquire during the time they had been immersed in the virtual environment. It would provide information about whether the addition of the spatialized sounds to the environment had an impact on the amount or type of spatial knowledge that participants were able to acquire.

Guilford-Zimmerman Orientation Test

The Guilford-Zimmerman Orientation test was given to participants so that there would be some indication of the relative spatial ability of participants in each group. Participants were given as long as they liked to read over the three pages of instructions for the test and were told they could ask the experimenter any questions if they were confused about something. Once participants were satisfied that they understood how the test would work they were given an answer sheet and told they had exactly ten minutes to do as many questions as possible. After the ten minutes were up the test was taken away from them.

Final Questionnaire

The final task participants were asked to perform was to complete a questionnaire[12] asking several questions about their experiences navigating around in the virtual environment. Participants were asked questions about:

1. Ease of movement through the virtual environment.
2. Ability to recognize objects from the virtual environment.
3. Usefulness of the sounds in helping to locate objects.
4. Disruptiveness of the sounds.
5. Difficulty in locating objects after the sound had been removed from the environment.

Participants were asked to ignore questions that did not apply to them. There was also a section for participants to write down comments.

3.5.4 Measurements Taken

The following measurements were recorded for each subject:

Participant Demographics The purpose of the participant demographics was to ensure that the three experimental conditions were more or less balanced with

[12] Shown in Appendix B.

respect to any participant characteristics that might influence performance on some of the tasks in the experiment. For example, the amount of computer experience a person had and the number of hours they had spent playing DOOM might have influenced how well they performed on the object search tasks.

Exploration and Object Search Paths The exploration and search paths of participants were recorded so that a qualitative analysis could be made to determine if the addition of spatialized sounds caused any noticeable differences in either exploration or object search strategies.

Time to Locate Each Object The time to locate each object was recorded to see if the addition of spatialized sounds to a virtual environment allowed participants to locate objects more quickly.

Number of Objects Recalled The number and types (Auditory or NonAuditory) of objects a participant was able to recall from the environment was taken as a measurement so that it would be possible to determine if the addition of spatialized sounds influenced the type or amount of landmark knowledge that a participant was able to acquire.

Number of Objects Correctly Placed on the Environmental Map The number of objects correctly placed on the map was used as an indication for the amount of survey knowledge a participant was able to acquire during their immersion in the virtual environment.

Guilford-Zimmerman Orientation Test Score The Guilford-Zimmerman test score was used to indicate if the three experimental conditions were comparable with respect to spatial ability.

Subjective Ratings Each participant was asked to complete a final questionnaire. The purpose of the final questionnaire was to measure how difficult participants found moving in the environment and recognizing objects from the environment. It was also used to get a subjective impression of how useful the spatialized sounds were and how much participants relied on them.

Chapter 4

Results

4.1 Preparing Data for Analysis

A total of 50 people participated in this experiment. Three participants experienced motion sickness while immersed in the virtual environment so the data for these participants was discarded. As well, the data for any participant whose time to locate objects was more than two standard deviations away from the mean was considered an outlier and was dropped. This caused the data for two participants to be discarded, leaving the data for 45 participants to be used in the analysis.

Table 4.1 shows the breakdown of participants in each of the conditions. These conditions were described in Section 3.5.2.

All statistical data analysis was conducted with DataDesk© version 3.0 running on a Macintosh II. All of the results reported can be assumed to have met the assumptions for analysis of variance unless otherwise stated. The significance for the results of all ANOVAs will be reported to at least 2 decimal places and the significance for the results of all t-tests will be reported as either significant at the

	NoSoundA	NoSoundB	PartialSoundA	PartialSoundB	FullSoundA
Male	6	7	7	6	6
Female	3	1	4	1	4
Total	9	8	11	7	10

Table 4.1: Breakdown of Participants

The above table shows the number of male and female participants in each condition

0.05 or 0.01 level.

4.2 Participant Demographics

4.2.1 Background Experiences

As described in Section 3.5.3, at the beginning of a session each participant was asked to complete a background questionnaire. Table 4.2 lists all of the recorded demographics and the mean scores for each experimental condition. Figure 4.1 displays a graph of the mean scores for each condition and Figure 4.2 shows a graph of the mean scores for males and females.

A one-way ANOVA was run on each of the factors to determine if there were any significant differences in participant demographics among conditions. Table 4.3 summarizes the results of these ANOVAs and includes the degrees of freedom (df), F-ratios, and probabilities that any differences occurred by chance. The results of the ANOVAs indicate that there were no significant differences in participant demographics among any of the three experiment conditions.

Factor	**Scale**	**Group**	**Mean Score**
Self-Assessment of Navigational Ability *(NAV)*	1-6	NoSound	3.6
		FullSound	3.0
		PartialSound	3.9
Average hrs/week spent playing DOOM *(DOOM)*	0-10	NoSound	0.62
		FullSound	0.56
		PartialSound	0.92
Average hrs/week spent listening to music *(MLIS)*	0-10	NoSound	6.4
		FullSound	6.9
		PartialSound	7.2
Average hrs/week spent playing a musical instrument *(MPLY)*	0-10	NoSound	1.5
		FullSound	1.1
		PartialSound	0.69
Average hrs/week spent using a computer *(CMPEX)*	0-10	NoSound	9.1
		FullSound	8.3
		PartialSound	9.0
Familiarity with Virtual Reality *(VR)*	1-6	NoSound	2.8
		FullSound	2.4
		PartialSound	2.6

Table 4.2: Participant Demographics

These are the six categories that were used to describe the participant demographics relevant to the study. The label in brackets is used to denote each factor in subsequent graphs

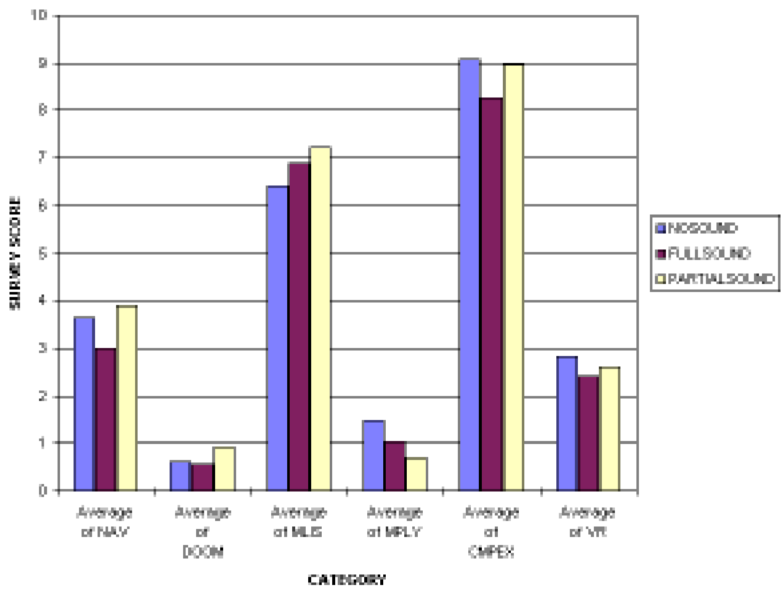

Figure 4.1: Demographics for Participants in Each Condition

These are the mean scores of the recorded participant demographics for each of the three experimental conditions. **NAV** denotes self-assessment of navigational ability, **DOOM** denotes hrs/week spent playing DOOM, **MLIS** denotes hrs/week spent listening to music, **MPLY** denotes hrs/week spent playing a musical instrument, **CMPEX** denotes hrs/week spent using a computer, and **VR** denotes familiarity with Virtual Reality.

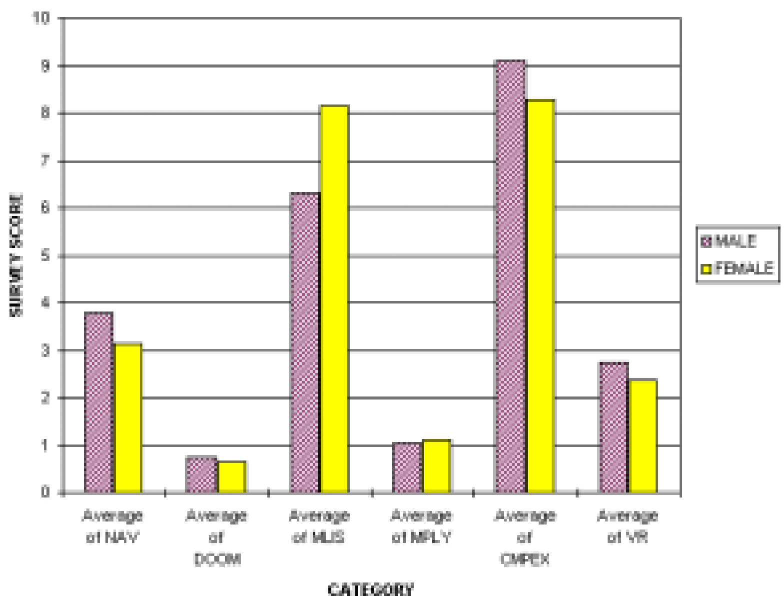

Figure 4.2: Demographics for Males and Females

These are the mean scores of the recorded participant demographics for both males and females. **NAV** denotes self-assessment of navigational ability, **DOOM** denotes hrs/week spent playing DOOM, **MLIS** denotes hrs/week spent listening to music, **MPLY** denotes hrs/week spent playing a musical instrument, **CMPEX** denotes hrs/week spent using a computer, and **VR** denotes familiarity with Virtual Reality.

Factor	**df**	**F-Ratio**	**Probability**
NAV	(2,42)	2.2	0.13
DOOM	(2,42)	0.27	0.77
MLIS	(2,42)	0.24	0.80
MPLY	(2,42)	0.61	0.55
CMPEX	(2,42)	0.46	0.63
VR	(2,42)	0.33	0.72

Table 4.3: One-Way ANOVA Results of Participant Factors Against Condition
The above table lists the results of the one-way ANOVAs that we ran on all of the participant factors against condition. The degrees of freedom (df), f-ratio, and probability that the differences were due to chance are given for each factor.

4.2.2 Spatial Ability

All of the participants wrote the Guilford-Zimmerman Aptitude Survey (GZAS) test of spatial orientation. This test had a total of 60 questions and participants were instructed to complete as many questions as they could within a 10 minute time period. The test was scored using the Guilford-Zimmerman formula: (total number correct) - (total number wrong/4). Figure 4.3 shows a graph of the mean GZAS score for participants in each condition, with each condition separated by gender. Table 4.4 lists the mean and standard deviation for each condition.

Figure 4.3 shows that there were no major differences in GZAS scores among the three experiment conditions, although there appeared to be substantial male-female differences. To determine if there were significant differences in GZAS scores between conditions or gender two one-way ANOVAs were run; the first on GZAS score against condition, and the second on GZAS score against gender.

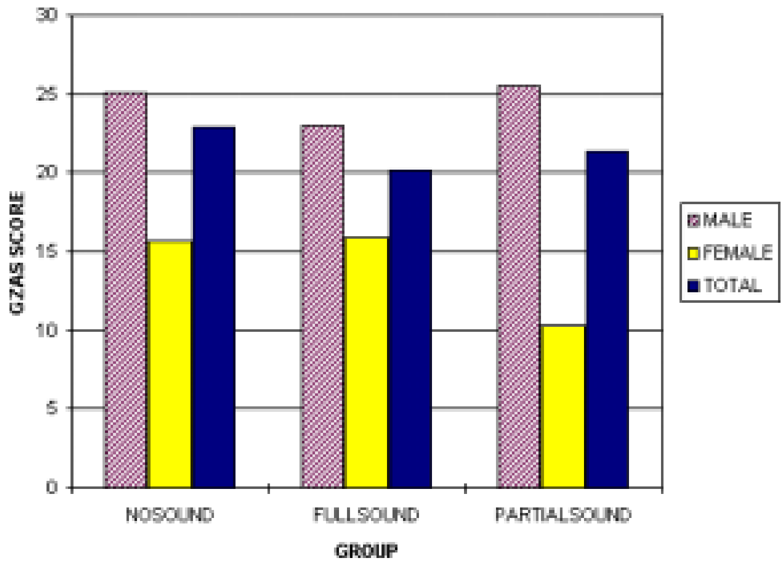

Figure 4.3: GZAS Means for each Condition

This graph displays the mean GZAS score for the males and females in each experiment condition. In all three conditions males scored higher than females.

Condition	**Mean**	**Standard Deviation**
NoSound	22.9	9.98
FullSound	20.2	7.99
PartialSound	21.3	14.2

Table 4.4: GZAS Mean Scores

The above table lists the mean and standard deviation of the GZAS score for each experiment condition.

The first ANOVA showed no significant difference in GZAS score among conditions; $F(2,42)=0.18$, $p=0.83$. The second ANOVA showed a significant gender difference; $F(1,43) = 11.1$, $p=0.002$. The means[1] for gender were: male (n=32), M=24.9, SD=11.2; female (n=13), M=13.7, SD=7.01. These are similar to the descriptive statistics for college students given by the Guilford-Zimmerman report; male, M=20.5, SD = 10.32; female, M=12.6, SD = 8.67.

4.3 Object Search Task

As described in Chapter 3, after exploring the environment for eight minutes each participant was required to find eight objects from the environment as quickly as possible, and their time to locate each object was recorded. For each participant, a mean time was calculated for the time it took to locate any one of the first four objects, and a mean time was calculated for the time it took to locate any one of the second four objects.

Collapsing the times for four individual object searchs into one mean time has the disadvantage that the analysis will not take into consideration any individual differences among the object search trials. It may have been the case that some of the objects were more difficult to locate than others, and that the effect of the addition of sound varied depending on the location of each object. For example, objects in the interior of the maze may have been more difficult to locate than objects in the exterior of the maze and therefore the addition of sound may have had a greater effect on performance when looking for those objects. However, analyzing each individual object search time was not considered appropriate because the route

[1]Throughout this thesis n denotes number of scores, M refers to Mean, and SD denotes Standard Deviation.

a participant took when attempting to find one object would have an effect on how they would perform on subsequent object search trials. For example, if their route on one object location task took them past object x, and the next object they were required to locate happened to be object x, then they would probably find object x very quickly. Analyzing a mean search time for a group of objects avoided this problem, although it will not be possible to determine from the results whether the addition of sound had the same effect on all types of object search trials.

As well, there were two reasons why the data for the first and second set of objects was analyzed separately: first, it was felt that the addition of sound would have less of an impact after participants had spent time looking for the first four objects, second, so that the effect of removing all sound from the environment before participants in the PartialSound group began looking for the second four objects could be examined.

4.3.1 Time to Locate First Four Objects

The experimental treatment for participants in the PartialSound and FullSound conditions was identical up to the end of the object search task for the first four objects. Therefore, these two conditions were considered to be one condition during the analysis on the time to locate one of the first four objects. Henceforth, we refer to this condition as the Sound condition.

For each participant, the average time it took them to locate one of the first four objects was calculated (i.e., *(total time to locate all 4 objects)/4*). Then the mean time for both conditions was computed: NoSound (M = 90.2 sec, SD = 37.0 sec) and Sound (M = 66.9 sec, SD = 27.9 sec).

Object Order Effects

As discussed in Section 3.5.3, each participant was asked to locate objects from one of two object sets. Figure 4.4 shows the mean time for participants in both the NoSound and PartialSound conditions to locate an object from each of the two object sets. The mean time for participants from the FullSound condition is not shown in Figure 4.4 because the first set of objects was always ObjectSetA for these participants.

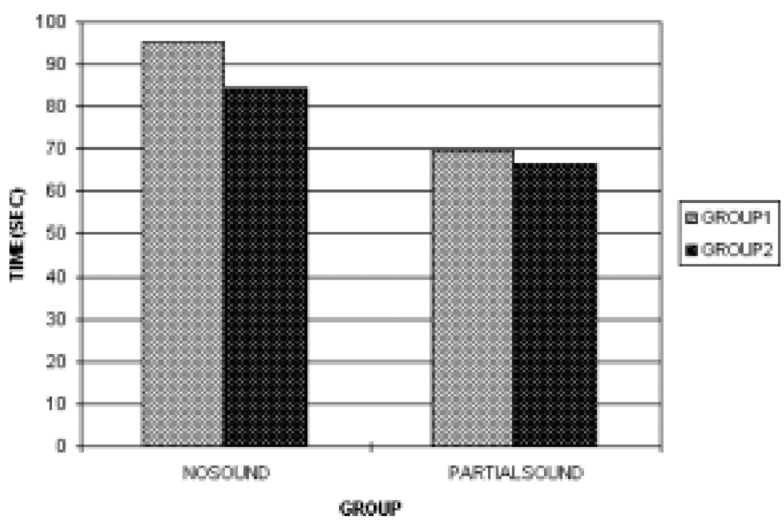

Figure 4.4: Mean Time to Locate One of the First Four4 Objects: Condition & Object Order

The average time to locate one of the first four objects was calculated for each participant. This graph shows the mean time for participants in each condition to find ObjectSetA and ObjectSetB objects.

It appeared that participants in both the NoSound and Sound conditions were able to find ObjectSetB objects slightly more quickly than ObjectSetA objects. A

2x2 ANOVA[2] was run to test if there was any significant difference in the time to locate objects from the two different object sets and to see if there was an interaction between condition and object set. The results of this ANOVA did not show any significant time difference between sets of objects; F(1,41)=0.27, p=0.60. As well, there was no significant interaction between condition and object set; F(1,41)=0.25, p=0.62.

Experiment Condition Effects

Figure 4.5 shows a graph of the first four object location time means for both the NoSound and Sound conditions, as well as the mean times for both male and female participants in each condition.

Participants in the Sound condition were able to find objects approximately 26% faster than participants in the NoSound condition. As well, in both conditions males were able to locate objects more quickly than females. This was expected since males scored higher than females on the Guilford-Zimmerman test of spatial ability.

A 2x2 ANOVA was run to test whether the time difference between the two conditions was significant, and also to test if the male-female time difference was significant. The results of the ANOVA[3] showed a significant time difference between condition; F(1,41)=4.5, p=0.04. The gender difference was not as strong as the time difference; F(1,41)=3.2, p=0.08. There was no significant interaction between condition and gender; F(1,41)=0.18, p=0.68.

[2]Figure C.2.

[3]Figure C.1.

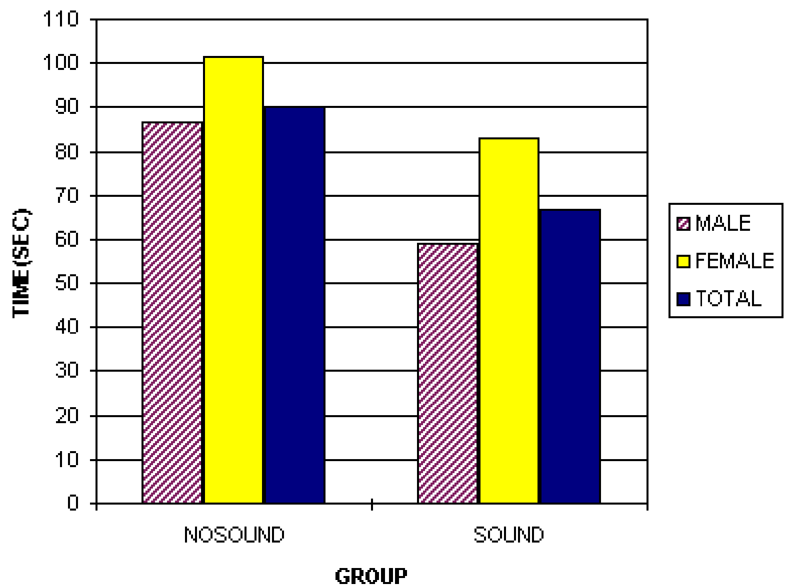

Figure 4.5: Mean Time to Locate One of the First Four Objects: Condition & Gender

The average time to locate one of the first four objects was calculated for each participant and this graph shows the mean time to locate an object for both the Sound and NoSound conditions. The mean times for both the males (Sound, n = 13 : NoSound, n = 19) and females (Sound, n = 4 : NoSound, n = 9) in each condition are shown as well.

Analysis of Male Data

Although the gender difference was not significant at the 0.05 level, it was still substantial (especially given the relatively low number of female participants), p=0.08, so it was decided to do further analysis on only the male data. A second 2x2 ANOVA[4] was run on search time (Male Data) against condition and object order. The significance of the time difference between the NoSound (M=86.7 sec, SD = 37.1 sec) and Sound (M=59.2 sec, SD = 23.3) conditions was stronger once the female data had been removed; $F(1,28)=6.9$, $p=0.01$. Male participants in the Sound condition were able to find objects approximately 32% faster than males in the NoSound condition. There was still no significant time difference between object order; $F(1,28)=1.5$, $p=0.23$ and there was no significant interaction between condition and object order; $F(1,28)=1.3$, $p=0.27$.

These results indicate that the addition of spatialized sound to a visually sparse virtual environment can decrease the amount of time it takes a person to locate objects that are associated with particular sounds in the environment.

4.3.2 Time To Locate Second Four Objects

Before participants in the PartialSound condition began looking for the final four objects all of the sound was removed from the environment. The average time it took each participant to locate one of the second four objects was calculated (i.e., *(total time to locate all 4 objects)/4*) and then the mean times for all three conditions were found, NoSound (M = 68.2 sec, SD = 32.4 sec), Sound (M = 62.7 sec, SD = 17.5 sec), and PartialSound (M=99.4 sec, SD = 53.2 sec).

[4]Figure C.4.

Object Order Effects

Figure 4.6 shows the mean time for participants in both the NoSound and PartialSound conditions to locate an object from each of the two object sets. The mean time for participants from the FullSound condition is not shown in Figure 4.4 because the second set of objects was always ObjectSetB for these participants.

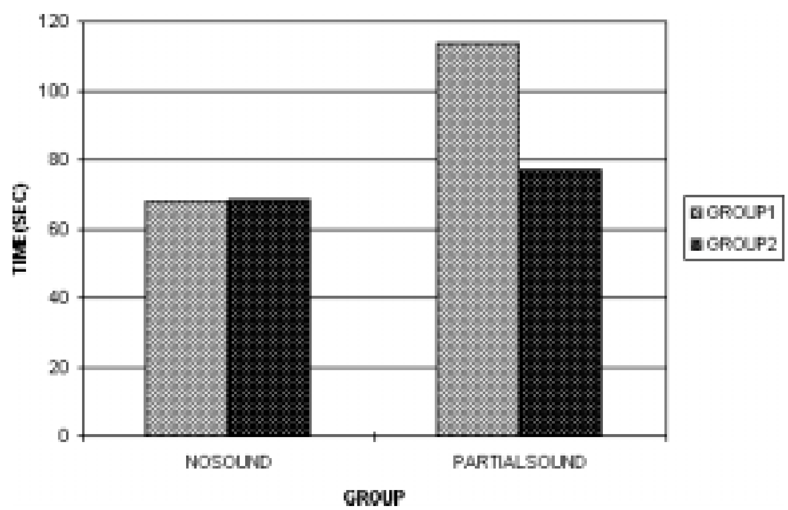

Figure 4.6: Mean Time to Locate One of the Second Four Objects: Condition & Object Set

The average time to locate one of the second four objects was calculated for each participant. This graph shows the mean time for participants in each condition to find ObjectSetA and ObjectSetB objects.

Participants in the NoSound condition appeared to take the same amount of time locating an object from both sets of objects. However, participants in the PartialSound condition were able to find objects from ObjectSetB more quickly than objects from ObjectSetA. Some of this difference may be due to the fact that there were more females in the PartialSound condition which searched for objects from ObjectSetA, and participants in this condition had a lower mean Guilford-Zimmerman spatial orientation test score (M=19) than participants in the PartialSound condition (M=24.9) who searched for objects from ObjectSetB. The mean Guilford-Zimmerman scores for the two NoSound groups of participants were approximately the same, (M=23.1) and (M=22.5).

A 2x2 ANOVA[5] was run to test if there was any significant difference in the time to locate objects from the two different object sets and to see if there was an interaction between condition and object set. The results of this ANOVA did not show any significant time difference between the two sets of objects; $F(1,39)=1.3$, $p=0.26$. As well, there was no significant interaction between condition and object set; $F(2,39)=0.89$, $p=0.42$.

Experiment Condition Effects

Figure 4.7 shows a graph of the second four object location time means for the three experiment conditions, as well as the mean times for both male and female participants in each condition.

Participants in the FullSound condition were only 8% faster at locating one of the second four objects than were participants in the NoSound condition. However, participants in the PartialSound condition were 46% slower at locating objects than were participants in the NoSound condition, and 59% slower than participants in the Sound condition. In all three conditions males took less time than females, although only in the PartialSound condition did this difference appear to be very large.

[5] Figure C.3.

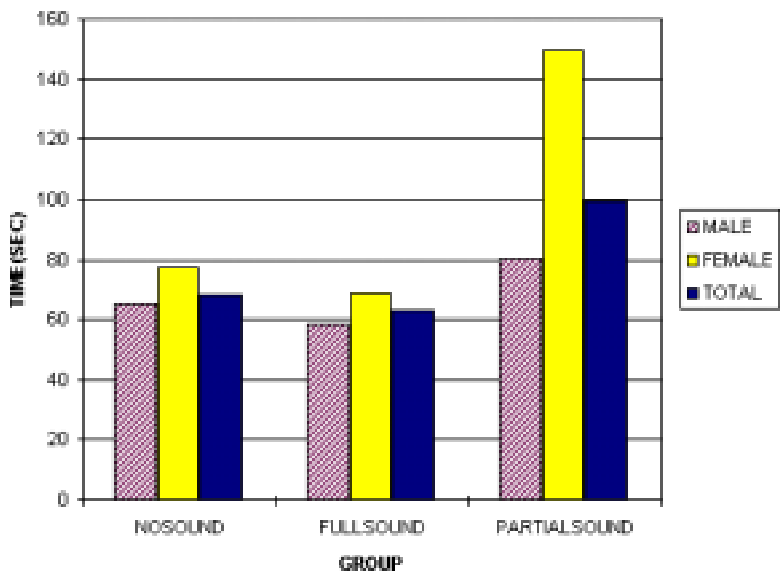

Figure 4.7: Mean Time to Locate One of the Second Four Objects: Condition & Gender

The average time to locate one of the second four objects was calculated for each participant and this graph shows the mean times to locate an object for the three experiment conditions. The mean times for both the males and females in each condition are shown as well.

A 2x2 ANOVA[6] was run on search time against condition and gender. The ANOVA showed a significant difference in search time among the three conditions; $F(2,39)=7.6$, $p=0.002$. There also appeared to be differences in search time between males and females; $F(1,39)=6.5$, $p=0.02$; and, although not significant at the 0.05 level, an interaction between condition and gender; $F(2,39)=2.9$, $p=0.07$. The interaction between condition and gender is shown in Figure 4.8.

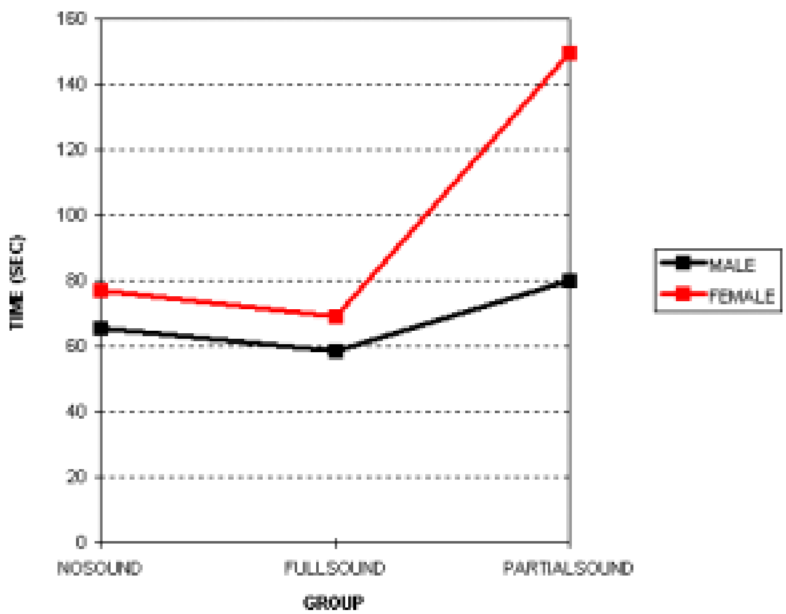

Figure 4.8: Interaction of Gender and Condition with Respect to Search Time for Second Four Objects

This graph illustrates the interaction between condition and gender on the search time for one of the second four objects. It appears that female participants were more affected by the removal of sound than were male participants.

A t-test on the average search time for second four objects was run between the

[6] Figure C.5.

FullSound and PartialSound conditions; $t(26)=-2.1$, $p < 0.05$; and also between the NoSound and PartialSound conditions; $t(33)=-2.1$, $p < 0.05$. There was not any significant time difference between the NoSound and FullSound conditions.

Analysis of Male Data

Once again, because of the strong gender differences, further analysis was done on only the male data. A 3x2 ANOVA[7] was run on search time (Male Data) against condition and object set.

Now there was no longer any significant difference among the time differences for the NoSound (M=65.4 sec, SD=35.1 sec), FullSound (M=58.5 sec, SD=17.3 sec), and PartialSound (M=80.1 sec, SD=49.2 sec) conditions.

Analysis of Female Data

Since there was a noticeable difference in the search time for the second four objects among the females in the NoSound (M=77.2 sec, SD=23.0 sec), FullSound (M=69.0 sec, SD=18.3 sec), and PartialSound (M=149.7, SD=21.2 sec) conditions it was decided to do some analysis on only the female data. A one-way ANOVA[8] was run on search time (Female Data) against condition and a significant difference in time among the conditions was found; $F(2,10)=20.7$, $p=0.0003$. Female participants in the PartialSound condition were 94% slower at locating objects than were female participants in the NoSound condition and one 117% slower than female participants in the Sound condition.

[7] Figure C.8.

[8] Figure C.10.

A t-test was done on the time difference between the FullSound and PartialSound conditions; $t(7)=-6.0$, $p < 0.01$; and also between the NoSound and PartialSound conditions; $t(7)=-4.9$, $p < 0.01$. There was not any significant difference between the NoSound and FullSound conditions.

4.3.3 Search Time Difference Between First and Second Four Objects

The data from the object search task was also analyzed by calculating, for each participant, the difference in search time between the mean search time for each of the first four objects and the mean search time for each of the second four objects. This was to provide some indication of whether participants were finding objects faster as a result of having spent more time in the virtual environment. The analysis was also done to examine the effect that removing sound from the environment had on participants in the PartialSound condition could be examined.

The mean time difference for all three conditions was computed, NoSound (M=-22.0 sec, SD=48.6 sec), FullSound (M=-1.2 sec, SD=22.5 sec), and PartialSound (M=30.9, SD=44.2). A negative time indicates a decrease in time for the second group of objects and a positive time indicates an increase in time.

Object Order Effects

A 2x2 ANOVA[9] was run to test if there was any significant difference in the time difference that was caused by the two different object sets, and to see if there was an interaction between condition and object set. The results of this ANOVA

[9] Figure C.7.

did not show any significant time difference between sets of objects; $F(1,39)=0.44$, $p=0.51$. As well, there was no significant interaction between condition and object set; $F(2,39)=1.2$, $p=0.32$.

Experiment Condition Effects

Figure 4.9 shows a graph of the mean time differences for all three conditions, as well as the mean times for both male and female participants in each condition.

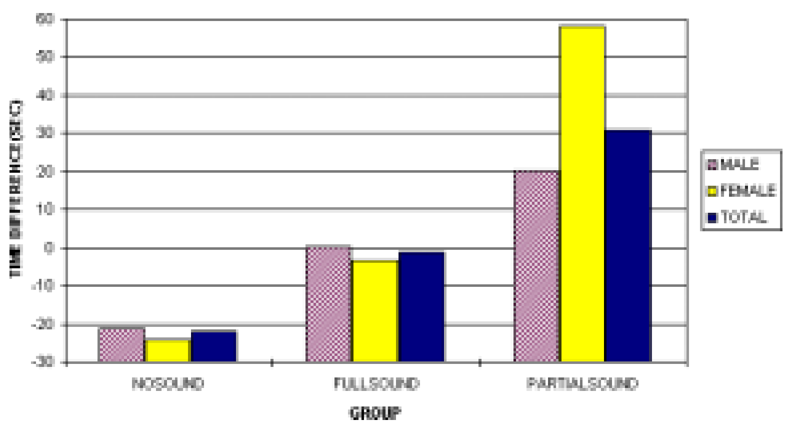

Figure 4.9: Mean Time Difference Between First Four and Second Four Objects
The difference in time to locate one of the first and second four objects was calculated for each participant and this graph shows the mean time difference for the three conditions. The mean times for both the males and females in each condition are shown as well.

Participants in the NoSound condition improved their average search time by 24% for the second four objects and participants in the PartialSound condition performed, on average, 45% worse on the second four objects. The search time for participants in the FullSound condition was approximately the same for both the first and second four objects. A 3x2 ANOVA[10] was run on time difference against condition and gender. The ANOVA showed a significant difference among conditions; $F(2,39)=7.4$, $p=0.002$ but no significant difference between males and females.

A t-test on the time difference between the FullSound and PartialSound conditions was run; $t(26)=-1.2$, $p < 0.05$; and also between the NoSound and PartialSound conditions; $t(33)=-3.4$, $p < 0.01$. There was not any significant difference between the NoSound and FullSound conditions.

Paired samples t-tests were also run to determine if there was a significant difference between the First4 time and the Second4 time within each condition. The results of the t-tests showed that there was a significant difference in time for both the NoSound condition; $t(16)=1.9$, $p< 0.05$; and the PartialSound condition; $t(17)=-3.0$, $p < 0.01$. There was no significant time difference for participants in the FullSound condition.

Analysis of Male Data

Again, because of the strong gender differences, further analysis was done on only the male data. A 3x2 ANOVA[11] was run on the mean search time difference (Male Data) between a first and second four object for each condition. The results of this ANOVA were not significant at the 0.05 level but were still strong; $F(2,26)=3.2$, $p=0.06$.

A t-test was run on the search time difference between the NoSound (M=-21.3 sec, SD = 51.4 sec) and PartialSound (M=20.5 sec, SD = 39.3 sec) conditions and

[10] Figure C.6.

[11] Figure C.9.

a significant difference was found; t(24)=-2.3, $p < 0.05$. There were no significant differences at the 0.05 level between the FullSound (M=0.3 sec, SD = 22.8 sec) condition and either the NoSound or PartialSound conditions.

Individual Object Search Times

For each of the eight objects that participants were required to locate the mean search time was calculated for all participants. Figure 4.10 shows the mean search times for the different groups of participants.

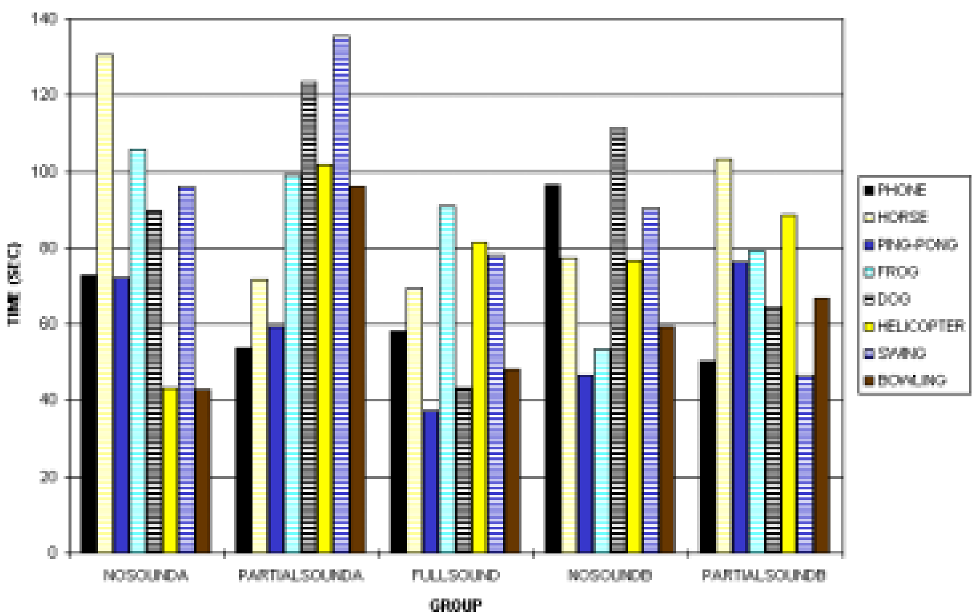

Figure 4.10: Mean Search Times For the Eight Objects Participants Were Required to Locate

This figure shows the mean search times for each of the Eight objects that participants were required to locate.

Figure 4.10 shows that participants in the NoSoundA group had the most trouble finding the Horse and Frog objects, and participants in the NoSoundB group had the most difficulty finding the Phone, Dog, and Swing objects. The Horse and Frog objects may have been difficult for participants in the NoSoundA group to find because they were both among the first four objects those participants were required to locate. As well, the Horse was located in a room in the middle of the environment so participants were unable to associate it with one of the colored exterior walls. The Frog was located in the southeast corner of the environment, and since participants began each object search task facing north they usually began moving to the north when they were unsure where an object was located. This meant that they would often have to backtrack to locate the Frog object. The Phone, Dog, and Swing objects may have been difficult for participants in the NoSoundB group to find because they were among the first five objects that participants in the NoSoundB group had to locate.

Participants in the FullSound group had the most trouble locating the Frog, Helicopter, and Swing objects. They may have had trouble locating the Frog object for the same reason as mentioned above, and their difficulty in locating the Helicopter object is explained in Section 5.1.3. It was unclear why the Swing object was difficult to find.

The Dog and Swing objects were the most difficult objects to find for participants in the PartialSoundA group and the Horse object was the object participants in the PartialSoundB group had the most trouble locating. The Dog and Swing objects were probably difficult for PartialSoundA participants to find because they were among the second four objects that those participants were required to find and all sound had been removed from the environment before they began searching for the second four objects. The Horse was the most difficult object for participants in the PartialSoundB group to find probably because the sound had been removed from the environment before they began searching for the Horse, and also because the Horse was located in a middle room of the environment.

4.4 Object Recall Task

After each participant completed the object search trials they were asked to recall as many objects as possible from the environment and place each object in the correct location on a map of the environment. Generic objects (e.g., dogs, horses, frogs, etc.) were treated as a single object during our analysis because it was not considered relevant to test whether participants could recall how many of a particular 'type' of object they could recall. This gave a total of twelve Auditory objects and twenty-three NonAuditory objects. Figure 4.11 shows the mean number of Auditory and NonAuditory objects recalled by participants in each of the three conditions.

The data was also analyzed to determine if there was a gender difference in the number of objects that a participant was able to recall. Figure 4.12 shows the mean number of Auditory and NonAuditory objects recalled by both males and females, as well as the total number of objects recalled by each gender.

From the two graphs it appears that the total number of objects recalled by participants was the same among the three conditions, with females (Auditory objects: M=10.2, SD=1.3 and NonAuditory objects: M=9.8, SD=4.1) recalling slightly more (approximately 15%) NonAuditory objects than males (Auditory objects: M=10.5, SD=1.1 and NonAuditory objects: M=8.5, SD=4.1).

However, it seems that there was a difference among condition in the type of objects which were recalled by participants. Fewer NonAuditory objects were

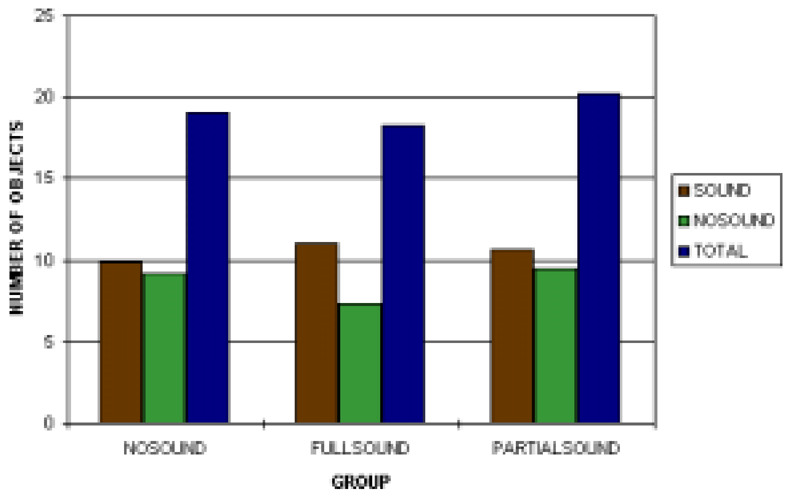

Figure 4.11: Mean Number of Objects Recalled by Participants in each condition
Participants in the NoSound condition recalled fewer Auditory objects than participants in the two sound conditions, and participants in the FullSound condition recalled the smallest number of NonAuditory objects.

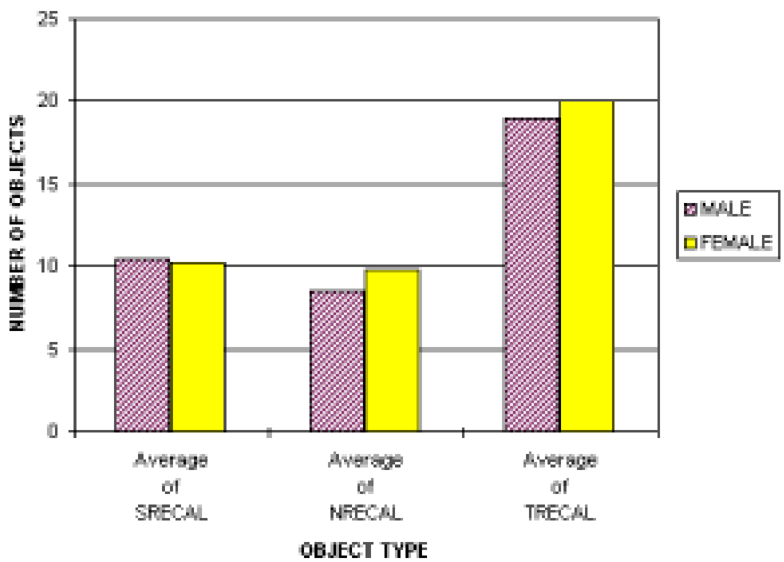

Figure 4.12: Mean Number of Objects Recalled by Both Males and Females

This graph shows the mean number of Auditory and NonAuditory objects recalled by males and females. *SRECAL* denotes Auditory objects, *NRECAL* denotes NonAuditory objects, and *TRECAL* denotes total number of objects. The graph shows that females recall slightly more NonAuditory objects than males.

recalled by participants in the FullSound condition (M=7.3, SD=5.52) than by participants in the NoSound (M=9.2, SD=3.0) and PartialSound (M=9.5, SD=4.1) conditions.

As well, approximately one less Auditory object was recalled by participants in the NoSound condition (M=9.8, SD=1.0) than by participants in the FullSound (M=11.0, SD=0.82) and PartialSound (M=10.7, SD=1.2) conditions.

Two 3x2 ANOVAs were run, one on the number of Auditory objects recalled against condition and gender[12], and the other on the number of NonAuditory objects recalled against condition and gender[13].

There was a significant difference in the number of Auditory objects recalled among the three conditions; $F(2,39)=4.2$, $p=0.027$. A t-test was run on the number of Auditory objects recalled between the NoSound and FullSound conditions; $t(25)=-3.1$, $p < 0.01$; and also between the NoSound and PartialSound conditions; $t(33)=-2.3$, $p < 0.05$. There was no significant difference between the FullSound and PartialSound conditions.

There was no significant difference in the number of NonAuditory objects recalled among the three conditions, but there did appear to be a slight gender difference; $F(1,39)=3.1$, $p=0.086$. Although not significant at the 0.05 level, there was also an interaction between condition and gender, which is shown in Figure 4.13.

[12]Figure C.11.

[13]Figure C.12.

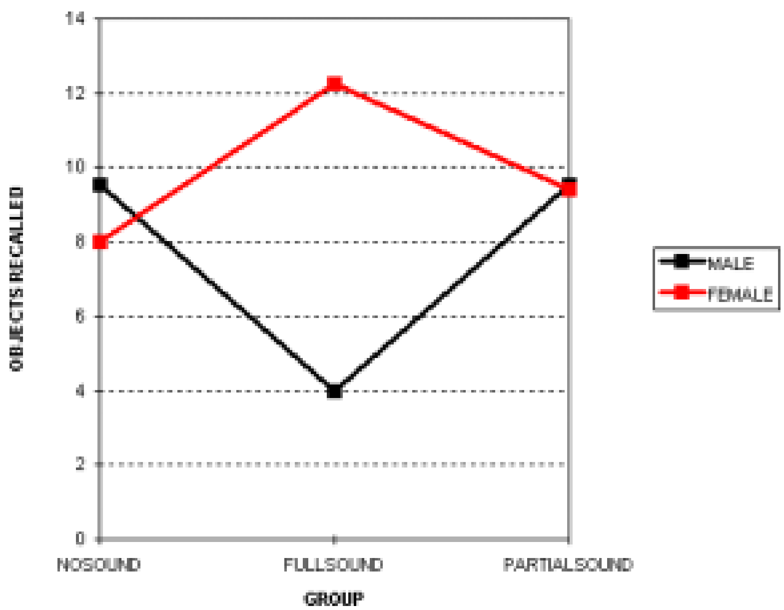

Figure 4.13: Interaction of Gender and Condition with Respect to NonAuditory Object Recall

This graph illustrates the interaction between condition and gender on the number of NonAuditory objects recalled. It appears that in the NoSound and PartialSound conditions females recalled approximately the same number of NonAuditory objects as males, but in the FullSound condition they recalled significantly more NonAuditory objects than males.

4.4.1 Analysis of Male Data

Since there were meaningful gender differences, it was again decided to do analyze only the male data (n=32). A one-way ANOVA[14] was run on the number of Auditory objects recalled (Male Data) against condition and the difference in the number of Auditory objects recalled among conditions was found to be stronger once the female data had been removed; $F(2,31)=5.1791$, $p=0.0119$.

A t-test was run on the number of Auditory objects recalled between the NoSound (M=9.8, SD=1.1) and FullSound (M=10.8, SD=0.75) conditions; $t(17)=-2.029$, $p < 0.05$; and also between the NoSound and PartialSound (M=11, SD = 0.92) conditions; $t(24)=-2.961$, $p < 0.05$. Male participants in the NoSound condition recalled one less Auditory object than male participants in the FullSound and PartialSound conditions. There was no significant difference between the FullSound and PartialSound conditions.

A one-way ANOVA (Figure C.14) was also run on the number of NonAuditory objects recalled (Male Data) against condition and this showed a significant difference in the number of NonAuditory objects recalled among conditions; $F(2,31)=5.9$, $p=0.007$.

A t-test was done on the number of NonAuditory objects recalled between the NoSound (M=9.5, SD=2.4) and FullSound (M=4, SD=3.0) conditions; $t(17)=4.3$, $p < 0.01$; and also between the FullSound and PartialSound (M=9.5, SD=4.6) conditions; $t(24)=-2.7$, $p < 0.05$. Male participants in the FullSound condition recalled approximately 58% fewer NonAuditory objects than male participants in both the NoSound and PartialSound conditions. There was no significant difference between the NoSound and PartialSound conditions.

[14] Figure C.13.

4.5 Object Placement Task

As well as recalling objects, participants were also asked to place as many objects as possible in the correct rooms on a map of the environment. The data examined for this analysis was the number of rooms each participant correctly identified. There were a total of twelve rooms in the environment. A room was considered to be correctly identified if at least one of the objects from a room was placed in the room on the map of the environment. The number of errors a participant made was not taken into consideration, but participants rarely attempted placing the same object in more than one room. The mean number of rooms identified by participants in the NoSound (M=6.9, SD=3.1), FullSound (M=3.5, SD=3.3), and PartialSound (M=5.2, SD=4.3) conditions were calculated. A graph of these means, as well as the male and female means, is shown in Figure 4.14.

Participants in the FullSound condition were only able to identify half as many rooms as participants in the NoSound condition, and 33% fewer rooms than participants in the PartialSound condition. Participants in the NoSound condition were able to identify the greatest number of rooms and males seemed to be able to identify more rooms than females, although this was not the case for participants in the FullSound condition. A one-way ANOVA[15] was run on the number of rooms identified against condition. There appeared to be a difference in the number of rooms identified among conditions, although this difference was not quite significant at the 0.05 level; $F(2,42)=2.8$, $p=0.07$. However, the main area of interest was in the difference between the NoSound and FullSound conditions so a t-test was run on the number of rooms identified by each of these conditions; $t(25)=2.7$, $p < 0.05$.

[15] Figure C.15.

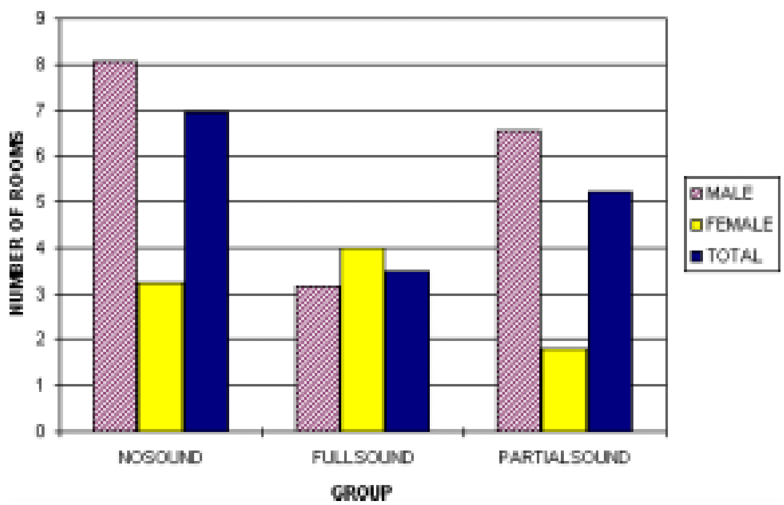

Figure 4.14: Mean Number of Rooms Identified by Participants in Each Condition
This graph shows the mean number rooms identified by participants in the three different conditions. Participants in the FullSound condition were able to identify fewer rooms than participants in the other two conditions. As well, other than in the FullSound condition, males were able to identify more rooms than females.

4.5.1 Analysis of Male Data

Once again, the female data was dropped and only the data for the males from the three conditions was analyzed, NoSound (M=8.1, SD=2.5), FullSound (M=3.2, SD=3.1), and PartialSound (M=6.5, SD=4.3). A one-way ANOVA[16] was run on the number of rooms identified (Male Data) against condition. Now that the female data was removed the difference among conditions was significant at the 0.05 level; $F(2,29)=4.1$, $p=0.03$.

A t-test was run on the number of rooms identified between the NoSound and FullSound conditions; $t(17)=3.7$, $p < 0.01$. Male participants in the FullSound condition identified 60% fewer rooms than male participants in the NoSound condition. There were no significant differences at the 0.05 level among any of the other conditions.

Two more one-way ANOVAs were run on only the male data. The first[17] was on the number of Auditory objects correctly placed (Male Data) against condition; $F(2,29)=3.5$, $p=0.04$. The second[18] was on the number of NonAuditory objects correctly placed (Male Data) against condition; $F(2,29)=4.2$, $p=0.03$.

t-tests were done on the number of Auditory objects placed on the map by participants in the NoSound (M=7.7, SD=2.7), FullSound (M=3.2, SD=3.1), and PartialSound (M=6.4, SD=4.3) conditions. There was a significant difference between the NoSound and FullSound conditions; $t(17)=3.3$, $p < 0.01$. There were no significant differences at the 0.05 level between the PartialSound condition and either of the other two conditions.

[16] Figure C.16.

[17] Figure C.17.

[18] Figure C.18.

Since the ANOVA on the number of NonAuditory objects placed also showed a significant difference among the conditions, t-tests were also run on the number of NonAuditory objects placed on the map by the NoSound (M=6.8, SD=3.8), FullSound (M=0.83, SD=1.2), and PartialSound (M=6.2, SD=5.6) conditions. There was a significant difference between the NoSound and FullSound conditions; $t(17)=3.7$, $p < 0.01$, and also the FullSound and PartialSound conditions; $t(17)=2.3$, $p < 0.05$. There was no significant difference between the NoSound and PartialSound conditions.

4.5.2 Analysis of Search Order

The participant data was also examined to see if the search order of the rooms had any effect on which rooms a participant was able to identify on the map. For each participant the twelve rooms were broken down into three groups of four rooms each, First4, Second4, and Other. The First4 group consisted of those rooms which contained an object that was among the first four objects the participant had to locate. The Second4 group consisted of those rooms which contained an object which was among the second four objects the participant had to locate, and the Other group contained those rooms that did not contain any objects that the participant had to find. Then the mean number of each type of room correctly identified by participants in the different conditions was calculated. Figure 4.15 shows a graph of these means.

As expected, in all three conditions, the rooms from the Second4 group were most accurately identified, followed by rooms from the First4 group. Rooms that did not contain an object that had to be found were the most difficult to identify.

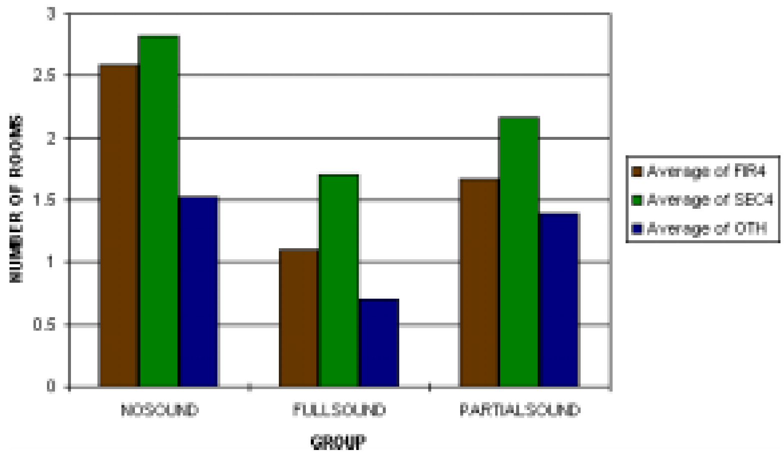

Figure 4.15: Mean Number of Different Room Types Identified by Participants in Each Condition

This graph shows the mean number of different room types identified by participants in the three different conditions. *FIR4* denotes the four rooms that a participant had to find an object from in the first object search phase and *SEC4* denotes the four rooms that a participant had to find an object from in the second object search phase. *OTH* denotes the rooms that a participant was never asked to find an object from. As expected, participants were more successful at identifying rooms from which they had to find an object.

4.6 Movement Data

The movement of each participant was recorded during the exploration and object search phases of the experiment. As described in Section 3.5.3, the virtual environment was divided up into thirty-one different sections, and each time a participant moved into a section the number of that section was recorded. Figure 3.5 in Section 3.5.3 illustrates how the environment was broken down into thirty-one different numbered sections. As well, the environmental coordinates of each participant were continually saved to a file so that a record of each participant's precise movement through the environment would be created.

The movement data from the exploration phase was examined in several ways to determine if the movement of participants who explored the sound enhanced environment was different from the movement of participants who were not exposed to any sounds.

The addition of sound to the environment may have had an effect on how quickly participants moved through the environment, so for each participant the total number of times they moved through a section of the environment was calculated. There was no significant difference in the number of sections that participants moved through between the Sound (M=58.0, SD=10.9) and NoSound (M=61.1, SD=15.4) conditions. However, there was a significant difference between males (M=62.7, SD=12.0) and females (M=50.4, SD=10.3); $t(43)=3.2$, $p < 0.01$. Males moved through twenty percent more sections than females. The fact that during the exploration phase males moved through the environment more quickly than females may have been one of the reasons why males performed better on the object search tasks than females.

As well, the addition of sound may have influenced the parts of the environment that participants explored during the exploration phase. Because there were less visual cues in the interior of the environment those participants who explored the silent environment may have spent more time in the exterior sections of the environment than did those participants exposed to the sound enhanced environment. To determine if this was the case each section of the environment was classified as either a room or a hallway. Rooms were those sections of the environment that contained objects, and hallways were all of the remaining sections other than the start location. There were a total of twelve room sections and eighteen hallway sections. Each section was further classified as either an internal or an external section. External hallways were those hallways that were next to at least one of the colored external walls. External rooms were those rooms that had either a colored wall as one of their walls, or had an entrance leading out into an external hallway. The remaining hallways and rooms were classified as interior. Table 4.5 lists the section numbers (shown in Figure 3.5) for all of the exterior and interior rooms and hallways.

Section Type	**Section Numbers**	**Total Number**
Interior Room	4,5,11,12	4
Exterior Room	2,3,6,7,8,9,10,13	8
Interior Hallway	14,15,18,20,21,24,26,27,28,29,30,31	12
Exterior Hallway	16,17,19,22,23,25	6
Start Location	1	1

Table 4.5: Environment Section Types

The above table lists all of the different sections in virtual environment. Figure 3.5 shows a map of the environment with each of the different sections labeled by a number.

For each participant the number of times they passed through an external hallway was calculated and also the number of times they passed through an internal hallway. The ratio of external to internal hallways that a subject passed through was also calculated. The Sound and NoSound means for each of these measures were then compared, but no significant difference was found between the two conditions. There were also no significant gender differences.

Participants in the NoSound condition may have entered more rooms than participants in the Sound condition because unlike participants in the Sound condition, who could hear what was in a room, participants in the NoSound condition were unable to ascertain what was in a room unless they looked into it. To determine if this was the case, the total number of times each participant entered a room was calculated, as well as the number of exterior and interior rooms each participant entered. The ratio of exterior to interior rooms entered was also calculated for each participant. The Sound and NoSound means were then compared for these measures and no significant differences were found for any of the measures between the two conditions. As well, there were no significant gender differences.

It appeared that the addition of sound to the virtual environment did not have any effect on how participants explored the environment. However, the data recorded for each participant did not capture how often a participant looked into a room. The only information recorded was when a participant actually walked into a room, and most participants only walked into each room once during the exploration phase. The rest of the time they just looked into rooms through the doorways and this information was not recorded. It may have been the case that participants in the NoSound condition looked into rooms more often than did participants in the Sound condition, although if so it did not slow them down at all since they were able to move through the same number of sections as were participants in the Sound condition.

The participant movement data that was collected during the object search phase of the experiment was also examined to check whether participants in the NoSound condition entered more rooms while looking for objects than did participants in the Sound condition. No significant difference was found between conditions, although once again it was the case that participants usually looked into a room rather than walking all of the way into it.

A more detailed discussion about the search and exploration strategies of participants can be found in Section 5.1.2.

4.7 Subjective Ratings

At the end of the experiment all participants were given a final questionnaire so that they could evaluate their ability to navigate around in the virtual environment. Participants in the FullSound and PartialSound conditions were also asked to answer several questions regarding the impact of the sounds. Figure 4.16 shows the mean scores of the different categories for participants in each of the conditions. Some of the questions did not apply to participants in the NoSound and FullSound conditions. The end of Section 3.5.3 described each of the five questions that participants were asked.

Participants had little difficulty moving around or recognizing objects from the environment. Participants also seemed to find the sounds very useful and found it difficult locating objects once the sounds had been removed. Participants did not appear to find the sounds overly disruptive.

Gender differences were also examined and Figure 4.17 shows the means for the different ratings according to gender. Males found it slightly easier than females

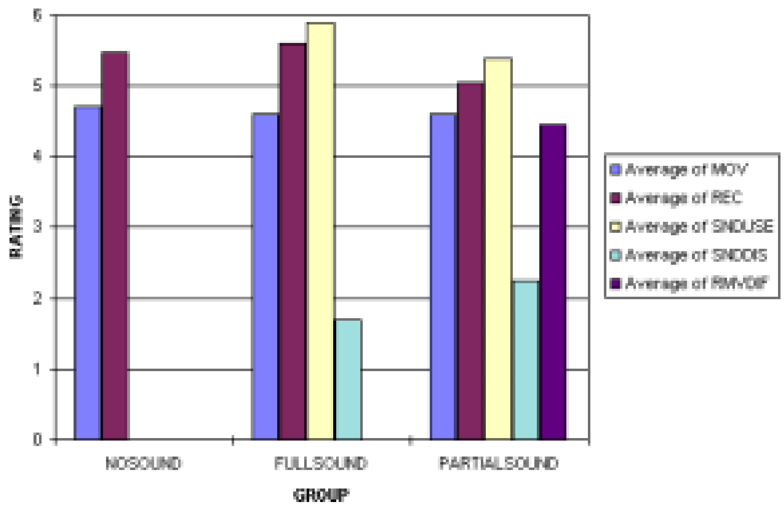

Figure 4.16: Mean Subjective ratings by Participants in Each Condition

This graph shows the mean ratings given by participants in each condition for the questions they were asked at the end of the experiment. Some of the questions did not apply to participants in the NoSound and FullSound conditions. The end of Section 3.5.3 described each of the five questions that participants were asked. From the graph it is clear that participants had little difficulty moving around *(MOV)* or recognizing objects *(REC)* from the environment. Participants also seemed to find the sounds very useful *(SNDUSE)* and found it difficult locating objects once the sounds had been removed *(RMVDIF)*. Participants did not appear to find the sounds overly disruptive *(SNDDID)*.

moving around and recognizing objects from the environment. It also looks like females found it much more difficult than males locating objects after the sound had been removed from the environment.

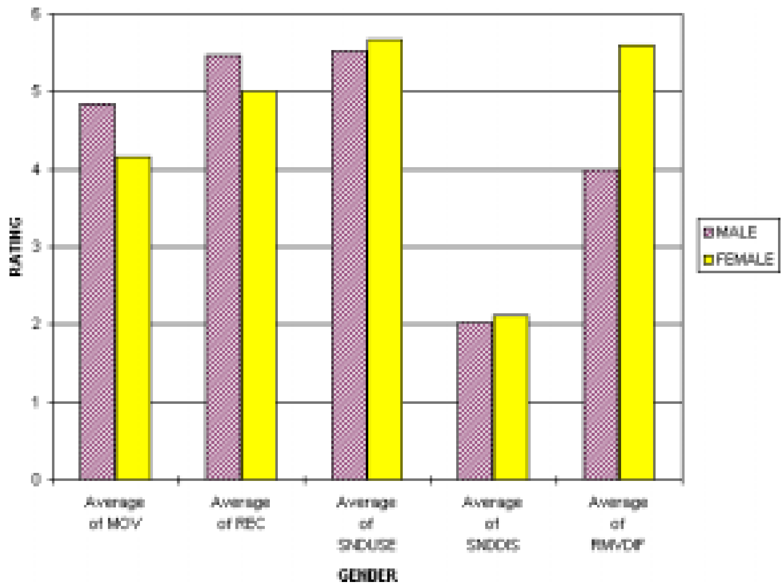

Figure 4.17: Mean Subjective ratings by Gender

This graph shows the mean ratings given by male and female participants for the questions they were asked at the end of the experiment. The end of Section 3.5.3 described each of the five questions that participants were asked. From the graph it appears that males found it slightly easier than females moving around *(MOV)* and recognizing objects *(REC)* from the environment. It also looks like females found it much more difficult than males locating objects after the sound had been removed from the environment *(RMDIF)*.

Chapter 5

Discussion

5.1 Experiment

5.1.1 Object Search Task

Search Time

The results of this experiment showed that including spatialized sounds in a visually sparse virtual environment helped a person locate objects that were associated with one of the sounds. Initially, participants in the FullSound condition were able to locate objects 26% more quickly than were participants in the NoSound condition. However, after a short period of time participants in the FullSound condition were only able to find objects 8% more quickly than participants in the NoSound condition were able to. Participants in the NoSound condition improved their search times by 22 seconds (a 24% improvement) over a short time span, unlike participants in the FullSound condition whose search times only improved by 1 second over the same time period.

Also, as shown by the large standard deviations in search times within each condition, there appeared to be very large individual differences in the ability of participants to locate objects. In fact, the deviations from the mean within each condition were larger than the magnitude of the effect observed between conditions. This suggests that the addition of 3D sound to a virtual environment did not cause 'poor' navigators to perform just as well as 'good' navigators on object search trials. However, the standard deviation for the search time of the first four objects within the Sound condition was approximately 9 seconds (24%) less than the standard deviation within the NoSound condition. This implies that to some extent the addition of 3D sound did reduce differences in search time caused by individual differences in navigational and spatial ability.

Because the length of the experiment was short, it was not clear if participants in the NoSound condition would eventually begin to perform better than participants in the FullSound condition. A very simple environment was used in this study; perhaps with a more complex environment the addition of spatialized sounds would have had a greater and longer lasting effect on navigation.

The average search times of participants in the PartialSound condition were 45% longer after the spatialized sounds were removed from the environment. In fact, after the sounds were taken away from the environment participants in the PartialSound condition took 46% longer (approximately 31 seconds) to locate objects than did participants in the NoSound condition. As well, there was a significant difference, $p < 0.01$, between the search times of the first and second four objects for participants in the PartialSound condition. They took approximately 31 seconds longer locating one of the second four objects than they did locating one of the first four objects. This seems to indicate that the addition of sound caused participants to pay less attention to the visual cues, and as a result they may have become dependent on the auditory cues. A number of participants made comments indicating that they had come to rely on the sounds and felt 'lost' once the sounds were removed from the environment.

The search time data also showed that males were 27% (approximately 206 seconds) faster at locating objects than were females. This was not surprising since males scored higher than females on the Guilford-Zimmerman test of spatial orientation. However, it is interesting that the removal of sound from the environment had a greater effect on the search times of females than it did on the search times of males. Females in the PartialSound condition performed 63% worse when the sound was removed and males in the PartialSound condition only suffered a performance decrease of 34%. However, since females scored lower than males on the test of spatial ability it is unclear whether the removal of sound affected females more than males, or whether it affected people with low spatial ability more than people with high spatial ability.

5.1.2 Observations

After observing all of the participants and listening to their verbalizations some of the ways that participants used the visual and auditory cues became clear, as well as some of the common strategies that were being used during the exploration and object search phases of the experiment. The following sections discuss a) the different ways that participants appeared to use the visual and auditory cues and b) some of the common exploration and search strategies of participants.

Use of Visual Cues

The virtual environment used in this experiment was visually sparse. To help participants orient themselves in the environment each of the exterior walls was given a different color, and colored geometric shapes were placed on the walls of the environment at most of the intersection points. However, comments made by participants during the experiment indicated that the colored geometric shapes were not being used as navigational aids. The only geometric shape that participants frequently reported as having used to obtain their orientation was the red diamond located on the wall directly north of the initial starting location. This shape was used by participants to orient themselves so that they were facing in the same direction as they were at the beginning of each object location task. There appeared to be two problems with the geometric shapes that limited their use as landmarks: they were not distinctive enough, and they had no obvious association with any of the objects in the different rooms.

Most participants seemed to find the differently colored external walls to be very useful navigational aids. There appeared to be three main ways that the external walls were used: First, the majority of participants had learned the colors of the north, south, west, and east walls from the map they were shown so they were able to use the color of an external wall to determine which way they were facing. Second, the colors of the walls allowed participants to determine which section of the maze they were currently exploring, i.e., north-west, south-east, etc.. Third, some participants were able to associate room objects with the color of the closest external wall and this helped them locate objects during the object search trials.

Use of Sound Cues

The responses by participants on the final questionnaire indicated that participants did not find the spatialized sounds distracting, and that they thought the sounds were useful in helping to locate objects. The verbalizations made by participants indicated that they were using two different aspects of the spatialized sounds to help them locate objects.

1. **Volume:** One sound cue participants found useful was the relationship between their distance from a sound source and the volume of the sound source. As the distance increased the volume of the sound source decreased, and as the distance decreased the volume of the sound source increased. This allowed participants to determine whether they were moving towards a particular sound source or away from it. Participants seemed to find this auditory cue the most useful one.

2. **Right & Left Orientation:** The majority of participants exposed to the spatialized sounds seemed to be able to determine whether a sound source was located to the left or right of them. This agrees with the study performed by Loomis et al., (Loomis, Hebert, and Cicinelli 1990) in which they found that people were able to readily distinguish right from left orientation for sounds in a virtual condition. Participants appeared to use this ability to help them locate objects. However, it took a while for some participants to realize that they could determine the direction of a sound source, and from watching them it appeared that most participants had trouble distinguishing if a sound was located in front of them or behind them.

Search Strategies

There were several common object search strategies used by participants in all three of the experiment conditions. When participants had no idea where to find an object they would often move towards an external wall and then walk around the perimeter of the environment, looking inside each each room that they moved past. Only after they had searched all of the exterior rooms did they begin moving into the interior of the environment. It appears that participants attempted to partition the environment into different areas. This is what Darken observed of people in the study he conducted on navigation in a large virtual environment (Darken and Sibert 1996).

Some participants reported that they were able to remember the map of the environment and they visualized the map as they were exploring. They felt that visualizing the map helped them find their way through the environment. However, the majority of participants were unable to remember very much of the map, and several participants mentioned that it would have been useful to have seen the map again after the exploration phase of the experiment.

A frequent strategy used by participants in both of the Sound conditions was to walk down each of the four hallways leading out of the initial starting location. As they walked down each hallway they would listen for the sound of the object they were currently looking for, and if they began to hear it they would try to home in on it. If they failed to hear the sound by the time they reached the end of the hallway they backtracked to the starting location and began moving down another hallway. Participants in the NoSound condition almost never backtracked which is what Peponis observed in a study he ran on navigation in a real world environment (Peponis, Zimring, and Choi 1990).

Some participants, especially women, seemed to develop very good route knowledge, but limited survey knowledge. For example, one female subject in the NoSound condition knew exactly how to find the room with the horse in it, but was unaware that the horse room was located immediately to the right of the initial starting location. Instead of turning right and moving towards the horse room she took a long route through the northern part of the environment, although her comments made it clear that she knew exactly where she was going.

5.1.3 Design Issues

After observing all of the participants perform the object search trials and listening to their comments it became apparent that there was a problem with the design that was limiting the usefulness of the auditory cues. The fact that the walls were transparent with respect to sound was causing many participants to become confused about where a sound was coming from. They assumed that when a sound became extremely loud the entrance to the room containing the sound source must be near by. However, due to the auditory transparency of the walls this was not always the case. This problem was made worse because of the way that the sound sources were placed in each room. Every sound was assigned to a specific object in each room, and sometimes this object was placed directly alongside one of the walls of the room. This meant that when a participant was located outside a room and directly beside one of these walls they heard the sound source at almost full volume, and expected to see the entrance to the room near by. However, for rooms such as the helicopter and horse rooms the entrance to the room was not always near by.

This problem could have been reduced slightly by placing sound sources in the middle of the rooms rather than assigning them to specific objects in the rooms. An even better solution would be to no longer treat the walls as transparent with respect to sound, and take things such as reflection and refraction into account. Unfortunately, at present no real time solution exists that models these properties of sound. However, even doing something crude such as increasing the volume of a sound source as a person moves closer to an entrance of the room containing the sound source might improve performance.

However, even given this problem, the addition of the spatialized sounds to the virtual environment significantly decreased the search time of participants. Perhaps with more realistic auditory modeling of the environment the time it takes people to locate objects would decrease even further.

5.1.4 Object Recall Task

Auditory Objects

Not surprisingly, participants who experienced the sound enhanced environment were able to recall slightly more Auditory objects than participants who did not hear any of the sounds. Participants in the FullSound condition recalled twelve percent more Auditory objects than participants in the NoSound condition and participants in the PartialSound condition recalled nine percent more Auditory objects NoSound participants. There were probably two main reasons why this was the case. First, the participants who experienced the sounds had two sources of information to help them recall each auditory object, visual and auditory. If they did not remember seeing an object they might have remembered hearing it, or vice-versa. The participants who were not exposed to any sounds only had visual information to aid them in recalling an object.

Second, the participants who were exposed to the sound enhanced environment could hear the auditory objects so they were aware of those objects even when they were not in rooms looking at them. They had more exposure to the auditory objects than did those participants who never heard any sounds.

NonAuditory Objects

The opposite was true when participants tried to recall objects that were not associated with any sounds in the environment. In this case, the participants who experienced the sound enhanced environment throughout the entire experiment were able to recall fewer NonAuditory objects than the participants in the other two conditions. There were probably three main reasons why this was the case. First, it was never necessary for the participants who could always hear sounds to look into a room to determine which room it was. This meant that they probably looked into the rooms of the environment much less than those participants who could not always hear sounds, and thus were not exposed to the NonAuditory objects as much as those participants were.

Second, participants in the FullSound condition took considerably less time to find all of the objects (23% less time than NoSound participants and 25% less time than PartialSound participants) so they had less exposure to all of the objects in the environment than did participants in both the PartialSound and NoSound conditions.

Third, the sound may have distracted them from paying attention to the NonAuditory objects because when they looked into a room their attention could have been drawn to the object emitting the sounds. This may have interfered with their ability to recall these objects.

5.1.5 Object Placement Task

On the object placement task participants who experienced sound during the entire experiment were able to identify significantly fewer rooms on the map than were participants in the the other two conditions. Participants in the FullSound condition were only able to identify half as many rooms as participants in the NoSound condition, and 33% fewer rooms than participants in the PartialSound condition. As was mentioned in Section 4.5, a room was considered to have been identified if at least one object was correctly placed in the room. The participants in the NoSound condition were able to identify the greatest number of rooms.

It appeared that although the addition of spatialized sounds made it easier for participants to locate objects, it did not help them acquire survey knowledge. This was illustrated by their lack of understanding of room and object locations within the environment. In fact, the addition of spatialized sounds may even have inhibited the acquisition of survey knowledge. Comments made by participants suggest that the sounds did in fact prevent them from acquiring survey knowledge. Most of the participants in the FullSound condition reported that they had been relying totally on the sounds and had not been paying any attention to the visual surroundings so they had no idea where most objects in the environment were located. Participants in the PartialSound condition made similar comments when the sound was turned off after the first object search phase of the experiment.

However, another factor contributing to the poor performance of participants in the FullSound condition may have been the fact that they took less time to locate all of the objects so they were exposed to the virtual environment for less time than participants in the other two conditions. Another experiment would have to be conducted to determine if the spatialized sounds were actually interfering with the acquisition of survey knowledge, although it is clear from this experiment that the sounds did not improve the amount of survey knowledge a participant was able to acquire.

Strategies

There were several strategies used by participants when they were attempting to place objects on the map of the environment. Generally, in all three of the conditions, participants who performed well on this task looked at rooms on the map as they attempted to recall which objects were in each room. Most participants who did poorly first attempted to recall as many objects as possible, and only looked at the map after they had finished recalling objects.

A fairly common strategy, which did not work very well, was to assume that those objects that had taken a long time to find were located in the rooms furthest away from the start location. If an object had been found quickly it was assumed to have been in a room near the start location.

Another strategy used by some participants was to attempt remembering the size and dimensions of rooms that specific objects had been in, e.g., the ping-pong table was in the long skinny room. Participants also tried to remember how many exits each of the rooms had.

Some of the participants who experienced the spatialized sounds attempted to remember which sounds they had been able to hear from the start location and assumed that these objects must have been in rooms near the start location.

Males were able to identify more rooms than females, which is not surprising given that males scored higher than females on the test of spatial ability.

5.2 Conclusions

Very little previous research has studied the effects of auditory cues on navigation. This study was one of the first to investigate the effects of spatialized sounds on navigation in a virtual environment. The results of this study suggest that the addition of spatialized sounds to a virtual environment can help a person locate objects in a virtual environment faster than without any sound cues ($p < 0.05$). However, the addition of spatialized sounds only caused performance on the object search tasks to increase by 26%, while at the same time decreasing the number of rooms identified on the object placement task by 50%. This indicates that the addition of spatialized sounds do not aid in the acquisition of survey knowledge, and may even suppress its development.

The performance of participants who were exposed to the spatialized sounds only marginally improved over time. This was unlike the performance of participants who never experienced any sounds. The performance of these participants showed marked improvement, ($p < 0.05$), and in a short period of time they were able to perform as well as those participants navigating in the sound enhanced environment.

Previous research by Darken (Darken and Sibert 1993) found that the addition of a spatialized sound to a virtual environment had the effect of enlarging the target area for the object that the sound was associated with. The spatialized sounds in this experiment appeared to have a similar effect and perhaps this explains why the acquisition of survey knowledge may have been suppressed. It may have been that during the object search trials when participants came within range of the enlarged target area of an object their attention to the visual details of the environment was reduced to the point where the acquisition of survey knowledge was inhibited. If this was the case then the addition of spatialized sounds to a virtual environment has the same affect that Moeser (Moeser 1988) observed when signs were present in a real world environment.

Females appeared to be more dependent on the spatialized sounds than males. It was unclear whether this was related to the fact that they scored lower than males on a test of spatial ability and needed the sound cues to help with navigation, or whether it was due to some other gender difference.

In conclusion, the results of this work indicate that the addition of 3D sounds to a virtual environment does improve the ability of people to perform object retrieval tasks, but causes a decrease in the amount of survey knowledge they are able to acquire.

5.3 Future Work

The environment used in this study was a maze with limited visual cues. To determine if the results of this study apply to all virtual environments further work needs to be done looking at more complex environments. In a virtual world less visually sparse than the one used in this experiment the visual navigational cues may be so dominant that the addition of spatialized sounds has no improvement on navigation. It would also be interesting to study the usefulness of spatialized sounds in environments where it is possible to move in three dimensions instead of just two. It is important to explore the use of auditory cues in more complex environments because in the near future hardware will have improved to the point where it is possible to create complex virtual environments.

Another interesting area to explore is whether the type of sound has an effect on the ability of a person to navigate in a virtual environment. In this study all of the sounds were associated with specific objects. It is unclear how useful sounds without any obvious associations to objects in an environment would be. As well, it would be interesting to determine if nonspatialized sounds can provide useful navigational information since it requires much less processing power to generate nonspatialized sounds than it does to generate spatialized sounds.

In this study the walls and objects of the environment were transparent with respect to sound. It would be interesting to do some acoustic modeling of the environment so that sounds reflected off surfaces as they do in our natural environment. Perhaps this would make the sounds even more useful as navigational aids and it might help determine how sounds affect navigation in the real world. It would also be worthwhile to determine the maximum number of distinct sounds that can be heard concurrently and still provide useful spatial information.

It is unclear from this study whether spatial ability or gender, alone or in combination, had an effect on how a person was able to use sound as a navigational aid. It appeared that females or people with low spatial ability became more dependent on sound than males or people with high spatial ability. More work needs to be done to determine whether there are individual differences in how people are able to use sound as a navigational aid. It is also important to determine how well visually impaired people are able to navigate with spatialized sound, and whether their search strategies and performance differ from those of sighted people. If visually impaired people are able to navigate with the use of spatialized sounds, then enhancing virtual environments with sound would provide access to these environments for those people who are visually impaired.

Another area to explore is whether the type of equipment used to display the virtual environment has an effect on the usefulness of auditory cues as navigational aids. In our study a HMD with a low resolution and narrow field of view was used.

As well, the head tracking had a slight lag time and this seemed to reduce the amount of head movement made by participants. It would be interesting to run a similar experiment with a more powerful HMD. This might increase the sense of presence and realism for participants, which could influence how well they are able to navigate and the amount of survey knowledge they acquire.

Bibliography

Batteau, D. W. (1968). *The Neuropsychology of spatially oriented behavior*, Chapter Listening with the naked ear. Nomewood, IL: Dorsey Press. Cited in Begault(Begault 1994).

Beaumont, P. B., J. Gray, G. T. Moore, and B. Robinson (1984). Orientation and wayfinding in the tauranga department building: a focused post-occupancy evaluation. *Environmental Design Research Association Proceedings 15*, 77–91.

Begault, Durand R. (1993). Head-up auditory displays for traffic collision avoidance system advisories: a preliminary investigation. *Human Factors 35*(4), 707–717.

Begault, Durand R. (1994). *3-D Sound For Virtual Reality and Multimedia*. Cambridge: Academic Press.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology 59*, 960–970.

Boff, K. R. and J. E. Lincoln (1988). *Engineering Data Compendium: Human Perception and Performance*. Ohio: Wright-Patterson AFB.

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

Brooks, F. P. Jr. (1992). Six generations of building walkthroughs. final technical report, walkthrough project. Technical Report TR92-026, National Science Foundation. Cited in Satalich(Satalich 1995).

Bryant, K. J. (1982). Personality correlates of sense of direction and geographical orientation. *Journal of Personality and Social Psychology 43*, 1318–1324.

Burgess, David A. (1992). Techniques for low cost spatial audio. Technical Report 92-09, Georgia Institute of Technology.

Butler, Darrell L., April L. Acquino, Alicia A. Hissong, and Pamala A. Scott (1993). Wayfinding by newcomers in a complex building. *Human Factors 35*(1), 159–173.

Darken, Rudolph P. and John L. Sibert (1993). A toolset for navigation in virtual environments. In *UIST 93 Conference Proceedings*, pp. 157–165. ACM Press.

Darken, Rudolph P. and John L. Sibert (1996). Wayfinding strategies and behaviors in large virtual worlds. In *CHI 96 Conference Proceedings*, pp. 142–149. ACM Press.

Evans, G. W. and K. Pezdek (1980). Cognitive mapping: knowledge of real-world distance and location information. *Journal of Experimental Psychology : Human Learning and Memory 6*, 13–24.

Galea, L. A. M. and D. Kimura (1993). Sex differences in route-learning. *Personality and Individual Differences 14*, 53–65.

Gardner, M. B. (1973). Some monaural and binaural facets of median plane localization. *Journal of the Acoustical Society of America 54*, 1489–1495. Cited in Burgess(Burgess 1992).

Guilford, J. P. and W. S. Zimmerman (1981). *The Guilford-Zimmerman Aptitude*

Survey: Manual of Instructions and Interpretations (Revised ed.). Consulting Psychologists Press.

Harris, L. J. (1981). *Spatial Representation and Behavior Across the Lifespan*, Chapter Sex-related variations in spatial skill, pp. 83–125. New York: Academic Press.

Hendrix, Claudia and Woodrow Barfield (1995). Presence in virtual environments as a function of visual and auditory cues. In *IEEE Virtual Reality International Symposium*, pp. 74–82.

Holding, C. S. and D. H. Holding (1989). Acquistion of route network knowledge by males and females. *Journal of General Psychology 116*, 29–41.

Howell, David C. (1987). *Statistical Methods for Psychology* (Second ed.). Boston: Duxbury Press.

Infield, S. D. (1991). *An investigation into the relationship between navigation skill and spatial abilities*. Ph. D. thesis, University of Washington.

Kozlowski, L. and K. Bryant (1977). Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology 3*, 590–598.

Labs, QSound (1996). *QMixer95 v2.1 API Guide*. Calgary, Alberta: QSound Labs.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles 30*, 765–779.

Lawton, Carol A., Stephanie I. Charleston, and Amy S. Zieles (1996). Individual- and gender - related differences in indoor wayfinding. *Environment and Behavior 28*(2), 204–219.

Levine, M., I. Marchon, and G. Hanley (1984). The placement and misplacement of you-are-here-maps. *Environment and Behavior 16*, 139–157.

Loomis, Jack M., Chick Hebert, and Joseph G. Cicinelli (1990, October). Active localization of virtual sounds. *Journal of the Acoustical Society of America 88*(4), 1757–1764.

Lynch, K. (1960). *The Image of the City*. Cambridge: MIT Press.

Martin, David W. (1977). *Doin Psychology Experiments*. Monterey, California: Brooks/Cole Publishing Company.

Miller, L. K. and V. Santoni (1986). Sex differences in spatial abilities: Strategic and experimental correlates. *Acta Psychologica 62*, 225–235.

Moeser, S. D. (1988). Cognitive mapping in a complex building. *Environment and Behavior 20*, 21–49.

Montello, D. R. and H. L. Pick (1993). Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior 25*, 457–484.

Moore, F. R. (1990). *Elements of Computer Music*. Englewood Cliffs, NJ: Prentice Hall. Cited in Burgess(Burgess 1992).

on Virtual Reality Research, National Research Council Committee and Development (1995). *Virtual Reality: Scientific and Technological Challenges*. Washington, D.C.: National Academy Press.

Passini, R. (1984). *Wayfinding in Architecture*. New York: Van Nostrand Reinhold Company Inc.

Peponis, John, Craig Zimring, and Yoon Kyung Choi (1990, September). Finding the building in wayfinding. *Environment and Behavior 22*(5), 555–590.

Rheingold, Howard (1991). *Virtual Reality*. New York, New York: Summit Books.

Sadalla, E. K. and D. R. Montello (1989). Remembering changes in direction. *Environment and Behavior 21*, 346–363.

Satalich, Glenna A. (1995). Navigation and wayfinding in virtual reality: finding the proper tools and cues to enhance navigational awareness. Master's thesis, University of Washington.

Searle, C. L., L. D. Braida, M. F. Davis, and H. S. Colburn (1976). Model for auditory localization. *Journal of the Acoustical Society of America 60*, 1164–1175. Cited in Burgess(Burgess 1992).

Sense8 (1996). *World Up User's Guide* (May ed.). Mill Valley, California: Sense8 Corporation.

Siegel, A. W. and S. H. White (1975). *Advances in Child Development and Behavior*, Chapter The development of spatial representation of large-scale environments. New York: Academic Press.

Streeter, Lynn A. and Diane Vitello (1985). A profile of driver's map reading abilities. *Human Factors*.

Streeter, Lynn A., Diane Vitello, and Susan A. Wonsiewicz (1985). How to tell people where to go: comparing navigational aids. *International Journal of Man-Machine Studies 22*, 519–562.

Thorndyke, Perry W. and Sarah E. Goldin (1983). *Spatial Orientation: Theory, Research, and Application*, Chapter Spatial Learning and Reasoning Skill, pp. 195–224. New York and London: Plenum Press.

Thorndyke, P. W. and B. Hayes-Roth (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology 14*, 560–589.

Thurlow, W. R., J. W. Mangels, and P. S. Runge (1967). Head movements during sound localization. *Journal of the Acoustical Society of America 42*, 489–493. Cited in Begault(Begault 1994).

Thurlow, W. R. and P. S. Runge (1967). Effects of induced head movements on localization of direct sound. *Journal of the Acoustical Society of America 42*, 480–487. Cited in Begault(Begault 1994).

Tierney, J. (1993, Spetember). Jung in motion, virtually, and other computer fuzz. *The New York Times 16*, B1–4. Cited in Begault(Begault 1994).

Tlauka, Michael and Paul N. Wilson (1996, September). Orientation-free representations from navigation through a computer-simulated environment. *Environment and Behavior 28*(5), 647–664.

Velleman, Paul F. and Agelia Y. Velleman (1988). *Data Desk Professional Statistics & Reference Guides*. Odesta Corporation.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology 27*, 339–368. Cited in Begault(Begault 1994).

Ward, S. L., N. Newcombe, and W. F. Overton (1986). Turn left at the church, or three miles north. *Environment and Behavior 18*, 192–213.

Warren, D. H., T. E. Scott, and C. Medley (1992). Finding locations in the environment: the map as a mediator. *Perception 21*, 671–689.

Wenzel, Elizabeth M., Marianne Arruda, Doris J. Kistler, and Frederic L. Wightman (1993, July). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America 94*(1), 111–123.

Wetherell, A. (1979). Short-term memory for verbal and graphic route information. In *Proceedings of the 23rd annual meeting on the Human Factors Society*, Santa Monica, CA. Human Factors Society.

Wickens, Christopher D. (1992). *Engineering Psychology and Human Performance - Second Edition*. HarperCollins.

Wickens, C. D., S. Todd, and K. Seidler (1989). Three-dimensional displays: Perception, implementation, and applications. Technical Report CSERIAC SOAR-89-01, Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory.

Wightman, F. L. and D. J. Kistler (1989a). Headphone simulation of free-field listening i : stimulus synthesis. *Journal of the Acoustical Society of America 85*, 858–867. Cited in Burgess(Burgess 1992).

Wightman, F. L. and D. J. Kistler (1989b). Headphone simulation of free-field listening ii : psychophysical validation. *Journal of the Acoustical Society of America 85*, 868–878. Cited in Burgess(Burgess 1992).

Appendix A

Technical Information

General Description	64-bit 3D Windows accelerator card PCI 2.1 system interface 4MB VRAM frame buffer configuration 4MB DRAM local buffer Support for Intel, Alpha
Rendering Features	Gouraud shading Z-buffering Alpha-blending Anti-aliasing(4x4 and 8x8 sub-pixel) Texture mapping Fog, atmosphere, depth cueing Stenciling Windowed and full-screen rendering
Local-buffer	4MB DRAM SIMM 32-bit and 24-bit Z-buffer 8-bit and 4-bit stencil buffer 8-bit and 4-bit fast clear
Performance	Up to 300k triangles/sec (25 pixel,24-bit Gouraud Shaded,Anti-aliased,Z-buffered,Alpha-blended) Up to 750k 3D vectors/sec.
Software Support	Windows NT 3.5, 3.51 OS/2, Windows 95, Apple OS 7.X(Under Development)

Table A.1: Fujitsu Sapphire 2SX Technical Information

Key Features	Supports VRML 1.0(with plans to support VRML 2.9) Includes 3D sound support A GUI-based interactive development environment Fast prototyping with BasicScript language Call DLL's/DSO's from script Cross-platform portability without recompiling Integrated modeler Atmospherics, including fog ODBC support Dynamic texturing of objects Resource management tools 3D text objects Dialog construction tool Level of detail in nodes scene graph Drap and drop editing in scene graph
Platforms	Windows NT SGI workstations Windows 95
System Requirements	Pentium 90 MHz 24MB RAM Open GL-based graphics accelerator board
Price	$3500 (US)

Table A.2: WorldUp Tecnical Specifications

Optics	Heads-up distortion-free display Field of view: 30 degrees each eye Fixed focus at 13 feet, and converged at 8 feet, to minimize eye strain Requires no IPD adjustment 100% stereo overlap Can be worn with eyeglasses
Displays	Two full-color LCDs Resolution: 180,000 pixels per LCD panel
Mechanical	Ergonomically designed for comfort Foldable Frame Weight: 12 ounces Clip-on immersion visor Fully adjustable earphones
Electrical	Power supply: 110 VAC input/9 VDC output Power consumption: 3 watts
Audio	Stereo RCA connectors, 1 volt peak-to-peak Frequency response 20Hz-20kHz Excellent 3-D audio spatialization
PC Head Tracker	3 degrees of freedom: pitch, roll and yaw 250 Hz sample rate Serial interface: RS-232C
PC Interface Module	Video: Single channel RCA input VGA interface-input and pass-through output No internal card needed 60-72 Hz operations Both field and line sequential 3-D formats
User Controls	Volume
Suggested Price	$499 (US)

Table A.3: Virtual i-glasses Technical Specifications

Arbitrary number of simultaneous playback channels
Mixes 4 channels at less than 5% CPU (P90,DirectSound,11KHz/8bit)
Mixes 8 channels at less than 7% CPU (P90,DoirectSound,11KHz/8bit)
11.025 k, 22.050 k, and 44.1 kHz sample rates
Variable playback frequency
8 and 16 bit resolution
Mono and stereo input files
Queuing of waves on a channel
Includes libraries for Borland, Microsoft Visual C++, and Watcom compilers

Table A.4: QSound Technical Information

Appendix B

Experiment Forms

Script for Researcher

The use of 3D Sound as a Navigational Aid in Virtual Environments

General Instructions to Participants

Script for Researcher

Read phrases in bold print only to subjects in one of the sound groups

1.0 INTRODUCTION

We are conducting an experiment designed to determine the effects that adding 3D sound to a virtual environment will have on navigation throughout the environment.

This experiment will take approximately 2 hours to complete and will require you to perform several different tasks. First, you will be asked to complete a short background questionnaire and then you will be given an equipment training session. After you feel comfortable using the equipment you will be taken on a tour of a virtual environment. Then you will be required to explore this virtual environment on your own for 8 minutes. Once you have finished the exploration you will be required to locate 8 different objects from the virtual environment you just explored. After you have located all of these objects you will be asked to perform a paper and pencil object recall and placement task, followed by a very short final questionnaire. The last thing you will be asked to do is a 20 minute test of spatial ability called the Guilford-Zimmerman Orientation Test. During some tasks you will be encouraged to verbalize your thoughts and these will be recorded on an audio tape for future analysis. A short rest period will be provided after the completion of each task, during which the instructions for the next stage will be explained in detail.

At the beginning of the session you will be assigned a numeric number. Any data collected during the session will only be identified with this numeric code to ensure that your identity is kept confidential. The only people who will have access to the original data will be the researchers that are directly involved with this project. All data will be kept for a minimum of one year. After I am finished with the data, all written records will be destroyed and all audio tapes erased.

There is a potential risk of which you need to be aware. Sometimes, people experience some "motion sickness" (a sense of dizziness and nausea) when trying to navigate through virtual environments. Should you experience any symptoms of motion sickness, or should you decide to stop a trial for whatever reason, simply inform me and I will immediately stop the trial.

Participants who complete the entire experiment will be paid $ 15 following the last task. Do you have any questions or concerns at this point?

Now that I have given you a basic overview of the experiment, I would like you to read over the Information Letter and sign 2 copies of the Consent Form if you are willing to participate. Please keep one copy of the Information Letter and Consent Form for yourself.

GIVE PARTICIPANT 2 COPIES OF INFORMATION LETTER AND CONSENT FORM AND WAIT FOR THEM TO READ IT OVER AND SIGN IT.

2.0 BACKGROUND QUESTIONNAIRE AND TRAINING SESSION

Now please take a few minutes to complete this background questionnaire. The purpose of this questionnaire is to allow us to describe the general demographics of the participants involved in this study.

HAND BACKGROUND QUESTIONNAIRE TO PARTICIPANT AND WAIT FOR THEM TO COMPLETE IT.

Several tasks in this experiment require you to navigate around in a virtual environment using a head mounted display and a mouse. This training session will introduce you to the equipment you will be using and allow you to practice using the head mounted display and mouse to move around in a virtual environment. You will be viewing the virtual environment with this light-weight head mounted display.

SHOW PARTICIPANT HEAD MOUNTED DISPLAY AND DEMONSTRATE PUTTING IT ON.

Sometimes people experience feelings of motion sickness, such as dizziness or nausea, when trying to navigate through virtual environments. You can minimize the chance of motion sickness by not moving or changing direction too quickly. If you are experiencing motion sickness, you can stop the trial at any time simply by removing the head mounted display.

The knob on the back of the display can be turned to adjust the head mounted display to fit properly on your head. It should be snug enough so that you can turn your head without the display slipping. The head mounted display is also equipped with a pair of stereo headphones **through which you will hear 3D sounds coming from objects located in the virtual environment. You will not hear any sounds during this training session.** Turning your head left or right will cause your view of the environment to change, tilting your head up or down will have no effect. The mouse is also used to navigate in the environment.

SHOW PARTICIPANT MOUSE AND ASK THEM IF THEY ARE LEFT OR RIGHT HANDED. PLACE MOUSE ON SIDE OF DOMINENT HAND.

Holding down the left mouse button has the same effect as turning your head to the left, and holding down the right mouse button is equivalent to turning your head to the right. When you put on the head mounted display you will notice a small blue sphere in the center of the screen. This sphere indicates your direction of forward motion, and will always be centered in the middle of the screen. Holding down both mouse buttons at the same time will allow you to move forward in this direction. It is not possible to move backwards or sideways. You will now be given several minutes to practice using the controls and navigating around in a virtual environment. If at any time you begin to feel sick or nauseous simply remove the head mounted display . Do you have any questions?

HELP PARTICIPANT PUT ON HMD AND ALLOW THEM TO EXPLORE THE PRACTICE ENVIRONMENT UNTIL THEY ARE COMFORTABLE WITH THE EQUIPMENT.

ASK PARTICIPANT TO REMOVE HMD AND PLACE BACK ON HOLDER.

3.0 TOUR AND EXPLORATION PHASE

Now that you have had some time to become familiar with the equipment you will begin exploring a new virtual environment. **As you explore the environment you will be able to hear many different sounds coming from the different locations of the environment.** Here is a map showing the layout of the environment.

GIVE PARTICIPANT A MAP OF THE VIRTUAL ENVIRONMENT TO LOOK AT.

You will notice there are 12 distinct locations on this map. Although they are not shown on this map, there are several objects contained in each location. You should also see an arrow in one of the middle locations. This arrow indicates your starting position and orientation at the beginning of the tour, exploration phase, and each object location task. As well, you can see that the exterior, boundary walls are each a different color. Red, green, blue, and yellow. There are also colored geometric shapes indicated on the map. These shapes represent geometric objects that are located on some of the walls in the environment. Do you have any questions? Now, please spend two minutes studying this map.

ALLOW PARTICIPANT TO STUDY MAP FOR 2 MINUTES AND THEN TAKE IT AWAY FROM THEM.

During the tour you will be wearing the head mounted display to view the virtual environment. While on the tour you will be transported from location to location so that you will be able to examine all of the objects **and hear the sound** in each location. You will be placed at the entrance of each location and given 14 seconds to view all of the objects in that location. You will be unable to change position, but you will be able to turn your head and look around each location. After spending 10 seconds in each location you will hear a bell, at which point you should look straight ahead so that you will have the correct orientation for the next location. 4 seconds after hearing the bell you will be transported to the next location. The order of the locations is not significant. The purpose of this tour is only to familiarize you with the objects **and sound** in each location. Please attempt to verbalize your thoughts as much as possible throughout this tour. For example, if you are unable to recognize a particular object **or sound** please verbalize this. I will be recording what you say on an audio tape and I may remind you periodically to verbalize your thoughts. The starting location of the tour is the middle room with the arrow in it and this will also be the final location of the tour. When you have returned to this location please remove the head mounted display. Do you have any questions?

ALLOW PARTICIPANT TO WEAR HMD AND THEN PRESS <SHIFT> T TO START THE TOUR.

Now that you have had a chance to view all of the objects and locations in the environment you will be given 8 minutes to explore the environment on your own. You will be wearing the head mounted display to do this and movement will be exactly the same as it was during the practice session. Your initial starting location and orientation will be the same as at the start of the tour. Once again, please try to verbalize your thoughts as much as possible. I will be making an audio tape of what you say and the computer will be recording the exact path that you take through the environment. This is so that your exploration strategy can be later analyzed. A bell will ring when you have 2 minutes of exploration time left. Once the 8 minutes are up a message will appear on the screen indicating that the exploration time is over. When you see this message please remove the head mounted display. Keep in mind that after you have finished exploring the environment you will be asked to locate 8 objects from the environment. As well, you will later be asked to recall objects from the environment and place them in their correct locations on a map. Do you have any questions?

ALLOW PARTICIPANT TO WEAR HMD AND THEN PRESS <SHIFT> E TO START THE EXPLORATION PHASE.

4.0 OBJECT SEARCH PHASE

Now that you have had some time to become familiar with the environment I am going to ask you to locate 8 objects from the environment. Before beginning to locate each object you will be placed back at the initial starting location with the same orientation as you had when you began the exploration phase.

SHOW PARTICIPANT THE SPACE BAR ON THE KEYBOARD.

Pressing the space bar on the keyboard will signal that you are ready to begin looking an object and an image of the object to look for will appear on the screen.

As soon as this image disappears please attempt to locate the shown object as quickly as possible. To signal that you have found the object move into it. The easiest way to do this is to center the blue sphere on the object and then move forward. A message will appear on the screen saying you have found the object once you have moved into it. After this message appears you will be placed back at the initial starting location and pressing the space bar will signal that you are ready to begin looking for the next object. Pressing the space bar at any other time will also cause an image of the object you are currently looking for to be displayed for about 2 seconds. If, after 4 minutes you have been unable to find the object, a message will appear saying that your time is up and you will be placed back at the start location. You should then press the space bar to begin looking for the next object. Once you have found 4 objects please remove the head mounted display and take a short break before attempting to locate the last 4 objects. I will remind you to do this. Once again, please verbalize your thoughts as much as possible. I will be recording what you say on an audio tape and the computer will be recording your path through the environment and time to locate each object. Do you have any questions?

ALLOW SUBJECT TO WEAR HMD AND TELL THEM TO PRESS THE SPACE BAR AS SOON AS THEY ARE READY TO BEGIN.

ASK SUBJECT TO REMOVE HMD ONCE THEY HAVE FOUND THE 4TH. OBJECT.

Now take a short 2 minute rest before you begin searching for the final 4 objects. The procedure is exactly the same as it was for the first 4 objects, **except that**

all sound will be removed from the environment.

ALLOW SUBJECT TO TAKE A 2 MINUTE BREAK AND THEN ASK THEM TO PUT ON THE HMD AND BEGIN SEARCHING AGAIN.

5.0 OBJECT RECALL AND PLACEMENT TASK

Now I would like you to do a paper and pencil task which will determine how well you are able to remember objects from the environment, as well as your memory of where they were located. You will be given a map of the virtual environment you just explored and a piece of paper with a list of numbers. Try to recall as many objects from the environment as possible, writing down the name of each object beside a number. For each object you recall please try to place it in the correct location on the map. Do this by writing the number of the object in the location you think it is located in. It does not matter where in a room you place it, just try to place it in the correct room. If you have no idea which room an object is in, you should still write it down in the list of objects, but you do not have to attempt placing it in a room. However, please try to place as many objects as possible on the map, even if you have to guess which location to put them in. There is no time limit for this task, however your total time will be recorded. It is much more important to be accurate and write down as many objects as possible then it is to be fast. Again, please verbalize your thinking process as you work on this task. We will be recording what you say on an audio tape. Do you have any questions?

GIVE PARTICIPANT A MAP AND A NUMBERED PIECE OF PAPER.

6.0 GUILFORD-ZIMMERMAN ORIENTATION TEST AND FINAL QUESTIONNAIRE

Now please complete the Guilford-Zimmerman Orientation test. You will have exactly 20 minutes to complete the test. The purpose of writing this test is so that we can balance our groups according to spatial ability. We will not be providing you with any scores or interpretations from this test.

GIVE PARTICIPANT GUILFORD-ZIMMERMAN TEST.

Thank you very much for participating in this study. Before leaving could you please fill out this short questionnaire. If you have any questions please do not hesitate to contact me.

Letter of Information

The Use of 3D Sound as a Navigational Aid in Virtual Environments

Letter of Information

Student Researcher: Ryan Gunther

Dept. of Computer Science

University of Waterloo

885-1211 ext. 3290

rtcgunth@cgl.uwaterloo.ca

Title of Research: Examining the Use of 3D Sound as a Navigational Aid in Virtual Environments

STUDY OBJECTIVES

We are conducting an experiment designed to determine if incorporating 3D sound into a virtual environment will help with navigation throughout such an environment. Previous research has shown that some of the primary aids people use when navigating throughout an environment are visual landmarks. This experiment will help determine if 3D auditory cues can be used as navigational aids when there are minimal visual cues available.

TASKS

All participants will be asked to complete a pencil and paper test of general spatial ability (the Guilford-Zimmerman Orientation test) as well as a general information background sheet. The purpose of these tasks is so that we can create groups of participants which are balanced according to computer experience and spatial skills. This should take approximately 20 minutes to complete.

The rest of the testing time involves navigating through a virtual environment. In order to view and hear the virtual environment, you will be asked to wear a light-weight head mounted display with head phones which will be adjusted to fit your head. You will find that the view of the environment changes as you move your head. A mouse will be used to allow you to move within the environment. Some participants will be navigating a virtual environment enhanced with sound. The session will end with a pencil and paper object recall and map labeling task as well as a short questionnaire.

RISKS

Sometimes, people experience "motion sickness" (a sense of dizziness and nausea) when trying to navigate through virtual environments. One way to minimize such effects is to avoid moving or changing directions too quickly. Participants can stop a trial at any time by removing the head mounted display. This will shut down the visual display of the environment.

TIME COMMITMENT AND PAYMENT

For this experiment, participants are asked to sign up for one session. The length this session will be 2 hours. Total payment for participation will be $ 15. Participants will be paid at the end of the session.

RIGHT TO WITHDRAWAL

Participants can decide to withdraw from this study at any time, at which point all data for that participant will be destroyed

CONFIDENTIALITY

Any data collected will be given a numeric code, so that a participant's identity is kept confidential. The only people who will have access to the original data will be the researchers that are directly involved with this project. All data will be kept for a minimum of one year. Once we are finished with the data, all written records will be destroyed and all audio tapes will be erased.

ETHICS REVIEW

This project has been reviewed and received ethics approval from 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.

Consent Form

The Use of 3D Sound as a Navigational Aid in Virtual Environments

Consent Form

I agree to participate in the study entitled "Examining the Use of 3D Sound as a Navigational Aid in Virtual Environments". I have read over the information letter and have had the opportunity to receive additional details about my participation in this study.

I understand that full participation in this study involves one session, lasting two hours. and that compensation for full participation is $ 15, to be paid at the end of the session.

I understand that I will be asked to complete a test of general spatial ability as well as a series of object search tasks that will involve navigating through a virtual environment displayed via a lightweight head-mounted display with stereo headphones.

I understand that I will be asked to verbalize (speak out loud) my decision making while carrying out the navigation tasks. I also understand that audio tapes of my verbalizations will be made for the purposes of analyzing general strategies of navigation.

I understand that I will be asked to answer questions concerning my impressions and evaluations of the usability (helpfulness) of the auditory and visual cues used in this experiment.

I understand that I will be asked to label a map of the virtual environment that I will encounter as part of the experimental trials.

I understand that all information obtained as a result of my participation in this experiment will be kept confidential, and that I will not be individually identified in any reports or presentations pertaining to this research.

I understand that I will not be given any feedback about my individual performance on any of the tasks I perform during this study.

I understand that I have the right to withdraw my consent to participate in this experiment at any time, and that upon doing so any data collected relating to myself or my performance will be immediately destroyed.

Participant's Name:
Participant's Signature:

Name of Witness:
Signature of Witness:

Date:

Participant Background Questionnaire

PARTICIPANT BACKGROUND QUESTIONNAIRE

This information will be used for the purposes of creating general descriptions of the participants involved in this study.

Gender: ______ Male ______ Female

Age: _______ years

Dominant Hand (used for writing): ______ RIGHT ______ LEFT

Do you have any hearing impairments? ____ NO ____ YES
If yes, then please explain.

Do you have any visual impairments? ____ NO ____ YES
If yes, then please explain.

Do you typically have problems distinguishing between colours?____ NO ____ YES
If yes, which colours do you have trouble telling apart?

Over the past year, how many hours per week did you spend playing video games similar to DOOM? (Skip question if unfamiliar with DOOM)

Please circle the most appropriate number.

```
0 hrs     1 - 2 hrs    3 - 4 hrs   5 - 6 hrs    7 - 8 hrs   10 + hrs
  |------------|-----------|------------|------------|------------|
```

Over the past year, how many hours per week did you spend using a computer. (Including any time spent playing video games on a computer)

Please circle the most appropriate number.

```
0 hrs     1 - 2 hrs    3 - 4 hrs   5 - 6 hrs    7 - 8 hrs   10 + hrs
  |------------|-----------|------------|------------|------------|
```

Over the past year, how many hours per week did you spend listening to music.

Please circle the most appropriate number.

```
0 hrs     1 - 2 hrs    3 - 4 hrs   5 - 6 hrs    7 - 8 hrs   10 + hrs
  |------------|-----------|------------|------------|------------|
```

Over the past year, how many hours per week did you spend playing a musical instrument?

Please circle the most appropriate number.

```
0 hrs     1 - 2 hrs    3 - 4 hrs   5 - 6 hrs    7 - 8 hrs   10 + hrs
  |------------|-----------|------------|------------|------------|
```

How familiar would you say you are with virtual environment technology?

Please circle the most appropriate number.

```
  1            2           3            4            5            6
  |------------|-----------|------------|------------|------------|
Not at all Familiar                                          Very Familiar
```

How would you rate your navigational ability compared to that of other people?

Please circle the most appropriate number.

```
  1            2           3            4            5            6
  |------------|-----------|------------|------------|------------|
Poor                                                          Excellent
```

Final Questionnaire

The Use of 3D Sound as a Navigational Aid in Virtual Environments

FINAL QUESTIONNAIRE

Would you please take the time to answer a few questions about your experience navigating around in the virtual environment.

How easy did you find it moving around in the virtual environment?

Please circle the most appropriate number.

```
1          2          3           4           5           6
|-----------|-----------|------------|------------|------------|
Difficult                                                       Easy
```

How easy was it to recognize objects from the virtual environment?

Please circle the most appropriate number.

```
1          2          3           4           5           6
|-----------|-----------|------------|------------|------------|
Difficult                                                       Easy
```

If you were exposed to the sound enhanced environment, did you find the sounds useful for helping to locate objects?

Please circle the most appropriate number.

```
1          2          3           4           5           6
|-----------|-----------|------------|------------|------------|
Not at All                                                     Very
```

If you were exposed to the sound enhanced environment, how disruptive did you find the sounds?

Please circle the most appropriate number.

```
1          2          3           4           5           6
|----------|----------|-----------|-----------|-----------|
Not at All                                                 Very
```

If you were exposed to the sound enhanced environment, how difficult did you find it locating objects after the sounds were removed from the environment?

Please circle the most appropriate number.

```
1          2          3           4           5           6
|----------|----------|-----------|-----------|-----------|
Not at All                                                 Very
```

GENERAL COMMENTS:

Object Recognition Form

List of 36 Objects

Instructions: Below is a list of 36 objects listed in alphabetical order. For each object, decide whether or not it was located in the environment you just explored and check off the appropriate box. (IN or OUT) If you think the object was located in the environment then try to place it in the correct location on the map by writing it's number where you feel it is located. As well, for each object you place on the map rate your confidence on how accurately you were able to place it by circling the appropriate number. (0 to 4, where 0 means very little confidence and 4 means extremely confident.) You may go through the objects in any order you wish and there is no time limit. Please notify me when you are finished. If you have any questions please ask them at any time.

1) Balloon	IN ☐	OUT ☐	Confidence 0 1 2 3 4
2) Barbell	IN ☐	OUT ☐	Confidence 0 1 2 3 4
3) Baseball Bat	IN ☐	OUT ☐	Confidence 0 1 2 3 4
4) Bed	IN ☐	OUT ☐	Confidence 0 1 2 3 4
5) Brown Bench	IN ☐	OUT ☐	Confidence 0 1 2 3 4
6) Bird	IN ☐	OUT ☐	Confidence 0 1 2 3 4
7) Bowling Alley	IN ☐	OUT ☐	Confidence 0 1 2 3 4
8) Car	IN ☐	OUT ☐	Confidence 0 1 2 3 4
9) Cat	IN ☐	OUT ☐	Confidence 0 1 2 3 4
10) Clock	IN ☐	OUT ☐	Confidence 0 1 2 3 4
11) Dart Board	IN ☐	OUT ☐	Confidence 0 1 2 3 4
12) Dog	IN ☐	OUT ☐	Confidence 0 1 2 3 4
13) Drums	IN ☐	OUT ☐	Confidence 0 1 2 3 4
14) Flute	IN ☐	OUT ☐	Confidence 0 1 2 3 4
15) Frog	IN ☐	OUT ☐	Confidence 0 1 2 3 4
16) Garbage Can	IN ☐	OUT ☐	Confidence 0 1 2 3 4
17) Hammer	IN ☐	OUT ☐	Confidence 0 1 2 3 4
18) Helicopter	IN ☐	OUT ☐	Confidence 0 1 2 3 4
19) Horse	IN ☐	OUT ☐	Confidence 0 1 2 3 4
20) Lamp	IN ☐	OUT ☐	Confidence 0 1 2 3 4
21) Lion	IN ☐	OUT ☐	Confidence 0 1 2 3 4
22) Mailbox	IN ☐	OUT ☐	Confidence 0 1 2 3 4
23) Piano Keyboard	IN ☐	OUT ☐	Confidence 0 1 2 3 4
24) Ping-Pong Table	IN ☐	OUT ☐	Confidence 0 1 2 3 4
25) Plant	IN ☐	OUT ☐	Confidence 0 1 2 3 4
26) Pool Table	IN ☐	OUT ☐	Confidence 0 1 2 3 4
27) Present	IN ☐	OUT ☐	Confidence 0 1 2 3 4
28) Rifle	IN ☐	OUT ☐	Confidence 0 1 2 3 4
29) Saxophone	IN ☐	OUT ☐	Confidence 0 1 2 3 4
30) Slide	IN ☐	OUT ☐	Confidence 0 1 2 3 4
31) Sunglasses	IN ☐	OUT ☐	Confidence 0 1 2 3 4
32) Swing Set	IN ☐	OUT ☐	Confidence 0 1 2 3 4
33) Teapot	IN ☐	OUT ☐	Confidence 0 1 2 3 4
34) Telephone	IN ☐	OUT ☐	Confidence 0 1 2 3 4
35) Television	IN ☐	OUT ☐	Confidence 0 1 2 3 4
36) Violin	IN ☐	OUT ☐	Confidence 0 1 2 3 4

Object Recall Form

OBJECT RECALL TASK

1.	21.
2.	22.
3.	23.
4.	24.
5.	25.
6.	26.
7.	27.
8.	28.
9.	29.
10.	30.
11.	31.
12.	32.
13.	33.
14.	34.
15.	35.
16.	36.
17.	37.
18.	38.
19.	39.
20.	40.

Appendix C

Statistical Tests

Analysis of Variance For **FIR4TM**

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
SGP	1	4291.89	4291.89	4.5134	0.0397
GNR	1	3050.64	3050.64	3.2081	0.0807
SGP*GNR	1	169.221	169.221	0.17795	0.6753
Error	41	38987.9	950.924		
Total	44	48887.2			

Figure C.1: ANOVA of First 4 Time Against Condition and Gender

This ANOVA shows that there was a significant difference in search time for the first four objects between participants in the Sound and NoSound conditions. There was no significant difference in search time between males and females, although the probability was low that the gender differences were due to chance. There was no interaction between condition and gender. *SGP* denotes condition and *GNR* denotes gender.

Analysis of Variance For **FIR4TM**

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
SGP	1	4978.29	4978.29	4.7850	0.0345
OBR	1	283.033	283.033	0.27205	0.6048
SGP*OBR	1	259.033	259.033	0.24898	0.6205
Error	41	42655.8	1040.38		
Total	44	48887.2			

Figure C.2: ANOVA of First 4 Time Against Condition and Object Order

This ANOVA shows that there was a significant difference in search time for the first four objects between participants in the Sound and NoSound conditions. There was no significant difference in search time between object set and there was no interaction between condition and object set. *SGP* denotes condition and *OBR* denotes object set.

Analysis of Variance For **SEC4TM**

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	9977.09	4988.54	3.1253	0.0551
OBR	1	2086.87	2086.87	1.3074	0.2598
GRP*OBR	2	2856.83	1428.41	0.89489	0.4169
Error	39	62251.7	1596.20		
Total	44	80035.9			

Figure C.3: ANOVA of Second 4 Time Against Condition and Object Order

This ANOVA shows that there was no significant difference in search time between object sets and there was no interaction between condition and object set. *GRP* denotes condition and *OBR* denotes object set.

Analysis of Variance For FIR4TM

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
SGP	1	5959.07	5959.07	6.9217	0.0137
OBR	1	1310.40	1310.40	1.5221	0.2276
SGP*OBR	1	1094.24	1094.24	1.2710	0.2691
Error	28	24105.9	860.926		
Total	31	32093.4			

Figure C.4: ANOVA of First Four Time Against Condition (Male) and Object Order

This ANOVA shows that there was a significant difference in search time for the first four objects between male participants in the Sound and NoSound conditions. There was no significant difference in search time between object set and there was no interaction between condition and object set. *SGP* denotes condition and *OBR* denotes object set.

Analysis of Variance For **SEC4TM**

Source	**df**	**Sum of Squares**	**Mean Square**	**F-ratio**	**Prob**
GRP	2	19442.1	9721.05	7.6224	0.0016
GNR	1	8258.25	8258.25	6.4754	0.0150
GRP*GNR	2	7388.86	3694.43	2.8969	0.0671
Error	39	49737.6	1275.32		
Total	44	80035.9			

Figure C.5: ANOVA of Time to Locate One of Second 4 Objects Against Condition and Gender

This ANOVA shows that there was a significant difference in search time for the second four objects among the three conditions. There was also a significant difference in search time between males and females. Although not significant at the 0.05 level, it also appears there was an interaction between condition and gender. *GRP* denotes condition, *OBR* and *GNR* denotes gender.

Analysis of Variance For **TMDIFF**

Source	**df**	**Sum of Squares**	**Mean Square**	**F-ratio**	**Prob**
GRP	2	26563.5	13281.7	7.3596	0.0019
GNR	1	930.036	930.036	0.51535	0.4771
GRP*GNR	2	3629.99	1814.99	1.0057	0.3751
Error	39	70382.8	1804.69		
Total	44	100296			

Figure C.6: ANOVA of Search Time Difference Against Condition and Gender

This ANOVA shows that there was a significant search time difference among the three conditions. *GRP* denotes condition and *GNR* denotes gender.

Analysis of Variance For **TMDIFF**

Source	**df**	**Sum of Squares**	**Mean Square**	**F-ratio**	**Prob**
GRP	2	20896.6	10448.3	5.7981	0.0062
OBR	1	792.981	792.981	0.44005	0.5110
GRP*OBR	2	4190.80	2095.40	1.1628	0.3232
Error	39	70279.3	1802.03		
Total	44	100296			

Figure C.7: ANOVA of Search Time Difference Against Condition and Object Set

This ANOVA shows that there was a no significant search time difference between Object Sets and there was no significant interaction between condition and object set. *GRP* denotes condition and *OBR* denotes object set.

Analysis of Variance For SEC4TM

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	2181.54	1090.77	0.65786	0.5264
OBR	1	298.943	298.943	0.18030	0.6746
GRP*OBR	2	1837.43	918.714	0.55409	0.5812
Error	26	43109.5	1658.06		
Total	31	47756.1			

Figure C.8: ANOVA of Time to Locate One of Second 4 Objects Against Condition (Male) and Object Set

This ANOVA shows that there was no significant difference in search time for the second four objects among the three different conditions of male participants. *GRP* denotes condition and *OBR* denotes object set.

Analysis of Variance For TMDIFF

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	11536.1	5768.03	3.1721	0.0585
OBR	1	179.292	179.292	0.09860	0.7560
GRP*OBR	2	5276.92	2638.46	1.4510	0.2527
Error	26	47277.1	1818.35		
Total	31	64118.5			

Figure C.9: ANOVA of Search Time Difference Against Condition (Male) and Object Set

This ANOVA shows that there was a search time difference for males among the three conditions. *GRP* denotes condition, and *OBR* denotes object set.

Analysis of Variance For SEC4TM

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	18167.5	9083.74	20.712	0.0003
Error	10	4385.75	438.575		
Total	12	22553.2			

Figure C.10: ANOVA of Time to Locate One of Second 4 Objects Against Condition (Female)

This ANOVA shows that there was a significant difference in search time for the second four objects among the three different conditionss of female participants. *GRP* denotes condition.

Analysis of Variance For SRECAL

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	8.34565	4.17282	3.9862	0.0266
GNR	1	0.757950	0.757950	0.72406	0.4000
GRP*GNR	2	4.19418	2.09709	2.0033	0.1485
Error	39	40.8256	1.04681		
Total	44	56.9778			

Figure C.11: ANOVA of Auditory Objects Recalled Against Condition and Gender

This ANOVA shows that there was a significant difference in the number of Auditory objects recalled among the three different conditions. There was no significant gender difference. *GRP* denotes condition and *GNR* denotes gender.

Analysis of Variance For NRECAL

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	10.6371	5.31857	0.39106	0.6790
GNR	1	42.3369	42.3369	3.1129	0.0855
GRP*GNR	2	147.176	73.5881	5.4108	0.0084
Error	39	530.412	13.6003		
Total	44	734.444			

Figure C.12: ANOVA of NonAuditory Objects Recalled Against Condition and Gender

This ANOVA shows that there was no significant difference in the number of NonAuditory objects recalled among the three different conditions. Although not significant at the 0.05 level, there did appear to be a gender difference and there was a significant interaction between condition and gender. *GRP* denotes condition and *GNR* denotes gender.

Analysis of Variance For SRECAL

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	9.47436	4.73718	5.1791	0.0119
Error	29	26.5256	0.914677		
Total	31	36			

Figure C.13: ANOVA of Auditory Objects Recalled Against Condition (Male)

This ANOVA shows that there was a significant difference in the number of Auditory objects recalled by males among the three different conditions. *GRP* denotes condition.

Analysis of Variance For NRECAL

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	149.538	74.7692	5.8530	0.0073
Error	29	370.462	12.7745		
Total	31	520			

Figure C.14: ANOVA of NonAuditory Objects Recalled Against Condition (Male)
This ANOVA shows that there was a significant difference in the number of NonAuditory objects recalled by males among the three different conditions. *GRP* denotes condition.

Analysis of Variance For NUMRMS

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	76.6922	38.3461	2.8427	0.0695
Error	42	566.552	13.4893		
Total	44	643.244			

Figure C.15: ANOVA of Number of Rooms Identified Against Condition
This ANOVA shows that, although not quite significant at the 0.05 level, there was a difference in the number of rooms identified by participants in the three different conditions. *GRP* denotes condition.

Analysis of Variance For NUMRMS

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	98.9816	49.4908	4.1363	0.0263
Error	29	346.987	11.9651		
Total	31	445.969			

Figure C.16: ANOVA of Number of Rooms Identified Against Condition (Male)
This ANOVA shows that there is a significant difference in the number of rooms identified by male participants in the three different conditions. *GRP* denotes condition.

Analysis of Variance For SPLACE

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	84.0970	42.0485	3.4803	0.0442
Error	29	350.372	12.0818		
Total	31	434.469			

Figure C.17: ANOVA of Number of Auditory Objects Correctly Placed Against Condition (Male)

This ANOVA shows that there was a significant difference in the number of Auditory objects correctly placed on the map by male participants in the three different conditions. *GRP* denotes condition.

Analysis of Variance For NPLACE

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
GRP	2	158.426	79.2131	4.1809	0.0254
Error	29	549.449	18.9465		
Total	31	707.875			

Figure C.18: ANOVA of Number of NonAuditory Objects Correctly Placed Against Condition (Male)

This ANOVA shows that there was a significant difference in the number of NonAuditory objects correctly placed on the map by male participants in the three different conditions. *GRP* denotes condition.

Appendix D

Pilot Study

Purpose

Before running the actual experiment, a pilot study was performed to test several things: equipment usability, recognizability of objects and sounds, environmental complexity, usability of visual and auditory cues, and appropriateness of paper and pencil tasks. The procedure followed during the pilot study was very similar to the procedure used for the final experiment, which is explained in detail in Section 3.5. Only the results and methodological issues, as well as any changes in procedure, will be discussed in this appendix.

All participants of the pilot study were asked to verbalize their thoughts throughout their interaction with the virtual environment so that their exploration strategies could be examined, and any problems with the experimental design could be detected. During the main experiment participants were also asked to verbalize their thoughts and an audio cassette player was used to record everything said by a participant.

Participants

Eight people volunteered to participate in the pilot study, all except for one were members of the Computer Graphics Laboratory at the University of Waterloo. Seven participants were males and one was female. All of the participants had extensive computer experience, and six were particularly familiar with Computer Graphics.

Experimental Conditions

Participants were placed into one of two conditions, with each condition consisting of four participants. One group of participants was exposed to the virtual environment with all of the sounds removed (NoSound Condition). The other group of participants were exposed to the sound enhanced environment for almost the entire experiment. The sounds were only removed during the final object search task (Sound Condition).

Methodological Issues

After running the pilot study several methodological issues were noted and some changes were made to the procedure for the final experiment. This section summarizes the observations from the pilot study and explains the procedural changes that were made. A detailed description of the experimental procedure will not be given here, only the parts of the procedure that are necessary to understand why specific procedural changes were made will be discussed.

Object Recognition

People had little difficulty recognizing objects from the environment. Participants in the Sound condition gave an average rating of 5.25/6 on their ability to recognize objects, and participants in the NoSound condition gave an average rating of 5/6. However, there were several objects that people had trouble identifying. These are listed in Table D.1, along with the steps taken to improve the recognition of each object during the final experiment.

Object	**Solution**
Pogo Sticks	Removed from environment
Keyboard	Changed rotation
Sunglasses	Increased size
Violin	No change (people thought was a guitar)
Wrapped Gift	Increased size
Bench	Nothing
Birds	Increased size
Hammer	Changed position and rotation

Table D.1: Difficult Objects to Recognize

The above table lists all of the objects that participants in the pilot study had a difficult time recognizing. The changes we made to each object for the experiment are shown in the table as well. The violin was left unchanged because everyone thought it was either a violin or a guitar, and since there were no guitars in the environment, if people reported seeing a guitar we would know which object they meant.

Sound Recognition

All of the participants who were exposed to the sounds from the environment reported that the sounds were easily identifiable, although people did have more difficulty identifying the bowling and seagull sounds.

Movement

For the pilot study it was decided to design movement to be similar to that of a person in a wheelchair. Participants were asked to sit in a chair during navigation, where performing a head rotation would change the view of the environment, but not the direction of forward motion. A mouse, sitting on a table in front of the participants, was used to change the direction of forward motion. Holding down the left mouse button would execute a left turn, and holding down the right mouse button executed a right turn. Holding down both mouse buttons simultaneously allowed a person to move forward. When using the mouse to turn, both the view of the environment and the direction of forward motion were changed. A small blue sphere was used to indicate the current direction of forward motion. For example, when a subject was looking straight ahead the blue sphere would be centered in the middle of the screen, and if they rotated their head to the left then the sphere would shift to the right with respect to their field of view. Figure D.1 shows the field of view of a person after a left head rotation. The blue sphere (shown as dark grey in the image) can be seen to the far right in their field of view.

Unfortunately, this metaphor for movement did not seem to work very well, and after several minutes most participants only used the mouse to control their movement. After observing participants and listening to their comments several reasons why the metaphor was not working as well as anticipated were identified.

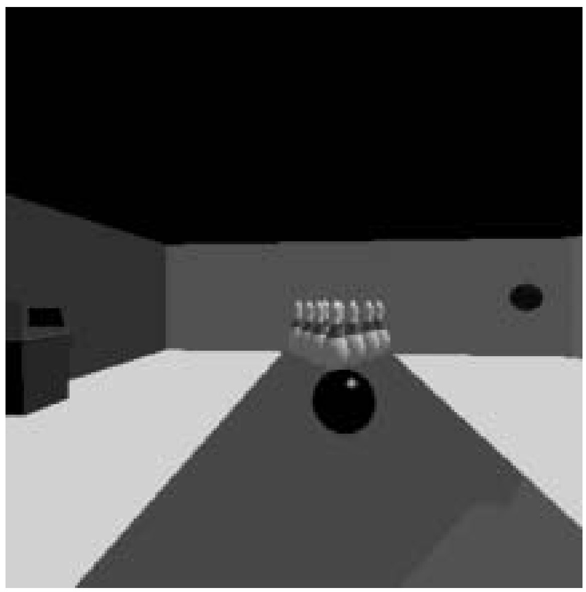

Figure D.1: Image Illustrating Movement Metaphor

This picture illustrates the movement metaphor for the pilot study. The picture shows the field of view for a person after they have turned their head to the left. The small, black sphere to the far right of the image indicates the direction of forward motion.

One problem was that people always wanted to be able to move forward in the direction they were looking, and this was not possible once they had rotated their head. Another problem was caused by a slight lag in response time that made it challenging to align the blue sphere in the center of the screen after a head rotation. It was also felt that the small field of view of the head-mounted display limited the amount of head movement because participants never saw anything out of the corner of their eye. Finally, all of the participants in the pilot study were very experienced using a mouse, but not a head-mounted display. It makes sense that they would use the device that was more familiar to them.

However, even given the design flaws in the movement mechanism, people did not have much trouble moving around the environment. The average rating given to the mouse as a tool for moving around in the virtual environment was 4.25/6. None of the participants reported feeling motion sickness, although a couple of people felt a little disoriented after the experiment was over. One person mentioned that they only began feeling dizzy half an hour after the experiment had ended.

Several people commented that it would have been nice to be able to move backwards or make a 180 degree turn. Some people also found that when turning the rate of movement was a little slow. For the final experiment the method of movement was changed so that people would always move in the direction they were looking. To accomplish this the effects of head rotations were altered so that they would change both the view of the environment and the direction of forward motion. This meant that both a head rotation and a mouse button push performed equivalent actions and the blue sphere was permanently centered in the middle of the screen, since this was now always the direction of forward motion. It was felt that this would make movement easier, and might encourage people to use the headtracking capabilities of the head-mounted display more. An upper bound of ten frames per second was placed on the framerate so that the rate of movement would not change drastically near areas of high visual complexity.

As well, an attempt was made to reduce individual differences in the ability of participants to move through the environment by adding an equipment training session to the beginning of the experiment. Each participant was placed in a small virtual environment and allowed to practice using the head-mounted display and mouse until they felt comfortable moving around. This was not considered necessary for the pilot study because all of the participants in the pilot study had a strong computer background.

Environmental Design

The configurational complexity of the environment seemed appropriate. Most people were able to develop some configurational knowledge, but no one found the object search or placement tasks trivial. Participants were given eight minutes to explore the environment and this was enough time for participants to visit every location at least once. At the end of the experiment participants were asked if they felt that the amount of time they were given to explore the environment was adequate. All of the participants stated that it was a reasonable amount of time, although one participant mentioned that some indication of how much time they had left would have been nice. It was decided that to avoid participants becoming distracted over the amount of time they had left a beep would be added to signal that two minutes of exploration time were left.

After both observing participants navigate through the environment and listening to their verbalizations it appeared that all of the participants relied on the colored walls to maintain their orientation and location in the environment. As expected, the participants in the NoSound condition appeared to use them more than the participants in the Sound condition. As well, a common strategy among participants in both conditions was to associate each exterior room with the nearest colored wall. This worked very well when objects in a room had a strong association with the color of this wall. For example, the room with the frogs and plants was located beside the green wall and as a result participants were able to find this room very quickly. It took participants an average of only 31 seconds to locate the frog. This was the second fastest time for all of the objects. Since the intent of this study was to focus on the auditory cues and not the visual cues the green and yellow walls were swapped for the final experiment to avoid this unintended cue.

The colored geometric objects that had been placed at intersection points were not useful navigational aids and appeared to be ignored by most of the participants. However, it was decided to leave them in the environment for the final experiment since it may have been the case that participants were subconsciously using them.

Tour

Prior to exploring the environment on their own for eight minutes, participants were taken on a tour which brought them to each room of the environment. During the tour participants were unable to control movement, although they could change their view of the environment by turning their head. The tour consisted of participants being transported to each room entrance, where they were given twelve seconds to examine all of the objects and sounds in the room. The order of the rooms on the tour was the same for each participant. The purpose of providing the tour was to ensure that each participant was exposed to all of the objects and sounds at least once.

The design of the tour appeared to be adequate, although several participants complained that the amount of time spent in each room was too short. It was decided to increase the length of time spent in each room during the final experiment to fourteen seconds so that participants would feel less rushed.

Object Search Tasks

After participants had explored the virtual environment they were asked to locate twelve objects from the environment. Before beginning to locate each object they were placed at the start location which was shown in Figure 3.3. As a measure of wayfinding, the time to locate each object was recorded. Participants rarely took more than 4 minutes to locate an object.

Participants were required to find two types of objects, ones that had a sound associated with them and ones with no associated sound. The objects were selected randomly for each subject, given the constraint that a minimum of six Auditory objects and three NonAuditory objects had to be chosen, as well as exactly one object from each room. In the Sound condition, all sound was turned off in the environment for the last 3 object location trials.

There were large differences in the amount of time it took participants in the Sound and NoSound conditions to locate the first nine objects. Participants in the Sound group were able to find Auditory objects much more quickly than participants in the NoSound condition. The opposite was true for NonAuditory objects, participants in the NoSound group were able to find these objects much faster than participants in the Sound group. Figure D.2 shows the times it took participants to find the first nine objects. Both groups of participants performed equally well locating the final three objects.

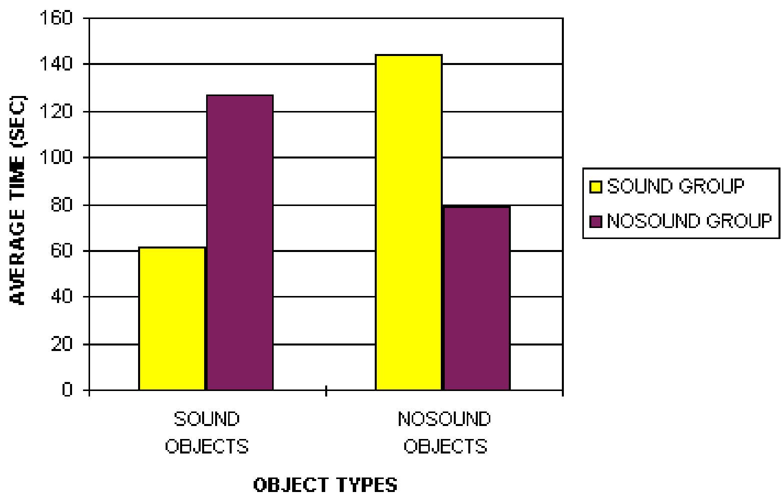

Figure D.2: Time For Pilot Participants to Find Objects

The above graph shows the times for participants in both the Sound and NoSound conditions to find Auditory and NonAuditory objects. Participants in the Sound condition took less time to locate Auditory objects than did participants in the NoSound condition. The reverse was true for NonAuditory objects.

It was thought that by having participants search for Auditory and NonAuditory objects it would be possible to determine if those objects associated with a sound were easier to locate than those objects with no associative sounds. However, after running the pilot study it became apparent that there were two problems with the design. The main problem was that it appeared that what was being tested was how well people were able to remember which objects were in the same room together. When asked to locate an object with no associative sound most people attempted to remember which Auditory object was in the same room, and then look for that object by listening to the spatialized sounds. Another problem was caused by the way rooms in the environment were designed. The dominant object in each room was the object that was associated with the sound for the room. Even participants who were never exposed to any sounds referred to rooms by the names of the Auditory objects. For the final experiment it was decided to have participants search only for objects that were associated with sounds.

As well, for the final experiment it was decided not to randomize the order of the objects that participants were asked to look for. This would make it possible to analyze and compare the different routes taken by participants as they attempted to find each object. It would also ensure that all of the objects a participant was required to locate were not clustered together in one area of the environment.

Some participants, especially those in the NoSound condition, became frustrated and tired near the end of the experiment. For the final experiment the total number of objects each participant was asked to locate was reduced from twelve to eight. Participants were also given a maximum of four minutes to locate each object. After four minutes were up a message would appear on the screen indicating that time was up and the participant would be transported back to the start location to begin looking for the next object.

Object Recognition and Placement Task

After locating all twelve objects, participants were given a map of the environment and a list of thirty-six numbered objects, shown in Appendix B. Twenty-four of these objects were from the virtual environment, twelve Auditory and twelve NonAuditory. The other twelve objects were not from the environment, although they were similar to the types of objects found in the environment. Beside the name of each object on the list were two boxes, one labeled IN and the other OUT. Participants were required to decide if an object was from the virtual environment and then check the appropriate box. For each object that a participant indicated as IN they were asked to write it down on the map in the room they believed it was located in. They were also asked to indicate their confidence in their placement by circling a number on a scale from one to five. The purpose of this task was to test how well participants were able to recognize objects from the environment, as well as how much survey knowledge they had been able to acquire.

After looking at the data from this activity it was obvious that the object recognition task was too easy. Figure D.3 shows a graph of the number of objects identified by participants. Participants in the Sound condition correctly identified an average of 21.25 objects as being from the environment and participants in the NoSound condition correctly identified 21.50 as in. As well, participants in the Sound condition correctly identified an average of 11 objects that were not from the environment. The average was 10.5 for participants in the NoSound condition. To make this task more challenging it was changed from an object recognition task into an object recall task. Instead of providing participants with a list of objects, they were given a blank piece of paper and asked to write down the name of as many objects from the environment as they could remember.

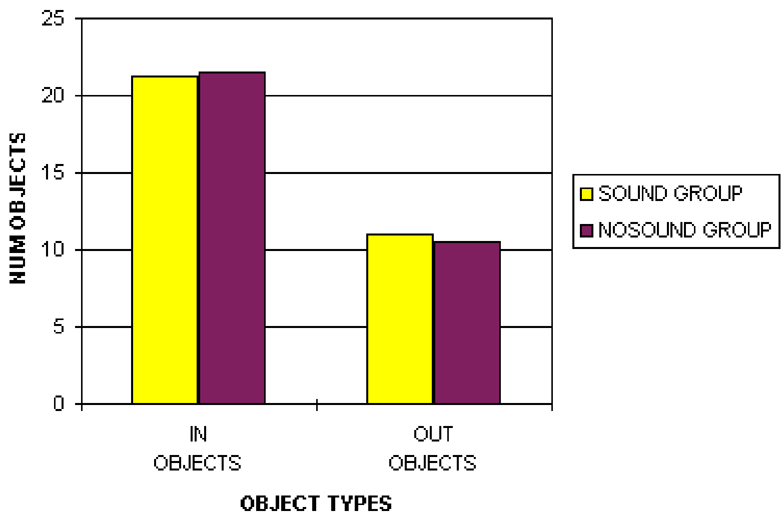

Figure D.3: Number of Objects Recognized by Pilot Participants

The above graph shows the numbers of IN and OUT objects identified by participants in both the Sound and NoSound conditions. Participants in both conditions identified approximately the same number of objects. IN objects were those which were from the environment and OUT objects were those which were not from the environment.

The object placement task seemed to be much more difficult than the object recognition task. Participants in the Sound condition were only able to correctly place an average of 6.5 Auditory objects and 3.5 NonAuditory objects. Participants in the NoSound group performed slightly better, correctly placing an average of 6 Auditory objects and 4.75 NonAuditory objects. Figure D.4 shows the number of objects correctly placed by participants. This task was left unchanged for the final experiment.

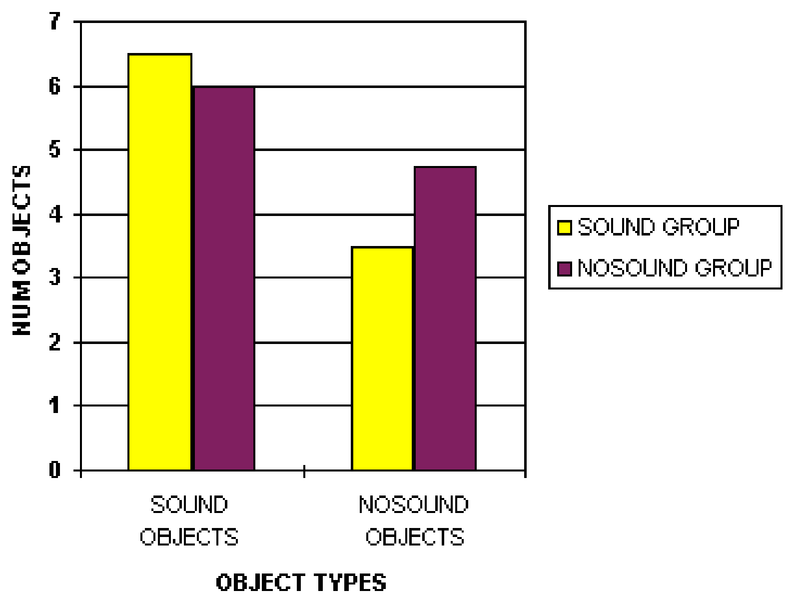

Figure D.4: Number of Objects Correctly Placed by Pilot Participants

The above graph shows the numbers of Auditory and NonAuditory objects correctly placed on the map of the environment by participants in both the Sound and NoSound conditions. Participants in the Sound condition were able to correctly place more Auditory objects than participants in the NoSound condition. The reverse was true for NonAuditory objects.

Spatial Ability

After running eight participants through the pilot study it was clear from observing people that there were large individual differences in spatial ability, which affected performance on the different tasks. It was decided to have all participants write the Guilford-Zimmerman Spatial Orientation test during the final experiment so that it would be possible to determine how balanced the different conditions were with respect to spatial ability. This was a ten minute, multiple choice test, and was described in Section 2.5.2.

Appendix E

Experiment Data

Table	Label	Description
Object Search Data	GRP	condition (1=NSoundA, 2=FSound, 3=PSoundA, 4=NSoundB, 5=PSoundB
	SEX	gender (1 = Male, 2 = Female)
	OBn	time (in sec) to locate the nth object
Background Questionnaire Data	AGE	participant's age
	DH	dominant hand (1 = Right, 2 = Left)
	HI	hearing impairments (1 = No, 2 = Yes)
	VI	visual impairments (1 = No, 2 = Yes)
	DC	trouble distinguishing colors (1 = No, 2 = Yes)
	DM	average hours per week spent playing DOOM
	CX	average hours per week spent using a computer
	LM	average hours per week spent listening to music
	PM	average hours per week spent playing music
	VR	familarity with virtual reality
	NA	self assessment of navigational ability
Final Questionnaire Data	MV	ability to move around the environment
	RE	ability to recognize objects in the environment
	SU	usefulness of the sounds
	SD	disruptiveness of the sounds
	SR	difficulty locating objects after the removal of sound
Object Recall & Placement Data	GZAS	Guilford-Zimmerman test score
	SREC	number of Auditory objects recalled
	NREC	number of NonAuditory objects recalled
	SPLC	number of Auditory objects correctly placed on map
	NPLC	number of NonAuditory objects correctly placed on map
	NRMS	number of rooms on map correctly identified
Participant Movement Data	ETS	total environmental sections moved through during exploration
	ETR	total rooms entered during exploration
	EXR	total exterior rooms entered during exploration
	EIR	total interior rooms entered during exploration
	ST	total number of times start location was ebtered during exploration
	EXH	total exterior hallways entered during exploration
	EIH	total interior hallways entered during exploration

Table E.1: Description of Data Labels

Data with a (*) was dropped from the analysis.

ID	GRP	SEX	OB1	OB2	OB3	OB4	OB5	OB6	OB7	OB8
001	3	1	31	70	17	51	130	28	23	26
002	3	1	72	150	13	162	135	119	240	240
003	1	1	52	217	164	50	118	45	53	53
004	3	1	92	90	66	160	143	145	164	25
005	1	2	32	73	60	139	137	28	89	43
006	3	2	30	34	62	102	42	221	152	168
007 (*)	1		240	240	46	218	29	240	200	58
008	1	1	20	95	221	37	36	27	125	26
009	3	2	116	135	99	104	133	94	240	88
010	3	1	54	29	87	36	108	34	25	21
011	1	1	78	44	22	26	114	72	50	51
012	1	2	127	76	15	240	145	61	183	53
013	2	1	88	79	36	63	48	56	49	80
014	2	2	23	58	14	38	20	31	109	53
015	2	1	31	48	19	70	19	131	106	36
016	2	2	39	72	59	133	26	47	104	35
017	2	1	54	50	17	153	56	40	30	46
018	1	1	27	240	41	91	103	52	39	24
019	2	2	179	190	25	53	67	140	73	56
020	3	2	47	33	20	130	240	49	219	240
021	1	1	240	144	34	201	41	45	240	27
022	3	2	38	37	233	147	56	122	240	128
023	1	1	33	240	64	33	16	27	45	26
024	3	1	33	67	17	91	31	28	27	27
025	2	1	26	27	87	129	86	147	77	31
026	3	1	31	47	17	29	178	40	25	24
027 (*)	2		86	122	38	122	211	115	73	55
028	3	1	45	96	23	76	161	240	135	68
029	2	2	58	70	76	72	50	62	149	82
030 (*)	2		86	225	240	99	127	240	82	113
031	2	1	62	71	21	125	28	86	37	37
032	2	1	23	28	17	73	30	74	48	25
033	4	1	117	127	104	164	135	24	166	26
034	5	1	57	87	26	25	23	37	63	34
035	1	2	47	45	26	135	99	33	42	81
036 (*)	5		67	196	218	74	46	69	19	201
037	4	1	54	31	43	34	25	72	13	25
038	4	1	59	66	36	30	26	36	17	47
039	5	1	19	29	23	25	20	93	142	46
040 (*)										
041	5	1	82	78	51	105	22	30	28	63
042	5	1	57	27	49	28	43	26	15	75
043	4	1	42	30	36	23	203	241	84	30
044	4	1	148	104	35	90	59	28	14	67
045	5	2	88	240	55	83	34	240	221	66
046	4	1	103	102	142	47	23	17	12	81
047	4	2	240	87	240	42	60	42	21	118
048	5	1	60	48	51	97	132	170	43	57
049	4	1	129	65	89	45	240	159	46	32
050	5	1	90	112	72	104	78	127	22	214

Table E.2: Object Search Data

ID	AGE	DH	HI	VI	DC	DM	CX	LM	PM	VR	NA
001	22	1	1	1	1	1.5	10	3.5	1.5	3	6
002	25	1	1	1	1	0	10	0	0	2	3
003	24	1	1	1	1	1.5	10	7.5	1.5	5	4
004	22	1	1	1	1	0	10	10	0	1	5
005	25	1	1	1	1	0	7.5	3.5	0	2	4
006	24	1	1	1	1	0	5.5	10	1.5	1	3
007 (*)	23	1	1	1	1	0	7.5	1.5	0	1	5
008	22	1	1	1	1	1.5	10	5.5	0	2	4
009	22	1	1	1	1	1.5	10	10	0	3	3
010	32	1	1	1	1	1.5	10	10	0	3	4
011	23	1	1	1	1	0	1.5	7.5	1.5	1	2
012	24	2	2	1	1	0	10	10	10	4	4
013	31	1	1	1	1	0	10	5.5	0	2	2
014	27	1	1	1	1	5.5	10	10	1.5	4	5
015	16	1	1	1	1	0	10	7.5	0	4	3
016	25	1	1	1	1	0	10	10	1.5	2	4
017	24	2	1	1	1	0	10	10	0	3	3
018	23	1	2	1	1	0	10	1.5	0	4	4
019	27	1	1	1	1	0	1.5	1.5	0	1	1
020	18	1	1	1	1	0	7.5	10	0	2	1
021	27	1	1	1	1	0	10	7.5	0	2	5
022	25	1	1	1	1	0	10	10	0	3	4
023	30	1	1	1	1	1.5	10	10	0	5	4
024	24	1	1	1	1	1.5	10	7.5	1.5	2	3
025	28	1	1	1	2	0	10	5.5	0	1	3
026	29	1	1	1	1	7.5	10	1.5	0	4	5
027 (*)	26	1	1	1	1	0	10	10	0	2	5
028	27	1	1	1	1	0	5.5	7.5	1.5	2	4
029	21	1	1	1	1	0	5.5	10	0	2	3
030 (*)	27	1	1	1	1	0	10	3.5	0	2	4
031	27	1	1	1	1	0	10	1.5	0	4	4
032	22	1	1	2	1	0	5.5	7.5	7.5	1	2
033	25	1	1	1	1	0	10	1.5	3.5	1	2
034	22	1	1	1	1	0	10	10	0	4	5
035	26	1	1	1	1	0	10	10	0	1	3
036 (*)	23	1	1	1	1	0	10	3.5	0	1	1
037		1	1	1	1	1.5	10	10	1.5	3	5
038	24	1	1	1	1	1.5	10	7.5	0	4	3
039	24	1	1	1	1	1.5	10	10	1.5	2	5
040 (*)											
041	31	1	2	2	1	1.5	10	10	1.5	5	5
042	22	1	1	1	1	0	7.5	7.5	0	4	4
043	26	1	1	1	1	0	10	10	1.5	5	3
044	24	1	1	1	1	1.5	5.5	1.5	0	1	4
045	24	1	1	1	1	0	10	5.5	0	2	2
046	26	1	1	1	1	0	10	0	0	1	4
047	28	2	1	1	1	1.5	10	5.5	0	4	4
048	31	1	1	1	1	0	5.5	5.5	3.5	2	4
049	24	1	1	1	1	0	10	10	5.5	3	3
050	26	1	1	1	1	0	10	1.5	0	2	4

Table E.3: Background Questionnaire Data

ID	MV	RE	SU	SD	SR
001	6	6	5	2	5
002	6	6	6	1	6
003	2	6			
004	5	6	6	1	4
005	5	5			
006	3	6	6	2	6
007 (*)	6	6			
008	5	5			
009	4	3	6	2	6
010	5	6	5	4	3
011	5	5			
012	6	6			
013	2	5	6	4	
014	5	6	6	1	
015	4	6	6	1	
016	5	5	6	1	
017	5	6	6	3	
018	4	6			
019	4	6	5	2	
020	2	1	6	3	6
021	5	5			
022	6	6	6	4	5
023	5	5			
024	4	5	5	2	2
025	5	6	6	1	
026	6	6	6	1	5
027 (*)	5	6	6	2	
028	5	4	6	2	5
029	5	6	6	1	
030 (*)	5	4	5	4	
031	5	4	6	2	
032	6	6	6	1	
033	5	6			
034	5	5	4	2	3
035	4	6			
036 (*)	4	5	6	1	5
037	5	5			
038	5	4			
039	5	5	5	2	3
040 (*)					
041	5	5	4	1	3
042	4	5	5	3	3
043	5	6			
044	6	6			
045	2	5	4	3	5
046	5	6			
047	3	4			
048	5	5	6	2	4
049	5	6			
050	5	6	6	3	6

Table E.4: Final Questionnaire Data

ID	GZAS	SREC	NREC	SPLC	NPLC	NRMS
001	51.75	12	15	12	14	12
002	1.25	11	1	1	0	1
003	30.25	9	5	5	3	5
004	27.25	10	5	2	0	2
005	10.25	9	10	2	0	2
006	7.75	9	5	0	0	0
007 (*)	20.0	8	11	6	8	8
008	40.75	12	11	12	10	12
009	22.50	10	11	4	2	4
010	22.50	11	9	10	6	10
011	18.5	11	7	9	5	10
012	16.75	11	11	2	3	2
013	28.0	12	6	0	0	0
014	14.25	11	9	3	2	3
015	29.75	11	1	8	1	8
016	9.5	12	14	10	10	10
017	34.0	11	8	5	3	5
018	17.75	9	9	8	8	8
019	20.75	10	8	1	0	1
020	7.75	8	12	0	2	1
021	11.75	9	12	4	2	5
022	2.0	12	9	2	2	2
023	35.75	10	13	11	13	11
024	28.0	9	9	4	3	5
025	14.25	10	5	0	0	0
026	26.75	12	12	10	10	10
027 (*)	18.0	9	5	1	0	1
028	11.50	11	8	3	5	3
029	19.0	12	18	2	4	2
030 (*)	9.0	7	3	0	1	1
031	13.75	10	0	2	0	2
032	18.25	11	4	4	1	4
033	30.0	10	12	6	8	7
034	31.50	12	17	10	13	11
035	26.50	10	1	4	1	5
036 (*)	5.75	11	5	0	0	0
037	11.50	10	11	10	11	10
038	15.0	11	10	10	10	10
039	27.25	11	16	11	16	11
040 (*)						
041	50.0	10	7	9	7	9
042	12.50	11	8	7	6	8
043	29.0	8	10	6	8	8
044	25.25	10	6	5	1	5
045	11.50	10	10	2	1	2
046	23.0	10	8	9	6	9
047	9.25	9	10	2	5	4
048	17.75	11	6	3	1	3
049	37.25	9	10	5	3	5
050	24.0	12	11	0	0	0

Table E.5: Object Recall & Placement Data

ID	ETS	ETR	EXR	EIR	ST	EXH	EIH
001	60	13	8	5	4	10	33
002	56	13	10	3	6	11	26
003	60	13	9	4	6	12	29
004	73	14	10	4	5	13	41
005	51	10	7	3	2	12	27
006	60	7	6	1	3	12	38
007 (*)	44	10	7	3	3	6	25
008	81	12	11	1	3	18	48
009	45	10	7	3	2	9	24
010	59	9	5	4	5	12	33
011	52	7	4	3	4	8	33
012	50	12	9	3	5	10	23
013	57	4	3	1	4	14	35
014	63	11	8	3	9	12	31
015	68	14	9	5	3	10	41
016	50	11	6	5	3	8	28
017	73	9	7	2	4	15	45
018	66	12	8	4	2	8	44
019	41	7	4	3	2	9	23
020	46	8	6	2	3	10	25
021	50	11	8	3	3	11	25
022	60	13	8	5	2	15	30
023	85	12	7	5	2	16	55
024	62	4	2	2	3	13	42
025	56	11	8	3	4	10	31
026	66	8	6	2	3	14	41
027 (*)	33	7	3	4	2	6	18
028	48	11	7	4	2	10	25
029	51	10	6	4	3	10	28
030 (*)	43	10	7	3	4	10	19
031	72	13	12	1	6	17	36
032	52	3	1	2	7	9	33
033	27	5	3	2	2	8	12
034	58	12	9	3	2	12	31
035	48	11	9	2	2	15	20
036 (*)	82	8	6	2	5	15	54
037	51	12	8	4	3	9	27
038	61	12	8	4	2	9	38
039	74	14	10	4	4	12	44
040 (*)							
041	62	15	11	4	1	13	33
042	57	12	9	3	4	12	29
043	88	14	9	5	2	9	63
044	66	8	5	3	5	12	41
045	26	6	4	2	1	9	10
046	63	15	10	5	5	12	31
047	64	10	7	3	3	10	41
048	71	13	9	4	4	10	44
049	75	14	9	5	6	10	45
050	57	11	7	4	5	7	34

Table E.6: Participant Movement Data