Design and evaluation of auditory spatial cues for decision making within a game environment for persons with visual impairments

Michael Anthony Oren
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Computer Sciences Commons

Recommended Citation

This Thesis is brought to you for free and open access by Digital Repository @ Iowa State University. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Digital Repository @ Iowa State University. For more information, please contact digirep@iastate.edu.
Design and evaluation of auditory spatial cues for decision making within a game environment for persons with visual impairments

by

Michael Anthony Oren

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Human Computer Interaction

Program of Study Committee:
Christopher Harding, Major Professor
Christopher Hopkins
Stephen Gilbert

Iowa State University
Ames, Iowa
2008

Copyright © Michael Anthony Oren, 2008. All rights reserved.
UMI Number: 1460090

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform 1460090
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346
## TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ iv

LIST OF TABLES .......................................................................................................... v

ACKNOWLEDGEMENTS ............................................................................................... vi

ABSTRACT ................................................................................................................... viii

INTRODUCTION .......................................................................................................... ix

Thesis Organization ................................................................................................... x

CHAPTER 1. BACKGROUND AND LITERATURE REVIEW ...................................... 1

1. Background ......................................................................................................... 1

2. Selective Literature Review ............................................................................. 2

2.1. Visual Impairments ...................................................................................... 2

2.2. Audio Games ............................................................................................... 3

2.3. Spatial Mapping ........................................................................................... 6

CHAPTER 2. SPEED SONIC ACROSS THE SPAN: A PLATFORM AUDIO GAME .... 8

Abstract .................................................................................................................. 8

1. Introduction ....................................................................................................... 8

1.1. Project and Motivation .............................................................................. 8

1.2. Background: Audio Games .................................................................... 9

1.3. Sound Design ........................................................................................... 11

2. User Study ....................................................................................................... 16

2.1. User Study Design .................................................................................... 16

2.2. Results ....................................................................................................... 16

3. Conclusion ....................................................................................................... 18

CHAPTER 3. AUDIO PLATFORM GAME (APG) DESIGN FOR PLAYERS WITH VISUAL IMPAIRMENTS ........................................................... 20

Abstract ............................................................................................................... 20

1. Introduction ................................................................................................... 20

2. Pilot User Study for Game Prototype .......................................................... 22

2.1. Method ....................................................................................................... 22

3. User Experiment for Final Game .................................................................. 29

3.1. Method ....................................................................................................... 29

3.2. Procedure ................................................................................................... 30

4. Results ........................................................................................................... 31

4.1. Game Performance .................................................................................... 31

5. Discussion ....................................................................................................... 34

CHAPTER 4 MAPPING METROID: EVALUATION OF 2D SPATIAL NAVIGATION ABILITIES IN PERSONS WITH VISUAL IMPAIRMENTS USING AN AUDITORY PLATFORM GAME ...................................................... 36

Abstract ............................................................................................................... 36
1. Introduction ........................................................................................................ 36
2. Background ....................................................................................................... 38
  2.1. Spatial Ability / Spatial Maps .................................................................... 38
  2.2. Sonification ............................................................................................... 40
3. Design and Development .................................................................................. 41
  3.1. Hardware and Software ............................................................................ 41
  3.2. Design ...................................................................................................... 42
4. Methods .......................................................................................................... 44
  4.1. Participants ............................................................................................... 44
  Procedure ........................................................................................................ 45
5. Results ........................................................................................................... 47
  5.1. Mapping Scores ....................................................................................... 47
  5.2. Follow-up Survey .................................................................................... 50
6. Conclusion ....................................................................................................... 53
7. Acknowledgements .......................................................................................... 54

CHAPTER 5 SUMMARY AND FUTURE WORK .................................................. 55
Future Work ....................................................................................................... 57

REFERENCES ...................................................................................................... 58
LIST OF FIGURES

Figure 1 A screenshot of the game with the primary object labeled “You”................................................................. 11
Figure 2 A screenshot of the game with the primary object labeled “You”................................................................. 11
Figure 3 A cartoon showing a collision with floating platform (top) and with a solid platform (bottom)................................................................. 13
Figure 4 A screen shot of the game with the primary objects labeled................................................................. 25
Figure 5 The four levels used in the game for platforms........................................................................................................... 25
Figure 6 A cartoon showing a collision with floating platform (top) and with a solid platform (bottom)................................................................. 26
Figure 7 A screenshot from Super Mario Bros........................................................................................................... 38
Figure 8 The four floors of the game................................................................................................................................. 43
Figure 9 A screenshot showing all objects used in the game........................................................................................................... 43
Figure 10 Left image is the map of level 3, right is the map of level 7. These are the two levels participants were asked to map........................................................................................................... 47
Figure 11 The results from one of the sighted participants in the audio-only group................................................................. 48
# LIST OF TABLES

Table 1 Primary objects in the game world and their auditory representations .................. 12  
Table 2 Primary Objects and Actions in the Game World with Their Auditory Representations .............................................................................................................. 24  
Table 3 Means and Standard Deviations for Survey Response for Participants with Normal Vision and with Visual Impairments .............................................................................. 33  
Table 4 Primary Objects and Actions in the Game World with Their Auditory Representation ......................................................................................................................................... 43  
Table 5 Breakdown of participants by game version played, visual disability and gender .... 44  
Table 6 Means and Standard Deviations for Survey Response for Participants with Normal Vision and with Visual Impairments ......................................................................................................................... 51
ACKNOWLEDGEMENTS

This project has been several years in the making. Many people throughout the process have aided me that I would like to acknowledge them. In particular, I would like to thank Kyle Shipley who helped me build the initial version of this game by writing the code that handles the graphic display and the collisions as well as helping me make arrangements with the Indiana School for the Blind to speak with their students and test our prototypes. In addition, I would like to thank Doug Harms and Dave Berque of DePauw University for introducing me to research, and especially Dave for introducing me to the field of Human Computer Interaction. I’d also like to thank Kelly VanBusum, former Associate Professor at DePauw University, for directing me to several different audio modules for Python and for recommending the use of pySonic.

I am also very grateful to Al Lovati and the students of the Indiana School for the Blind without whom this project would not have been possible as it was a design partnership with them that led to this project’s fruition. Mike Hooley and the Iowa Braille and Sight Saving School’s students and staff were also instrumental in evaluating this project and I am extremely grateful for having had the opportunity to work with them. I would also like to thank the Iowa State Disability Resource Center for helping me recruit participants on campus. My appreciation is also extended to all the participants of all stages of my study who were patient and willingly gave me an hour of their time despite little (ice cream) to no compensation.

Most importantly, I would also like to thank Chris Harding for serving as my advisor and introducing me to the classic sonification research as well as his help in getting this work accepted to conferences. I would also like to thank both Christopher Hopkins and Stephen Gilbert for agreeing to be on my committee and for pointing me in the direction of several concepts in electronic music and psychology that helped articulate some concepts and
introduced me to new ideas that might further extend this project in the future. Also, Terri Bonebright of DePauw University is not officially on my committee, I would like to thank her for the much needed aid she provided in helping to plan the study and run the statistical analysis as well as her tireless effort to help revise conference and journal paper submissions about the project.

My thanks also extends to the members of the audio gaming community, particular AudioGames.net, Audyssey e-zine, members of the game-accessibility.com forums, and the over one thousand people who downloaded copies of this project for helping to motivate me and for appreciating the work not just from an academic point of view but from a creative side as well. These are the people whom I originally envisioned making the game for and these are the people who will motivate me to continue to improve upon the project in the future despite its official completion on the academic side.

My family and friends have always offered me their support whether I knew I needed it or not and even when I thought I did not want it, and for that I am forever grateful. In particularly I’d like to thank my parents, Yolanda Jordan and Ronald Oren, for their encouragement and support as well as for buying the video game systems and computers that first led to my interest in video games and programming.

Finally, I would like to thank Danielle Skwierczynski, my girlfriend of almost five years now, for inspiring my interest in music and sound work that led to the idea of creating an audio game as well as her support over the years.
ABSTRACT

An audio platform game was created and evaluated in order to answer the question of whether or not an audio game could be designed that effectively conveys the spatial information necessary for persons with visual impairments to successfully navigate the game levels and respond to audio cues in time to avoid obstacles. The game used several types of audio cues (sounds and speech) to convey the spatial setup (map) of the game world. Most audio-only players seemed to be able to create a workable mental map from the game’s sound cues alone, pointing to potential for the further development of similar audio games for persons with visual impairments. The research also investigated the navigational strategies used by persons with visual impairments and the accuracy of the participants’ mental maps as a consequence of their navigational strategy. A comparisons of the maps created by visually impaired participants with those created by sighted participants playing the game with and without graphics, showed no statistically significant difference in map accuracy between groups. However, there was a marked difference between the number of “invented” objects when we compared this value between the sighted audio-only group and the other groups, which could serve as an area for future research.
INTRODUCTION

The goal of this research was to overcome the interface design challenges that are involved in creating auditory spatial displays. In order to meet these interface design challenges, several sonification strategies proposed by Flowers, Blattner, Kramer, and Gaver were explored as the foundation of this research (Flowers, 2005; Blattner, 1989; Kramer, 1999; Gaver, 1988 and 1997). This research also complements work by Golledge, Klaztzky, Loomis, and Marston on how persons with visual impairments interpret audio cues for the purpose of route finding (Loomis, 2005; Marston, 2007; Klatzky, 1995; Golledge, 2004). In contrast to Golledge et al. (2004), who found that persons with visual impairments prefer verbal feedback to non-verbal sonification techniques for route finding, this work takes place in the reality of the virtual game world and emphasizes non-verbal techniques over verbal feedback to enhance player immersion. Similar to the guideline for the creation of simple audio games outlined in (Gärdenfors, 2002), the results of this research may help the development of more complex, especially spatially based audio games. Using the audio-only version, several persons with visual impairments were able to play this game at a level on par with sighted participants playing the audio-visual version in terms of the number of failures to avoid obstacles and in terms of the accuracy of their mental maps; however, persons with visual impairments, as well as sighted persons playing the audio-only version of the game, had a mean level completion time roughly three times longer than sighted individuals playing the audio-visual version of the game.

Creating a platform game for visually impaired students required a solution to the difficult problem of representing the 2D spatial map inherent to these games via audio cues alone. Such maps do not translate easily to audio games because of the difficulty of creating good spatial audio cues. In addition to creating effective audio cues to represent the game’s objects, it was also necessary to find methods to convey the nature of the spatial relationships
Besides the benefits to audio gamers, the results from this investigation may also be applicable to non-game applications that need to encode spatial information (distance) into audio, for example multi-modal information systems or virtual environments.

**Thesis Organization**

The study of sonification, audio games, and mental mapping of spatial data motivated the work presented in this thesis. Through a set of three manuscripts (chapters 2-4), this thesis explores the development of the sonification methods used in the game (chapter 2), the design process and evaluation of the interface (chapter 3), and the results of performing a comparative analysis of the spatial mapping data (chapter 4). The result is a look at the design and use of sonification techniques to create a platform game and, more generally, factors that need to be taken into consideration when building auditory interfaces to display spatial data to the visually impaired.

Chapter 2 begins with an overview of the sonification techniques used in the game as well as the overall auditory design. This chapter details an example of an auditory interface for platform games that is based on the sonification techniques that have been developed over the past few decades in auditory display research. In addition, this chapter presents the results of an initial usability study of sighted participants comparing a group that played an audio-visual version with another group that played an audio-only version of the game.

Chapter 3 describes the iterative design process used in the initial phase of the game’s development. This process involved students at the Indiana School for the Blind and other visually impaired users throughout the process. Specifically this focuses the design changes and compromises made in order to ensure that the game was both easily accessible and still retained the primary elements of the platform game genre. Results are provided from the usability study of the game with visually impaired participants.
Chapter 4 explains the need for mental maps to navigate through the game and provide a comparative analysis of the mental maps created by all three participant groups. This chapter explains the methods for collecting the mental mapping data and compares the mapping results from the three groups. It also explores possible differences in the mental mapping strategies between participants with visual impairments and sighted participants.
CHAPTER 1. BACKGROUND AND LITERATURE REVIEW

1. Background

Boarding schools exist throughout the United States to provide education for students with visual impairments (Council for Schools for the Blind). The first such school opened in 1832 when Dr. Samuel Howe began teaching students at what eventually became Perkins Institution for the Blind (Perkins Museum). The passage of the Individuals with Disabilities Act (IDEA) in 1970 called for the proper education of students with disabilities, including the visually impaired (US Department of Education, 2005). Passage of the 2004 reauthorization of IDEA called for increasing technology access for students with visual impairments in section 674 (US Government Printing Office, 2004). In section 602 of IDEA, there is specific mentions recreational activities as related services that should be supplied to students with disabilities. The visually impaired community has long bemoaned the lack of entertainment options for persons with visual impairments. Specifically, in 2001 Charlie Crawford, then executive director of the American Council of the Blind, stated: “Blind people are tired of waiting for access to entertainment and information that others in our society, including those who are deaf or hard of hearing, take for granted” (United Cerebral Palsy, 2001). To help meet this need for adequate entertainment options, discussions about possible solutions were held with students and staff at the Indiana School for the Blind in 2006. The focus group resulted in the decision to develop an audio game in the platform genre.

Although many audio-only games have been developed over the past decade which have brought computer gaming entertainment to visually impaired gamers, there have been no audio-only versions of platform games that retain the basic elements of the genre (see AudioGames.net for a list of currently available audio games). Conceptually, platform games are very simple — inside a 2D, side-scrolling world, a player hops on or over objects,
such as obstacles and platforms, to reach a goal (Wikipedia). To add to the choices available to visually impaired gamers, this research resulted in the development a platform video game that offers an audio interface in addition to the traditional video game interface. A platform audio game was created because several of the visually impaired users who participated in a focus group had fond memories of playing the Super Mario Bros. videogame (see figure 1 in chapter 2), either by using their limited vision or by memorizing the layout of levels through trial and error.

The game was evaluated through a series of usability studies consisting of both participants with visual impairments and sighted participants. The sighted participants were split into two groups; one that played the audio-only version of the game and one that played the traditional (graphical) version. In addition to evaluating the interface of the game, this research also explored the formation of mental maps through auditory cues by comparing the results of participants’ verbal descriptions of their map of levels.

2. Selective Literature Review

The following subsections detail work relevant to the understanding of this research as well as the previous work that this research builds upon.

2.1. Visual Impairments

For the purpose of this research the term “visually impaired” to refer to individuals who are classified as legally blind. Legally blind is defined as a having a maximum visual acuity of 20/200 in the better eye with the best possible correction or a field of vision that is no greater than 20 degrees (American Foundation for the Blind, 2002). As of a 2002 world population estimate on visual impairments, there are 1.386 million totally blind individuals under the age of 15; 5.181 million individuals are between the ages of 15 and 49; while a staggering 30.308 million individuals are 50 and over (Resnikoff, 2004). A further 124.264 million individuals in the world are classified as having low vision (3/60 maximum visual...
acuity or a field of vision below 20 degrees). Cataract and Glaucoma are the two leading causes of visual impairment accounting for 47.7% and 12.3% of worldwide visual impairments respectively while 3.7% of visual impairments are caused by childhood blindness (Resnikoff, 2004).

In terms of computer usage, a minimal estimate of 1.5 million persons with visual impairments in the United States use computers with just under one million visually impaired over the age of 15 reporting that they use a computer regularly (American Foundation for the Blind, 2002). Persons with visual impairments are the second least likely group to have Internet access in their homes within the United States with 78.9% lacking Internet access at home compared to 43.3% of the general population (National Telecommunications and Information Administration, 2002). This indicates a serious digital divide that needs to be rectified by making interfaces more accessible to persons with visual impairments as well as by finding ways to get them interested in using computers at a young age.

2.2. Audio Games

One way of engaging persons with visual impairments with computers at an early age is through computer-based entertainment. There is a notable lack of computer games for the visually impaired created by large mainstream game companies. The majority of accessible games are still created by small independent game companies, and by researchers (International Game Developers Association, 2006). In the early days of computer games, people with visual impairments were on a more or less equal playing field with sighted gamers as there were a plethora of text-based adventures and multi-user dungeons (MUDs) that required no visual perception to play (Röber, 2005). However, over the last several decades a transition from simple text-based experiences to the more and more complex visuals has sidestepped the needs of gamers with visual impairments.
Several papers have been written on the subject of audio games, from the creation of specific games to guidelines for creating audio-only games (McCrindle, 2000; Winberg, 2000; Ossman, 2006; and Gärdenfors, 2002). The game examples tend to be fairly simplistic, such as the Towers of Hanoi, Memory, or Space Invaders and borrow heavily from the field of auditory interfaces in order to map graphical information to auditory cues. More complex games, such as Shades of Doom, have been created but research on the design and creation of these more complex games has not been published (GMA Games). Andresen describes the design process and problems encountered while creating a top-down adventure game where players navigate through mazes of rooms by using point-source audio so that players know which direction the center of a room or hallway is located (Andresen, 2002). However, Andresen never appeared to have run a formal evaluation of the game to determine if there were any problems in distinguishing audio cues or presenting too many simultaneous cues for the user to properly process.

Recognizing the limitations of audio games, Gärdenfors proposed several design frameworks for audio games that seek to maintain a level of musicality that is missing in many audio games; however, the tradeoff for increased musicality tends to be either a game that cannot be played in real-time or a limitation of player controls (Gärdenfors, 2002). Others, such as Ossman, provide guidelines for designing accessibility within games, whether they are intended for a specialized audience or an audience with a specific disability, and these guidelines focus primarily on which options to include in the game (e.g. ability to turn music on/off, ability to adjust volume of individual objects, etc.) as well as guidelines for level progression that includes tutorial options and difficulty progression/options (Ossman, 2006).

While these research projects looked at specific games and design guidelines, they seemed to ignore one of the most crucial aspects of any game: how fun it is to play. Without this aspect of fun, or entertainment, users of these audio games will quickly grow bored and
the laudable goal of bringing computer entertainment options to persons with visual impairments and thus increasing active use of computers ends up falling short. Targett noticed this lack of studying the “fun factor” and conducted a study of two simple audio games (Tic-Tac-Toe and Mastermind) with a goal of determining how fun participants of a usability study found these games (Targett, 2003). Despite the simplicity of these games, participants found these games to be “quite fun”, which was right under the maximum score of “extremely fun” on the Likert scale used in their study. It should be noted, however, that during the study participants only had a chance to play each of these games one to three times and no study seems to have been conducted on the long term entertainment levels of auditory games. In addition to examining the “fun factor”, Targett also examined whether or not participants felt they were able to acquire skills or knowledge through playing the games and all participants indicated that they felt the game would be useful in acquiring skills (although the skill acquired through playing the games was not unanimous among participants).

Targett’s research also demonstrates that games are not simply a means of entertainment or a gateway to computer usage, but also a means of teaching skills and concepts. While this paper does not explore the possibilities of audio games for the use in education, the International Game Developer’s Association (IGDA) whitepaper on game accessibility lists the need to make educational options equal to all students as one reason for creating accessible games (IGDA, 2006). This is important as such games may become increasingly integrated into the education system as means of inducing student interaction and conveying both simple and complex concepts (Brown, 2008).
2.3. Spatial Mapping

As the game created for this research relies heavily on the user being able to navigate and remember the location of a series of objects in the virtual world, it is necessary to explore the concepts of spatial mapping and mental maps.

Epstein’s 2005 study determined that people with better navigational abilities were better able to tell when a scene they were viewing was from a different angle or in an entirely different location (Epstein, 2005). In the past, games have been created to convey spatial knowledge and skills, such as Volbracht’s “City Game” that was used to teach spatial orientation to players in a 3D world (Volbracht, 1998). Spatial orientation and understanding are critical skills for people both in daily life and in their understanding of geography, geometry, and many other spatially based subjects. While many of these studies focus primarily on understanding a 3D location or orientation within a building or city, spatial comprehension of 2D (top-down) maps may be a crucial first step in learning how to navigate 3D virtual environments.

While the spatial tasks listed above dealt with visual perception, studies have also been conducted to determine the spatial abilities of persons with visual impairments (primarily the early blind) using auditory cues. Després found that early blind individuals (individuals who lost their vision at a young age) actually have an enhanced self-localization using auditory cues to properly place the location of the sound on a 2D map of the room as compared to the performance of fully sighted individuals attempting to complete the same task (Després, 2005). Furthermore, while it is often assumed that persons with visual impairments have an inferior ability to complete spatial tasks, Eardley determined that while persons with visual impairments perform slightly worse on 2D tasks compared to fully sighted participants, there is no performance difference in the completion of 3D tasks (Eardley, 2007). Eardley’s study concluded that the most likely reason for the poor performance of persons with visual impairments during the two-dimensional tasks had to do
with lack of experience in two-dimensional environments. Thus, it may be beneficial for persons with visual impairments to play an audio game set in a 2D environment in order to gain experience working with 2D spatial data.

The act of understanding the abilities of persons with visual impairments is important; however, equally important is understanding ways of improving spatial memory and recognition for those with limited spatial knowledge (such as the early blind). Brunyé conducted a series of experiments to examine this issue and determined that reducing the spatial details through graphic generalization in verbally presented information improves spatial memory and results in more accurate recall of spatial relations (Brunyé, 2007). No studies have been conducted to examine how this principle may apply to spatial data conveyed aurally through non-verbal cues, but the concepts of simplification of the spatial detail and keeping the idea of human working memory constraints in mind most likely also applies to the design of the auditory interfaces.
CHAPTER 2. SPEED SONIC ACROSS THE SPAN: A PLATFORM AUDIO GAME

Modified from the proceedings of the 13th International Conference on Auditory Displays (ICAD), July 2007, Montreal, Canada.
Michael Oren¹, Chris Harding², Terri Bonebright³

Abstract

We describe the sound design and initial user study of an audio game created for gamers with visual impairments. Despite the wild popularity of platform games such as Super Mario (Wikipedia) and the development of many audio games over the past decade, the platform genre has so far been all but ignored by audio game designers. To fill this gap and to add to the limited entertainment choices visually impaired gamers have, we developed a platform game that can be played via an audio-only interface. We conducted a study to test our game and audio design choices, to measure the users’ performance and to find out if the game was fun to play. This usability study used 9 participants who played the game as a traditional video game (audio-visual input) and 9 participants who played the game as a pure audio game (audio-only input). The results show that, although it took the audio-only group considerably longer to play through the game, they did not make significantly more mistakes and they seemed to find the audio-only version challenging, yet enjoyable.

1. Introduction

1.1. Project and Motivation

Boarding schools exist throughout the United States to provide education for students with visual impairments. Teachers and administrators of these schools have bemoaned the fact that visually impaired students have few options to entertain themselves compared to fully sighted students. We initially discussed the lack of entertainment options with students
and staff at the Indiana School for the Blind, and after a couple of focus group meetings with students we decided to make an audio game in the platform genre.

Conceptually, platform games are very simple — inside a 2D, side scrolling world, a player hops on or over objects, such as obstacles and platforms, to reach a goal (Wikipedia). We chose to create a platform audio game because several of the visually impaired users we spoke with had fond memories of playing the Super Mario Brothers video game, either by using their limited vision or by memorizing the layout of levels through trial and error (Wikipedia).

Our primary goal was to create a platform game for visually impaired students to enhance their entertainment options and quality of life. This required us to investigate the difficult problem of representing the 2D spatial map inherent to platform games via audio cues alone. Such maps do not translate easily to audio games because of the difficulty of creating good spatial audio cues. Not only did we have to create audio cues to represent the game’s objects, but we also had to find methods to convey the spatial relationships (the height, the distance, and the length) of objects – which are critical relationships for the gameplay.

Besides the benefits to audio gamers, the results from this investigation may also be applied for some non-game applications, such as encoding spatial information (distance) into audio as part of a multi-modal information system or for use in virtual environments.

1.2. Background: Audio Games

Audio games have been around for over a decade. Many “classic” games have been converted to audio games, such as Space Invaders and Doom (McCrindle, 2000; GMA Games). In addition, guidelines have been written regarding the design of simple audio games (Gärdenförs, 2002).
The key issue for sound-based game design is to effectively represent game objects via audio cues. For example, objects can be represented by the kind of sound they would naturally make (called auditory icons), or be represented with an abstract sound (such as a simple piano chord, called earcons) (Gärdenfors 2002; Gaver, 1988; Blattner, 1989). Using the “natural connections” of auditory icons as opposed to more abstract auditory mappings of earcons thought to make it easier to recognize and remember the sounds since no new semantic links need to be learned (Brewster, 1993). However, for some types of game objects, such as pits (bottomless holes) and platforms (floors that can be walked on) there is no good natural audio cue, For these objects the only other option is to use abstract audio cues, which need to be taught to the user and need to be designed so that the user is able to remember them easily (Gärdenfors, 2002).

Currently available audio games run the gamut from simple memory games to first person shooters such as Shades of Doom (GMA Games; Gärdenfors, 2002). While games such as Shades of Doom and several others are fairly complex and require an understanding of spatial relationships, the majority of audio games are fairly simple maze or puzzle games. More complex games require that a more realistic spatial environment can be presented via audio; however the concept of distance (or spatiality) ultimately requires a 2D or 3D rendering of the sound and hence binaural cues. Without these binaural cues, relationships are difficult to determine (Gärdenfors, 2002).

Typical platform games require that multiple game objects and their positions are presented to the user while simultaneously allowing him or her to react to these objects in an instant, such as jumping over a moving enemy to reach a platform at a higher floor. This spatial context makes the design of audio cues particularly challenging. The only example of an audio platform game that is currently available is Super Liam (AudioGames.net). Super Liam uses simplified platform game rules by only having one platform without any higher
platforms or pits that the user would have to jump to or jump over; therefore it can be played without requiring the user to learn complicated spatial audio cues (AudioGames.net).

1.3. Sound Design

Our game consists of four primary objects: the player (a representation of the user on the screen), platforms, enemies, and pits [Figure 1]. Each type of object has its own set of sounds and set of rules that it follows to convey its information to the user.

![Figure 1](image1.png)  
Figure 1 A screenshot of the game with the primary object labeled. The player is labeled “You”.

Table 1 lists the game objects and the methods used to convey their locations. The game has the player move through several game levels. The objective of each level in the game is to move to the right through the 2D game world until a certain location is reached - the player then starts at the next game level and moves again to the right. The player is always at a central position in the game. The relative positions of all other objects are rendered (via stereo panning) with regard to the player.

In our game the player can stand still, walk or jump. We indicate the player’s walk via footprint sounds. There are four floors at fixed elevations (called 1 to 4) that the player can walk on that correspond to the vertical placement of platforms [Figure 2]. When the
player jumps, a pitched jump sound is played, which symbolically “draws” the curve of the jump via its sound. Jumps will always have the same maximum height - just enough to jump up one floor. Upon landing again on the ground, the user hears a pitched landing sound (drum) plus a vocal cue: a voice reading the current floor number (“one”, “two”, etc.). Using this vocal cue is meant to reduce the cognitive load on the user, who already has to focus on the spatial cues and react to impending dangers. The vocal cues are monosyllabic and are as short as a typical auditory icon or an earcon.

Platforms come in two sub-types: floating and solid platforms. The player can walk underneath floating platforms, but will run into a “wall” when approaching a solid platform forcing him/her to jump on top of it to continue. Jumping while under a floating platform will result in an “ow” sound when hitting the “ceiling”; running into the wall of a solid platform will play a “d’oh” sound [Figure 3].

<table>
<thead>
<tr>
<th></th>
<th>Voice (numbers)</th>
<th>Voice (icon)</th>
<th>Auditory Icon (Pitched, Panned)</th>
<th>Auditory Icon (Pitched)</th>
<th>Earcon (Pitched, Panned)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player Jump</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Player Land</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Player Walk</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floating Platform</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solid Platform</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End Of Platform Warning</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Pit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bees</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dogs</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Primary objects in the game world and their auditory representations
We chose some of the sounds based on popular cultural cues (the jump sound from Super Mario Brothers, Homer’s “d’oh” for running into an object, etc) to build on these existing semantic links for our target audience.

As there is no good semantic link for an approaching platform, we decided to design the sound as an earcon. The critical information for a player is whether an approaching platform could be jumped onto from the current position. We decided to use a deliberately simplified game world where a higher platform is always exactly one floor higher than the current position. However, we decided that it would be possible for the player to descend one floor.
or more floors down and safely land on a lower platform. In both cases, the new floor’s number is conveyed via the vocal cue of the landing sound. This narrows down the needed information for any platform the player approaches to a) higher, b) lower and c) same floor (i.e. a simple gap).

The player uses a built-in directional sonar device to get information about the position of a platform ahead (How far away? Up or Down?). In order to “hear” the sound that is bounced off the platform, the player must be facing the platform. This sound contains information about the platform’s distance and relative position.

To answer the “Up or down?” question we used a glissando modification of the platform sound: glissando up for a higher platform and glissando down for a lower platform. No glissando is used for a platform of equal elevation. To answer the “How far away?” question, the sound is panned left and right accordingly, with the sound in the center when the player is very close to the platform. Traveling any further at this point would result in being unable to reach the platform by jumping onto it as this would result in the player hitting the bottom of the platform [Figure 3]. In addition to the panning, the speed of platform sound’s repetition and its loudness help the user decide when to jump – they both increase as the player approaches the platform’s edge. A similar approach has been successful in auditory graph design (Flowers, 2005).

In our initial pilot study we found that it was difficult to know when the end of a platform was close. We therefore added a platform-end warning sound. We use a loud ding tone as our earcon, similar to the warning sound on Windows computers, to warn that the end of the current platform is approaching and that the player will either have to jump to another platform or fall to a lower platform (Blattner, 1989). Since the ding sound is already used by Windows to call for the user’s attention, we hope to again leverage an already existing semantic link for this sound (although this will only work for Windows users).
In order to make the game more enjoyable we introduced two elements designed to add a “danger factor” to the game. As typical for the platform genre, we decided to add obstacles: a pit and two “enemies”.

The pit is always placed on floor 1 and forces the player to jump over it or fall into it and start the game level over. When the player approaches a pit, a panned organ sound is played. If the player is on floor 2 or higher when approaching the pit (on floor 1) the pit warning replaces the end of platform warning. An organ sound was chosen because its “spooky” tone might leverage a semantic link between a bottomless (dangerous) pit and the “spooky” organ sound often associated with Halloween and funerals.

Enemy objects need to be avoided as coming into contact with them will force the user to restart the game level. The enemy objects’ sounds can be heard no matter which direction the player is facing. These sounds are again panned according to the location of the enemy - left when the enemy is left of the player and right when the enemy is on the right of the player. “Enemy” sounds also get louder as the player gets closer to them.

We use two types of enemies: bees and dogs. Bees only travel vertically—moving between the top of the game world and a platform. The bee makes a buzzing sound in order to enforce the semantic link that serves as an auditory icon (Battner, 1989). The bee’s sound is pitch shifted to indicate its relative elevation – listening to the bee’s pitch rising and falling will allow the user to move the player under the bee (when it is at it is highest point) or jump over it (when it is at it’s lowest point).

The dog enemy object only travels horizontally. Since determining the proper horizontal position of the dog is critical for the user in order to time the jump over it, the dog’s distance is tied to the repetition speed of its bark. When the player gets within a certain horizontal (very close) range of the dog, the bark changes to a growl, telling the player to jump immediately or be bitten. The bark and growl tones were chosen because of their semantic links to dogs.
2. **User Study**

In order to test our application, we conducted a user study consisting of 18 undergraduate and graduate students (8 males and 10 females). We worked with sighted participants but blindfolded half of them. The participants were randomly assigned to play either the audio-only version (blindfolded) or a traditional video-game version (audio-visual).

### 2.1. User Study Design

With a researcher present, participants first read the instructions from a Microsoft Word document with embedded sound files. The instructions were divided into several sections, each of which introduced a new game element. Guided by the instructions, the participants played through ten increasingly complex training levels to familiarize themselves with the game mechanics and the audio cues. Participants in both groups played with a game controller and with headphones, however only the audio-visual group was able to see the game on a monitor. The participants read through the instructions at their own pace, played the sample sounds as often as they desired, and at the end of each section played through a training level that dealt with each newly introduced game element in isolation.

The participants then played through the seven levels of the actual game. The participants’ game performance was scored based on the speed of level completion (faster being better) and the number of times they were sent back to the beginning of the level (more restarts being worse). After completing the main game, the participants filled out an exit survey in which they rated their ability to recognize the various objects, spatial relationships, etc. and generally commented on the game experience.

### 2.2. Results

We performed an ANOVA between the two groups using the answers from the questionnaires and the game performance data. In terms of relevant background experience
that might affect their performance the analysis results showed no significant differences between the audio-only and the audio-visual participants or between males and females. Most participants played computer (video) games on occasion. All participants used computers very frequently, while only very few of the participants (17%) used programs where auditory signals provided the primary means of obtaining information (such as screen readers). Most participants had some familiarity with platform games.

We asked the participants about audio in computer games and they reported the following: 5% simply ignore the sounds; 61% use sound as an enhancement of the realism of the game, but they do not actually use it in a functional (guiding) way; 28% use sound to warn them of critical game events; such as an incoming attack; and 39% use sound to obtain spatial information (e.g. about enemies and objects behind them).

Most participants in both groups found the sounds for dogs, the bees and the pit very easy to identify, while the platform sounds were more difficult to recognize. This may have to do with the more abstract sounds assigned to platforms.

The majority of participants in the audio-only group indicated that they were able to determine both the vertical and horizontal location of objects via the sound (via pitch shift and via panning); however a few participants indicated problems with recognizing the vertical axis (pitch). Conversely, those in the audio-visual group found on average that the sound was beneficial for locating objects along the horizontal plane but not for locating objects along the vertical axis. This might indicate that the sound cues failed to provide any added benefit to what the players could already see on the screen.

Most of audio-only participants, and some audio-visual participants, also used audio cues in determining the distance, direction, height, and type of objects. However, it was also mentioned that audio cues did not help to determine the speed or length of objects, which is to be expected, given that we did not provide any dedicated cues for speed or length.
With regard to game performance, the number of restarts did not vary significantly between groups. The audio-visual participants needed a mean of 0.38 restarts (SD = 0.17); audio-only participants needed a mean of 0.76 restarts (SD = 0.18). It took the audio-only group, on average, nearly three times as long to complete levels as the audio-visual group: participants in the audio-visual group needed a mean time of 36.54 seconds (SD = 5.05), while participants in the audio-only group needed a mean level completion time of 102.34 seconds (SD = 5.73). Whether this is due to their inexperience with audio interfaces or perhaps problems with using the auditory cues provided is unclear at this point.

3. Conclusion

This project allowed us to take what is graphically a very simple video game and translate it into a challenging audio game that is accessible to people with visual impairments and blind users. Despite the slightly modified platform game rules and, what to the participants in the blindfolded group must have been an unusual gaming experience, the majority of participants found the game enjoyable. Several participants from each of the groups asked to try out the alternative version of the game after completing the exit survey – many of the audio-visual users found the audio-only version to be more challenging and more enjoyable.

This not only shows that an audio game requiring complex 2D spatial relationships can be built and played with a relatively good success rate, but it also shows that audio games can also be enjoyed by sighted gamers. The audio techniques we used to create this 2D game may also be useful as a teaching tool for helping visually impaired students understand 2D maps and floor plans. It may be possible to apply some of our techniques to existing audio map projects that help the visually impaired with using spatial orientation (such as the BATS project) (Parente, 2003).
We are currently repeating this study with visually impaired and blind participants to see how their performance of the audio-only version compares to that of sighted participants (our audio-only group). Besides some accommodations, such as assisting the participants with reading the instructions, playing the example sound files and filling out the questionnaires, the procedure is the same as for the audio-only group. As the participants with visual impairments are likely to have day-to-day experience with auditory interfaces, a comparison of their results with the results from our current study may help us determine if the difference between the speed of completion between our audio-only group and our audio-visual group may be the result of lack of experience with auditory interfaces on the part of the audio-only group. We will also allow the visually-impaired participants to play the game for several weeks and solicit their feedback after this time. This will determine if our game is indeed an enrichment of the entertainment options for visually impaired students.
CHAPTER 3. AUDIO PLATFORM GAME (APG) DESIGN FOR PLAYERS WITH VISUAL IMPAIRMENTS

Michael Oren, Chris Harding, Terri Bonebright

Abstract

A novel audio platform game (APG) that creates a spatial, interactive experience via audio cues was evaluated. A pilot user study with players with visually impairments and an experiment comparing the visual and audio game versions using both players with normal vision and with visual impairments revealed that all participants played APG successfully.

1. Introduction

Aided by advances in computer graphics technology the video game industry has experienced strong growth in recent years, earning $17.9 billion dollars in 2007 alone (Boyer, 2008). Similar improvements in audio technology have encouraged the development of a growing number of dedicated audio-only games (AudioGames.net), which use text, button presses, joysticks or speech as inputs and provide recorded or synthesized sound and/or speech output to “display” the game content. However, despite efforts by the United Nation (1994), to encourage universal access to recreation and leisure activities for persons who are blind or visually impaired, most of the games developed for this population are still fairly simplistic; blind and visually impaired persons would welcome a wider variety and higher complexity of entertainment options (van Tol, 2006). In addition to their potential entertainment value, games may also provide a framework to better motivate student learning, such as the historically based computer games created by Muzzy Lane, and these games must remain accessible or risk leaving persons with disabilities behind (International Game Developers Association, 2004).
Originally developed in the 1980’s, the platform game genre has considerable potential for a conversion to an audio-only format. Platform games are set in a 2D, side-scrolling world with a user-controlled character that hops on or over objects, such as obstacles and platforms, to reach a goal. The player needs to use strategic thinking to traverse multi-floor platforms to reach a goal and quick reactions to complete such tasks as jumping over an approaching opponent. While state-of-the-art platform games employ 3D graphics that often bars persons with visual impairments from enjoying them, many players with visual impairments are able to enjoy playing 2D platform video games, such as Super Mario Bros., despite being forced to adapt to the game by using their limited vision or by memorizing the layout of levels through trial and error.

Creating an audio-centric platform game requires translating the inherently 2D spatial game world into audio cues that effectively represent the game’s objects (the characters, opponents and obstacles) and their spatial relationships. An audio-only platform game (Super Liam) already exists; however it significantly simplifies many of the elements typical of this game genre that makes them challenging and interesting to play. For the design of a more challenging audio-only platform game, we turned to guidelines for the creation of audio games proposed by Gärdenfors (2002), and to the field of auditory displays and interfaces.

One type of auditory display, sonification, is based on research in auditory perception (Bregman, 1990), and a general theory for this technique has emerged over the last decade (Stevens, 1996; Barrass, 2003; Walker, 2005). In sonification, non-speech audio is used to translate or map information into sound (Kramer, 1999). For example, a series of numbers can be mapped onto a series of pitches for auditory graphs with the upward or downward shifting of a sound source’s pitch mapped as a rise or fall of the source’s elevation (Flowers, 2005). Another example of sonification is auditory icons (Gaver, 1988), which use pre-recorded sound “caricatures” of real-world sounds that link a concept to the sound it produces; for example, a barking sound communicates the general concept of a dog.
Auditory icons tend to be learned easily; however, such links to natural sounds do not exist for all concepts, and therefore abstract concepts are especially difficult to represent using this method. However, such concepts can be successfully represented using earcons (Blattner, 1989), another sonification technique, which employ artificial sounds (e.g. a short melody of piano tones) to encode the abstract information (Brewster, 1993). For example, changing the repeat rate of an audio cue can encode changes in distance to a source object. Finally, a simple technique using 2-D spatial audio effect known as panning, a subtle left to right speaker change (Blauert, 1997) can be used convey the left-right movement of a sound source.

2. Pilot User Study for Game Prototype

APG can be played as a traditional, albeit audio-enhanced video game (audio-visual interface), or via an audio-only interface. Using the aforementioned auditory display methods, the audio interface was designed to convey the spatial relationships of the game objects and their changes over time to the player via audio cues. Using an iterative, participatory design process resulted in several critical changes to the auditory interface. A pilot study was conducted to evaluate the design of APG’s audio cues for players with visual impairments.

2.1. Method

2.1.1. Participants

Six students with visual impairments from the Indiana School for the Blind participated in the design process and the pilot study. Three participants were blind (2 males, 1 female, one was born blind, two became blind before the age of five) their average age was 17 (SD=0.58). Three participants (all females) had severe visual impairments with an average
age of 16 (SD=0.58). The pilot was conducted with approval from the Institutional Review Board at DePauw University.

### 2.1.2. Hardware and Software

Although more modern sound cards and other hardware exist on the market, we chose to test this application on a laptop with a Pentium IV 2.0 GHz processor, 512 MB of RAM, and an Intel AC'97 onboard sound card. We purposely chose to use this older machine with an onboard sound card as our work with boarding schools for the visually impaired indicated that they very rarely purchase separate sound cards for their machines and often use computers that are three or four years old. Participants played the game using standard, over-the-ear Sony MDR-V150 headphones, which are readily available in the consumer market. A Logitech Dual Action game pad was used as an input mechanism for the game.

We used the Python programming language to write the game and used the pyGame module to handle rendering of the visual components as well as the detection of object collisions (i.e., to determine if a player had defeated a dog or bee, landed on top of a platform, etc.) We also used the pySonic module, a Python wraparound for the FMOD library, in order to play and manipulate the sounds we used to aurally depict the game objects.

### 2.1.3. Game and Sound Design

There are four primary game objects: the character (the player’s on-screen representation), platforms, opponents, and pits (Figure 4). Each type of object has its own set of audio cues to convey information, as shown in Table 2. The character can walk left and right producing footstep sounds and can jump up and down, which produces a jumping sound. As the character moves horizontally, the game world adjusts accordingly.
<table>
<thead>
<tr>
<th></th>
<th>Voice</th>
<th>Auditory Icon (pitch, pan, repeat)</th>
<th>Auditory Icon (pitch)</th>
<th>Earcon (pitch, pan, repeat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player Jump</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Player Land</td>
<td>Numbers</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Player Walk</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floating Platform</td>
<td>Sound byte</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Solid Platform</td>
<td>Sound byte</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>End Of Platform Warning</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Pit</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Bees</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dogs</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are also four imaginary “floors” at fixed elevations (called 1 to 4) with possible platforms for the player to walk on (Figure 5). The character is only allowed to jump up one floor at a time, and when the character lands a “landing” sound (drum) is produced. Platforms can either be floating or solid. The character can walk underneath the floating
platforms but will run into a “wall” when approaching a solid platform that forces the character to jump on top of it to continue. Jumping while under a floating platform will result in an “ow” sound when hitting the “ceiling”; running into the wall of a solid platform will produce a “d’oh” sound (Figure 6). Some areas of the game world present fairly complex combinations of floating and solid platforms that force the character to backtrack to find a way on top of the platform, this process requires a good sense of the spatial layout of the level.

![Figure 6 A cartoon showing a collision with floating platform (top) and with a solid platform (bottom).](image)

It is quite common that the character reaches the end of the current and needs to jump over a gap onto a new platform. It is important to convey the relationship of this new platform to the character’s position before the jump. The game uses a “sonar sense” that points in the direction the character faces. The sonar uses a repeating sequence of abstract sounds to indicate the spatial relationship of the player’s elevation to the elevation of the next platform in range. A glissando-up sequence indicates that a higher platform ahead, a
glissando-down sequence indicates a lower platform ahead, and a same-pitch sequence indicates that the next platform is at the same floor. The left or right position of the platform is encoded via panning with a middle position indicating that the player is close enough to the next platform to reach it with a jump. Similarly, the repetition rate of the platform sound increases as the player approaches the best “jump-off” point.

To make the game more challenging and enjoyable, some obstacles typical for platform games were added: bottomless pits and opponents. The pit is always placed on floor 1 and forces the character to jump over it or fall into it, which causes the game level to start over. When the character approaches a pit, a panned organ sound is played. Opponents come in two types: bees and dogs, and if the character comes into contact with them results in a restart. The opponents’ sound cues can be heard beyond the area that is shown on the screen from both the left and right directions, and they increase in intensity as the opponent approaches the character. Bees only move vertically between the top of the game world and a platform; they emit a buzzing sound that is pitch-shifted to indicate their relative elevation. Listening to the bee’s pitch rising and falling allows the player to run under the bee when the bee is at its highest point or jump over it when it is at its lowest point. Dogs only travel horizontally and emit a barking sound that is tied to the distance to the player so that the dog barks faster as it gets closer.

2.1.4. Procedure

During the pilot the six participants played through different iterations of the prototype of the game. Each of the game sessions lasted for 30 minutes and occurred every three weeks. All participants wore headphones and used a game pad. Participants were asked to think aloud while playing the game and to report any difficulties they encountered to the researchers. After each session a focus group style discussion with all six players was held to capture their feedback and suggestions for improvements for the design of the next iteration.
2.1.5. Results & Discussion

Only one participant was able to actually complete the first of two game levels, three participants struggled with navigating the first level and needed help, and one participant gave up completing the first level and gave up in trying to complete the second level. None of the participants were able to complete the second, more complex, game level.

The fact that the participants with some intact vision did much better using the audio-visual interface, pointed to a serious flaw in this implementation of the auditory interface, especially with regard to conveying information on the distance and the position of game objects quickly and accurately.

Initially, the player heard sound cues for all platforms as they were positioned around the character, which means that there could be multiple sounds coming simultaneously both from the left and right of the character. After experimenting with vocal cues to display the location and type of objects, we decided to use non-vocal cues and panning as the participants commented that it seemed more natural and immersive in the game world. Initially sounds for jumping, landing and footsteps were pitched according to the character’s current elevation, so that the pitch would become higher or lower as the character moved up and down. Participants indicated that these methods for locating the character resulted in auditory overload and made it very difficult to establish and maintain a situational awareness of the spatial layout. Thus, we decided to focus on conveying only those game objects that are essential to the player. Therefore, a sonar scheme was employed where the player “hears” platforms only from the current direction of movement (i.e. only from right or from left) that are close enough to impact the character. Floor numbers (1 to 4) were also introduced and all platforms were arranged to fit into this scheme, which allows the player to determine the elevation via a spoken number rather than forcing the player to interpret pitched sounds for the floors. This helped to clarify the level’s layout without significantly reducing the complexity of the game.
While observing the participants, we found that a combination of poor level layouts and insufficient audio cues led the character to jump off a platform too soon and, as a result, to miss subsequent platforms. Thus a platform-end warning was added as a signal to players to indicate when they need to make a decision to jump. Similarly, the design of the levels was changed so that game objects with the same sounds, such as two dogs of equal distance that are in opposite directions from the character, are better separated and less likely to be confused as a single sound source. Finally secondary tones were added to the enemies, such as a dog’s growl, in order to provide proper feedback to the gamers so that they could make critical decisions about movements to avoid the opponents.

In summary, the study of the initial game design provided strong evidence that simply overlaying sounds on top of an existing visual platform game model would not work when creating an audio-only version of the game. Instead, the game needed to be designed by focusing on a good audio solution for the most important game elements. This helped to reduce auditory clutter and generally improved the “playability” of the audio-only game while still keeping the primary elements and essence of the platform game intact.

3. User Experiment for Final Game

A formal usability study was conducted to evaluate the new version of the game. The researchers studied three scenarios: sighted participants playing the audio-visual version; sighted participants playing the audio-only version; and participants with visual impairments playing the audio-only version.

3.1. Method

3.1.1. Participants

Eighteen people (8 men and 10 women) who had normal vision were recruited from Iowa State University; their average age was 28 (SD = 6.9). Participants were pseudo-
randomly divided into an audio-only and audio-visual group with equal numbers of men and women in each. In addition, 9 people (7 men and 2 women) who were legally blind (1 blind after birth, 8 with severe visual impairments) were recruited from Iowa State University and Iowa Braille and Sight Saving School. These participants had an average age of 26 (SD = 11.2). The study was reviewed and approved by the Iowa State’s Institutional Review Board and the research was conducted within the guidelines of ethical treatment of participants.

3.1.2. Hardware and Software

The same hardware and software were used as in the pilot study.

3.2. Procedure

Participants were trained to play the game by reading through (or listening to) a series of instructions from a Microsoft Word document with embedded sound files at their own pace. During this training, they played the sample sounds as often as they desired, and at the end of each section of instructions, they played through a training level that dealt with each newly introduced element in isolation. These training levels moved from the very simple, such as moving around on a completely flat level, to the more complex elements, such as avoiding of pits, dogs, and bees. All participants played with a game controller and with headphones; those playing the audio-visual version were able to see the game on a monitor. Participants with normal vision assigned to the audio-only version of the game were blindfolded during training and while they played the game. Participants were free to pause the game at any time to take a break, ask a question, or adjust the headphone volume.

The participants then played through seven levels of the game. The participants’ game performance was based a) on the time it took to complete each level (shorter being better) and b) the number of times they were sent back to the beginning of the level (more restarts being worse). Participants were informed of this scoring mechanism but were not shown their time or number of restarts. When participants became “stuck” at a spatially
tricky part, they were encouraged to pause the game, to “think aloud” about their current position and to articulate their assumptions and problems. Although no direct help was given to participants, they were always able to continue through the game. During pauses, the game was halted so that this did not affect the time recorded for playing the game. Finally, participants completed a follow-up survey that included two general categories of items. First they were asked to assess how frequently they used computers and played computers games, how often they played platform games, and how familiar they were with screen readers using a 5-point rating scale from 1, never or almost never, to 5, very frequently. Second they provided feedback on their experience playing the game by commenting on how well they could recognize the various game objects and how well they could detect the spatial relationships in the game.

4. Results

There were two independent variables of interest: game version, and gender. Game version had three levels: audio-visual played by players with normal vision (n = 9) and audio-only played by players with visual impairments (n = 9) and those with normal vision (n = 9). We were also interested in whether there were any differences between male and female gamers, regardless of their type of vision. The dependent measures were the total number of restarts for the game, and the average amount of time to complete all levels in seconds. All tests were performed using ANOVA with .05 as the significance level.

4.1. Game Performance

The mean number of times participants restarted did not vary significantly between those with normal vision who played the audio-only game (M = 5.22, SD = 4.58), the audio-visual game (M = 2.67, SD = 1.73) and those with visual impairments playing the audio-only game (M = 6.44, SD = 6.58), F(1,24) = 1.49, p = .25). However, there were differences in the average amount of time taken to complete all levels, such that the participants with both
normal vision (M = 119.74, SD = 73.13) and visual impairments (M = 146.51, SD = 80.72) who played the audio-only version took at least three times longer than the participants with normal vision who played the audio-visual version (M = 36.54, SD = 8.75), F(1,24) = 7.44, p = .003, η²p = .26.

Men and women showed no differences for either average time taken to complete the levels (men: M = 105.81, SD = 79.40; women: M = 95.67, SD = 77.59), F(1,25) = .11, p = .74, nor for the total number of restarts for the game, (men: M = 5.50; SD = 5.27; women: M = 4.00; SD = 4.04), F(1, 25) = .64, p = .43, regardless of their visual abilities.

4.1.1. Follow-up Survey

Survey data on questions regarding computer and computer game usage that were rated on a 5-point scale, from 1, never or almost never, to 5, very frequently, revealed that participants with normal vision indicated that they very frequently used a computer and that they played computer games infrequently. These participants also reported that they had infrequently played platform games and that none of them had experience with screen readers, although they had heard of them. In contrast participants with visual impairments reported significantly less frequent use of computers than those with normal vision, F(1,22) = 12.50, p = .002, and a significantly higher level of familiarity with screen readers and other audio representations for visual objects, F(1,23) = 39.89, p < .001. There were no differences between the two groups of participants on frequency of playing computer games in general or platform games in particular (see Table 2 for means and standard deviations).
Table 3 Means and Standard Deviations for Survey Response for Participants with Normal Vision and with Visual Impairments

<table>
<thead>
<tr>
<th>Question</th>
<th>Normal Vision</th>
<th>Visually Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often do you play computer games?</td>
<td>2.81 (1.17)</td>
<td>2.88 (.35)</td>
</tr>
<tr>
<td>How often do you use a computer?</td>
<td>*4.88 (.34)</td>
<td>*3.50 (1.51)</td>
</tr>
<tr>
<td>How often do you have played platform games?</td>
<td>2.53 (.87)</td>
<td>2.25 (1.39)</td>
</tr>
<tr>
<td>How much experience have you had with a screen reader or other representations of visual objects through sounds or voice?</td>
<td>*1.82 (1.07)</td>
<td>*4.50 (.76)</td>
</tr>
</tbody>
</table>

Rating scale used was from 1, never or almost never, to 5, very frequently. Standard deviations are in parentheses. * indicates a significant difference.

Participants also reported on how they typically use audio in previous computer games they had played. Regardless of their visual abilities, 11% reported that they simply ignore the sounds; 50% said that they use sound as an enhancement of the realism of the game, but they do not actually use it in a functional way; 8% revealed that they used sound to warn them of critical game events; such as an incoming attack; and 32% reported that they use sound to obtain spatial information (e.g. about opponents and objects behind them when playing 3D games). When asked about their preferred method of audio feedback for audio interfaces in general, participants with normal vision reported no prior experience with auditory interfaces of any kind, while all participants with visual impairments reported that they used auditory interfaces that provide voices and/or tones.

In the open-ended comments, several participants with normal vision and with visual impairments in the audio-only group indicated a desire for more levels in the game so they could continue playing it. In addition, the most common request for an improvement from participants with visual impairments was for more and different kinds of objects (such as moving platforms). The majority of the comments on the follow-up survey indicated that the
audio game provided an enjoyable game experience. For example, one participant with normal vision stated, “I was surprised at how easy it was to play; I had expected it to be harder for me to learn, but the sound cues worked really well for the most part.”

5. Discussion

This project allowed us to take a visually simple video game and translate it into a challenging audio game that is accessible to gamers with visual impairments. It also shows that an audio game requiring complex 2D spatial relationships can be built and played with a relatively good success rate by both, visually-impaired and sighted gamers. Despite the complexity of the audio-only game and the unusual gaming experience, the majority of participants found the game enjoyable. One important lesson in audio platform design was that limiting the “view” of the player to relative cues rather than recreating the screen via audio cues reduces the player’s frustration and produces a more enjoyable game experience.

Whether the longer completion time for audio-only gamers is due to their inexperience with audio interfaces or perhaps problems with using the auditory cues provided is unclear at this point. However, some participants with visual impairments played the game with impressive competence. For example, to complete level six (the most complex of the total of seven levels played during the test), the three fastest of the nine participants with visual impairments needed roughly 10 seconds less than the average gamer of the audio-visual group.

We hope that this project will prompt others to create similar “spatial” audio games that can be enjoyed by users with visual impairments and sighted users equally. The researchers have made this project’s source code publicly available (http://sourceforge.net/projects/moosic) so others can modify it. The audio encoding techniques used for this game may be more generally applicable for other, non-gaming
applications in which spatial information (distance) needs to be translated into audio to persons with visual impairments. (e.g. see the BATS project; Parente, 2003).

It is relatively simple to edit existing levels and to create new levels by simply editing a text file that defines the location, type, and sound of the game objects (opponents, platforms, pits, etc.). We hope that this will encourage others, especially persons with visual impairments, to create their own levels of the game, and to exchange them with others. With a basic understanding of the Python programming language, the game’s code can be modified to create educational types of games. Elements, such as the opponents, could be replaced with sounds for numbers, letters of the alphabet, words or concepts (voice recordings or auditory icons), so that rather than simply traversing the level, players could have the task of hopping over a subset of game elements, possibly in a certain order. For example, to complete a level they may need to hop over all verbs in a sentence, or over the names of state capitals and avoid names of other cities, or over certain numbers that add together into a certain sum.
CHAPTER 4 MAPPING METROID: EVALUATION OF 2D SPATIAL NAVIGATION ABILITIES IN PERSONS WITH VISUAL IMPAIRMENTS USING AN AUDITORY PLATFORM GAME

Modified from an article under review for the Journal of Visual Impairment & Blindness. Michael Oren, Chris Harding, Terri Bonebright

Abstract

In this paper, we describe the results of a comparison of mental maps created by participants of a platform game created for persons with visual impairments. The study consisted of three groups with 9 participants each: sighted participants playing the audio-visual version as a control group, sighted participants playing the audio-only version of the game, and participants with visual impairments playing the audio-only version of the game. We used verbal map descriptions of the game levels and compared the accuracy of maps between the three groups. The results indicate that the strategy we used to convey audio cues allows users to create mental maps with the same accuracy as users who experience the graphical version. Our results also suggest that previous visual experience may affect the formation of mental maps formed from auditory data, although further study is necessary to confirm this finding.

1. Introduction

Over the past several decades many studies have been conducted to evaluate the spatial navigation abilities of persons with visual impairments (Golledge, 1989; Marston, 2007; and Loomis, 2005). These studies were designed to evaluate systems that were created to help persons with visual impairments navigate around the 3D world in which they live and work. The goal of these studies into route navigation is to create a system that will aid
persons with visual impairments when attempting to navigate a new environment for the first time (such as a college campus).

While it is undeniable that the research into 3D spatial navigation is critical to improving the lives of persons with visual impairments, we believe that there are potential benefits to improving researchers understanding of 2D spatial abilities and navigation as well. This is particularly relevant to designers of computer interfaces for persons with visual impairments as computer interfaces are still traditionally two dimensional by nature and it may be useful to help guide users to objects on the screen. We also feel that this is an important topic for research as previous work by Eardley has indicated that persons with visual impairments are at a disadvantage when trying to navigate 2D space, such as a spreadsheet (Eardley, 2007). Stockman and others have recently conducted research aimed at providing auditory cues for spreadsheets, but such work focuses primarily on the auditory display of data as opposed to spatial orientation and navigation (Stockman, 2005). It is our hope that this project may help recognize the difficulties some individuals with visual impairments experience and that it can lead to training programs that would help such people improve their 2D spatial orientation and mental mapping abilities. This training could prove extremely valuable, as Eardley believed that the performance of 2D spatial orientation tasks by persons with visual impairments suffered due to their lack of experience with such tasks (Eardley, 2007). Eardley’s explanation is supported by the fact that people with visual performed as well as sighted participants on 3D spatial tasks, while performing worse at 2D tasks.

We set out to try to create a simple game that could help familiarize persons with visual impairments with the concepts of 2D orientation and navigation that are common in diagrams and spreadsheets. For this project, we used the Audio Platform Game (APG) that we had created as a means of providing a more complex computer-based entertainment option for students in boarding schools for the visually impaired (Oren, 2007, May). The
game is based around the same concepts as the early Super Mario Bros. (Figure 7) games, in that the user navigates a character that must hop on platforms and avoid obstacles (such as pits and opponents). APG is set in a 2D, side-scrolling world with four distinct heights, called floors, that the user must recognize via audio cues and navigate in a timely manner in order to avoid the various obstacles in the game.

Using APG as a tool, we were able to conduct an evaluation of the spatial mapping abilities of persons with visual impairments and compare these results with sighted individuals that had played an audio-only version as well as an audio-visual version of the game. Our goal for this project was to determine if players experiencing the audio-only version of the game were able to develop a mental map of the game levels that was as accurate as those playing the graphical version of the game. As prior research conducted by Dodds suggested that persons who are blind from birth use primarily egocentric means of navigation, we thought it would be useful to ask participants to self-report the means they used to navigate through the level as well as the means they used to create their mental maps of levels in order to see if these findings still held true for the spreadsheet-like navigation present in the game (Dodds, 1982).

2. Background

2.1. Spatial Ability / Spatial Maps

There are two primary display methods that researchers have used to provide people with visual impairments with a greater sense of spatial awareness: auditory and tactile. While

Figure 7 A screenshot from Super Mario Bros.
we are primarily interested in the past work with auditory displays for assessing and assisting spatial navigation, work in tactile displays can also provide insight into the spatial abilities of persons with visual impairments. In a study using tactile input, Klatzky found that past visual experience had no effect on a participant’s ability to complete spatial tasks (Klatzky, 1995). These findings were reinforced by the work of Röder whose experiment utilizing haptics and sound illustrated that previous visual experience did not affect participants’ ability to map spatial relationships through imagined movement (Röder, 1998). It is clear from both of these studies that visual experience does not play a central role in a person’s ability to properly analyze spatial information and form mental images of spatial relationships.

Although tactile interfaces, which allow users to easily feel the space between objects, are an effective means of conveying spatial relationships, many of the solutions tend to cost more or be less dynamic than graphical counterparts. An alternative to tactile interfaces is auditory interfaces; however, they present some unique challenges for conveying spatial relationships. Part of this challenge comes from trying to properly convey distance information via auditory cues, as Loomis found in his study when participants consistently performed worse on interpreting distance via audio cues (Loomis, 1998). A 2005 study by Després provides evidence that although distance is normally hard for individuals to discern via audio, persons with visual impairment exhibit signs of enhanced auditory spatial abilities that may aid in their interpretation of auditory distance cues (Després, 2005). This supra-normal auditory ability was found to be true regardless of when the participant developed a visual impairment (Voss, 2004).

In his early work with auditory navigation, Golledge found that persons with visual impairments preferred verbal cues to tonal cues (Golledge, 2004). Further discoveries about verbal cues are explained by the findings of Röder, who used fMRI scans to discover that the visual cortex is activated through verbal cues in persons who are blind from birth (Röder, 2002). Despite this evidence in favor of verbal cues over tonal cues, we chose to use tonal
cues based on the work done by Kramer, Blauert, and other sonification researchers (Kramer, 1999; Blauert, 1997). We chose this divergent route because while findings have shown a preference for verbal cues and an activation of the visual cortex, persons with visual impairments that made up our focus group at the Indiana School for the Blind indicated a preference for non-verbal cues in order to increase their immersion in the game world (Oren, May 2007).

It is also noteworthy that Alfonso’s work has shown that previous experience with navigating a spatial environment plays a greater role in navigation than verbal directions (Alfonso, 2005). This indicates that it may be more important to provide persons with visual impairments with a way of virtually traveling an area rather than simply providing them with verbal directions before, or even as, they travel. It may, therefore, be beneficial for a person with a visual impairment to first navigate through a new area via a 2D auditory map that uses some of the spatial routing techniques explained in this paper or techniques similar to those of the BATS project (Parente, 2003).

Brunyé found that it is important to limit the level of detail in spatial information provided to users because there is a negative correlation between level of detail provided and spatial memory, which would adversely affect participant mental maps (Brunyé, 2007). With this in mind, we chose to design the game so that it focused on representing spatial relationships (higher or lower) as opposed to the actual location of objects on the screen. We also limited the “view” of the user to objects they could immediately navigate to.

### 2.2. Sonification

We chose to borrow from the field of auditory displays in order to create our non-verbal route guidance system. One type of auditory display, sonification, is based on research in auditory perception (Bregman, 1990), and a general theory for this technique has emerged over the last decade (Stevens, 1996; Barrass, 2003; Walker, 2005). In sonification, non-
speech audio is used to translate or map information into sound (Kramer, 1999). For example, a series of numbers can be mapped onto a series of pitches for auditory graphs with the upward or downward shifting of a sound source’s pitch mapped as a rise or fall of the source’s elevation (Flowers, 2005). Another example of sonification are auditory icons (Gaver, 1988), which use pre-recorded sound “caricatures” of real-world sounds that link a concept to the sound it produces; for example, a barking sound communicates the general concept of a dog. Auditory icons tend to be learned easily but, such links to natural sounds do not exist for all concepts, and therefore abstract concepts are especially difficult to represent using this method. However, such abstract concepts can be successfully represented using earcons (Blattner, 1989), another sonification technique, which employ artificial sounds (e.g. a short melody of piano tones) to encode the abstract information (Brewster, 1993). For example, changing the repeat rate of an audio cue can encode changes in distance to a source object. Finally, a simple technique using 2-D spatial audio to represent horizontal location is panning, in which subtle left to right speaker changes convey the left-right movement of a sound source (Blauert, 1997).

3. Design and Development

Here, we briefly sum up the design of the game. In our previous work we have gone into greater detail regarding the auditory interface design and our initial evaluation of the game, which found that participants playing the audio-only version of the game took longer to complete levels but completed them just as well as those playing the graphical version (Oren, June 2007).

3.1. Hardware and Software

Although more modern sound cards and other hardware exist on the market, we chose to test this application on a laptop with a Pentium IV 2.0 GHz processor, 512 MB of RAM, and an Intel AC'97 onboard sound card. We purposely chose to use this older PC with an
onboard sound card as our work with boarding schools for the visually impaired indicated that they very rarely purchase separate sound cards for their machines and often use computers that are three or four years old. Participants played the game using standard, over-the-ear Sony MDR-V150 headphones, which are readily available in the consumer market. A Logitech Dual Action game pad was used as an input mechanism for the game.

We used the Python programming language to write the game and used the pyGame module to handle rendering of the visual components as well as the detection of object collisions (i.e., to determine if a player had come into contact with an opponent, landed on top of a platform after a jump, etc.) We also used the pySonic module, a Python wraparound for the FMOD library, in order to play and manipulate the sounds we used to aurally depict the game objects.

3.2. Design

The game consists of seven levels that increase in difficulty as the player successfully completes each level. Within the levels, there are four primary objects: the character (the player’s on-screen representation), platforms, opponents, and pits (Figure 9). During the mapping levels of the game, opponents were not present since we found that they added undue stress that affected the users’ ability to recreate the map of the level. Each type of object has its own set of sounds and methods to convey information, as shown in Table 4. The character, that is assumed to be in the middle of the screen, can walk left and right or jump up and down, which moves the game world accordingly. There are also four imaginary “floors” at fixed elevations (called 1 to 4) with possible platforms for the player to walk on (Figure 8).
<table>
<thead>
<tr>
<th>Object</th>
<th>Voice (numbers)</th>
<th>Voice (icon)</th>
<th>Auditory Icon (Pitched, Panned)</th>
<th>Auditory Icon (Pitched)</th>
<th>Earcon (Pitched, Panned)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player Jump</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Player Land</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Player Walk</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Floating Platform</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Solid Platform</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>End Of Platform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Warning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pit</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bees</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dogs</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Primary Objects and Actions in the Game World with Their Auditory Representation

The character moves through the game world with sounds designating footsteps and jumping. The character is only allowed to jump up one floor at a time. Floors in the game consist of both floating and solid platforms. The character can walk underneath the floating platforms but will run into a “wall” when approaching a solid platform that forces the character to jump on top of it to continue. Some areas of the game world present fairly...
complex combinations of floating and solid platforms that force the character to backtrack to find a way on top of the platform, a process that requires the gamer to have a strong sense of the spatial layout of the level.

It is quite common that the character reaches the end of the current platform (i.e. reaches a gap) and needs to jump over the gap to a new platform. Glissando cues are used to indicate if an object is relatively higher or lower than the player while panning is used to indicate the relative horizontal position of objects, although only objects directly in front of the player are represented via audio cues. Similarly, the repetition rate of the platform sound increases as the player approaches the best “jump-off” point and a “platform end warning” is played if a player is near the edge of a platform.

4. Methods

4.1. Participants

Eighteen people (8 men and 10 women) who had normal vision were recruited from Iowa State University; their average age was 28 (SD = 6.9). Participants were pseudo-randomly divided into an audio-only and audio-visual group with equal numbers of men and women in each. In addition, 9 people (7 men and 2 women) who were legally blind (1 blind after birth, 8 with severe visual impairments) were recruited from Iowa State University and Iowa Braille and Sight Saving School. These participants had an average age of 26 (SD = 11.2). Participant breakdown by gender and game version is shown in Table 5.

<table>
<thead>
<tr>
<th>Audio-only version</th>
<th>Audio+Video version</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (6 male, 2 female)</td>
<td>9 (4 male, 5 female)</td>
</tr>
<tr>
<td>18 (8 male, 10 female)</td>
<td></td>
</tr>
<tr>
<td>Participants with visual impairments</td>
<td>Sighted participants</td>
</tr>
</tbody>
</table>
Procedure

Participants were trained to play the game by reading through (or listening to) a series of instructions from a Microsoft Word document with embedded sound files at their own pace. During this training they played the sample sounds as often as they desired and at the end of each section of instructions they played through a training level that dealt with each newly introduced element in isolation. These training levels moved from the very simple, such as moving around on a completely flat level, to the more complex elements, such as avoiding of pits, dogs, and bees.

All participants played the game with a game controller and headphones; those playing the audio-visual version were able to see the game on a monitor. Participants with normal vision who were playing the auditory-only version of the game were blindfolded during training and while they played the game. During training (and the game phase), participants were free to pause the game at any time to take a break, ask a question, or adjust the headphone volume.

After three of the training levels participants were asked to map the levels and rate the difficulty of creating the map.

Participants then played through a series of seven levels of increasing complexity. At the end of levels three and seven, each of which lacked enemy obstacles (although they did both contain a pit), participants were asked to map the level. As we needed to ensure that sighted persons did not have an advantage in creating the map, we chose to use a verbal mapping technique.

The mapping procedure used for the training levels was repeated for levels three and seven of the actual game, but in this case the researcher, with a piece of paper hidden behind a blind to hide the map from the participant’s view, drew an image of the map based on the participant’s description. This procedure consisted of the researcher asking the participants to state the floor that they started on. From there, the researcher asked the participants to travel
from left to right (the same direction they traveled through while playing the level) and recall the objects they encountered regardless of whether they chose to try to jump on to it or not. If the participants indicated the next object had been a platform, they were asked if the object was floating or solid. Participants were then asked if the floor number of the object was higher or lower than their current recalled location. Along with this information, participants were asked to state the floor number of the objects they encountered. After fully describing the object in this manner, participants were asked if there were any other objects that had been encountered at the currently recalled location. If not, they were asked to mentally move on to the next object and repeat this process. By the end of the third training map, participants were sufficiently proficient with the procedure that they no longer needed prompting by the researcher, although occasionally the researcher would still need to ask the participant to clarify vague descriptions (e.g. when a participant simply indicated that a platform was encountered without specifying the relative height).

While both Gordon and Tobler proposed methods for the creation of mental maps via numerically encoded matrices, we decided to forgo this complication as the maps we sought to create were more simplistic since they did not require the user to estimate exact distances of objects (Gordon, 1989; Tobler, 1976). This simpler procedure for creating and scoring images of the participants’ mental maps was done because unlike Tobler and Gordon, we were not interested in how accurate participants’ mental distance measure was for objects within the game world.

Participant’s verbal descriptions were translated into diagrams of the levels’ maps (Figure 10), including missed objects and made up objects (i.e., objects the participant had erroneously claimed to have encountered). These maps were scored so that a perfect score was zero (meaning no mistakes), while one point was added for each missed object or made up object and a half point was added for each object out of place or identified as the wrong type of a similar object (e.g. they indicated an object was a solid platform when it was
floating). Screen and audio recordings of the participants’ traversal through the level were analyzed so that points would not be counted against participants in either of the audio-only groups if the participant did not hear an object.

5. Results

There was one independent variable of interest (game version) with three levels: audio-visual played by sighted participants (n=9), audio-only played by sighted participants (n=9), and audio-only played by participants with visual impairments (n=9). The dependent measures were the mapping scores for level three and level seven, and the number of objects the participant “invented” (objects that did not exist in the actual level that the participant placed in her mental map). A mapping score of zero indicated that the participant had perfectly described the actual level.

5.1. Mapping Scores

The mean mapping scores did not vary significantly for level three between those with normal vision playing the audio-visual version of the game (M=3.11, SD=1.60), normal vision playing the audio-only version of the game (M=4.17, SD=2.08), and visually impaired playing the audio-only version of the game (M=2.39, SD=1.50), $F(2,24) = 2.37, p = .12$. The
mean mapping scores also did not vary significantly for level seven between those with normal vision playing the audio-visual version of the game (M=3.06, SD=2.11), normal vision playing the audio-only version of the game (M=3.11, SD=1.41), and visually impaired playing the audio-only version of the game (M=3.00, SD=1.85), $F(2,24) = .008, p = .99$.

The mapping score results suggest that the audio-only interface was able to properly convey the spatial relationships and routing information that participants required to form a proper mental map of the level. Samples of the map results can be seen in Figure 11.

![Figure 11](image.png)

Figure 11 The results from one of the sighted participants in the audio-only group.
The mean number of invented objects for level seven did not vary significantly between those with normal vision playing the audio-visual version of the game (M=.33, SD=.500), normal vision playing the audio-only version of the game (M=.67, SD=.866), and visually impaired playing the audio-only version of the game (M=.44, SD=.882), \(F(2,24) = .44, p = .65\). However, there was significant variance in the number of invented objects for level 3 between those with normal vision playing the audio-visual version of the game (M=.67, SD=.867), normal vision playing the audio-only version of the game (M=2.11, SD=1.965), and visually impaired playing the audio-only version of the game (M=.11, SD=.333), \(F(2,24) = 6.09, p = .007, \eta^2 p = .34\).

While it is not clear why participants with normal vision using the audio-only interface invented significantly more objects for level three than the other participants, we do have two possible theories. Our first hypothesis is based on the work of Després and Voss that indicates that persons with visual impairments develop a supra-normal auditory spatial abilities that may have simply allowed them to better understand the spatial relationship of objects in the auditory game world (Després, 2005; Voss, 2004). However, the lack of a significant difference for the overall map score between groups speaks against this hypothesis. There may be another clue to this peculiarity since only the third level of the game, which placed the goal on the third floor as opposed to the first floor, resulted in a significant difference between the number of “invented” objects. Participants had encountered a similar situation in a training level of the game, but the majority of levels placed the goal on the first floor. Similarly, most popular 2D platform games such as Super Mario Bros. and Sonic the Hedgehog, place the ending of the level on the first floor—although the original Super Mario Bros. encouraged players to jump up to the goal in order to gain more points. Participants with normal vision may simply have been affected by their previous visual experience with platform games and may thus have falsely perceived the goal to be on the first floor of the level. This seems to be supported by the fact that four
participants in each of the groups with normal vision indicated that the goal on level 3 was placed on the first floor, while only one participant with a visual impairment placed the goal on the first floor of level three. For the participants with normal vision playing the audio-only version of the game, this often led to a description of platforms going down from the actual end of the level (floor three) to the platform on floor one.

5.2. Follow-up Survey

Survey data on questions regarding computer and computer game usage that were rated on a 5-point scale, from 1, never or almost never, to 5, very frequently, revealed that participants with normal vision indicated that they very frequently used a computer and that they played computer games infrequently. These participants also reported that they had infrequently played platform games and that none of them had experience with screen readers although they had heard of them. In contrast participants with visual impairments reported significantly less frequent use of computers than those with normal vision, $F(1,22) = 12.50, p = .002$, and a significantly higher level of familiarity with screen readers and other audio representations for visual objects, $F(1,23) = 39.89, p < .001$. There were no differences between the two groups of participants on frequency of playing computer games in general or platform games in particular (see Table 8 for means and standard deviations).

The response from participants regarding their previous experience with platform games may indicate that our hypothesis about previous experience playing platform games led to more invented objects. However, the question only asked participants about how often they currently play platform games and did not inquire about past experience.
Table 6 Means and Standard Deviations for Survey Response for Participants with Normal Vision and with Visual Impairments

<table>
<thead>
<tr>
<th>Question</th>
<th>Normal Vision</th>
<th>Visually Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often do you play computer games?</td>
<td>2.81 (1.17)</td>
<td>2.89 (.33)</td>
</tr>
<tr>
<td>How often do you use a computer?</td>
<td>*4.88 (.34)</td>
<td>*3.67 (1.50)</td>
</tr>
<tr>
<td>How often do you have played platform games?</td>
<td>2.53 (.87)</td>
<td>2.44 (1.42)</td>
</tr>
<tr>
<td>How much experience have you had with a screen reader or other representations of visual objects through sounds or voice?</td>
<td>*1.82 (1.07)</td>
<td>*3.89 (1.69)</td>
</tr>
</tbody>
</table>

Rating scale used was from 1, never or almost never, to 5, very frequently. Standard deviations are in parentheses.

* indicates a significant difference

Participants also reported on how they typically use audio in previous computer games they had played. Regardless of their visual abilities, 11% reported that they simply ignore the sounds; 50% said that they use sound as an enhancement of the realism of the game, but they do not actually use it in a functional way; 8% revealed that they used sound to warn them of critical game events; such as an incoming attack; and 32% reported that they use sound to obtain spatial information (e.g. about opponents and objects behind them when playing 3D games). When asked about their preferred method of audio feedback for audio interfaces in general, participants with normal vision reported no prior experience with auditory interfaces of any kind, while all participants with visual impairments reported that they used auditory interfaces that provide voices and/or tones.

Using the self-reported mapping strategies of participants, we scored their strategy on a 7-point scale, from 1, very exocentric, to 7, very egocentric. These numerical scores were based on which of four choices participants chose to describe their mapping strategies; they
were allowed to choose multiple options to describe their strategies. The strategies available for participants to choose from were (A) created a picture in your head where you placed the objects to get a picture of the overall scene; (B) remembered the relative audio cues as you went along and figured out things based on their position in relation to the player character; (C) used your path through the level and drew objects that your character jumped on/over; and (D) other (please explain). Although two participants chose option D, the description of their strategy was essentially the same as the currently available strategies. These scores were mapped to numerical values by placing strategy A at the beginning of the spectrum (1) and strategy C at the end of the spectrum (7-most egocentric). Combinations of the strategies were then scored accordingly to how they fit between those two extremes (a view of the whole map—exocentric—vs. a view of only what came next—egocentric).

We performed a Pearson’s correlation to determine if the level of egocentricity had a linear relationship with the mapping score and determined that there was not a significant correlation between mapping strategy and mapping score, \( r(25) = .12, p = .55 \). We also performed a three group one-way analysis of variance and compared the mapping strategies used between groups to determine if Dodds’s earlier finding about mental maps of persons with visual impairments being more egocentric in nature held true for the side-view maps of our game (Dodds, 1982). For this analysis, marginal significance was found suggesting that participants in the visually impaired group (\( M=544, SD=1.667 \)) had the most egocentric mapping strategy followed by participants in the audio-only group (\( M=3.11, SD=1.616 \)) and then those in the audio-visual group (\( M=4.33, SD=2.5 \)), \( F(2,24) = 3.16, p = .06 \). We believe that these findings support Dodds’s earlier findings, although it may also be possible that our presentation of the audio-only version of the game world through relative cues may have influenced these results.
6. Conclusion

The results found in this study provide a good indication that the auditory interface created for this game can be used to properly convey the spatial relationships needed for users to create proper mental maps of a 2D interface regardless of whether users have normal vision or have a visual impairment. It appears likely that these interface techniques can be adapted from a side-view to a top-down view, and may thus provide benefits to route planning for persons with visual impairments. However, due to the small sample size used, the results of the ANOVA may not have enough power to conclude that all groups produced maps with equal levels of accuracy. An expansion of this study would allow us to draw stronger conclusions that would be able to confirm the null hypothesis and provide stronger evidence that auditory interfaces can lead to mental maps that are as accurate as those create via visual input. While we would have liked to have more participants, we ran into difficulties finding persons with visual impairments available to spend an hour to an hour and a half participating in this study that were located in or around central Iowa. The majority of our participants with visual impairments were students at the Iowa Braille and Sight Saving School, and we tested all of the students there who were able to get parental consent to participate in the study.

We are currently unable to fully explain why, in map level 3, participants with normal vision playing the audio-only version of the game tended to invent game objects more frequently than participants from the two other groups and what possible role prior experience with platform games may have played. A follow-up study using only participants with normal vision, half of whom have played platform games in the past and half whom have not, would help to clarify this question. Significant differences found in this follow-up study could provide evidence for the more general effects of previous visual experience for the construction of mental maps from auditory data.
7. Acknowledgements

We thank the students and staff of the Indiana School for the Blind, particularly Al Lovati who served as our direct contact. We would also like to thank Carol Farrenkopf, for her advice and guidance in working with the visually impaired and blind. Our appreciation also extends to the Iowa Braille and Sight Saving School students, staff, and faculty who helped us evaluate this project, particularly Mike Hooley who served as our liaison there. In addition, we thank the students at Iowa State University who served as participants for this study. Finally, we would like to thank Kyle Shipley who worked on the first iteration of this project with us.

Part of the research has been funded by NSF grant IIS-534681.
CHAPTER 5 SUMMARY AND FUTURE WORK

There is a long history of exploring auditory interfaces and sonification, from the earlier research of Gaver and Blattner to more recent studies by Flowers and Kramer. The audio-centric game system described in this thesis’ research is based on such pioneering work. This game system conveys 2D spatial information through audio cues alone and requires only low-cost consumer-grade equipment. While this research also relates to the work by Golledge, Klatzky, Loomis, and Marston in the field of auditory route guidance it focuses on routes inside a two-dimensional side-view game world and favors non-verbal cues.

This research expands upon Gärdenfor’s guidelines for audio games to allow for more complex, spatially based games. While this effort still does not advance audio games up to the current state of video games, it is hoped that this research helps narrow the entertainment divide by providing more entertainment options for persons with visual impairments and that it may more generally narrow the digital divide by increasing their use of computers.

Chapter 2 explored the basic research in auditory displays and sonification that served as the basis for the audio cue design used in this research. The results presented in this paper showed how auditory icons and earcons are used to create an immersive virtual world. This interface allowed blindfolded sighted participants to navigate through the levels with equal adeptness (in terms of no statistically significant difference in the number of restarts) as participants playing the game with graphics.

Chapter 3 explained the results of the participatory design sessions conducted with students at the Indiana School for the Blind. It details the results of a usability study comparing the results of persons with visual impairments with a group of sighted participants playing the audio-only version of the game and sighted participants playing the graphical version of the game. No difference was found in the performance within the game between
persons who had a visual impairment, were blindfolded, or were able to see while playing the game. However, the results showed that both audio-only groups took three times as long to complete levels as compared to the sighted audio-visual group. While it is unclear why the audio-only groups took longer to complete the levels of the game than those able to see the game’s graphics, it is possible that participants were more cautious using a sense they had less experience using with which to navigate.

Finally, chapter 4 reviewed past work done in the field of spatial route navigation and mental mapping and presented the results of a mental mapping study conducted using the level layouts of our research. This revealed some interesting findings in the differences between the mental maps formed by persons who are visually impaired compared to sighted persons, as sighted individuals playing the audio-only version of the game tended to invent more objects that were not present in the level. While the results for the mapping strategy used by participants were only marginally significant, they did reveal evidence to support Dodds’s findings that persons with visual impairments are more likely to use egocentric mapping strategies.

In general, this research focused on the study of the first audio platform game that maintains all primary components of games within the platform genre. Based on previous research conducted in the fields of auditory displays, and audio games this work serves as an example in the design of an interface to convey spatial relationships via audio cues that allows users to easily interpret and quickly react to objects within the virtual environment. It is hoped that these results are able to increase the understanding of the differences in the spatial abilities of persons with visual impairments and the sighted because this may lead to further awareness of the role auditory perception and visual experience play in human development.
**Future Work**

Possible future work includes the use of non-verbal information for providing route information to persons with visual impairments (such as while in a mall) and their use in guiding users through a spreadsheet application. The software engine created for the game could be used to develop additional platform games for persons with visual impairments, including educational games that improve educational equality with the education opportunities already available to people with normal vision. It is also hoped that future studies can be conducted using participants who are blind from birth, who would not have the previous experience of navigating with their visual sense that may provide some evidence for the hypothesis that individuals who are not used to navigating via auditory cues may exhibit a greater level of caution and slow their performance on navigational tasks.

The results of the current study were unable to explain the anomaly of sighted participants in the audio-only group inventing more objects in their mental map of the third game level; this opens up an area of future work where past visual experience is explored as it relates to the interpretation of auditory spatial information.
REFERENCES

Alfonso, Amandine; Katz, Brian FG; Blum, Alan; Jacquemin, Christian; and Denis, Michel (2005, July). "A Study of Spatial Cognition in an Immersive Virtual Audio Environment: Comparing Blind and Blindfolded Individuals". ICAD ’05. Limerick, Ireland.


AudioGames.net. “Super Liam”.


Brunyé, Tad; Taylor, Holly; and Worboys, Michael (2007). “Level of Detail in Descriptions and Depictions of Geographic Space”. Spatial Cognition and Computation, 7(3), 227-266.


Röder, Brigitte; Stock, Oliver; Bien, Siegfried; Neville, Hellen; Rösler, Frank (2002). "Speech processing activates visual cortex in congenitally blind humans". European Journal of Neuroscience, Vol. 16, 930-936.


Voss, Patrice; Lassonde, Maryse; Gougoux, Frederic; Fortin, Madeleine; Guillemot, Jean-Paul; and Lepore, Franco (2004, October). "Early- and Late-Onset Blind Individuals Show Supra-Normal Auditory Abilities in Far-Space". Current Biology, 14, 1734-1738.

