Designing Accessible Nonvisual Maps

by Brandon Biggs

Submitted to OCAD University in partial fulfillment of the requirements for the degree of Master of Design in Inclusive Design

Toronto, Ontario, Canada, April, 2019

Copyright Notice

Author's Declaration

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revisions, as accepted by my examiners.

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

I understand that my MRP may be made electronically available to the public.

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

Signature ____

Abstract

Access to nonvisual maps has long required special equipment and training to use; Google Maps, ESRI, and other commonly used digital maps are completely visual and thus inaccessible to people with visual impairments. This project presents the design and evaluation of an easy to use digital auditory map and 3D model interactive map. A co-design was also undertaken to discover tools for an ideal nonvisual navigational experience. Baseline results of both studies are presented so future work can improve on the designs. The user evaluation revealed that both prototypes were moderately easy to use. An ideal nonvisual navigational experience, according to these participants, consists of both an accurate turn by turn navigational system, and an interactive map. Future work needs to focus on the development of appropriate tools to enable this ideal experience.

Acknowledgements

There are too many people to thank individually, so if you are not below, thank you for helping me realize everything that we did in this project. Thank you to:

- Everyone at the Perceptual Artifacts Lab for being a fantastic team.
- ESRI for my internship and support with the ESRI API.
- The NSERC CREATE Data Visualization program for funding me.
- James, thank you for jumping in during my second year and really helping me to get CamIO working, then agreeing to be on my committee, and contributing so much to this project.
- Peter, thank you for being my adviser. You really encouraged me to stretch myself both conceptually and academically. I learned so much from our brainstorming sessions and I can't say how important being a part of the Perceptual Artifacts Lab has been. Thank you for everything you did and being such an amazing advisor and mentor.
- Everyone at Sonja Biggs Educational Services for supporting this project.
- Jay, for all the work you did on both prototypes and helping run the study.
- Olenka, for being so supportive and giving this project meaning.
- Both my parents for listening to my crazy ideas.
- Josh, for the critical math in the map navigation system.
- My parents-in-law for everything that allowed me to focus on this project.
- Claudia, for being such an understanding, loving, and supportive partner through these last two years. You are the best alpha tester and partner I could have ever asked for.

Table of Contents

List of Figures

Preface

statement of contributions

Chapter 2: The Audio Game Laboratory was previously published as:

Biggs, B., Yusim, L., & Coppin, P. (2018). The audio game laboratory: Building maps from games. Retrieved from http://icad2018.icad.org/wp-content/uploads/2018/06/ICAD2018_paper_51.pdf

Chapter 3: Design and Evaluation Of An Audio Game-Inspired Auditory Map Interface has been accepted to the International Conference of Auditory Display 2019 and will be published in June 2019 as:

Biggs, B., Coughlan, J., & Coppin, P. (2019). Design and Evaluation Of An Audio Game-Inspired Auditory Map Interface.

Chapter 1: Introduction

Navigation can be very challenging for nonvisual users, but despite these challenges, an effective nonvisual navigational solution has not been developed (Jacobson, 1999). According to the co-design with 16 nonvisual participants in chapter 3, an ideal navigational experience is an extremely accurate turn-by-turn navigational solution connected with an interactive map of some kind. Accurate localization, which enables turn-by-turn navigation, has been researched by Fusco and Coughlan (2018), Sato et al. (2017), and Serrão et al. (2015). This thesis focuses on a more effective interactive map.

Extensive research has been done on raised line and 3D model tactile maps such as Jones and Sarter (2008), Paladugu, Wang, and Li (2010), Brock and Jouffrais (2015), and Holloway, Marriott, and Butler (2018), but a digital 3D tactile map has not been developed. There has been some research into digital auditory maps such as Loeliger and Stockman (2014) and Heuten, Henze, and Boll (2007), but digital auditory maps have not entered the consideration of nonvisual users (Butler, Holloway, Marriott, & Goncu, 2017). There has been a few studies creating interactive maps (tactile maps with auditory labels, such as Brock and Jouffrais (2015) and Holloway et al. (2018)), but no studies connecting a digital auditory map with a tactile map have been conducted. This thesis presents the design and evaluation of an interactive 3D tactile map and a digital auditory map based on the same underlying map data. When used together, these two prototypes create an interactive cross-sensory map experience.

The following chapters present the design and evaluation of both a digital auditory map and a 3D model map, shows how to connect the digital and physical maps, and presents a co-design with blind participants asking them to describe their ideal navigational experience. Chapter 1 describes the inclusive design approach underlying the designs in this project, Chapter 2 presents the theoretical basis for a digital auditory map, Chapter 3 presents a study evaluating a prototype using the interface elements from Chapter 2, Chapter 4 presents a study evaluating a system to connect the digital interface in Chapter 3 to a tactile map in the first half, and and undertakes a co-design to ask users their ideal digital map experience in the second half, and Chapter 5 presents further research that needs to be done to build an ideal nonvisual navigational experience.

There is currently a “mismatch” for nonvisual users in the mapping space. Google maps and ESRI products are the standard mapping solutions for governments, schools, and other institutions (*Cal fire*, 2019), and they are completely visual. There are some extremely basic digital auditory maps from Google Maps and ESRI, but they do not show important information such as streets, terrain, shapes, or sidewalks (ESRI, 2018; Google, 2019). Maps are fundamental tools in data analytics, navigation, science, and many other fields (Tufte, 1983; Yau, 2011). Despite this information being so critical, the existing nonvisual mapping solutions are expensive and often require a manual transcriber and a special set of machines to produce (Butler et al., 2017). This situation where a particular group has been completely excluded from a design is called a “mismatch” (Holmes, 2018).

Inclusive Design

inclusive Design Research Centre (2018) and Holmes (2018) characterize inclusive design as design that considers people of all abilities. Three “dimensions” of inclusive design are proposed by Treviranus (2018a), Treviranus (2018c), Treviranus (2018b), and inclusive Design Research Centre (2018): “(1) Recognize, respect, and design for human uniqueness and variability. (2) Use inclusive, open & transparent processes, and co-design with people who have a diversity of perspectives, including people that can’t use or have difficulty using the current designs. (3) Realize that you are designing in a complex adaptive system.” Treviranus (2018a) and Inclusive Design Research Centre community (2018) propose that designs that follow the inclusive design principles are generally more extensible and adaptable by default and can more easily be improved to reach more users. This inclusive design approach means creating a unique solution that matches the user’s preference, or a one size fits one design (Inclusive Design Research Centre community, 2016). Donovan (2018) proposes that creating designs that are built to continuously allow for more users means there is a constantly growing user base which increases revenue (W3C Web Accessibility Initiative, 2019). In business, often the first step to scaling a product is to make a “handcrafted” experience that a few people love, then expand that service to as many users as possible (Masters of Scale, 2017). Inclusive Design Research Centre community (2016) proposes that an inclusively designed product should have the modularity and extensibility to allow a handcrafted experience for every user. The process for designing that hand-crafted experience needs to be in the hands of the user so they can constantly be a part of the solution they are using (Charlton, 1998; Treviranus, 2018c). There are over a billion users in the world with a frequent mismatch between the standard designs employed by most companies and the user's ability to use the product (Charlton, 1998; Donovan, 2018). These are users who would probably be customers but for the lack of adaptability in the platform (Treviranus, 2018b). This project managed to follow both the inclusive design principles and create an extendable minimum viable product.

Inclusive Design in this Project

This project followed many of the principles of inclusive design, which could be why the prototypes were so well received, although there were some areas that need more focus in the future. The digital auditory map was developed using a modular web design that allowed for multiple views, such as first-person and grid, and more interfaces can easily be added to the design as users request or make them. The web interface means that any mobile and computer platform with a modern web browser can use this prototype. The system also included the infrastructure for a keymap which, when fully implemented, will allow users to route any keyboard, mouse, touch gesture, console controller, accelerometer, or any other input event to an action. This means that a user could potentially map their particular device to do any action in the prototype.

The design of the digital auditory map interface came from a natural laboratory of audio games (games that can be played completely using audio). These audio games contained maps that were already being used by some members of the nonvisual community, so rather than exclusively relying on past academic research, conducting numerous studies, running a co-design, or following standard web accessibility practice, the initial design came from an analysis of audio game conventions present in existing maps that had been tested by the demands of a consumer nonvisual market. This technique of observing working conventions or solutions outside academic literature is referred to as a natural laboratory by Biggs, Yusim, and Coppin (2018). A prototype was then built following these conventions. The feedback on the prototype was extremely positive, although the prototype would have been even more effective if a co-design session had been conducted asking users about their ideal navigational experience before the initial prototype was built. There were two justifications for holding the co-design session later: first, it was more economical on a low budget project to only have participants show up once rather than twice (once for the co-design, and once to evaluate the initial prototype). Second, the lead researcher was blind and had lived the experience of having problems navigating, so he initially focused on building a solution he would find useful.

The team first built the initial prototype and as users were brought in to test the prototype, they were first asked to co-design their ideal navigational experience before viewing the prototype. Asking participants to design a solution before viewing the prototype meant that there was no design fixation on matching the prototype. Some of the features participants mentioned that they wanted in a digital map were a mobile version, more details and pathways, and a tactile interface. It would have been possible to provide more detail and a mobile interface in the initial prototype if this had been the designer’s priority. Thanks to the extensible design of the prototype, adding these features will be extremely easy in the next iteration.

The tactile prototype was also made before doing a co-design, which meant that more focus was given to the objects rather than pathways. Participants loved the objects, but found the map too broad to use. During the co-design, they mentioned pathways and different terrain were the most important information to have on a map. They also described a turn-by-turn navigational solution that could have also been evaluated along with the tactile map if the co-design had been the first step in the design process.

Data is Universal

One of the major objectives with building the prototypes presented in this paper was to create a redundant experience in both the auditory and tactile sensory modalities that was called “cross-sensory” (Walker, 2016). The fundamental premise that shaped the research in each modality was that data had no inherent sensory properties and the design of each representation was dictated by the point, goal, or story that each modality needed to convey in the most effective way (Yau, 2011). For example, one of the goals was to convey a sense of an object’s shape. In the audio modality, this was done by creating a set of tiles represented by a short sound and speech clip that the user could move in and out of as they pressed the arrow keys. In the tactile modality, a 3D model of the object was used. The only relation between the two modalities was the underlying data that underpinned the representation. This approach allowed for multiple interface designs in each modality that all conveyed the same information.

A cross-sensory representation allows users to access the data in the modality they prefer. Users with a sensory impairment can use the modality they need and users who can access more than one modality can choose the modality that they feel most comfortable with. Currently producing a single modal representation seems less expensive in the short-term, but once the cost of adaptation is considered, it would have been less expensive to build a cross-modal representation to begin with (Butler et al., 2017; W3C Web Accessibility Initiative, 2019). For example, producing a map in a picture or paper form means the user is forced to use that one modality. It costs between $5 and $15 to transcribe a single braille page of a tactile graphic, which means universities and schools end up not being able to provide an equal experience to blind students because it is too expensive, and blind people are pressured to major in fields with fewer graphics (Amanuensis Braill, 2019; Butler et al., 2017). In education and psychology, there are several contentious theories, such as the theory of multiple intelligences (Gardner, 1993), and the VARK modalities (Fleming, 2014), that claim learners have a particular intelligence they operate best in. A cross-sensory representation allows one to use their “best intelligence” without worrying about missing information. Studies such as Brock et al. (2015) have found that participants did have a strong preference for accessing information through either audio or tactile interfaces. There have been many neurological studies, such as Wolbers, Klatzky, Loomis, Wutte, and Giudice (2011), and Loomis, Klatzky, and Giudice (2013), that have found spatial information is inherently amodal and other studies, such as Dell’Aversana (2017), Brock et al. (2015), Ferris and Sarter (2008), Bălan et al. (2017), Yeung (1980), Jacobs and Yildirim (2012), and Papadopoulos, Barouti, and Koustriava (2018) that found adding multiple sensory modalities was more effective than just using one sensory modality. The advantage of accessibility by default and the possibility of increased effectiveness of the interface for users who can access multiple senses is an extremely strong argument for creating cross-sensory interfaces for data representations.

Chapter 2: The Audio Game Laboratory

INTRODUCTION

Audio games are computer games that can be played completely with an audio interface. They have existed since 1972 and they span the spectrum from being completely speech-based to being almost exclusively based on non-speech sounds (Swarts, 2016). Within the context of this paper, audio games are what we refer to as a “natural laboratory”– a site of investigation where extensive iteration in culture, driven by selection pressures, has refined a set of artifacts and/or conventions to what are likely to be effective states. Elsewhere, this has been referred to as “artifact evolution” (Kirsh, 2006). Here, we focus on audio games because they demonstrate an emergence of interactive parameter mapping sonifications that potentially optimally display geographical information and a large number of simultaneous data variables (Grond & Berger, 2011). Our preliminary investigation of audio games is in response to a call for more research on parameter mapping sonifications, such as from Walker and Mauney, Krygier, and Flowers, Turnage, and Buhman (Flowers, Turnage, & Buhman, 2005; Krygier, 1994; Walker & Mauney, 2010). Commercially available interactive interfaces and audio games – that have been shaped and informally “tested” by the selection pressures of a demanding consumer market – can serve as examples of potentially effective conventions that can inform future work in the auditory display research community. Past research on audio game map interfaces has been limited, only appearing in a few studies that either do not exactly follow existing conventions such as Loeliger and Stockman (2014), or who do not detail the exact conventions they used (Sánchez, Espinoza, Borba Campos, & Merabet, 2013). Research using audio game interfaces has also focused mostly on navigational maps, and not focused on maps for data analytics. Through our preliminary analysis of the 15 audio games presented in this thesis, examples are provided to suggest solutions to some of the more difficult problems in sonification. This exploration of audio game interfaces aims to expand the research presented in chapter 20 of the Sonification Handbook by Brazil and Fernstrom and to encourage greater investigation into the potential value of audio games to the auditory display community (Brazil & Fernstrom, 2011).

BACKGROUND

There has been little research on utilizing audio game interfaces for data representations and mapping. For example, researchers have demonstrated that utilizing some of the techniques in audio games have enabled blind people to improve their sense of orientation and cognitive mapping ability in the physical world (Balan, Moldoveanu, & Moldoveanu, 2015; Sánchez et al., 2013). Zhao, Plaisant, Shneiderman, and Lazar (2008) utilized a few audio game conventions for their data analytics map, but they did not explicitly recognize audio game influences. However, none of the existing research has acknowledged the vast differences in auditory interfaces audio games present. Such discoveries within the audio game community may expose new opportunities, such as strategic uses of text-to-speech labels, that have not been the focus of auditory display researchers. For example, developments in sonification have particularly focused on non-speech sounds (Hermann & Hunt, 2005).

However, for auditory mapping displays, the strategic use of text-to-speech is fundamental. Edler and Lamnert- Siepmann, who have researched the addition of an auditory dimension to cartographic maps, recommend understanding what makes current map types successful in helping to structure auditory interfaces (Edler & Lammert-Siepmann, 2017). In visual maps, there are several purposes for using text when an image will not suffice. Titles, legends, and labels are examples where the text is more useful than an image. Furthermore, in a visual map scenario, a sidebar containing a written overview may enhance the reader's understanding of the map by providing context for the data being represented. Similarly, there will be instances in sonification where text, in addition to nonspeech sounds, may prove useful (Walker et al., 2013).

Call for research and clues from audio games

Researchers of the auditory display community have raised many questions about representing complex data sets through sound that audio games may have an answer for. Here are three examples: Krygier asked how would one best design a sonic legend for a map (Krygier, 1994). Most audio games employ auditory icons – a type of representation that bears an analogical resemblance to the process or activity being represented – to reduce the need for a legend (Balan et al., 2015; Gaver, 1986; Peirce, 1902). An example of this may be found in the audio game Swamp, in which buzzing flies represent corpses containing loot and footsteps mixed with groans represent zombies (Kaldobsky, 2011b). If a map key is required, such as in Adventure at C or The Gate, an ever-accessible tree based menu presents a name of the different items and a keypress, such as enter, will play a sound corresponding to the selected item (VGStorm, 2016, 2017). The second issue identified by Krygier and Flowers, Turnage and Buhman, called for research on the ability for users to identify objects in spatial audio (Flowers et al., 2005; Krygier, 1994). The aforementioned Swamp utilizes 3D audio and requires users to center moving objects in their headphones through spinning around and either shoot, avoid or run over the objects. Users have become incredibly accomplished at this task, being able to quickly navigate through complex and varied terrain while defending against dozens of enemies, all while searching for a single item (Hannibal, 2017). A third question posed by Krygier concerns whether blind people can build up maps while only being able to view a single point at a time (Krygier, 1994). Balan, Moldoveanu, and Moldoveanu, and Sanchez et al., found that the highly immersive and attractive nature of audio games enables blind people to create a spatial representation of their environment that can be assigned to aid performing real-life navigation tasks (Balan et al., 2015; Sánchez et al., 2013). For example, strategy games such as Tactical Battle and Castaways are precisely based around the user being able to know exactly what is happening on all parts of the map at once (Kaldobsky, 2011a; Reed, 2013a). Flowers, Turnage & Buhman, Walker & Nees, and Walker & Mauney, comment that it is difficult to determine appropriate mappings between data and displays for blind audiences for several reasons, including the fact that many sonification studies are conducted with sighted participants and these studies may not account for perceptual differences due to expertise (Flowers et al., 2005; Walker & Mauney, 2010; Walker & Nees, 2011). However, given that the audio game community is mostly blind, all successful audio games have been extensively – though informally – “tested” by blind individuals (Swarts, 2016). Therefore, observing the data-to-display mappings in audio games can suggest generally accepted practices in the blind population. New types of experiments could perhaps be developed and deployed using games within the audio game community, similar to Loeliger and Stockman (2014).

INTERFACES

The authors propose that audio game interfaces fit into five distinct map types: Text based, tree based, grid based, side-scroller, and first-person 3D. Each of these interfaces have completely different ways of conveying the three levels of a digital interface and displaying data about the game world (Adams & Rollings, 2006; Jørgensen, 2013). The three levels are: physical input and output, WIMP (window, icon, menu, and pointer), and the game world itself (Jørgensen, 2013). The output interface of audio games is almost exclusively through headphones (audiogames.net, 2018). Each map type utilizes its own basic conventions, such as the mouse in first-person 3D and typing on a keyboard in text based maps (Balan et al., 2015; The Mud Connector, 2013). The WIMP (window, Icon, Menu, and pointer) interface is present in audio games, but should be changed slightly to replace pointer with message (WIMM) because audio games generally do not utilize a mouse pointer, instead using a mouse to change orientation (Christensson, 2014; Kaldobsky, 2011b; Out of Sight Games, 2019; Zhao et al., 2008). The WIMM elements, which contain information about the game, stats, health, and exact location are mostly presented through speech (Balan et al., 2015). Every aspect of the audio game experience consists of either: auditory icons, earcons, and/or speech (Balan et al., 2015). “The game world is represented as a landscape, topography, objects, and inhabitants” (Jørgensen, 2013). Each of these game world components and their interactions are represented by one or more audio elements which could be combined to represent different levels of realism, such as varying the timbre of footsteps based on the size and ground surface, as seen in Swamp (Friberg & Gärdenfors, 2004; Kaldobsky, 2011b). The most significant attribute of auditory maps is the representation of location, either as a set of coordinates or parent-child relationships within a tree-based map structure. The five auditory maps are taken from an analysis of 15 audio games, existing interface design literature, past studies, and the author’s observations through 11 years of audio game play. During the design process for interactive parameter mapping sonifications, the third step, after identifying the data and the message the interface should convey, is the choice of what interface to use. Almost every data set can be represented with any interface, so it is the purpose of the sonification that dictates what interface should be used.

Text based

Text based games utilize text for both input of commands and output of content (Adams & Rollings, 2006). Users interact with the world by typing commands on a keyboard such as “north” to move north and access system options through typing instructions such as “set exits on” (Adams & Rollings, 2006; The Mud Connector, 2013). Text based games are considered audio games because blind users can access their content through their screen reader which reads the output window and players will often add in custom auditory elements to represent events in the game world (Adams & Rollings, 2006; *VIP MUD*, 2017; Williams, 2016). Game world components are represented by short text descriptions such as “A black iron pot is here” after the room’s description (Rumble, 2018). Rooms have a title attribute, such as “Shrine of St. Wiseheart” (Materia Magica, 2017). Text based games can represent any type of data that can be put into text. They enable a user to zoom in on elements through commands such as “look pot” to get a detailed description and can represent a high number of variables at once, such as having weight, color, material, and size all present in an item’s description (Materia Magica, 2017; Rsl games, 2006). The disadvantages are the high level of linguistic competency required to understand the content (Balan et al., 2015). Interfaces where users need a strong analogical connection to the content do not work well in text based interfaces. The content is conceptually specific, so it will say a large dog, rather than making the sound of a large dog. The command line on computers is an example of a text based interface and is actually the graphical display for Rsl games (2006).

Tree based

Tree based maps are a set of parent child relationships that show hierarchical structure (Weisstein, 2018). The input to tree based maps is generally comprised of four parts: move forward, move back, move up a level, and move down a level. The only analogical relationship is hierarchical, every other representation is conceptually specific. Browser based games and list based games such as Sryth and Crafting Kingdom are organized as tree structures (MetalPop LLC, 2018; Yarrows, 2018). Game world components are represented as nodes, either pages or menu items, and often with accompanying speech and auditory icons that play when the user selects an object. Almost every game utilizes a tree based menu structure to allow the user access to actions such as start, exit, and save (DragonApps, 2017; Driftwood games, 2008; Friberg & Gärdenfors, 2004; Reed, 2013b). One advantage of tree based maps is that users don’t need to remember all the data points in a set; they just need to remember the location of the data they are looking for. Menus allow users to explore data, without needing to remember everything in order to navigate through the interface. However, trees do not represent distances between two nodes very well. They also tend to have a steep learning curve, and as the size of the tree structure increases, browsing through it becomes impractical (Tidwell, 2010). Representations showing groups of data and lists of points are best shown with tree maps. File explorer and most websites are tree based interfaces.

Grid based

Grid-based maps are based on a set of coordinates representing squares placed together in a column–row relationship (Soegaard, 2018). The most common example of a grid interface is the spreadsheet. The user navigates through a rectangular array of tiles or cells, and components are placed in different cells around the rectangular array to represent the map content. The input interface generally consists of the arrow keys to allow moving between cells, and keystroke combinations, such as Ctrl+Up arrow to jump upward through the tiles to the next object (Reed, 2013b). Game messages are frequently accessed through menus and by pressing letter keys on different tiles. Each game world component has a tile that it belongs to and the user can move around the squares and view component locations in relationship to one another. Auditory elements of a tile’s components play when the user enters a new tile. Strategy games (such as Tactical Battle and Castaways), where showing an overview of a map is important, employ grid-based maps (Kaldobsky, 2011a; Reed, 2013a). Zhao et al. found that grid-based maps allow blind users to easily explore unfamiliar geographical data and quickly understand adjacent relationships (Zhao et al., 2008). Grids are ideally suited to allowing the user to get an overview of the data being represented (Tidwell, 2010). Grids, however, do not show irregular shapes; users do not like the number of keystrokes it takes to move through the map; and it takes a while for users to understand everything that the map has to offer (Zhao et al., 2008). Grid interfaces lock the user's orientation and sometimes create rooms by having walls -users cannot pass (Driftwood games, 2008). Grid based interfaces are also called top down interfaces (Adams & Rollings, 2006), but because audio only allows for a first-person perspective, the fundamental defining characteristic is the tile relationships (Driftwood games, 2008; Kaldobsky, 2011a; Reed, 2013a). Many data sets, in comma separated values (CSV) format, can be represented using text in a grid based spreadsheet such as seen in Zhao et al. (2008). However, adding non-linguistic sonic cues and brief speech messages allow multiple variables to be represented in a single cell.

Side-scroller

Side-scroller maps utilize a side view of the map where the user primarily moves left and right and analogical object sounds are represented using stereo panning (Adams & Rollings, 2006; Oren, 2007; Techopedia, 2018). Games such as Adventure at C and The Gate utilize a side-scroller map (VGStorm, 2016, 2017). Users mostly use the arrow keys to move through the map and access conceptually specific status information through menus and letter key presses, such as h for health (Reed, 2013b). Oren performed a study designing a complex side-scroller game with multiple levels and different heights of objects (Oren, 2007). He demonstrated that three variables–height, position, and type–could be effectively represented simultaneously using pitch, stereo panning, and texture respectively. The conclusion was that side-scroller interfaces allow for quick recognition of several variables at once, and complexity above that leads to confusion. The interface is optimal for quick navigation through a high level of analogical detail.

First-person 3D

First-person 3D maps are characterized by the use of 3D audio and are the most studied of the audio game map types (Balan et al., 2015; Friberg & Gärdenfors, 2004; Heuten et al., 2007; Klein & Staadt, 2004; Loeliger & Stockman, 2014; Podkosova, Urbanek, & Kaufmann, 2016; Rober & Masuch, 2005; Sánchez et al., 2013). The input interface to first-person 3D maps span the range of input devices, from “keyboard only” (as in Top Speed 3), to both mouce and keyboard as seen in Swamp and A Hero’s Call, to gyroscope as in cyclepath, to GPS location as presented by Rober and Masuch (DragonApps, 2017; Kaldobsky, 2011b; Out of Sight Games, 2019; Rober & Masuch, 2005; Ruijter, Ruijter, Duvigneau, & Loots, 2015). WIMM elements are both inside and outside the game world. The game world is incredibly detailed and tries to mirror reality (Adams & Rollings, 2006; Podkosova et al., 2016). The analogical game world components are positioned around the user and use short looping sound samples to let the user know where they are (Friberg & Gärdenfors, 2004; Rober & Masuch, 2005). Most first-person interfaces, such as Swamp, allow users to change orientation (Kaldobsky, 2011b). Terrain and topological features are represented by thuds as the user hits impassible obstacles and changes in footstep timbre or tone quality as the user moves over different surfaces (Friberg & Gärdenfors, 2004; Kaldobsky, 2011b; Out of Sight Games, 2019). Short speech messages are often played when the user enters and exits a terrain feature like a road or bridge (Kaldobsky, 2011b). First-person 3D provides the most detailed and analogical interface out of the five map types, allows the user to feel as if they are in the game world, and can allow for many different variables to be represented at once (Adams & Rollings, 2006). This makes first-person 3D perfect for data sets with many variables. The disadvantages of first-person 3D are the easy disorientation of the users and the length it takes to explore an interface.

Mixed Interfaces

This analysis of Audio Game interfaces has highlighted clear conventions for each interface, but some interfaces combine interface types, such as first-person and grid. For example, Swamp is a pure first-person interface, it uses 3D audio with short looping sounds, it allows the player to change orientation, it has different footsteps for terrain, there is a thud when a user hits an obstacle, and there are messages as the user enters and exits terrain features (Kaldobsky, 2011b). Cyclepath, on the other hand, uses 3D audio, but the user is locked facing north and there is only an engine sound of the user’s motorcycle since the road is always the same material (DragonApps, 2017). Cyclepath would be a side scroller that is going forward and back using 3D audio from first-person. Entombed is another game that uses a grid and first-person together (Driftwood games, 2008). In Entombed, the first-person interface features include: footsteps that change as the terrain changes, and messages that play as the user enters new terrain. The Grid interface elements include: locking the user's orientation, only playing sounds present in the current tile (door and hallway sounds are positioned in stereo using hard right, hard left, and center), hallways are one tile wide, and objects are placed in single tiles. This interface, due to the lack of coordinates and tile titles, makes it very easy to get lost, so there is a key command that tells you how to travel to unexplored areas. In general, it is best to stick with the conventional interface styles, unless there is a good reason, like motorcycles can't drive off the road, or the user is in a maze, to use a mixture of interface styles.

CONCLUSION

Audio games provide a large set of potentially effective sonic interfaces that can be used in designing data sonifications. Researchers can shape their hypotheses based off what seems to work in audio games. If one has access to the data in an audio games on how users play the game in practice, it would be possible to perform empirical studies on user tendencies. Currently most sighted users have had no experience with these interfaces and it may require some training before they become proficient. This means that there may be factors other than good design that may lead to a successful interface and there may be interfaces beyond what has been mentioned that may be employed in audio games.

Designing data sonifications using the framework set forth by Jørgensen will allow a more dynamic and intuitive interface and expand the capabilities of existing sonification practices, such as those presented in Brazil & Fernstrom (Brazil & Fernstrom, 2011; Jørgensen, 2013). The use of auditory elements, such as auditory icons, earcons, and speech should be ubiquitous in most sonic interfaces. Future research needs to validate and clarify the systematic application of the five audio game map interfaces to data sonifications. A larger set of audio games also needs to be evaluated to insure validity.

Analyzed Games

Figure 1: Analyzed Games.

Title	Interface Type	Reference
Swamp	First-Person 3D	Kaldobsky (2011b)
A Hero's Call	First-Person 3D	Out of Sight Games (2019)
Top Speed 3	First-Person 3D	Ruijter et al. (2015)
A Blind Legend	First-Person 3D	DOWiNO (2019)
Papa Sangre	First-Person 3D	Somethinelse (2010)
Cyclepath	First-Person 3D and Side-Scroller	DragonApps (2017)

Entombed	Grid-Based and First-Person 3D	Driftwood games (2008)
Castaways	Grid-Based	Kaldobsky (2011a)
Tactical Battle	Grid-Based	Reed (2013a)
Adventure at C	Side-Scroller	VGStorm (2016)
The Gate	Side-Scroller	VGStorm (2017)
Materia Magica	Text-Based	Materia Magica (2017)
World of Ledgends	Text-Based	Rsl games (2006)
Sryth the age of Igtheon	Text-Based and Tree-Based	Yarrows (2018)
Crafting Kingdom	Tree-Based	MetalPop LLC (2018)

Chapter 3: Design and Evaluation of an Audio Game-Inspired Auditory Map Interface

INTRODUCTION

Visual maps have been a part of civilization for many years, but it has only been in the last couple of decades that these visual maps have been turned into digital audio (Friberg & Gärdenfors, 2004; Parente & Bishop, 2003). Despite a number of digital auditory interfaces being presented in the academic literature (Heuten et al., 2007; Loeliger & Stockman, 2014; Parente & Bishop, 2003; Zhao et al., 2008), governments and large mapping companies still do not offer effective nonvisual digital maps commercially, and the Google Maps and ESRI interfaces do not follow auditory display conventions described in the literature (*Cal fire*, 2019; ESRI, 2018; Google, 2019; Logan, 2018). It is difficult to pinpoint why the digital auditory interfaces from the academic literature have not made it into commercial mapping products thus far, but some possible reasons include the need to train users to use an unfamiliar paradigm, an inability to customize the few auditory interfaces that exist, and a limited number of published interface evaluations.

Chapter 2 describes a “natural laboratory” in the form of audio games, games that can be played completely using audio, a domain in which extensive iteration in a commercial market has created a set of effective conventions for auditory digital maps that are already familiar to a community of nonvisual users. The present study examines what happens when experienced Audio Gamers interact with a complex digital map that utilizes familiar Audio Game interface conventions identified in Chapter 2. The hypothesis here is that participants would leverage their implicit knowledge of conventions from audio games and find the proposed interface faster and easier to use than the alternatives introduced thus far in the existing auditory display research literature. The findings of the study did not offer a valid comparison in many cases with other studies due to missing data in other studies or due to the data set used in this study not being the dataset used in other studies. This study did highlight that several audio game conventions, such as scan, allowing the use of a personal screen reader, having multiple interface types, and combining speech with audio, should be employed in future auditory map designs. Audio game interfaces often undergo rigorous beta testing, and users find the interfaces easy and fun enough to use. The evidence of this is their willingness to pay for the game (Adams & Rollings, 2006; audiogames.net, 2018; Friberg & Gärdenfors, 2004; Out of Sight Games, 2019). Biggs et al. (2018) outlined a set of interface conventions present in audio games utilized by the prototype in this study, similar to the audio game *A Hero’s Call* (Out of Sight Games, 2019). The objective of this study was to evaluate reactions and performance of blind participants on a map utilizing audio game conventions.

Definition of digital map

For the purposes of this paper, a digital map is conceptualized as a dynamic representation of items configured in spatial and topological relations with each other, represented in a virtual sensory format. This excludes much of the research on interactive maps that use a combination of digital and non-dynamic and non-refreshable physical displays, such as raised-line paper maps over the top of a touch screen and other examples that can be found in Brock and Jouffrais (2015).

BACKGROUND

Several promising studies report on auditory digital maps that utilize multiple interfaces such as first-person and grid, but the influence of audio game conventions remains limited.

The map presented in Loeliger and Stockman (2014) and Feng, Stockman, Bryan-Kinns, and Al-Thani (2015) is the most promising, given that it is a downloadable Windows application and follows many audio game conventions. Loeliger and Stockman (2014) utilizes a first-person interface and a tree interface, along with a “scan function” to “scan” through points of interest around the player. In the first-person view, looping auditory icons convey the spatial location of points of interest, like the clinking of dishes for restaurants and a fast-moving stream for rivers, that are placed using 3D audio and that change as the user moves around the map. The menus representing different locations one can go to is in a tree interface.

Feng et al. (2015) utilized an automatic orientation adjustment to keep participants on a path. In contrast, most first-person interfaces in audio games do not have an automatic orientation adjustment because users can get extremely disoriented, and this is what the study found. The choice to use earcons rather than footstep sounds also could have contributed to the difficulties they had with distance estimation.

Other studies, such as Parente and Bishop (2003), Heuten et al. (2007), and Milne, Antle, and Riecke (2011), attempted to utilize a first-person interface, but their systems were often considered complex by participants, even though these studies also found that utilizing auditory icons through 3D audio allowed participants to develop a mental map of a location.

Zhao, Plaisant, and Shneiderman (2005) and Zhao et al. (2008) presented *ISonic*, a grid-based interface that allowed users to observe trends in data across different geographical regions by listening to speech and musical sounds while the participant arrowed around a grid of the U.S. The most significant feature they found was that participants loved the ability to switch between viewing a table of regional data and switching to the current region on the map, allowing multiple modes for navigation. Their interface, however, differed significantly from that used in audio games (Reed, 2013a). For example, the participant did not jump a fixed distance when moving around the map; instead they jumped region by region. When a participant pressed the up arrow while on Washington state, they went to Alaska; but when they pressed the down arrow to go back to Washington, they landed in Hawaii instead. Their interface also had a training time of 1.82 hours, which is much longer than the 2.5 minutes it takes to read (with a screen reader) the three-page user guide for the audio game *Tactical Battle* with a grid interface and/or get used to the interface in the tutorial levels (Reed, 2013b).

It is difficult to quantify the effectiveness of many of these interfaces, such as Parente and Bishop (2003), Heuten et al. (2007), and Loeliger and Stockman (2014), because these papers contain limited results that can be used to compare across studies. Customizability for navigation modes, platform preferences, and synthesizer choice remain extremely limited in all the above prototypes.

MATERIAL

Platform

One of the major objectives of the prototype design was to allow participants to use their own computer and screen reader. This was a deliberate choice that was contrary to most studies, which restrict use to a self-voicing feature (provides a voice that talks by default) and single platform (Brock & Jouffrais, 2015; Feng et al., 2015; Loeliger & Stockman, 2014; Walker et al., 2013). The reason for this choice was to allow participants to focus completely on the interface, rather than being required to split their attention by learning an unfamiliar synthesizer, although self-voicing was provided by default. The prototype presented in this study was programmed in Javascript and React (*React*, 2018) to be used in the web browser. Audio was played using the Web Audio API and text to speech was obtained either through triggering the participant’s screen reader through using ARIA live regions, or used the Web Speech API. The prototype only allowed for keyboard access.

Map data

The map data was compiled from a combination of measuring shapes from Google Earth and manual measurements taken at the Magical Bridge Playground in Palo Alto, California (Ofiesh & Poller, 2018). The playground map was based off a rectangle that encompassed an area that was 76 meters wide by 62 meters long.

Interface design

The auditory interface prototype utilized three modes of navigation: a first-person view, a grid view, and a tree view. The grid view and first-person view utilized the same position and step size settings, so there was no disorientation when alternating between views. It was expected that participants would utilize the tree interface to quickly move between objects, the grid interface to get shape information and spatial relationships between objects, and first-person to walk routes between objects. Each interface had a particular specialty and it was expected participants would utilize the most effective interface for each task. It was not possible to complete the tasks with the tree interface, because there was no information on route information, object shapes, or distance. Allowing these tasks to be completed with the tree interface will be work for future iterations of this project. All modes used the same data from the array of objects. The first-person and grid interfaces used data from the participant’s current location to construct their experience.

The first-person interface had a locked orientation with the participant facing the top of the playground. When the participant pressed the arrow keys, the character used footsteps to walk a specified distance every 0.3 seconds. When the participant entered a polygon (i.e., a 2D polygonal region defining an object on the playground), a recorded label would play saying the name of the object. Several of the objects, such as the long ramp, had a material attribute set, such as “wood”. Footsteps of that material would play when the participant walked over the objects.

The grid interface had more speech and auditory feedback. Every time a participant moved to a new square in the grid interface, a spearcon (a short speech message Walker et al. (2013)) would say the name attribute of the polygon followed by the coordinates. The default spearcon was called “Playground Walkway”. Several of the objects had short, less than 0.7 second, auditory icons that would play when the participant entered the square with the polygon. The auditory icons were unique identifying clips from the recordings of the object being used. The spearcon and auditory icon would play together. The default sound was an unobtrusive scuff sound.

The tree interface listed the items all together in the object menu, where the name attribute of the object was read out as a spearcon as the participant moved through the menu (Reed, 2013a; Walker et al., 2013). The object menu was effectively the map key. Pressing Enter on each object brought up a submenu with the options:

1. Go: take the player to the center of the object polygon.
2. Listen: hear the sound associated with the object in isolation from the other sounds.
3. Description: Hear the textual description of the object, if any.
4. Directions: Say where the object was in relationship to the participant's current position and the nearest point. The key "d" would then be set to quickly replay updated directions relative to the player's current location.

The main menu brought up a list of most commands that could be done in the game along with their key shortcut. For example, "Toggle Sounds, t" was the first item. Both the menus were closed by pressing Escape.

METHOD

Structure

The qualitative study comprised two phases: the first was an interview asking participants about their experience with maps and technology, and the second was to show participants a prototype and evaluate their usage and comments on the prototype. The whole study was estimated to take approximately one hour. The studies were all conducted remotely over Skype. Skype was a deliberate choice as it is widely used by the blindness community and allows users to share system audio on Windows. Participants were asked to make sure they had Skype, an updated browser, and headphones.

Study

All the participants were asked to complete eight tasks (listed below), then rate their performance on the NASA Task Load Index (Human Performance Research Group (NASA Ames Research Center), n.d.; NASA, 2018). The NASA Task Load Index is an established method of obtaining a subjective assessment for human-computer interactions and provides a simple numeric score for comparison across multiple tests and interfaces. The eight tasks were chosen to explore the aspects of navigation identified in Heuten, Wichmann, and Boll (2006) and Brock et al. (2015) such as getting an abstract overview of a map, getting an overview of what is around a location, getting routes between locations, and the exact placement of specific locations. Most of the tasks revolved around participants developing and demonstrating route, landmark, and survey knowledge of the map (Brock et al., 2015). Tasks 6 and 7 were used to evaluate if this type of map could be used for scatterplots, heat maps, or other types of representations that require the identification of trends such as those in Zhao et al. (2008). Each task was timed starting from when the participant began to complete the task and finished when they completed the task or when they verbally indicated they were done with the task. All the participants were able to ask for the task to be repeated. The headings in the results section were the text that the interviewer said. If the participant asked for clarification a short description or reiteration of the task was given. For example, “Locate the climbing giraffe” could be described as: “Go to the climbing giraffe in any way you wish”. The clarification was mostly used by the four of the ten English as a second language participants. Participants were not given the definition of each object before starting the task. The eight tasks participants were asked to complete are as follows and are described further in the results section: 1. Locate the climbing giraffe. 2. Describe the route from the stepping sounds to the roller slide. 3. Describe the shape of the KinderBells. 4. What are the objects on both ends of the long ramp? 5. Describe the shape of the long ramp. 6. What is the smallest item on the map? 7. Where is the highest density of items? And 8. Describe the overall layout of the map.

Participants

Ten congenitally blind male participants were recruited from a forum post on audiogames.net. The study was approved through the institutional review board from OCAD University and no compensation was given for the study. The participants ranged from 16 to 43 years old. The participants were from many different countries including India, South Africa, Romania, Canada, United States, and Iran. All the participants had audio game experience and all of them had used a screen reader for at least five years. All but one user used Nonvisual

Desktop Access (NVDA) (2017), and one participant used JAWS for Windows (n.d.). Six participants used Firefox and four used Chrome. None of the participants were familiar with the Magical Bridge playground in Palo Alto. Seven of the participants had no vision, one participant had light perception, and two participants were considered very low-vision, to the point where they used a screen reader to read rather than large print (one participant said their vision was 20/800 and the other did not know). The analysis of results showed no difference in the performance of the different participants, so they were all aggregated together in the results section.

RESULTS

Exploration Phase: Please explore the map and let me know when you feel comfortable with the interface.

During the exploration time, the researcher gave hints of buttons to press to insure every participant explored the entire interface. The main hints were to press t to toggle the sounds, backslash to toggle between text to speech and the screen reader, escape to bring up the main menu, dash and equals to zoom in and out, and to make sure each participant explored grid view and the objects menu. When the participant finished exploring each part of the interface, the researcher prompted: “Let me know when you feel comfortable using this interface, then we can move on to the tasks.” There are three methods that have been explored in the literature for map exploration: Feng et al. (2015) and Brock et al. (2015) gave a time limit of 15 and 10 minutes respectively to explore the interface before starting the tasks. Zhao et al. (2008) had a tutorial that took 1.82 hours on average to complete. The approach in this study was similar to Heuten et al. (2006) that took between 5-10 minutes where they let participants say when they felt comfortable with the interface.

On average, the participants in this study spent 9.87 minutes (SD 6.07) exploring with the fastest being 2.6 minutes and the longest being 19.5 minutes. Five of the participants took less than eight minutes to explore the interface and the other five took more than eleven minutes. It’s important to note that the participant who took the longest to explore the interface went to all 43 objects on the map before saying they were comfortable. The fastest participant quickly moved through all the features. There was no major difference between the performance of the slower explorers and the faster explorers. The Faster explorers accomplished 7/8 of the tasks 3 minutes faster on average than the slower explorers. Finding the climbing giraffe took the faster explorers 1.2 minutes and the slower explorers 0.9 minutes. Future studies should compare the performance of slow explorers when timed on a tutorial vs allowing them to feel comfortable with the interface. This exploration method seems faster than the other methods of exploration. There were 43 objects on this map, 8 objects in Heuten et al. (2006), and 50 objects in Zhao et al. (2008) and the other studies did not indicate the number of objects on their maps.

Task 1: Locate the climbing giraffe.

The climbing giraffe is a giraffe leaning over with its neck horizontally curved covered in handholds and toys for kids to play with. The climbing Giraffe was randomly selected from the list of 16 objects that contained sounds and that was not the “Stepping Sounds” which is the first object participants encounter on the map. Participants were asked this question after they felt comfortable using the interface and had explored all the interface features. This task was to evaluate how a participant would find a specific location/landmark on the map. Finding an object by name is similar to the tasks given in Loeliger and Stockman (2014) where they asked participants: " Your starting point is the fast food restaurant ‘Subway’. Make your way across the railway line to the pub ‘Jamies Wine Bar’ in Fleet Place." Although this set of instructions gave both a starting and ending location as well as some brief description of what was in between, these instructions were not indicative of a user visiting a map for the first time. The expected use case for this map included the user knowing the name of an object and wanting to find that object. This is similar to a participant knowing an address and needing to find the address. This task was also going to be repeated for tasks 2 through 5, so it was critical participants knew how to quickly locate items on the map.

There were three methods participants could have used to complete this task: 1. First, they could have moved around in either grid or first-person view and found the object by hearing the sound or hearing the label announced while exploring the map. One of the 10 participants accomplished the task in first-person view doing this method. It took 2.32 minutes. 2. They could have used the Object Menu to get "directions" and walked to the object using the directions. Six of the 10 participants used this method with their times in minutes being: 1.43, 1.18, 6.83, 0.83, 1.5, and 0.97. The participant who took 6.83 minutes tried finding the object first through exploring, then gave up and used the object menu to get directions. 3. They could have used the "go" option to jump to the object. Three of the 10 participants used this method with their times in minutes being: 0.65, 0.47, and 0.38.

The results of this task were not necessarily predictive of future behavior. Nine of the 10 participants used both the "go" and "directions" option at least once during the study with the sole exception being the participant who only moved in first-person during the study. The average time to find the object was 1.66 minutes (SD 1.91).

Task 2: Describe the route from the stepping sounds to the roller slide.

Stepping sounds are an art installation with a speaker that plays different footstep sounds as users walk in front of a motion sensor. The roller slide is a slide made out of long rotating dowels that spin under the person sliding. This task assessed the ability of users to find a route between two objects. Many map studies use a task to travel between objects as one of the major factors in assessing the effectiveness of a map (Brock et al., 2015; Feng et al., 2015; Heuten et al., 2006; Loeliger & Stockman, 2014). Feng et al. (2015) describes “decision points” participants encountered during the exploration which were basically intersections or turns. This map had no barriers, so intersections were not applicable. Participants did need to choose the method for travel between objects and identify the objects between the start and end of the route. These two objects were chosen because they both had a sound, and they were relatively far apart (from the nearest point they were 39 squares diagonally apart) with most of the objects between. Loeliger and Stockman (2014) had success with blind participants describing routes using “free text”. The theory was that verbal descriptions and free text would yield similar results, but verbal would be faster and give more detail as participants did not need to type every obstacle and turn they made.

There were three methods participants used to find the route between the two objects: 1. Seven of the 10 participants used the “go” option in the menu to get to one of the objects, then used the “directions” option in the menu to get to the other object. The times in minutes it took to complete the task were: 5.8, 5.32, 4.23, 3.07, 2.65, 3.68, and 6.28. 2. Two of the 10 participants used “go” to get to an object and relied on both the scan function and their memory to locate the second object. The times in minutes it took were: 9.78 and 4.6. 3. One of the 10 participants used first-person to navigate between the objects from memory. It took 3.75 minutes for them to walk to the stepping sounds and find the roller slide.

On average it took all the participants 4.92 minutes (SD 5.93) to navigate and describe the route. In Feng et al. (2015) it took participants 16 minutes on average to navigate their route, although there was no number of squares given between the start and end points, so a comparison is difficult to make. They also indicate interruption time separate from navigation time. In this study, participants gave feedback while navigating, so it was not possible to separate navigation from interruption times. Feng et al. (2015) also stated their participants had five types of keyboard error: Orientation errors, Omitting error, Unintentional pressing, Incorrect keystrokes while self-orienting, and Miss-keying. None of these errors occurred with the participants in this study. Three of the 10 participants did get lost during the study, but they were able to complete the task with minimal prompting: One of the three participants was prompted “You can use the menu to navigate” when they verbally expressed they were lost and they were able to “go” to the object and make their way to the other object without further prompting (this was the participant that took 9.78 minutes to complete the task). One of the other participants suggested they thought in routes rather than a map, so this task was very easy.

All of the participants managed to navigate between the objects, but all of the routes were slightly different from one another. Each participant was able to articulate the objects they passed and the route they took. For example (starting from the stepping sounds): " Go up, passed the mini slide, go a few steps up (maybe 5 or 6), then go right. You pass the disk swings and keep going right, you pass a slide, then you're there." (This participant took 4.23 minutes and used “directions” eight times.) This description is very similar to the text descriptions given in Loeliger and Stockman (2014): “Leave Shakespeare’s Globe Theatre and turn right along the river. Walk on until you reach your destination, Pizza Express”. Future studies should evaluate how participants physically navigate between the objects. Three of the 10 participants expressed their route was not realistic because of needing to cross over the ramp which could not be crossed in real life. This interface should also evaluate the same route in Feng et al. (2015), although there is no mention of the start and end points they evaluated on.

Task 3: Describe the shape of the KinderBells.

KinderBells are a set of bells children can bang with a ball to ring them. It is not clear how important shape recognition is in digital maps. Heuten et al. (2006) and Heuten et al. (2007) attempted shape recognition in a 3D auditory landscape, but the “shape of the drawn objects often differs clearly from the real shapes”. This description is also valid for the findings in this study. More focused auditory shape recognition has been investigated in several studies such as Uno, Suzuki, Watanabe, Matsumoto, and Wang (2018), Bermejo, Di Paolo, Hüg, and Arias (2015), and Rice, Jacobson, Golledge, and Jones (2005), and several applications for auditory shape recognition and creation have been developed such as Greve (2009), Sysop-Delco (2017), Sysop-Delco (2015), and Balanced (2017). For this task, participants were asked to verbally describe the shape of an irregular symmetrical shape. Most studies ask participants to draw shapes or ask participants to describe recognizable shapes such as stars or squares (Heuten et al., 2006; Rice et al., 2005). Physically drawing on swell paper was not possible through the remote medium this study employed and utilizing an application such as Sysop-Delco (2015) would have defeated the cross-platform ability of the study.

The grid medium in this modality meant that the descriptions were all tile based. A slant or curve would look like “steps”. The Kinderbells are small, so participants were required to zoom in to the highest level to view the shape. The below squares are at the highest zoom level. The exact description of the Kinderbells set by the researcher was: “A symmetrical 4-step object with 2 tiles on the top and 2 tiles on the bottom with a single tile nob on either end on the second level. Starting from the top, the horizontal width of the levels are 2, 5, 4, 5, 2. The length of each level from the top, going to the right is: 2, 2, 1, 2, and the top level has a single square step going to the left.” None of the participants gave this level of a description. Five of the 10 participants expressed they did not know how to describe the shape. Two of the 10 participants did not want to switch to the grid view which, in this version, was the only way to get the 2D shape. Three of the 10 participants were able to describe a basic shape: “It’s like a sideways rectangle with points on each end. The points are 1 wide… They are off-set…

They are at an angle… It's like a crescent with a thicker end and a thinner end. It curves to the bottom of the map."

What should improve the result is the addition of optional borders to object polygons, so that users are able to stay in a polygon if they wish, rather than needing to exit and reenter the polygon every time they move past the edge. Future work needs to incorporate a better shape description system, either using something like Sysop-Delco (2015), or having participants list the points of the polygon.

Task 4: What are the objects on both ends of the long ramp?

The long ramp is a 44 square long ramp that outlines the bottom right edge of the play area and slants up to the right 13 squares. It has 11 steps and ranges from one to four squares wide. This task tested the ability of participants to follow a path and getting an overview of what is around a location. Feng et al. (2015) had participants follow a route, but it was not a single path. Heuten et al. (2007) has "following paths" as future work that needs to be done.

Seven out of ten participants were able to identify both objects on either end of the long ramp. One participant suggested that along with borders along the edge of the path, earcons of beeps and buzzes representing openings, doors, and objects should be used, similar to those in Out of Sight Games (2019). There were three methods that participants used to accomplish this task: 1. Four out of seven participants followed the ramp landings until they went out of the object, then they checked if the ramp went up or down from their current location until they reached the end of the ramp. They all started by using the “go” option to get to the center of the ramp. 2. One out of seven participants read the description of the long ramp to answer the question. 3. Two of the seven participants remembered objects from past exploration.

Task 5: Describe the shape of the long ramp.

Seven out of 10 participants were able to follow the ramp from start to finish and described the ramp as “steps going up to the right”. The other three out of 10 participants followed the ramp at least 13 squares to the right and five squares up (four out of 11 “steps”).

Task 6: What is the smallest item on the map?

This question was to evaluate the effectiveness of this map in dealing with something like a scatter plot such as in Zhao et al. (2008). Only one out of 10 participants was able to answer this question correctly. This is because he systematically used the “go” option in the Objects Menu on the highest zoom setting and explored the size of objects in grid view. Once he reached the first object that was one square, he stopped and said that object was the smallest. It took him 6.97 minutes. Seven out of 10 participants started doing this task correctly, but gave up around the 13th (out of 43) object. It would have been much more efficient to have a sound mapped to the area of each object and play that sound as participants arrowed through the Object Menu, or had a sorting option for the Object Menu, similar to Zhao et al. (2008). There was no task completion time given in Zhao et al. (2008), and participants were not identifying the size of objects, so it is difficult to compare the two studies, but the above methods would reduce the amount of steps currently required to review size.

Task 7: Where is the highest density of items?

This question was to test how effective the map is at conveying clusters of data points. Nine out of 10 participants found one of the two areas with the highest density of items (average minutes = 1.51, SD = 1.13). Three of those nine participants employed scan to count the number of items that were nearby (Average minutes = 2.46, SD = 0.96), five of the nine participants mentioned that they listened for the highest number of sounds clustered together (average minutes = 1.53, SD = 0.95), and one participant used their past knowledge of the map to identify the highest density of items in 0.02 minutes. Seven of the nine participants expressed uncertainty with their choice " I wouldn’t say if it is the most clustered, but there is a lot going on".

Task 8: Describe the overall layout of the map.

This is the first task sighted users do when viewing a map and it is one of the most important uses of a map (Heuten et al., 2006). Both Heuten et al. (2006) and Heuten et al. (2007) evaluate sketches participants drew after hearing their auditory map. The sketches in Heuten et al. (2006) showed all eight objects properly identified and spatially placed correctly. The sketch method was not possible in this study, so a free verbal description was asked for.

One problem that made itself apparent very quickly was that the participants did not have the vocabulary or chunking skills to systematically describe the map. A common sentiment was: “I don’t know how to put all that into words, how things are located.” Or “I wouldn’t be able to tell you exactly where something is”. This response meant that the participants needed a framework to put their responses. The researcher broke the playground into nine squares: Top right, top middle, top left, middle right, center, middle left, bottom right, bottom middle, and bottom right. The researcher then asked the participant to describe generally what was in each area one section at a time using chunking (Miller, 1956). It was not practical for participants to remember all 43 objects, especially if the chunks were not extremely clear. This meant that accuracy was evaluated on the percentage of objects correct in each chunk. Five out of 10 participants were able to give a 100% accurate overview with all correct objects in each chunk, four of the 10 participants were able to give a pretty accurate overview with only one or two items incorrect, and one participant was unable to describe any overview. When participants were exploring the interface to get an overview, seven participants switched to grid view and held down the keys so they only heard the auditory icons in each tile. When they heard a sound they didn’t know, they would stop, investigate the items, then continue moving as fast as possible to the edge. They performed this action in a grid pattern so they could get what was in each tile.

Several comments were that there needed to be sounds for each object to maximize the effectiveness of this strategy. One participant even turned off his screen reader completely and just used the sounds to get an overview of the playground. The average time in minutes for getting an overview was 6.12 (SD 3.19).

This method of evaluation was not ideal as it was difficult to quantify. Future work needs to explore better methods of getting an overview of large-scale landscapes.

Other Results

5. Participants were asked to rate their comfort level physically navigating between two objects that were on either ends of the map. The mean score was 46 (SD = 30.89) with the min score of 0 and a max score of 90, a median of 35 and a mode of 30. 0 was not at all confident and 100 was very confident. The participant with the highest score admitted that he would need his mobility equipment which included his white cane and Sunu band, a wrist band that uses haptic feedback to alert users of obstacles to their upper body (*Sunu band*, 2019).
6. Eight of the participants used all three interface types to accomplish the tasks and two participants never used the grid interface past the initial exploration stage despite it being the best interface for getting the shape of an object. All the participants also expressed a preference for either grid or first-person for the majority of their navigation. This means that users have a preference for a mode and some will stick with their preference, even if it may not give the information they need. This means it’s important that each interface convey the same level of information, such as object shape, spatial relations, and texture.
7. All the participants elected to use their own screen reader to accomplish the study. It took less than a minute for all the participants to get the prototype running on their machine. Prior testing showed the prototype working perfectly with Macintosh and Windows platforms, both with self-voicing and

screen readers. Parente and Bishop (2003), Loeliger and Stockman (2014), Heuten et al. (2007), and Zhao et al. (2008) all require participants to use the self-voicing feature, rather than use their own screen reader. These results suggest participants prefer the ability to use their own screen reader, like they can do in games such as Out of Sight Games (2019) and Kaldobsky (2011b).

8. Nine out of 10 participants repeatedly used the Object Menu to either "go" to an object or get "directions" to an object. Zhao et al. (2008) presented a function they called a "spreadsheet" interface that listed objects in a list that could be navigated using up and down arrow keys and navigated focus to the selected object when focus was given to the map. Participants were very enthusiastic about this feature in Zhao et al. (2008), and most participants really liked the feature in this interface.
9. All participants made extensive use of the "scan" function. The suggestions were to make instructions more accurate, so rather than saying "far off, behind and to the left", it would say something similar to "4 meters behind and 10 meters to the left". Also, participants really wanted to adjust the distance of the scan function rather than having it locked at 10 meters.
10. The "directions" need to give more constant and accurate feedback. Although directions were extensively used by nine of the 10 participants, the usage pattern was quite excessive. Participants pressed the d key every three seconds when looking for an object. Using beacons similar to Kaldobsky (2011b) and Out of Sight Games (2019) would give a more steady source of the participant's current location relative to the target.

Task Load Index ratings

The overall workload score in all categories for the NASA TLX was an average of 39 (SD = 10.58), 38.3 median, 24 min, and 53.5 max. The NASA Task Load Index is a method of obtaining a subjective score for mental load when completing a task. Scores can be used as a baseline when evaluating future work on the same or similar projects (Agency for Healthcare Research and Quality, 2019; Meshkati, Hancock, Rahimi, & Dawes, 1995). Participants were asked to rate their experience in six subscales on a scale of 0-100, where 0 was as little as possible and 100 was as much as possible. The subscales and their scores are: mental demand: mean = 55.1 (SD = 20.58), median = 47.5, Mode = 40, min = 30, and max = 90. Physical demand: mean = 5.5 (SD = 7.52), median= 4, mode = 5, min = 0, and max = 25. Temporal demand: mean = 38.5 (SD = 19.59), median = 40, mode = 60, min = 10, and max = 60. Performance: 58.1 (SD = 21.39), median = 57.5, mode = 30, min = 30, and max = 92. Effort: mean = 50 (SD = 31.62), median = 42.5, mode = 25, min = 0, and max = 100. Frustration level: mean = 27.5 (SD = 22.88), median = 20, mode = 50, min = 0, and max = 60. Other auditory map interfaces have not been evaluated for mental task load.

Feedback on the prototype

Participants were asked their general thoughts on the prototype. Three participants said they "really liked it" and five said they liked it or thought it was cool because of the familiar interface, ability to get a detailed overview, and sounds. The users who were more moderate in their feedback said it was interesting, but of limited use, and they didn't think they could do anything with it. In general, participants said they found the controls intuitive and very easy because of their resemblance to audio games. All the participants liked the idea of allowing the user to dictate their mode of navigation, either through grid view or first-person, similar to Out of Sight Games (2019). Each participant was asked why they used each mode of navigation: tree, grid, or first-person. Their responses are summarized as follows:

- Tree was used for quick navigation through the map.
- Grid view was used to quickly navigate and get an overview of the map.
- First-person allowed users to “relate” to the space.

The final question asked users for any final thoughts they had about the prototype. Six of the participants reiterated that they wanted to see a map like this made for more locations: “It was quite fun. If this was released, I would be so happy and use it on a daily basis.” Another participant wanted first-person to match the exact navigation system (with ability to change orientation and earcons for surrounding items) as Out of Sight Games (2019).

CONCLUSION

The prototype in this study utilized common audio game conventions which outperformed many other studies, and created a set of baseline results future studies can use to compare against. There were three major findings from the tasks: the interface was extremely easy to learn and navigate, participants all had unique navigational styles, and participants needed user interface features that made it easier to understand and answer questions about spatial properties and relationships. Future studies need to figure out a more effective way of evaluating the shapes blind users recognize and create a better method for giving a general overview of the map.

Chapter 4: Design and Evaluation of an Interactive 3D Map

Introduction

Diagrams and maps have been used since cave paintings to represent ideas and display information about the world. When what is now Perkins School for the Blind began creating materials for blind students to become literate, a book of raised line tactile graphics (tactile graphics) and an atlas of the United States were some of the first materials to be created, even before the invention of braille (Weimer, 2017). Recently, researchers began comparing the use of “interactive” digital labels with braille labels on tactile graphics and have found a greater preference for and extensibility with the digital labels (Brock & Jouffrais, 2015). Braille labels are often embossed on clear labels with an adhesive back that are stuck to a surface. Digital labels are described later in the paper, but in brief, they are labels that utilize some kind of electronic output such as speech, LCD display, or braille. Holloway et al. (2018) recently published a study comparing tactile graphics with 3D tactile models and found a significant preference for the 3D models. The study presented in this paper followed both of these later findings and employed a 3D map combined with interactive digital audio labels to evaluate the reactions and goals of blind participants when navigating using the map. The second part of this paper presents the results of a co-design that investigated what a blind participant’s ideal navigational experience might be. Both projects were led by a blind designer who is the author of this thesis.

3D Model Maps and Graphics

The current practice for producing tactile graphics is by creating a raised line drawing, however, recent research suggests that 3D tactile models are more effective. The braille Authority of North America published a set of guidelines for creating tactile graphics and these are the standards followed by braille transcribers and teachers of the blind and visually impaired when creating tactile graphics (BANA & CBA, 2011). In 2003, 27 blind tactile graphic readers and tactile graphic transcribers were surveyed, revealing a number of difficulties with tactile graphics, with the main problem being their lack of availability (Rowell & Ongar, 2003; Rowell & Ungar, 2003). Other problems included: users not understanding the graphics, a lack of standardization for symbols, low tactile graphic experience and literacy among the readers, most readers requesting more information about how to read the graphic (so recordings or braille descriptions by the transcriber need to be given with the graphic), and a need for simpler graphics due to the limited resolution of tactile graphic production methods and the possible lower capacity of haptic perception (which unfortunately, reduces the amount of information that can be presented via raised line graphics) (Rowell & Ongar, 2003; Rowell & Ungar, 2003). Several recent studies, such as Jafri and Ali (2015), Siu (2014), Agarwal, Jeeawoody, and Yamane (2014), and Shi et al. (2017a), found 3D models, mostly created through 3D printing, do not pose as many difficulties for readers relative to raised line tactile graphics. Agarwal et al. (2014) and Holloway et al. (2018) both propose that, for blind readers, textures can serve as tactile correlates for colors. These authors also suggest that, because blind readers touch objects in three dimensions in their daily lives, three-dimensional models that resemble items that blind readers might have touched in the world would be much easier to recognize relative to drawings that utilize visual conventions such as perspective presented in flattened 2D scenes via raised line graphics. Holloway et al. (2018), Gual, Puyuelo, and Lloveras (2014), and Hasper et al. (2015) directly compared 3D models to tactile graphics and found that 3D models were preferred, were easier to remember, were able to convey more information, and allowed for easier to understand icons. These findings suggest that 3D models should be used whenever possible, especially for users with little or no experience with tactile graphics.

Interactive 3D Maps and Graphics

The current practice for labeling 3D tactile maps and graphics is to use braille either through sticky labels or by Brailling on the graphic itself (BANA & CBA, 2011), however, interactive graphics with digital labels have been found to be just as effective, faster, preferred by blind users, and more functional relative to braille labels alone (Brock et al., 2015). Adding braille labels can become difficult when graphics or models require long labels or a large number of labels. If there are few enough labels to allow for abbreviations, labels that say something similar to “DS” for Diamond Street are used on the graphic and a separate set of pages is used as a key (BANA & CBA, 2011). Needing to frequently switch between the graphic and key is disruptive to the map reading process and limited information can be stored on a key. Studies, such as Brock and Jouffrais (2015) and Götzelmann (2016), present raised line tactile maps that allow the participant to query a computer using gestures as they touch the map. For example, Brock and Jouffrais (2015) and Brock et al. (2015) have users double-tap the place they wish to hear the label of and the computer speaks the label. The Talking Tactile Tablet is a commercial product that facilitates the creation and use of interactive tactile graphics (Landau & Gourgey, 2001). Brock and Jouffrais (2015) ran a study with 24 blind participants that directly compared braille-labeled raised line tactile maps with interactive raised line tactile maps, finding that learning time was significantly reduced when using interactive maps, participants preferred the interactive maps over the braille-labeled maps, and there was no difference in the recall or comprehension between the braille maps and Interactive maps. None of these studies evaluated a system for both braille and audio labels. The small number of users who preferred braille labels over speech labels were expert braille readers who expressed a dislike for speech labels relative to braille labels. One of the major advantages to interactive maps is that they do not require knowledge of braille to be effective (Brock & Jouffrais, 2015). Only around 10% of the blind people in the United States are braille readers. In other words, braille labels are not accessible to 90% of the blind people in the United

States due to limited braille materials and lack of instruction in the blindness population (National Federation of the Blind, 2009). Combining a braille display with a speech-based interactive map would provide access to both tactile and auditory learners. The only groups who would be excluded from this braille-and-speech setup would be severely deafblind users who do not know braille, such as users with Usher's syndrome who become deafblind later in life. These users would not be able to read the braille or understand the speech. There is a display that represents text using ASL fingerspelling rather than braille that would make this display more usable to deafblind users (Fang, Dixon, & Wong, 2012). Shi et al. (2017a) and Holloway et al. (2018) evaluated blind participants using interactive 3D models and observed a greater preference for interactive models than braille labeled models. This was probably because of how difficult it was to attach braille labels to 3D models.

Adding Interactivity to Tactile Graphics and 3D Models

Interactivity has been added to tactile graphics and or tactile 3D models in multiple ways including: buttons, capacitive sensors, touchscreens, hand recognition, computer vision with markers, and QR codes. Holloway et al. (2018) utilized twelve capacitive touch points on their 3D models. When the user touched a point with their finger, the label played. They found that users preferred having the sensor in a location where they could not touch it accidentally while exploring. Participants did like having different information depending on the number of taps or length of a hold on the point. Götzelmann

(2016) had a 3D model that was no more than 1 mm high over the top of a touchscreen. This allowed for gesture recognition and detailed finger tracking, but it did require a small flat model that could lie on a touchscreen. Zhang, Laput, and Harrison (2017) created a 3D model that was covered in a capacitive surface that used a machine learning algorithm to detect finger gestures, effectively making any surface into a touchscreen. O'Sullivan, Picinali, Gerino, and Cawthorne (2015) used hand recognition to detect the participant's gestures, although they found the technology was too unstable for the precision needed to label areas on the tactile graphic. Coughlan and Miele (2017), and Shi et al. (2017b) used computer vision to detect the user's fingertip (through capturing the position of a colorful sticker on the user's fingernail) which allowed accurate labeling of both 3D models and tactile graphics. Baker et al. (2016) used QR codes read by an iPhone to manage the labels. They found participants had 80% accuracy and took around 30 seconds to identify labels. They eventually concluded that a hands-free device that didn't constantly require the user to point the camera would be better and implemented a system very similar to Coughlan and Miele (2017). The PenFriend is a commercial product that has a camera on the end of a pen-shaped device that reads custom-printed stickers with uniquely-bar-coded labels and there is no trouble with focusing the device. An important limitation of PenFriend, however, is that it's only able to play audio recordings uploaded or recorded to the device and only works with the uniquely-bar-coded labels from the company.

Approach

3D Map

The 3D map that is the focus of this paper was of the Magical Bridge Playground in Palo Alto, California (Ofiesh & Poller, 2018). This playground was created to be highly accessible to all visitors and to “remove the physical and social barriers of today’s typical playgrounds and give everyone a place to play.” The 3D map (see Fig. 1) was a scale model that consisted of a cardboard-covered foam board mounted with models made out of laser-cut wood and clay. The map measured 76 cm by 62 cm; one cm on the model represents one meter on the real playground. Several of the models, such as the swings, were fully functional, so participants could explore exactly how the object worked if they wished. Several of the models, such as the slide mound, were made out of clay. One object, which resembled a large marble, was actually represented by a marble to show both the shape and texture of the object. The use of 3D printing was not investigated in this study because of the difficulty the blind designer had with creating CAD models using OpenSCAD. Future iterations will use 3D modeling.

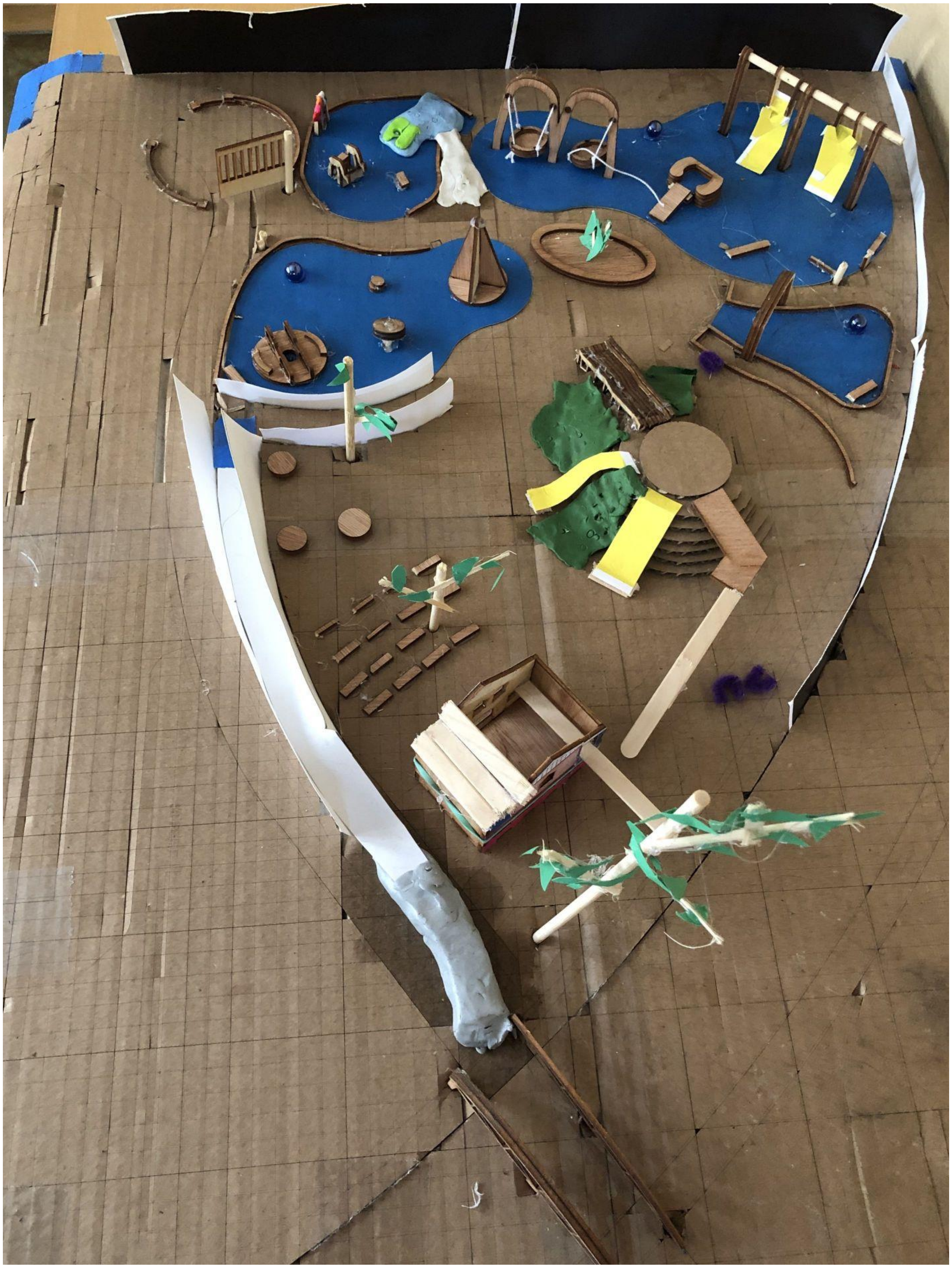

Figure 1. Photo of 3D map, which is a scale model of the Magical Bridge Playground.

Interactive System

This prototype utilized CamIO, a computer vision augmented reality system for annotating physical objects (Coughlan & Miele, 2017). CamIO (short for “Camera Input-Output”) is a computer vision based system to make physical objects (such as documents, maps, devices and 3D models) accessible to blind and visually impaired persons, by providing real-time audio feedback in response to the location on an object that the user is pointing to. The latest alpha version of CamIO uses a laptop computer connected to a webcam, which is fixed on a tripod with the webcam capturing the entire object of interest. (An alternate version of CamIO under development runs as a self-contained iOS app using the built-in iOS device camera.) Unlike an earlier version of CamIO that directly tracked the location of the user’s pointing finger (Coughlan & Miele, 2017), the new alpha version estimates the 3D location of the tip of a stylus that the user holds to select points of interest on the object. A pre-defined annotation file specifies the 3D coordinates of various hotspots, i.e., points of interest on the object, along with the corresponding audio label that is issued when the stylus enters the hotspot. The 3D coordinates are rigidly attached to the object, with the X and Y axes defining the horizontal plane that the object lies on and the Z axis defining the vertical direction. For the 3D playground map, the CamIO webcam is mounted rigidly relative to the map, so a simple calibration procedure allows the camera-centered coordinates to be converted to the map-centered coordinates that the annotations are expressed in; this obviates the need to track the 3D pose of the playground map over time. Five items were required to run the system, a laptop running Python, a Logitech Webcam HD Pro C920, a piece of paper with a barcode to mark the ground plane on which the map is placed, a tripod that was at least four feet above the map, and a stylus. The stylus (see Fig. 2) consisted of three parts joined together: two foam square cubes that were two inches to a side with barcode markers glued to the face, and a clay grip that was about an inch and a half thick that tapered to a point (the whole stylus was about seven inches long).

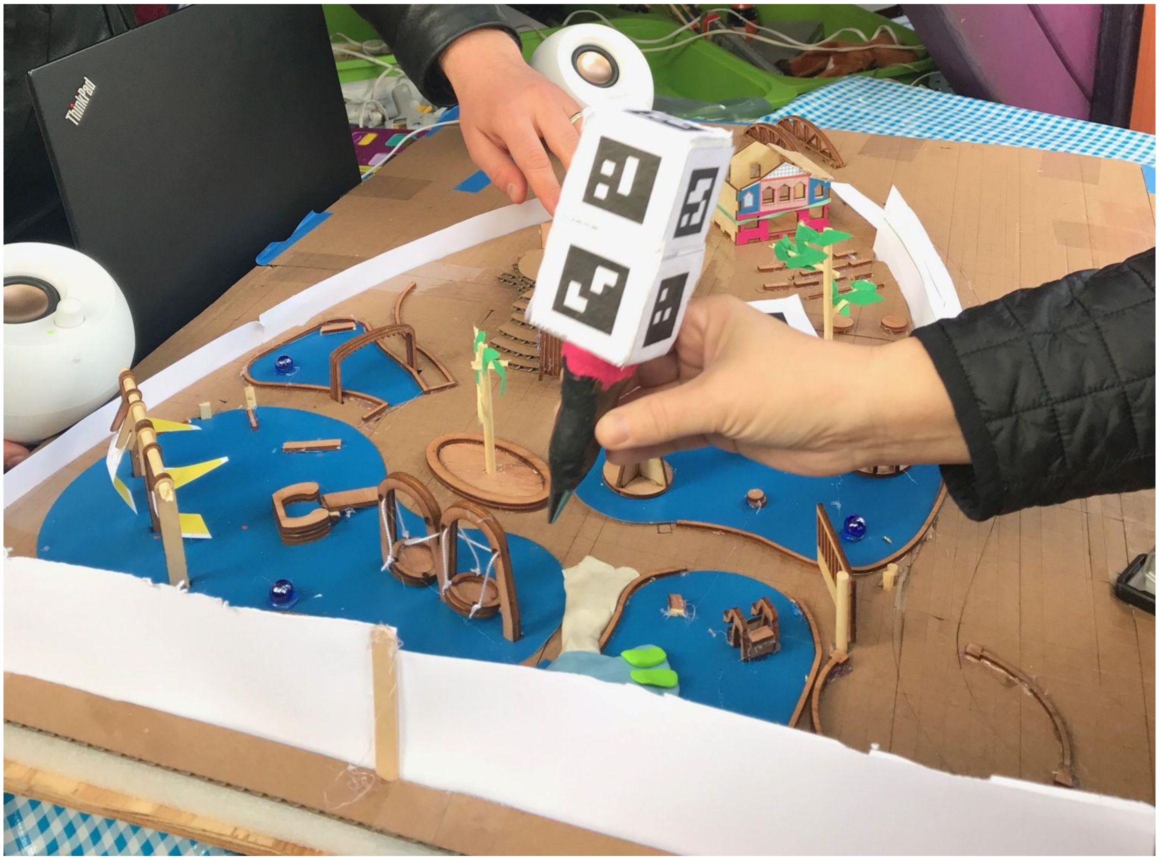

Figure 2. Photograph of participant holding stylus to point to features on the 3D map.

The CamIO interface was modified slightly to allow a spatial 3D soundscape and data file of label polygons (i.e., the 2D polygonal regions defining the objects on the playground). The audio soundscape consisted of looping sounds of the different playground objects, for example, a recording of someone going down a metal slide to represent a slide. The X and Y horizontal coordinates of the stylus tip were used to position the sound listener in relationship to the looping sounds. This meant that as the stylus moved around the map, one could hear the sounds of the objects that were around the stylus, as if the participant were actually walking around a busy playground. The digital grid was based on the 76 cm by 62 cm 3D model map and each polygon was the planar projection of the object as seen from above. A Z coordinate specifying the height of the object on the physical map was also given. Once the stylus entered the polygon, a speech label would play, saying the name of the object.

Participants

Seven blind and low-vision participants were recruited from both the volunteer community from Magical Bridge, the Vista Center, and the Silicon Council of the Blind. The participants ranged from 9 to 79 years old. There were four female and three male participants. None of the participants were extremely familiar with the layout of the playground, although five of the participants had been to Magical Bridge before. Three of the participants were low-vision and four were completely blind.

Results with Tactile Prototype

Finding Objects on the Map

Participants were first presented with the stylus and shown the map. They were then told to touch the tip of the stylus to an object on the map to hear the name of the object. Then they were asked to find different objects on the map, such as a swing or slide. The average time to find an object was 1.17 minutes. The participants were asked to find three objects, and the average times were 1.45

minutes for the first object, 1.5 minutes for the second object and 0.58 minutes for the third object. The longest it took someone to find an object was 4.75 minutes. This was because the blind participant circled the map several times, missing the object which was on the perimeter. Typically, participants used one hand to feel objects and the other hand to hold the stylus and touch the stylus to objects after the first hand felt them. None of the participants explored the prototype without the stylus and all the participants made extensive use of the labels, touching a new label every three or four seconds. One of the participants wanted something that could trace the route on the map from his current location to the object he was looking for.

The NASA Task Load Index (TLX) was used in obtaining a quantitative subjective assessment of the mental load on each participant (Human Performance Research Group (NASA Ames Research Center), n.d.; NASA, 2018). The NASA TLX is a method of obtaining a subjective score for mental load when completing a task. Scores can be used as a baseline when evaluating future work on the same or similar projects (Agency for Healthcare Research and Quality, 2019; Human Performance Research Group (NASA Ames Research Center), n.d.; Meshkati et al., 1995). Participants were asked to rate their experience in the six subscales of mental demand, physical demand, temporal demand, performance, effort, and frustration level with the tasks on a scale of 0-100, where 0 was as little as possible and 100 was as much as possible. To obtain an overall task load score, the six subscales were averaged together. The

NASA TLX was used to evaluate how difficult each participant found it to complete tasks with the map. For finding objects on the map, only one person had the overall average task load score slightly above 50. The average overall task load score for all seven participants together was 38.94 with a mode of 40.17, a max of 54.17 and a min of 24.17. One of the participants wanted the play zones of the playground more clearly marked so they could chunk the objects together more easily. They found the terminology such as “Sway Fun” difficult and wanted some way to figure out what zone the stylus was in at any time. Another participant found it very easy because they were able to match the incomplete mental model they had of the playground with what they were seeing on the prototype. The two participants with an overall average task load score less than 26 explicitly mentioned that the map was fun and repeatedly described using the map as easy. Two of the other participants expressed frustration that they didn’t systematically memorize the map at the beginning of their exploration. Two of the participants felt they spent a long time finding objects, one thought finding objects was fine, and the other did not like wandering around the map.

Physically Navigating to Objects

The next part of the study explored how participants might physically navigate to an object in the actual playground after feeling the corresponding route to it on the map. Three participants felt comfortable physically navigating to an object in the playground, two gave feedback on why they were not yet comfortable, one said no, and the other didn’t think they could navigate around the playground very well in their wheelchair. Comments from the participants who declined to navigate centered around needing more time and assistance with memorizing the route: “having something that would tell me what direction the bucket swing is from my location here” and “It would take me a good part of the afternoon because I’m still so new at this.”

Of the three participants who navigated to the object, two were low-vision and one was completely blind, although everyone touched the map to memorize the route. It took the low-vision participants 0.15 minutes and 1.42 minutes to memorize the route and the blind participant 3.45 minutes to memorize the route. What took the longest time was participants needing to find the object on the map they were finding the route for. Once they found the object, then two of them traced with their hand the route they were planning to walk between the objects.

Finding the object took the two low-vision participants 0.75 minutes and 2.48 minutes and the blind participant 5.7 minutes. During the walk, all the participants queried the researcher about features in their environment. One of the low-vision participants asked if a concrete strip was a path and they were told that no, it was a curb with plants in it. The blind participant asked for descriptions and names of objects as she hit them with her cane. She also asked for a scan of the area to find out the objects that were around at that point in time. Two of the participants, one low-vision and the blind participant, got lost where the path needed to go around a large planter. The blind participant asked what direction the object was from them as they circumnavigated the planter. The fastest navigator commented that the textures of the ground were extremely helpful in finding their way. The overall average task load score of all three participants for navigating to the object was 37.5 with a range of 3.34.

Navigational Goals When Entering the Playground

Four participants said they wanted a sighted person to take them around the playground initially to get them oriented. One participant wanted to show their sighted granddaughter around the playground and point out objects she could interact with. Two of the participants were worried about hurting other kids who were not paying attention while walking through the playground. The two participants with dog guides were worried about keeping their dog safe both from running children and while they were interacting with the objects on the playground. Two participants wanted to know about ground textures their wheeled devices could traverse.

It was reiterated that a map would be part of the orientation process. One participant said after the playground was memorized, the map would no longer be useful.

Feedback on the Prototype

Liked

Five of the participants were asked what they liked, what needed to be improved, and what they didn’t like about the prototype. All of the participants loved the 3D models and the fact the texture of the models often matched the textures of the real objects. “I like the 3 dimensionality, from some of the objects, I could tell what they were without needing any labels… It was really cool that the texture on the map matched the texture in real life… I was able to picture the objects around me as I was walking through them and that was very helpful.” Another participant mentioned they really liked the audio sounds along with the tactile models: “I was seeing things spatially, tactilely, and auditorily”.

What Could Be Improved

In general, the participants expressed they wanted more details with the map. Three participants explicitly mentioned they wanted more textures and pathways. When navigating, one of the participants found the textures extremely important for finding their way and they wanted to see those textures reflected on the map. Another participant had a walker that could only go on particular surfaces, so he needed a way to plan out what surfaces he could go on. Two of the participants got stuck, when navigating, on a feature that was not clear on the map. On the map it looked like one could walk through an area that was not walkable, so the detail needed to be as high as possible so participants could plan effectively. One of the participants detailed a system that would show routes between objects. He wanted a miniature robotic person that he could touch to walk from point A to point B, taking the fastest route.

Did not Work

When asked what they didn’t like about the prototype, three of the participants couldn’t think of anything. The other two participants expressed they wanted the audio labels to be more accurate, only triggering when they touched the object, rather than triggering a centimeter away from the object. One of the participants did not realize there was an object a centimeter away from her finger when she touched the stylus to the map. They also wanted one label to play at a time and they found the map to be too general with not enough details like paths and textures.

Co-Designing an Ideal Navigational Solution

Introduction

17 blind and low-vision participants were asked their experiences and preferences for maps, then they were asked to design an ideal navigational solution without worrying about feasibility or possibility. The participants were recruited from the Silicon Council of the Blind, Vista Center for the Blind, Magical Bridge volunteers, and a post on audiogames.net. The participants ranged from 9-89 years old with a mean age of 35.12, and were from countries including India, South Africa, Romania, Canada, United States, and Iran. The impetus for this co-design stemmed from the second dimension of inclusive design: “Use inclusive, open and transparent processes, and co-design with people who have a diversity of perspectives, including people that can’t use or have difficulty using

the current designs.” (Treviranus, 2018c). Even though the head designer on this project was blind and had their own opinions on the topic, every participant was different and brought a different set of perspectives to the table.

Experience with Maps

Before starting the study, all participants were asked what they thought of when they heard the word “map”. Responses ranged from “something that helps you navigate places” to raised line and 3D maps to “Something that gives you directions to go somewhere… a GPS”. Three of the participants explicitly expressed the word “map” brought up negative experiences: “Not pleasant, not informative to visually impaired people, not something I want anything to do with, not helpful, unless it is a mental map where I’ve walked a location. Maps in general have not been anything to me except as a deterrent to getting anything done.” “I think of something 2-dimensional that I looked at and said ‘wait a minute, how is this adding up to the street I just walked on’”. Eight of the participants thought of raised line 3D model maps. The broadest interpretation of map, that encompasses all the responses is: “something that helps you navigate places”. The next question asked participants when had they wished for a map, or wished for a map that would have been more useful to them. Two of the participants had “never wished for a map”. Five of the participants wanted maps of their neighborhood, four participants wished they had maps of enclosed areas like a mall, restroom, or school campus, three participants wanted a map to understand geography and the world, and one participant wanted a map right then of the area around where they were living.

Desired Information from a Map

The next set of questions had the participants outline what they would like to learn from maps. One participant summarized all the responses nicely: “The little explicit details is what I want”. The “explicit” details outlined by the participants included: size of the sidewalk, pillars, cars parked on the sidewalk, what side of the street something was on, the little streets and driveways without names, rooms and who works in that room, seats, restrooms, exits, benches, picnic tables, extremely accurate turn by turn directions from one location to another, ground textures and terrain, and plants. Ten of the participants wanted to bring the map around with them and have it track their location as they moved combined with an extremely accurate GPS system. They wanted the map to tell them exactly when to turn, tell what objects were around at any point in time, identify the object they were hitting with their cane, and have the digital map show the route they needed to travel.

Co-Design

Participants were asked: without worrying about feasibility or possibility, what would make their navigational experience amazing. Asking participants what their dream solution was allowed the participants to express what they wanted to see in the future, rather than having them evaluate a design. Eleven of the participants described a turn by turn navigation system combined with a digital map of some kind. The turn by turn system would be “something that spoke, that said ‘entering the park, turn to your right for such and such.’” “A person. They would guide me around, I just hang onto their arm, tell them where I want to go, and they take me there.” “I would want an implant that would show me what was going on around me, while I was walking, and something not super distracting. I would want it to be super accurate. Super up to date.” “I would love a map that would track me both indoor and outdoor… My ideal system would also work offline.” The researcher asked why something like Aira, an application that has trained operators who can remotely see through a pair of glasses worn by the user and give the user information and directions, would not work (*Aira*, 2018). One participant who used Aira explained: “I use Aira now, but to have the independence, that would be great. I love Aira, but Aira is very unaffordable for many people, you need to go out with glasses, so for most people it is not feasible. With a digital map, it wouldn’t matter if you were in the U.S. or UK, I could still use it, whereas Aira only works in the U.S.” Another user mentioned that something like Aira made them feel as if they were under time pressure to accomplish their task and they wanted to be able to take their time and explore. In general, a very accurate turn by turn navigation system that allowed querying of the environment was requested. Out of the ten participants who described a map, three major ideas emerged: (1) Four users wanted a full VR experience where they could hear their environment in 3D audio, feel objects and textures, and use their cane. (2) Three participants described a refreshable tactile hologram that could expand to about an 11 inch square that would have a miniature of the location with them in the center. (3) Two participants wanted an implant in their brain that would allow them to just know where to go, in an accurate, up-to-date way. In addition, one participant wanted a set of buttons they could press as they moved from location to location that would verbally tell what was around and how to get to different areas. All the described maps were dynamic and could change as new information was made available. Most of the ideal maps included redundant information through multiple senses. Three of the participants could not read braille, and five of the participants wanted the ability to read braille on the map, so presenting labels in multiple modalities was critical.

Conclusion

This study investigated the effectiveness of an interactive map made out of 3D models and presented the results of a co-design to create the ideal navigational experience for blind participants. In general, participants liked the interactive 3D map and how they were able to connect the textures and shapes of the models with their life-sized counterparts. There were two problems with the map: First, it took participants an average of 1.17 minutes to find an object on the map. Compared to other studies, such as Baker et al. (2016), the time to find an object should be closer to 0.5 minutes. Second, there was not enough detail for pathways and ground textures for participants with wheeled devices to effectively plan their routes, and the map only showed large objects without indicating where it was possible to walk. During the co-design, most participants described their ideal navigational solution to be a highly accurate turn by turn navigation system combined with a dynamic 3D audio and 3D model map (also known as in-navigation and pre-navigation systems (O'Sullivan et al., 2015)). All the participants agreed more detail was better. Rather than reducing detail, as one needs to do with tactile graphics BANA and CBA (2011), 3D models and audio allow for high definition tactile representations that bear an ecological resemblance to the environment. Future work will focus on developing a more accurate 3D map that better represents the playground in detail, adding functionality to the interactive 3D map to actively guide users to objects and along paths of interest, and implementing an accurate turn by turn navigation tool (e.g., similar to Fusco and Coughlan (2018)) to help users navigate the playground. Finally, our study highlights the need for long-term research into technologies such as dynamic VR and holographic tactile displays Soviak et al. (2016) that enable truly seamless and effortless navigation experiences.

Chapter 5: Conclusion and Future Work

The results presented in this thesis provide a direction for future nonvisual map displays. Using conventions from Audio Games, it is possible to create a completely digital and highly effective nonvisual map. Reactions to the interactive 3D tactile map were focused around needing more detail and information. Both these studies present a foundation for the map portion of the conclusions of the co-design, which was a highly accurate turn by turn navigation system combined with a dynamic audio and tactile 3D map. Future research needs to focus on developing technology for these two areas.

Accurate Turn by Turn Navigation System

Accurate turn by turn navigation systems are being investigated using a variety of different localization technologies, but the computer vision approaches seem to be the most promising, due to their minimal requirements (i.e., little or no new infrastructure) and ability to function indoors (Fusco & Coughlan, 2018; Sato et al., 2017; Serrão et al., 2015). Future research needs to focus on combining indoor and outdoor navigation, allowing users to query information about what is around them, provide users the ability to enter information into a geographic information system, and continue to increase accuracy.

Map Displays

The research on maps needs two areas of focus: creating high definition digital dynamic 3D tactile displays, and improving and integrating the digital auditory display with geographic information systems.

Digital Tactile Map Display

The three types of tactile displays suggested in the co-design were a tactile hologram, VR touch displays, and neural haptic transmission. For the hologram, research needs to focus on advancing ElectroRheological and MagnetoRheological substances, which are substances that can rapidly change shape through minimal electrical stimulation (Chouvardas, Miliou, & Hatalis, 2005). A display needs to be made with enough detail that textures can be represented. The virtual reality tactile displays are most likely going to be a glove of some kind that utilize multiple tactile display techniques listed in (Chouvardas et al., 2005). Currently force feedback and vibration have been developed with gloves (Soviak, 2015), but the fidelity has barely become good enough to represent braille, which is extremely important for a VR map. The material of the glove will need to be a substance that can quickly change texture as the user moves. Neural tactile transmission is the most promising of all the approaches, because it requires the fewest number of external devices and allows a full range of motion (Slopsema et al., 2018). In nerve stimulation, electrodes are placed on large nerves in the arm or leg and send electrical waves to the nerve. The current application is in simulating touch for amputees (Graczyk et al., 2016; Slopsema et al., 2018). Prototypes such as (Slopsema et al., 2018) have demonstrated participants can feel a range of sensations with this strategy. Future work should focus on creating a 2-way broadband multichannel interface to decode nerve impulses from both the brain to the nerve and from the nerve back to the brain, as well as controlling muscular response for force feedback simulation. Bergmeister et al. (2017) present how it is possible to decode messages from the brain to the nerve; now the micro electrodes need to be turned in the other direction and convey impulses from the nerve to the brain. Sensory transmission has already been done with visual and auditory stimuli with the Argus II Retinal Prosthesis System (Humayun et al., 2012), and cochlear implants, so tactile stimulation should be feasible.

Digital Auditory Map Display

The auditory display needs to have further testing done on shape recognition, getting an overview of an area, have a mobile interface developed, and integrations with geographic information systems (GIS) need to be created.

Several ideas were presented during the study for getting the shape of objects: The first idea was to have optional borders for objects. Each object was a polygon and participants wanted the ability to turn on borders around that polygon, so they could not exit the object's perimeter. The second idea was to create unique footstep sounds for each object. That way one could walk around the object and identify the borders by the changing footstep sounds. The third option was to create earcons defining the borders that were around using spatial audio and multiple timbres. If there was an edge to the left, a buzz would play in the left speaker, letting the user know there is a wall there. If a wall was behind the user, a muffled buzz could play in both speakers indicating there is an edge behind. If there are two edges, then two sounds could play at once, similar to the symbolic set of navigational queues in Out of Sight Games (2019).

Participants found getting an overview very difficult, but there were a few ideas to make this process easier. Currently, using the existing grid view, it would take five minutes to view the whole map. One participant suggested adding a bird's eye view, where sounds are closer together and the user can hear the relative positions of all the sounds together. Heuten et al. (2007) attempted this technique and showed it was not the most effective, especially if objects were close together or had similar sounds. Another option could be providing a gist, similar to Zhao et al. (2008), where a keypress triggers items playing in sequence using a spatial sweep from left to right and top to bottom. This could take around 30 seconds if all items were played, but chunking sounds together into six to nine chunks theoretically should be easier to remember (Miller, 1956). This particular playground actually contained around seven play zones that lend themselves to natural chunking (Ofiesh & Poller, 2018).

Most of the participants wanted a mobile interface and several examples were given for how a mobile interface could be developed. Most of the participants wanted the sound listener to be attached to their finger. They wanted to move their finger around the mobile screen and hear objects around their finger. Other navigational interfaces included the first-person interfaces in DOWiNO (2019) and Somethinelse (2010), which both approach navigation differently. For grid view and the menus, participants thought swiping would work.

The next step in making this digital auditory interface more applicable to users is to connect it with geographic information systems, such as Google Maps, ESRI, or OpenStreetMap. Now that research has been done to establish a set of digital auditory map interface conventions, the GIS application programming interfaces (APIs) need to be connected to the display. There are several difficulties that need to be overcome to allow this interface to connect to a GIS system: First, most map objects in APIs are based off visual coordinates rather than geographical coordinates. Some translation needs to be done to either: make the visual grid match the auditory grid, or convert the user’s movement to feet and meters, taking into consideration how latitude is shorter around the poles than around the equator. Lag also needs to be removed. In the study, a frequent method of navigation was to hold down the keys and move as fast as possible to another location. If there is a lot of space, large jumps to the next land mass should also be available. Quickly scanning what is around was also a frequently used feature. If the system is communicating with the server every keypress, these common modes of navigation will be much too slow and users will not want to use the map. Probably the best method of reducing lag is to download as much of the polygon information as possible to the user’s device. GIS APIs have many other functionalities, and these also should be added into the interface, such as route planning and restaurant review information. It is also important that nonvisual users get the ability to add new information to maps so they can become full productive members of the cartographic community. Crowdsourcing map data for blind travelers has been successfully used in apps such as BlindWays, and in more general contexts such as with OpenStreetMaps. Future mapping solutions need to allow crowdsourcing detailed information, such as ground texture, pole shapes, and driveway information from both the blind and sighted communities to allow both extended information about a location, similar to BlindWays, and better localization from machine learning localization algorithms such as from Fusco and Coughlan (2018).

Creating an Ideal Nonvisual Navigational Experience

The focus needs to be on combining the turn by turn navigation solution with a detailed cross-sensory map. It would be very useful if maps of spaces, such as the Art Gallery of Ontario, University Campuses, and Disney Land could have this full navigational solution available for their nonvisual users. A process needs to be developed to quickly and easily obtain polygon and 3D model data for each space. Then once this data is obtained, features such as the polygon's name, entrances and exits, opening and closing times, and room numbers need to be added. This experience also needs to have a graphical component, so sighted users can participate as well. With this detailed information of buildings and locations, any user will be able to use the digital map to find rooms they are looking for, robots will be able to deliver goods between locations using this set of tools, and virtual participation can become extremely realistic for both virtual and in-person participants.

References

Adams, E., & Rollings, A. (2006). *Fundamentals of game design*. Prentice-Hall.

Agarwal, A., Jeeawoody, S., & Yamane, M. (2014). *3D-printed teaching aids for students with visual impairments*. Retrieved from http://diagramcenter.org/wp-content/uploads/2014/06/ E110-Final-Report-Team-Walrus.docx

Agency for Healthcare Research and Quality. (2019). *NASA task load index*. Retrieved from https://healthit.ahrq.gov/health-it-tools-and-resources/evaluation-resources/workflow-assessment-health-it-toolkit/all-workflow-tools/nasa-task-load-index

Aira. (2018). Retrieved from https://aira.io/

Amanuensis Braill. (2019). *Amanuensis braill - transcription*. Retrieved from https://amanuensisbraille.com/transcription_services.php

audiogames.net. (2018). *AudioGames, your resource for audiogames, games for the blind, games for the visually impaired!* Retrieved from http://audiogames.net/

Baker, C. M., Milne, L. R., Drapeau, R., Scofield, J., Bennett, C. L., & Ladner, R. E. (2016). Tactile graphics with a voice. *ACM Transactions on Accessible Computing (TACCESS)*, *8*(1), 3.

Balan, O., Moldoveanu, A., & Moldoveanu, F. (2015). Navigational audio games: An effective approach toward improving spatial contextual learning for blind people. *International Journal on Disability and Human Development*, *14*(2), 109–118. https://doi.org/10.1515/ijdhd-2014-0018

Balanced, S. (2017). *New smarter balanced desmos calculators are free and accessible to students*. Retrieved from http://www.smarterbalanced.org/new-smarter-balanced-desmos-calculators-free-fully-accessible/

BANA, & CBA. (2011). *Guidelines and standards for tactile graphics*. Retrieved from http://www.brailleauthority.org/tg/web-manual/index.html

Bălan, O., Moldoveanu, A., Moldoveanu, F., Nagy, H., Wersényi, G., & Unnórsson, R. (2017). Improving the audio game-playing performances of people with visual impairments through multimodal training. *Journal of Visual Impairment & Blindness (Online)*, *111*(2), 148.

Bergmeister, K. D., Vujaklija, I., Muceli, S., Sturma, A., Hruby, L. A., Prahm, C., … others. (2017). Broadband prosthetic interfaces: Combining nerve transfers and implantable multichannel emg technology to decode spinal motor neuron activity. *Frontiers in Neuroscience*, *11*, 421. Retrieved from https://www.frontiersin.org/articles/10.3389/fnins.2017.00421/full

Bermejo, F., Di Paolo, E. A., Hüg, M. X., & Arias, C. (2015). Sensorimotor strategies for recognizing geometrical shapes: A comparative study with different sensory substitution devices. *Frontiers in Psychology*, *6*, 679. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00679

Biggs, B., Yusim, L., & Coppin, P. (2018). *The audio game laboratory: Building maps from games*. Retrieved from http://icad2018.icad.org/wp-content/uploads/2018/06/ICAD2018_paper_51.pdf

Brazil, E., & Fernstrom, M. (2011). Chapter 20 navigation of data. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The sonification handbook*. Retrieved from http://sonification.de/handbook/download/TheSonificationHandbook-chapter20.pdf

Brock, A. M., & Jouffrais, C. (2015). Interactive audio-tactile maps for visually impaired people. *ACM SIGACCESS Accessibility and Computing*, (113), 3–12.

Brock, A. M., Truillet, P., Oriola, B., Picard, D., & Jouffrais, C. (2015). Interactivity improves usability of geographic maps for visually impaired people. *Human–Computer Interaction*, *30*(2), 156–194. https://doi.org/10.1080/07370024.2014.924412

Butler, M., Holloway, L., Marriott, K., & Goncu, C. (2017). Understanding the graphical challenges faced by vision-impaired students in australian universities. *Higher Education Research & Development*, *36*(1), 59–72.

Cal fire. (2019). Retrieved from http://www.fire.ca.gov/general/firemaps

Charlton, J. I. (1998). *Nothing about us without us: Disability oppression and empowerment*. Univ of California Press.

Chouvardas, V. G., Miliou, A. N., & Hatalis, M. K. (2005). Tactile displays: A short overview and recent developments. *Proceedings of the 5th international conference on technology and automation (icta-05*.

Christensson, P. (2014). *WIMP definition*. Retrieved from https://techterms.com/definition/wimp

Coughlan, J. M., & Miele, J. (2017). Evaluating author and user experience for an audio-haptic system for annotation of physical models. *Proceedings of the 19th international acm sigaccess conference on computers and accessibility*, 369–370. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5714613/

Dell'Aversana, P. (2017). *Neurobiological background of exploration geosciences: New methods for data analysis based on cognitive criteria*. Academic Press.

Donovan, R. (2018). *Unleash different: The economics of inclusion*. ECW Press.

DOWiNO. (2019). *A blind legend*. Retrieved from http://www.ablindlegend.com/en/about-us/

DragonApps. (2017). *Cyclepath*. Retrieved from https://www.iamtalon.me/cyclepath/

Driftwood games. (2008). *Entombed*. Retrieved from http://www.blind-games.com/

Edler, D., & Lammert-Siepmann, N. (2017). Approaching the acoustic dimension in cartographic theory and practice. *Meta-Carto-Semiotics*, *3*(1), 1–15. Retrieved from https://www.researchgate.net/publication/267386172_Approaching_the_Acoustic_Dimension_in_Cartographic_Theory_and_Practice_An_Example_of_Mapping_Estonian

ESRI. (2018). *A11y-map*. Retrieved from https://github.com/Esri/a11y-map

Fang, B., Dixon, C., & Wong, T. (2012). *Robotic fingerspelling hand for deaf-blind communication*. Retrieved from https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1138&context=mesp

Feng, F., Stockman, T., Bryan-Kinns, N., & Al-Thani, D. (2015). An investigation into the comprehension of map information presented in audio. *Proceedings of the XVI International Conference on Human Computer Interaction*, 29. ACM.

Ferris, T. K., & Sarter, N. B. (2008). Cross-modal links among vision, audition, and touch in complex environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(1), 17–26. https://doi.org/10.1518/001872008X250566

Fleming, N. D. (2014). The vark modalities. *Vark-Learn*, *14*. Retrieved from https://web.archive.org/web/20150314235648/http://vark-learn.com/introduction-to-vark/the-vark-modalities/

Flowers, J., Turnage, K., & Buhman, D. (2005). Desktop data sonification: Comments on flowers et al., icad 1996. *ACM Transactions on Applied Perception (TAP)*, *2*(4), 473–476. https://doi.org/10.1145/1101530.1101545

Freedom Scientific. (n.d.). *JAWS, job access with speech*. Retrieved from http://www.freedomscientific.com/products/software/jaws/

Friberg, J., & Gärdenfors, D. (2004). Audio games: New perspectives on game audio. *Proceedings of the 2004 acm sigchi international conference on advances in computer entertainment technology*, 148–154. https://doi.org/10.1145/1067343.1067361

Fusco, G., & Coughlan, J. M. (2018). Indoor localization using computer vision and visual-inertial odometry. *International conference on computers helping people with special needs*, 86–93. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-94274-2_13

Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. New York, NY: BasicBooks.

Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, *2*(2), 167–177. https://doi.org/10.1207/s15327051hci0202_3

Google. (2019). *Accessibility in Google Maps*. Retrieved from https://support.google.com/maps/answer/6396990?co=GENIE.Platform%3DDesktop&hl=en

Götzelmann, T. (2016). LucentMaps: 3D printed audiovisual tactile maps for blind and visually impaired people. *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, 81–90. ACM.

Graczyk, E. L., Schiefer, M. A., Saal, H. P., Delhaye, B. P., Bensmaia, S. J., & Tyler, D. J. (2016). The neural basis of perceived intensity in natural and artificial touch. *Science Translational Medicine*, *8*(362), 362ra142–362ra142. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5713478/

Greve, D. (2009). *Blind paint*. Retrieved from http://www.tichnut.de/blindpaint/

Grond, F., & Berger, J. (2011). *The sonification handbook* (Hermann, Hunt, Neuhoff, Ed.). Retrieved from https://sonification.de/handbook/chapters/chapter15/

Gual, J., Puyuelo, M., & Lloveras, J. (2014). Three-dimensional tactile symbols produced by 3D printing: Improving the process of memorizing a tactile map key. *British Journal of Visual Impairment*, *32*(3), 263–278. https://doi.org/10.1177/0264619614540291

Hannibal. (2017). *Pattern for the strip mall*. Retrieved from http://www.blackscreengaming.com/swamp/warehouses/strip_mall/pattern/index.php

Hasper, E., Windhorst, R. A., Hedgpeth, T., Van Tuyl, L., Gonzales, A., Martinez, B., … Baluch, D. P. (2015). Methods for creating and evaluating 3D tactile images to teach stem courses to the visually impaired. *Journal of College Science Teaching*, *44*(6), 92–99.

Hermann, T., & Hunt, A. (2005). Guest editors' introduction: An introduction to interactive sonification. *IEEE Multimedia*, *12*(2), 20–24. Retrieved from http://eprints.whiterose.ac.uk/655/

Heuten, W., Henze, N., & Boll, S. (2007). Interactive exploration of city maps with auditory torches. *CHI'07 extended abstracts on human factors in computing systems*, 1959–1964. ACM.

Heuten, W., Wichmann, D., & Boll, S. (2006). Interactive 3D sonification for the exploration of city maps. *Proceedings of the 4th Nordic conference on Human-Computer Interaction: Changing roles*, 155–164. ACM.

Holloway, L., Marriott, K., & Butler, M. (2018). Accessible maps for the blind: Comparing 3D printed models with tactile graphics. *Proceedings of the 2018 chi conference on human factors in computing systems*, 198. ACM.

Holmes, K. (2018). *Mismatch: How inclusion shapes design*. MIT Press.

Human Performance Research Group (NASA Ames Research Center). (n.d.). *Task load index (nasa-tlx) v 1.0*. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021488.pdf

Humayun, M. S., Dorn, J. D., Da Cruz, L., Dagnelie, G., Sahel, J.-A., Stanga, P. E., … others. (2012). Interim results from the international trial of second sight's visual prosthesis. *Ophthalmology*, *119*(4), 779–788.

inclusive Design Research Centre. (2018). *What do we mean by inclusive design?* Retrieved from https://idrc.ocadu.ca/resources/idrc-online/49-articles-and-papers/443-whatisinclusivedesign

Inclusive Design Research Centre community. (2016). *One-size-fits-one*. Retrieved from https://guide.inclusivedesign.ca/insights/OneSizeFitsOne.html

Inclusive Design Research Centre community. (2018). *Virtuous tornado*. Retrieved from https://guide.inclusivedesign.ca/activities/VirtuousTornado.html

Jacobs, R. .., & Yildirim, I. (2012). A rational analysis of the acquisition of multisensory representations. *Cognitive Science*, *36*, 2.

Jacobson, R. D. (1999). *Exploring geographies of blindness: Learning, reading and communicating in geographic space.* (PhD thesis). Queen's University Belfast.

Jafri, R., & Ali, S. A. (2015). Utilizing 3D printing to assist the blind. *Proceedings of the 2015 international conference on health informatics and medical systems* *(hims 2015), july*, 27–30. Retrieved from https://pdfs.semanticscholar.org/33c6/8017f12eb2c4a30dc775dd96a1cffa7fad9b.pdf

Jones, L. A., & Sarter, N. B. (2008). Tactile displays: Guidance for their design and application. *Human Factors*, *50*(1), 90–111.

Jørgensen, K. (2013). *Gameworld interfaces*. https://doi.org/10.7551/mitpress/9780262026864

Kaldobsky, J. (2011a). *Castaways*. Retrieved from https://www.kaldobsky.com/ssl/audiogames.php

Kaldobsky, J. (2011b). *Swamp*. Retrieved from http://www.kaldobsky.com/audiogames/

Kirsh, D. (2006). *Explaining artifact evolution*. Retrieved from https://philpapers.org/archive/KIREAE.pdf

Klein, E., & Staadt, O. (2004). *Sonification of three-dimensional vector fields*. Retrieved from http://graphics.cs.ucdavis.edu/~staadt/download/p_Kle04.pdf

Krygier, J. B. (1994). Sound and geographic visualization. In *Modern cartography series* (pp. 149–166). Retrieved from https://makingmaps.net/2008/03/25/making-maps-with-sound/

Landau, S., & Gourgey, K. (2001). Development of a talking tactile tablet. *Information Technology and Disabilities*, *7*(2).

Loeliger, E., & Stockman, T. (2014). Wayfinding without visual cues: Evaluation of an interactive audio map system. *Interacting with Computers*, *26*(5), 403–416.

Logan, T. (2018). *Accessible maps on the web*. Retrieved from https://equalentry.com/accessible-maps-on-the-web/

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In *Multisensory imagery* (pp. 131–155). Retrieved from https://digitalcommons.library.umaine.edu/cgi/viewcontent.cgi?article=1013&context=sie_facpub

Masters of Scale. (2017). *Handcrafted*. Retrieved from https://mastersofscale.com/brian-chesky-handcrafted/

Materia Magica. (2017). *Materia magica*. Retrieved from https://www.materiamagica.com/

Meshkati, N., Hancock, P. A., Rahimi, M., & Dawes, S. M. (1995). *Techniques in mental workload assessment.* Retrieved from https://www.researchgate.net/profile/Peter_Hancock2/publication/232471724_Techniques_in_mental_workload_assessment/links/0c960532a1f2b630b0000000/Techniques-in-mental-workload-assessment.pdf?_sg%5B0%5D=q0nEQprB-SenPDVLF-1lKmAMgEncODDnUi2QU_FYKABkndaRNntiVniNcSFwIY7hGOONzl_82A8CgYeoq4SQjw.SS0mBhkUN4p

MetalPop LLC. (2018). *Crafting kingdom*. Retrieved from http://metalpopgames.com/forum/index.php?/forum/11-crafting-kingdom/

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81. Retrieved from http://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf

Milne, A. P., Antle, A. N., & Riecke, B. E. (2011). Tangible and body-based interaction with auditory maps. *CHI'11 extended abstracts on human factors in computing systems*, 2329–2334. ACM.

NASA. (2018). *TLX*. Retrieved from https://humansystems.arc.nasa.gov/groups/TLX/

National Federation of the Blind. (2009). *The braille literacy crisis in america*. Retrieved from https://nfb.org/images/nfb/documents/pdf/braille_literacy_report_web.pdf

NV Access. (2017). *NVDA 2017.4 user guide*. Retrieved from https://www.nvaccess.org/files/nvda/documentation/userGuide.html

Ofiesh, N., & Poller, L. (2018). *A playground for the entire community: The design of magical bridge playground*. Retrieved from https://slc.stanford.edu/sites/default/files/a_playground_for_the_entire_community_final_submitted_to_mb_and_pw_0.pdf

Oren, M. A. (2007). Speed sonic across the span: Building a platform audio game. *CHI'07 extended abstracts on human factors in computing systems*, 2231–2236. https://doi.org/10.1145/1240866.1240985

O'Sullivan, L., Picinali, L., Gerino, A., & Cawthorne, D. (2015). A prototype audio-tactile map system with an advanced auditory display. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, *7*(4), 53–75.

Out of Sight Games. (2019). *A hero's call*. Retrieved from https://outofsightgames.com/a-heros-call/

Paladugu, D. A., Wang, Z., & Li, B. (2010). On presenting audio-tactile maps to visually impaired users for getting directions. *CHI'10 extended abstracts on human factors in computing systems*, 3955–3960. ACM.

Papadopoulos, K., Barouti, M., & Koustriava, E. (2018). Differences in spatial knowledge of individuals with blindness when using audiotactile maps, using tactile maps, and walking. *Exceptional Children*, *84*(3), 330–343.

Parente, P., & Bishop, G. (2003). BATS: The blind audio tactile mapping system. *Proceedings of the ACM Southeast Regional Conference*, 132–137. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.6189&rep=rep1&type=pdf

Peirce, C. S. (1902). *Logic as semiotic: The theory of signs*. Retrieved from http://theory.theasintheas.org/wp-content/uploads/2013/02/Peirce-C-S-Logic-Semiotic.pdf

Podkosova, I., Urbanek, M., & Kaufmann, H. (2016). A hybrid sound model for 3D audio games with real walking. *Proceedings of the 29th international conference on computer animation and social agents*, 189–192. https://doi.org/10.1145/2915926.2915948

React: A JavaScript library for building user interfaces. (2018). Retrieved from https://reactjs.org/

Reed, I. (2013a). *Tactical battle*. Retrieved from https://blindgamers.com/Home/IanReedsGames

Reed, I. (2013b). *User Guide for Tactical Battle*. Retrieved from https://blindgamers.com/Home/IanReedsGames

Rice, M., Jacobson, R. D., Golledge, R. G., & Jones, D. (2005). Design considerations for haptic and auditory map interfaces. *Cartography and Geographic Information Science*, *32*(4), 381–391. Retrieved from https://www.researchgate.net/publication/228948257_Design_Considerations_for_Haptic_and_Auditory_Map_Interfaces

Rober, N., & Masuch, M. (2005). *Leaving the screen new perspectives in audio-only gaming*. https://doi.org/http://hdl.handle.net/1853/50168

Rowell, J., & Ongar, S. (2003). The world of touch: An international survey of tactile maps. Part 2: Design. *British Journal of Visual Impairment*, *21*(3), 105–110.

Rowell, J., & Ungar, S. (2003). The world of touch: An international survey of tactile maps. Part 1: Production. *British Journal of Visual Impairment*, *21*(3), 98–104.

Rsl games. (2006). *World of ledgends*. Retrieved from https://audiogames.net/db.php?id=World+of+Legends

Ruijter, L. D., Ruijter, P. D., Duvigneau, B., & Loots, D. (2015). *Top speed 3*. Retrieved from http://www.playinginthedark.net/topspeed3_e.php

Rumble. (2018). *The builder's tutorial*. Retrieved from https://www.tbamud.com/getting-started-menu

Sato, D., Oh, U., Naito, K., Takagi, H., Kitani, K., & Asakawa, C. (2017). Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. *Proceedings of the 19th international acm sigaccess conference on computers and accessibility*, 270–279. ACM.

Sánchez, J., Espinoza, M., Borba Campos, M. de, & Merabet, L. B. (2013). Enhancing orientation and mobility skills in learners who are blind through video gaming. *Proceedings of the 9th acm conference on creativity & cognition*, 353–356. https://doi.org/10.1145/2466627.2466673

Serrão, M., Shahrabadi, S., Moreno, M., José, J., Rodrigues, J. I., Rodrigues, J. M., & Buf, J. H. du. (2015). Computer vision and gis for the navigation of blind persons in buildings. *Universal Access in the Information Society*, *14*(1), 67–80.

Shi, L., Zhao, Y., & Azenkot, S. (2017a). Designing interactions for 3D printed models with blind people. *Proceedings of the 19th international acm sigaccess conference on computers and accessibility*, 200–209. Retrieved from https://www.researchgate.net/profile/Lei_Shi125/publication/320542748_Designing_Interactions_for_3D_Printed_Models_with_Blind_People/links/5a1ee1abaca272cbfbc09ea7/Designing-Interactions-for-3D-Printed-Models-with-Blind-People.pdf?origin=publication_detail

Shi, L., Zhao, Y., & Azenkot, S. (2017b). Markit and talkit: A low-barrier toolkit to augment 3D printed models with audio annotations. *Proceedings of the 30th annual acm symposium on user interface software and technology*, 493–506. Retrieved from https://www.researchgate.net/profile/Lei_Shi125/publication/320571164_Markit_and_Talkit_A_Low-Barrier_Toolkit_to_Augment_3D_Printed_Models_with_Audio_Annotations/links/

59fcb353a6fdcca1f2965909/Markit-and-Talkit-A-Low-Barrier-Toolkit-to-Augment-3D-Printed-Models-with-Audio-Annotations.pdf?_sg%5B0%5D=

Siu, Y. T. (2014). *3D printing for accessible materials in schools - final report*. Retrieved from http://diagramcenter.org/wp-content/uploads/2014/06/3D_FinalReport_SIU_3.docx

Slopsema, J. P., Boss, J. M., Heyboer, L. A., Tobias, C. M., Draggoo, B. P., Finn, K. E., … Polasek, K. H. (2018). Natural sensations evoked in distal extremities using surface electrical stimulation. *The Open Biomedical Engineering Journal*, *12*, 1. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5842387/

Soegaard, M. (2018). *The grid system: Building a solid design layout*. Retrieved from https://www.interaction-design.org/literature/article/the-grid-system-building-a-solid-design-layout

Somethinelse. (2010). *Papa sangre*. Retrieved from https://audiogames.net/db.php?id=papasangre

Soviak, A. (2015). Haptic gloves for audio-tactile web accessibility. *Proceedings of the 12th Web for All Conference*, 40. ACM.

Soviak, A., Borodin, A., Ashok, V., Borodin, Y., Puzis, Y., & Ramakrishnan, I. (2016). Tactile accessibility: Does anyone need a haptic glove? *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, 101–109. ACM.

Sunu band. (2019). Retrieved from https://www.sunu.io/en/index.html

Swarts, J. (2016). *A look at audio games, video games for the blind*. Retrieved from https://www.inverse.com/article/13331-a-look-at-audio-games-video-games-for-the-blind

Sysop-Delco. (2015). *BrushTone*. Retrieved from https://sysop-delco.itch.io/brushtone

Sysop-Delco. (2017). *AudiMesh3D*. Retrieved from https://sysop-delco.itch.io/audimesh3d

Techopedia. (2018). *What is a side scroller?* Retrieved from https://www.techopedia.com/definition/27153/side-scroller

The Mud Connector. (2013). *Getting started: Welcome to your first mud adventure*. Retrieved from http://www.mudconnect.com/mud_intro.html

Tidwell, J. (2010). *Designing interfaces: Patterns for effective interaction design* (2nd ed.). O’Reilly Media, Inc.

Treviranus, J. (2018a). *The three dimensions of inclusive design: Part one*. Retrieved from https://medium.com/fwd50/the-three-dimensions-of-inclusive-design-part-one-103cad1ffdc2

Treviranus, J. (2018b). *The three dimensions of inclusive design: Part three*. Retrieved from https://medium.com/@jutta.trevira/the-three-dimensions-of-inclusive-design-part-three-b6585c737f40

Treviranus, J. (2018c). *The three dimensions of inclusive design: Part two*. Retrieved from https://medium.com/@jutta.trevira/the-three-dimensions-of-inclusive-design-part-two-7cacd12b79f1

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CN: Graphics Press.

Uno, S., Suzuki, Y., Watanabe, T., Matsumoto, M., & Wang, Y. (2018). *Sound-based image and positon recognition system: SIPReS*. Georgia Institute of Technology.

VGStorm. (2016). *Adventure at C*. Retrieved from http://www.vgstorm.com/aac/about.php

VGStorm. (2017). *The Gate*. Retrieved from http://www.vgstorm.com/the_gate.php

VIP MUD. (2017). Retrieved from https://audiogames.net/db.php?id=VIP+MUD

W3C Web Accessibility Initiative. (2019). *The business case for digital accessibility*. Retrieved from https://www.w3.org/WAI/business-case/

Walker, B. N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D. K., Dingler, T., & Jeon, M. (2013). Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *55*(1), 157–182. https://doi.org/10.1177/0018720812450587

Walker, B. N., & Mauney, L. M. (2010). Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing (TACCESS)*, *2*(3), 1–16. https://doi.org/10.1145/1714458.1714459

Walker, B. N., & Nees, M. A. (2011). Chapter 2: Theory of sonification. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The sonification handbook*. Retrieved from http://sonification.de/handbook/download/TheSonificationHandbook-chapter2.pdf

Walker, P. (2016). Cross-sensory correspondences: A theoretical framework and their relevance to music. *Psychomusicology*, *26*(2), 103–116. https://doi.org/10.1037/pmu0000130

Weimer, D. (2017). To touch a sighted world: Tactile maps in the early nineteenth century. *Winterthur Portfolio*, *51*(2/3), 135–158. Retrieved from https://ocadu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=asu&AN=126713290&site=ehost-live

Weisstein, E. W. (2018). *Tree*. Retrieved from http://mathworld.wolfram.com/Tree.html

Williams, I. (2016). *How blind players made a text-only rpg more accessible*. Retrieved from https://waypoint.vice.com/en_us/article/kwz77n/how-blind-players-made-a-text-only-rpg-more-accessible

Wolbers, T., Klatzky, R. L., Loomis, J. M., Wutte, M. G., & Giudice, N. A. (2011). Modality-independent coding of spatial layout in the human brain. *Current Biology*, *21*(11), 984–989.

Yarrows, M. H. (2018). *Sryth the age of igtheon*. Retrieved from http://www.sryth.com/

Yau, N. (2011). *Visualize this: The flowingdata guide to design, visualization, and statistics*. John Wiley & Sons.

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

Zhang, Y., Laput, G., & Harrison, C. (2017). Electrick: Low-cost touch sensing using electric field tomography. *Proceedings of the 2017 chi conference on human factors in computing systems*, 1–14. Retrieved from http://chrisharrison.net/projects/electrick/electrick.pdf

Zhao, H., Plaisant, C., & Shneiderman, B. (2005). ISonic: Interactive sonification for non-visual data exploration. *Proceedings of the 7th international acm sigaccess conference on computers and accessibility*, 194–195. https://doi.org/10.1145/1090785.1090826

Zhao, H., Plaisant, C., Shneiderman, B., & Lazar, J. (2008). Data sonification for users with visual impairment: A case study with georeferenced data. *ACM*

Transactions on Computer-Human Interaction (TOCHI), *15*(1), 1–28.

https://doi.org/10.1145/1352782.1352786