DESIGN AND EVALUATION OF AN AUDIO GAME-INSPIRED AUDITORY MAP INTERFACE

Brandon Biggs
OCAD University,
Toronto, ON, M5T 1W1, Canada
3164451@student.ocadu.ca

James M. Coughlan
Smith-Kettlewell Inst.,
San Francisco, CA,
94115, United States
coughlan@ski.org

Peter Coppin
OCAD University,
Toronto, ON, M5T 1W1, Canada
pcoppin@faculty.ocadu.ca

ABSTRACT

This study evaluated a web-based auditory map prototype built utilizing conventions found in audio games and presents findings from a set of tasks participants performed with the prototype. The prototype allowed participants to use their own computer and screen reader, contrary to most studies, which restrict use to a single platform and a self-voicing feature (providing a voice that talks by default). There were three major findings from the tasks: the interface was extremely easy to learn and navigate, participants all had unique navigational styles and preferred using their own screen reader, and participants needed user interface features that made it easier to understand and answer questions about spatial properties and relationships. Participants gave an average task load score of 39 from the NASA Task Load Index and gave a confidence level of 46/100 for actually using the prototype to physically navigate.

1. 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 [1], [2]. Despite a number of digital auditory interfaces being presented in the academic literature [1], [3], [4], [5], 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 [6], [7], [8], [9]. 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.

[10] 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 [10]. 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 a scan function, 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 [11], [12], [13], [2]. [10] 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 [11]. The objective of this study was to evaluate reactions and performance of blind participants on a map utilizing audio game conventions.

1.1. 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 [14].

2. AUDIO GAME CONVENTIONS

The three types of audio game interfaces utilized in this prototype were grid-based, first-person, and tree-based. [10] presents these interfaces: "Grid-based maps are based on a set of coordinates representing squares placed together in a column-row relationship" that are navigated through using the arrow keys. When a user enters a cell, a spearcon (a short speech message [15]) along with a short auditory icon (an iconic sound of an object [16], [17]) play, followed by the cell's coordinates [11], [18]. Grid interfaces are best for getting an overview of a map such as in strategy games [18]. First-person interfaces utilize 3D audio to position objects around the player through looping auditory icons of an object. The use of footstep sounds tell the user what type of terrain they are walking on and how fast they are going. First-person is used to give the player a realistic connection to the real world because the cues presented bear an ecological resemblance to an experience in a real physical environment [19]. Tree interfaces are composed of hierarchical parent-child relationships showing in a hierarchy such as a menu. Games often use tree interfaces for complex menus [20]. Most games, such as [18], [19], and [11] use tree interfaces

https://doi.org/10.21785/icad2019.051

to list locations or options users can select, often with child menus with further options.

3. 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 [5] and [21] is the most promising, given that it is a downloadable Windows application and follows many audio game conventions. [5] 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.

[21] 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 [1], [3], and [22], 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.

[23] and [4] 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 [18]. 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 [24].

It is difficult to quantify the effectiveness of many of these interfaces, such as [1], [3], and [5], 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.

4. MATERIAL

4.1. 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 [14], [15], [21], [5]. 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 [25] 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.

4.2. 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 [26]. The playground map was based off a rectangle that encompassed an area 76 meters wide by 62 meters long.

4.3. 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. (The polygon shapes are shown in Fig. 1.) 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.

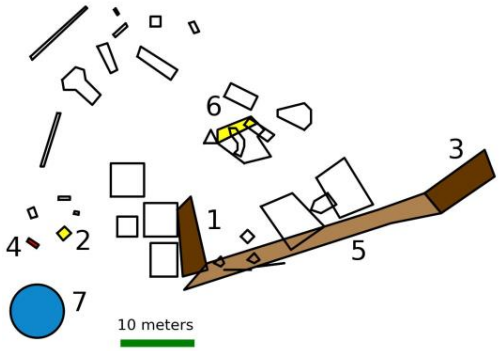

Figure 1. Polygon shapes shown on playground map. Each polygon is drawn with a black outline; polygons that were addressed in the participant tasks are filled in color, with a number label from 1 to 7 printed nearby. The number labels correspond to the following structures: 1 = Ava's Bridge, 2 = Climbing giraffe, 3 = creek bridge, 4 = KinderBells, 5 = long ramp, 6 = roller slide, 7 = stepping sounds. The green bar near the bottom indicates the scale of the map.

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 [15]) 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 [15], [18]. The object menu was effectively the map key. Pressing Enter on each object brought up a submenu with the options:

- Go: take the player to the center of the object polygon.
- Listen: hear the sound associated with the object in isolation from the other sounds.
- Description: Hear the textual description of the object, if any.
- 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.

5. METHOD

5.1. 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.

5.2. Study

All the participants were asked to complete eight tasks (listed below), then rate their performance on the NASA Task Load Index [27], [28]. 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 [29] and [30] 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 [30]. 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 [4].

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 instructions 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 participants for whom English was a second language. 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.

5.3. 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) [31], and one participant used JAWS for Windows [32]. 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 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.

6. RESULTS

6.1. 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: [21] and [30] gave a time limit of 15 and 10 minutes respectively to explore the interface before starting the tasks. [4] had a tutorial that took 1.82 hours on average to complete. The approach in this study was similar to [29] 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 [29], and 50 objects in [4] and the other studies did not indicate the number of objects on their maps.

6.2. 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. 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).

6.3. 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 [21], [29], [30], [5]. [21] 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. [5] 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 [21] 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. [21] 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, past 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 [5]: "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 [21], although there is no mention of the start and end points they evaluated on.

6.4. 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. [29] and [3] 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 [33], [34], and [35], and several applications for auditory shape recognition and creation have been developed such as [36], [37], [38], and [39]. 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 [29], [35]. Physically drawing on swell paper was not possible through the remote medium this study employed and utilizing an application such as [38] 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 "tiles" 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 tile width of the levels are 2, 5, 4, 5, 2. The tile 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 [38], or having participants list the points of the polygon.

6.5. 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. [21] had participants follow a route, but it was not a single path. [3] 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 [11]. 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.

6.6. 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").

6.7. 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 [4]. 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 [4]. There was no task completion time given in [4], 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.

6.8. 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".

6.9. 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 [29]. Both [29] and [3] evaluate sketches participants drew after hearing their auditory map. The sketches in [29] 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 into. 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 [40]. 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.

6.10. Other Results

- 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 [41].
- 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.
- 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. [1], [5], [3], and [4] 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 [11] and [19].
- Nine out of 10 participants repeatedly used the Object Menu to either "go" to an object or get "directions" to an object. [4] 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 [4], and most participants really liked the feature in this interface.
- 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.

- 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 [19] and [11] would give a more steady source of the participant's current location relative to the target.

6.11. Task Load Index ratings

The overall workload score in all categories for the NASA TLX was an average of 39 (SD = 10.58). 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 [42], [43]. 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 mean scores are: mental demand: 55.1 (SD = 20.58), Physical demand: 5.5 (SD = 7.52), Temporal demand: 38.5 (SD = 19.59), Performance: 58.1 (SD = 21.39), Effort: 50 (SD = 31.62), and Frustration level: 27.5 (SD = 22.88). Other auditory map interfaces have not been evaluated for mental task load.

6.12. 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 [11]. Each participant was asked why they used each mode of navigation: tree, grid, or first-person. Their responses are summarized as follows:

1. Tree was used for quick navigation through the map.
2. Grid view was used to quickly navigate and get an overview of the map.
3. 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 [11].

7. CONCLUSION

The prototype in this study evaluated the use of common audio game conventions to display topological objects on a map. There were several major findings from the tasks: the interface was extremely easy to learn and navigate, participants all had unique navigational styles and preferred using their own screen reader, 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.

8. ACKNOWLEDGMENTS

Dr. Coughlan would like to acknowledge funding support from NIH grant 1R01EY029033 and NIDILRR grant 90RE5024-01-00.

9. REFERENCES

[1] P. Parente and G. Bishop, "BATS: The blind audio tactile mapping system," in Proceedings of the ACM Southeast Regional Conference, 2003, pp. 132–137 [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.6189&rep=rep1&type=pdf

[2] J. Friberg and D. Gärdenfors, "Audio games: New perspectives on game audio," in Proceedings of the 2004 acm sigchi international conference on advances in computer entertainment technology, 2004, pp. 148–154.

[3] W. Heuten, N. Henze, and S. Boll, "Interactive exploration of city maps with auditory torches," in CHI'07 extended abstracts on human factors in computing systems, 2007, pp. 1959–1964.

[4] H. Zhao, C. Plaisant, B. Shneiderman, and J. Lazar, "Data sonification for users with visual impairment: A case study with georeferenced data," ACM Transactions on Computer-Human Interaction (TOCHI), vol. 15, no. 1, pp. 1–28, 2008.

[5] E. Loeliger and T. Stockman, "Wayfinding without visual cues: Evaluation of an interactive audio map system," Interacting with Computers, vol. 26, no. 5, pp. 403–416, 2014.

[6] "Cal fire." State of California, 2019 [Online]. Available: http://www.fire.ca.gov/general/firemaps

[7] T. Logan, "Accessible maps on the web," 2018 [Online]. Available: https://equalentry.com/accessible-maps-on-the-web/

[8] Google, "Accessibility in Google Maps," 2019 [Online]. Available: https://support.google.com/maps/answer/6396990?co=GENIE.Platform%3DDesktop&hl=en

[9] ESRI, "A11y-map," 2018 [Online]. Available: https://github.com/Esri/a11y-map

[10] B. Biggs, L. Yusim, and P. Coppin, "The audio game laboratory: Building maps from games," 2018 [Online]. Available: http://icad2018.icad.org/wp-content/uploads/2018/06/ICAD2018_paper_51.pdf

[11] Out of Sight Games, "A hero's call." 2019 [Online]. Available: https://outofsightgames.com/a-heros-call/

[12] audiogames.net, "AudioGames, your resource for audiogames, games for the blind, games for the visually

impaired!" 2018 [Online]. Available: http://audiogames.net/

[13] E. Adams and A. Rollings, Fundamentals of game design. Prentice-Hall, 2006.

[14] A. M. Brock and C. Jouffrais, "Interactive audio-tactile maps for visually impaired people," ACM SIGACCESS Accessibility and Computing, no. 113, pp. 3–12, 2015.

[15] B. N. Walker et al., "Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus," Human Factors: The Journal of Human Factors and Ergonomics Society, vol. 55, no. 1, pp. 157–182, 2013.

[16] E. Brazil and M. Fernstrom, "Chapter 13 auditory icons," in The sonification handbook, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Berlin, Germany: Logos Publishing House, 2011 [Online]. Available: http://sonification.de/handbook/download/TheSonificationHandbook-chapter13.pdf

[17] W. Gaver, "Auditory icons: Using sound in computer interfaces," Human-Computer Interaction, vol. 2, no. 2, pp. 167–177, 1986.

[18] I. Reed, "Tactical battle." 2013 [Online]. Available: https://blindgamers.com/Home/IanReedsGames

[19] J. Kaldobsky, "Swamp," 2011 [Online]. Available: http://www.kaldobsky.com/audiogames/

[20] MetalPop LLC, "Crafting kingdom." 2018 [Online]. Available: http://metalpopgames.com/forum/index.php?/forum/11-crafting-kingdom/

[21] F. Feng, T. Stockman, N. Bryan-Kinns, and D. Al-Thani, "An investigation into the comprehension of map information presented in audio," in Proceedings of the XVI International Conference on Human Computer Interaction, 2015, p. 29.

[22] A. P. Milne, A. N. Antle, and B. E. Riecke, "Tangible and body-based interaction with auditory maps," in CHI'11 extended abstracts on human factors in computing systems, 2011, pp. 2329–2334.

[23] H. Zhao, C. Plaisant, and B. Shneiderman, "ISonic: Interactive sonification for non-visual data exploration," in Proceedings of the 7th international acm sigaccess conference on computers and accessibility, 2005, pp. 194–195.

[24] I. Reed, "User Guide for Tactical Battle," 2013 [Online]. Available: https://blindgamers.com/Home/IanReedsGames

[25] "React: A JavaScript library for building user interfaces." Facebook Inc., 2018 [Online]. Available: https://reactjs.org/

[26] N. Ofiesh and L. Poller, "A playground for the entire community: The design of magical bridge playground," Magical Bridge Foundation, 2018 [Online]. Available: https://slc.stanford.edu/sites/default/files/a_playground_for_the_entire_community_final_submitted_to_mb_and_pw_0.pdf

[27] NASA, "TLX," 2018 [Online]. Available: https://humansystems.arc.nasa.gov/groups/TLX/

[28] Human Performance Research Group (NASA Ames Research Center), "Task load index (nasa-tlx) v 1.0" [Online]. Available: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021488.pdf

[29] W. Heuten, D. Wichmann, and S. Boll, "Interactive 3D sonification for the exploration of city maps," in Proceedings of the 4th Nordic conference on Human-Computer Interaction: Changing roles, 2006, pp. 155–164.

[30] A. M. Brock, P. Truillet, B. Oriola, D. Picard, and C. Jouffrais, "Interactivity improves usability of geographic maps for visually impaired people," Human–Computer Interaction, vol. 30, no. 2, pp. 156–194, 2015.

[31] NV Access, "NVDA 2017.4 user guide," 2017 [Online]. Available: https://www.nvaccess.org/files/nvda/documentation/userGuide.html

[32] Freedom Scientific, "JAWS, job access with speech" [Online]. Available: http://www.freedomscientific.com/products/software/jaws/

[33] S. Uno, Y. Suzuki, T. Watanabe, M. Matsumoto, and Y. Wang, "Sound-based image and positon recognition system: SIPReS," 2018.

[34] F. Bermejo, E. A. Di Paolo, M. X. Hüg, and C. Arias, "Sensorimotor strategies for recognizing geometrical shapes: A comparative study with different sensory substitution devices," Frontiers in psychology, vol. 6, p. 679, 2015 [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00679

[35] M. Rice, R. D. Jacobson, R. G. Golledge, and D. Jones, "Design considerations for haptic and auditory map interfaces," Cartography and Geographic Information Science, vol. 32, no. 4, pp. 381–391, 2005 [Online]. Available: https://www.researchgate.net/publication/228948257_Design_Considerations_for_Haptic_and_Auditory_Map_Interfaces

[36] D. Greve, "Blind paint." 2009 [Online]. Available: http://www.tichnut.de/blindpaint/

[37] Sysop-Delco, "AudiMesh3D." 2017 [Online]. Available: https://sysop-delco.itch.io/audimesh3d

[38] Sysop-Delco, "BrushTone." 2015 [Online]. Available: https://sysop-delco.itch.io/brushtone

[39] S. Balanced, "New smarter balanced desmos calculators are free and accessible to students," Smarter Balanced, 2017 [Online]. Available: http://www.smarterbalanced.org/new-smarter-balanced-desmos-calculators-free-fully-accessible/

[40] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information." Psychological Review, vol. 63, no. 2, p. 81, 1956 [Online]. Available: http://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf

[41] "Sunu band." Sunu, Inc., 2019 [Online]. Available: https://www.sunu.io/en/index.html

[42] N. Meshkati, P. A. Hancock, M. Rahimi, and S. M. Dawes, "Techniques in mental workload assessment." 1995 [Online]. Available: https://www.researchgate.net/profile/Peter_Hancock2/publication/232471724_Techniques_in_mental_workload_assessment/links/0c960532a1f2b630b0000000/Techniques-in-mental-workload-assessment.pdf

[43] Agency for Healthcare Research and Quality, "NASA task load index," 2019 [Online]. Available: https://healthit.ahrq.gov/health-it-tools-and-resources/evaluation-resources/workflow-assessment-health-it-toolkit/all-workflow-tools/nasa-task-load-index