Eyes-Free Art: Exploring Proxemic Audio Interfaces For Blind and Low Vision Art Engagement

KYLE RECTOR, University of Iowa
KEITH SALMON, Independent Artist
DAN THORNTON, Seattle University
NEEL JOSHI and MEREDITH RINGEL MORRIS, Microsoft Research Redmond

Engagement in the arts[1] is an important component of participation in cultural activities, but remains a largely unaddressed challenge for people with sensory disabilities. Visual arts are generally inaccessible to people with visual impairments due to their inherently visual nature. To address this, we present Eyes-Free Art, a design probe to explore the use of proxemic audio for interactive sonic experiences with 2D art work. The proxemic audio interface allows a user to move closer and further away from a painting to experience background music, a novel sonification, sound effects, and a detailed verbal description. We conducted a lab study by creating interpretations of five paintings with 13 people with visual impairments and found that participants enjoyed interacting with the artwork. We then created a live installation with a visually impaired artist to iterate on this concept to account for multiple users and paintings. We learned that a proxemic audio interface allows for people to feel immersed in the artwork. Proxemic audio interfaces are similar to visual because they increase in detail with closer proximity, but are different because they need a descriptive verbal overview to give context. We present future research directions in the space of proxemic audio interactions.

CCS Concepts: • **Human-centered computing →Auditory feedback • Human-centered computing → Accessibility technologies**

Additional Key Words and Phrases: Accessibility, art, proxemic interface, eyes-free, blind, low vision, depth camera

ACM Reference format:

Kyle Rector, Keith Salmon, Daniel Thornton, Neel Joshi, and Meredith Ringel Morris. 2017. Eyes-Free Art: Exploring Proxemic Audio Interfaces For Blind and Low Vision Art Engagement. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1, 93. 3 (September 2017), 20 pages.
DOI: https://doi.org/10.1145/3130958

1 INTRODUCTION

The arts are an important component of full participation in cultural and educational activities, where children and adults alike take an ownership in their learning [17]. Viewing visual art in person can leave a lasting memory, for example, noting that the Mona Lisa is much smaller than expected. Unfortunately, in-person exploration of visual art remains inaccessible for people who are blind or low vision. While accessible technology to fulfill basic needs is important, developing accessible technologies that help make life more

2474-9567/2017/September - 93 $15.00

https://doi.org/10.1145/3130958
Author's addresses: K. Rector (contact author), University of Iowa, 101K MacLean Hall, Iowa City, Iowa, 52240, USA; K. Salmon, independent artist (no affiliation); D. Thornton, Seattle University, Casey Building, 825 10th Ave, Seattle, WA, 98122, USA; N. Joshi and M.R. Morris, Microsoft Research Redmond, 14820 NE 36th Street, Building 99, Redmond, Washington, 98052, USA.

fulfilling, such as by allowing better access to the arts, is also an important aspect of disability rights [1]. There is an opportunity for ubiquitous technology to increase access to the arts for people who are visually impaired. The vision of ubiquitous computing is for technology to be seamlessly integrated into all aspects of life [46], and by introducing a new method of proxemic interaction (proxemic audio interfaces) into the museum space, we provide a novel way for people who are visually impaired to interact with artwork.

There are existing options to make art exploration accessible, including guidelines for verbal descriptions and accessible art tours or guides. Art Beyond Sight [3] creates accessible art programs and educational materials to help museums generate accessible programs. In addition, larger museums such as the Metropolitan Museum of Art [26] or the Museum of Modern Art (MoMA) [32] provide recorded audio guides for patrons who are visually impaired. There are in-person accessible art tours that provide detailed verbal descriptions or tactile art exhibits (e.g., Seattle Art Museum (SAM) [41]). However, these solutions are not yet pervasive; typical accessible tours are offered infrequently (e.g. SAM and MoMA offer such tours once per month). People who are blind do not have the ability to spontaneously visit the museum and have an accessible experience.

Many museums provide audio descriptions, but the descriptions are based on the premise that the user is sighted. The descriptions focus more on interpretation and historical context rather than on literal visual descriptions of the work. General-audience audio guides are also difficult for a person who is blind because they have not typically been designed with accessibility in mind. With these limitations, it is challenging to create an interactive experience for people who are visually impaired.

To address this, we created an alternative sensory experience, which, while not equivalent to the original visual artwork, aims to be aesthetically interesting and appealing. We present Eyes-Free Art as a design probe that provides an engaging, interactive, "on demand" sonic experience and an interpretation in complement to a visual art piece. The experience uses a proxemic interface built using a Microsoft Kinect to provide accessible audio interpretations of existing paintings. We determine where the user is in relation to the painting, and play an audio interpretation aloud based on their proximity (Fig. 1). Eyes-Free Art parallels sighted exploration of a proxemic visual interface; as a user moves closer to the target, more details are rendered [19].

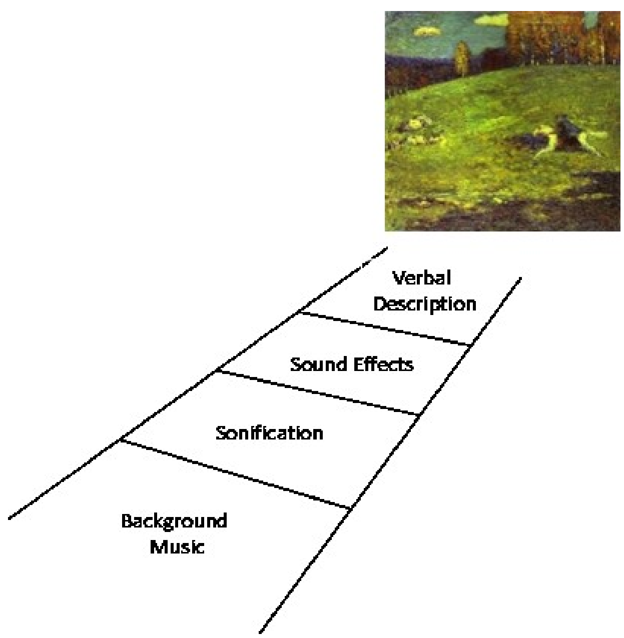

Fig. 1. Eyes-Free Art is a proxemic audio interface that changes the audio interpretation based on a user's distance from a painting. From furthest to closest the user hears: 1) background music, 2) a novel sonification technique, 3) sound effects, and 4) detailed verbal description. Image: The Blue Rider (Wassily Kandinsky, 1903).

The proxemic zones include background music, sonification, sound effects, and a verbal description. We chose this particular order of zones so that the content in each zone increases in detail, because this mirrors related work in visual proxemics. For example, in Harrison and Dey [19], the amount of visual detail of furniture construction directions increased on a computer display as the person leaned closer to the computer. The quantity and type of interpretations in our design probe serve as examples to explore the design of a proxemic audio interface. Blindness covers a wide range of conditions both from birth and later in life, and different zones may appeal to some users more than others based on their personal tastes and their personal history with their level of vision. Eyes-Free Art allows a customized experience since the user may adjust their time spent in each zone.

Further, we adapted this design probe into a live installation to account for multiple users and paintings. We developed this installation with a visually impaired artist. We created the new installation based on the lessons learned from our lab study, including to present the information in increasing detail, but to also give a stage-setting verbal description at the beginning of the experience. In addition, the sound effects in the installation conveyed a sense of physical presence in the landscape depicted by the visual artwork.

This research offers several contributions. We developed an accessible experience to help people who are blind or low vision explore paintings. We present findings and considerations that extend ubicomp interaction paradigms by considering how proxemic interactions can generalize beyond large visual displays and into the audio realm. We expand on prior work on proxemic interfaces to consider how this interaction style might apply in the realm of audio, rather than visual media, and contribute guidelines for proxemic audio design. Our work addresses the following research questions:

1. How can we design a proxemic audio interface to help patrons who are blind or low vision facilitate art exploration?
2. How does an interactive proxemic audio interface compare to a passive verbal description?

Below, we discuss related research in proxemic interfaces, sensory substitution, eyes-free museum opportunities, and interactive art. Then, we describe our Eyes-Free Art probe and the evaluation with 13 people who are blind or low vision. Next, we present our installation that occurred both in October 2016 and April 2017 and discuss patrons' reactions.

2 RELATED WORK

Our work builds on prior work in the areas of proxemic interfaces, sensory substitution, eyes-free museum opportunities, and interactive art.

2.1 Proxemic Interfaces

Anthropologist Edward T. Hall [18] introduced the idea of proxemics, where people are able to assess how much distance to place between themselves and others based on their relationship type: intimate, personal, social, or public. Proxemic visual interactions were initially designed to support interaction from different distances with large form-factor "board" type displays, such as those proposed in Weiser's vision of ubicomp [46]. Previous work in proxemics have considered the design of proxemics in the visual realm. Vogel et al. [43] incorporated the idea of proxemic distances into technology by exploring transitions of a large display's content using implicit and explicit interactions, for example, ensuring that the display is not displaying personalized content if there is a crowd present. Marquardt et al. used proxemic theory to inform the design of ubicomp technologies [24], creating four zones for ambient display, implicit interaction, subtle interaction, and personal interaction [16]. Dingler et al. [10] created a similar framework for using zones in front of larger screens, while Harrison et al. [19] employed a similar technique using a computer screen and web camera. Hello.Wall [36] employed a proxemic technique with a light display where the user progresses from an "ambient zone" shown to all users to a "cell interaction zone" where users can interact with the wall itself. Morris et al.'s "cooperative gestures" [29] adapted the concept of proxemics to gesture design.

Unlike proxemic visual interfaces, the space of proxemic audio interfaces is far less explored. There is prior work on *location-based* audio interfaces; for example, Sørensen et al. [42] developed a system where music follows a user from room to room based on their location. In Fosh et al.'s Sculpture Garden [11], audio is foregrounded when users engage and backgrounded when they reflect. Kabisch et al. [23] presented a *soundscape* of spatialized recorded audio while exploring a landscape panorama; this interaction involved both sight and sound. A navigation system could be designed as a proxemic audio interface if the level of detail in navigation information were to vary with distance from a target; however, current voice navigation systems are not designed in this fashion, to our knowledge. Exploring the impact of proxemic audio design on voice navigation utility and usability is an interesting avenue for future research. To the best of our knowledge, researchers have not explored the concept of a proxemic audio interface, where the detail of audio information increases across zones as a person moves closer to the target object.

2.2 Sensory Substitution

There have been several research efforts to use haptics for people who are blind to explore charts, paintings, or the physical world. For instance, Wall and Brewster [45] developed an approach using a tactile tablet and stylus for a person to explore using the pen. Giudice et al. [12] also used a Samsung Galaxy Tablet and provided both haptic and verbal feedback to convey information about charts. Morris and Joshi used the Phantom Joystick to convey haptic feedback about depth and contours in the physical world [28]. However, these smaller technologies limit the amount of interaction a person can have with a painting because they are unable to explore with their body.

Researchers have also created tactile reliefs of images and paintings. Reichinger et al. [40] developed a workflow to create tactile reliefs of paintings. While there is some manual curation required, most of the tasks are intuitive for the artist to complete. Later, Reichinger et al. [39] augmented the tactile reliefs by adding an interactive audio guide for people who are visually impaired to independently explore with their finger. Hernandez and Barner [20] created tactile reliefs by using watershed-based image segmentation, followed by merging similar regions of the image and identifying the clearest edges. However, tactile reliefs are more difficult to acquire and use, which is a factor in the abandonment of assistive technology [35].

Sound has been used for conveying edges and spatial locations of items in pictures. Goncu et al. [14] created an iPad application where a user explores a floor plan with their finger and listens to verbal and non-verbal sound feedback. O'Neill and Ng [33] developed a system for people to explore images using a Wii mote. The person hears auditory feedback conveying the mass and complexity of the image segment based on where the Wii mote is pointing, and haptic feedback through the Wii mote as edges are crossed. The users can also hear the path to the center of each segment and to the closest perimeter of the segment. Yoshida et al. [47] developed a phone application that allows the person to move their finger along the image and they hear a sonification of the local edge that their finger is touching, and a sonification to the nearest image edge. Our sonification technique is similar to these bodies of work in that the centroid of the user's hand is an input to the painting image, but we are different because we wanted the person to be facing the painting like others visiting an installation as opposed to facing a phone or tablet screen.

Researchers have also used sound to convey colors in an image or painting. Cavaco et al. [7] developed software to convey hue, saturation, and brightness using musical features including timbre, pitch, and volume. Meijer [25] developed a sensory substitution system to convey image brightness and height using pitch and volume. Pun et al. [37] represented the hues in an image with different musical instruments, and the depth of the elements using different rhythms. One commonality among these projects is that sounds are presented as a literal translation of the color, which may not produce aesthetically pleasing sounds. In our sonification, we only modulate the volumes of existing instrument music tracks to keep the music aesthetically pleasing.

2.3 Eyes-Free Museum Opportunities

A small subset of museums have developed audio guides that are accessible to people who are blind or have low vision. For example, the Metropolitan Museum of Art [26] and the Museum of Modern Art (MoMA) [32] provide recorded audio guides for patrons who are visually impaired. However, blind patrons are limited only to those museums and those locations within the museums that have accessible audio tours. There is another application called SFMOMA which provides an immersive audio experience while visiting the museum [8]. However, the target of SFMOMA is to "keep your phone in your pocket and your eyes on the art" [8], so it is designed with a sighted audience in mind. The Broad also offers an app [6], but if a patron wants an accessible art tour, they need to book at least two weeks in advance [5]. There are in-person accessible art tours that provide detailed verbal descriptions or tactile art exhibits (e.g., Seattle Art Museum (SAM) [41]). However, these solutions are not yet pervasive; typical accessible tours are offered infrequently (e.g. SAM and MoMA offer such tours once per month). In general, a person who is blind has a limited amount of museums that they can visit spontaneously and have an immersive experience. We hope that the concept of proxemic audio interfaces can provide a method in which to create an accessible experience for people who are blind to explore art.

2.4 HCI and Interactive Art

The HCI community has embraced interactive art and made contributions including how to evaluate art installations [13], evaluate enjoyment [15], and use vocabulary terms to increase our understanding of interaction and play [31]. Gonzales et al. [15] conducted a between-subjects study with a passive music installation or an interactive music installation with a Wiimote. They found a strength of interactive art is that it increases enjoyment in artistic spaces. Further, Wakkary and Hatala determined that it is important to bring a sense of liveliness to art displays [44] after exploring tactile and audio experiences. There is support to develop multi-sensory interactive experiences for art, but these efforts did not explore whether or not the audio or haptic elements could stand alone independently of the visual elements for people who are blind or low vision.

3 INITIAL INTERVIEWS AND DESIGN GOALS

Before determining our project direction, we interviewed people who are blind or low vision and museum/art domain experts about their experiences with accessible/inaccessible art and how technology may be involved while consuming art. We interviewed seven people who are visually impaired (VI1-VI7: 6 females, 28-65, median age = 53), five artists (A1-A5: 1 female, 29-68, median age = 48, two tactile artists), one museum curator (C1: female, 58), and one museum accessibility coordinator (ACC1: female, 37). These interviews were semi-structured and were conducted over the phone, and the themes were iteratively identified through open-coding-based techniques.

In the discussions with the people who are blind or low vision, we inquired about the benefits and challenges while consuming art, accessibility at art museums, ideal art experiences, and whether or not they use technology while consuming art. From the artists and curator, we wanted to learn about what patrons are intended to learn or experience while engaging with a piece of art and whether they have had experience developing or curating art that was accessible. From the accessibility coordinator, we wanted to learn about how the museum experience differs for people who are visually impaired as opposed to people who are sighted. In addition, we wanted to learn about the benefits and challenges of creating accessible experiences in the museum. These interviews shaped our design goals and inspired us to explore proxemic audio as a method for facilitating art's accessibility. We determined that the stakeholders wanted access to an experience that was aesthetically pleasing and conveyed both detailed and emotional aspects of a painting. Below, we explain each of the design goals we derived from these interviews in detail with supportive quotes:

First, we wanted to develop our experience with a **commodity technology**. Providing access to an off-the-shelf technology in a museum setting can increase the user's independence: "It makes it easier to schedule things, time things, [and] do things on my own schedule (VI6)." Further, VI2 found that not having access to an accessible technology in the museum hindered their independence: "They were like if you want a guided tour you have to prearrange 30 days in advance." Additionally, VI3 chose not to explore the Louvre despite traveling to Paris due to inaccessibility: "There's no point, they are going to keep us back behind barricades, and they will not let us get close to the Mona Lisa...Here I came all that way and had to say no, it's not worth it."

Second, we wanted to make the experience **subjectively satisfying and aesthetically pleasing**. The experience should be moving; C1 noted that: "The point of art in museums is to have a direct personal experience ... in your body." More concretely, the design of audio should sound appealing: "When the sound is good, you don't notice it that much. When the sound is bad, it annoys you (A2)." For Eyes-Free Art, we designed a custom sensory substitution experience by using already composed musical tracks and changing the volume of instruments to convey color. The existing composed music ensures that it is aesthetically pleasing, while the changes in volume provide information about the color.

Third, it is important to include **both a detailed verbal description and a presentation to convey emotion and mood**. According to the accessibility coordinator, it is important to include literal aspects such as: "A sense of scale, the colors involved (ACC1)," VI2 enjoyed hearing the background music on an audio guide because: "The music gives certain moods." Additionally, VI6 also mentioned the benefits of using background music: "It's a wonderful Andy Warhol piece - it's called 'Silver Elvis.' It is larger than life and it's taken from the publicity from his first movie, and the docent actually played him singing the title song from the move. It was a lot of fun." Providing literal facts and emotional aspects of a painting are distinct, and yet both are important:

> "There are two kinds of descriptions; one is informative and one is aesthetic ... you have things like the color, maybe the shapes that are being used ... but art tends to also convey a certain aesthetic ... in that respect it is very important to convey visually all of that information for the mood or the sentiment from the painting (VI1)."

We also did not want to develop another audio guide that a person has to operate. VI3 reported:

> "This is a pet peeve [...] I think you will hear it from other people as well [...] Even if all museums offered recorders - every recorder will operate in a different way. [...] Maybe it is super easy to learn, but what I have found is that sometimes when you are in a new situation, you are trying to orient yourself. You are either with a group of people or there are lots of people around you so you are trying to stay alert, make sure your cane or dog is not near anywhere else, make sure you are not inadvertently standing in the middle of everyone, [or] the pathway. Doing all of these other things, you know physical things, and then you have to be learning this machine and operating it in a timely way because you have to keep moving from art piece to art piece."

We do not want participants to have to learn a new application or physical device, so we designed Eyes-Free Art to be a proxemic audio interface, where the person hears information based on their proximity from the painting. We decided to build the experience using a Microsoft Kinect instead of developing customized hardware because commodity technology is easier to utilize in an installation or exhibit. They do not have to operate a small device because their hand will already be occupied by a cane or guide dog. Instead, they control the experience by moving around the museum space with their body. Eyes-Free Art also has multiple zones, so that patrons could move closer and further from the painting to hear a detailed verbal description and experience the emotion and mood of the painting.

4 EYES-FREE ART DESIGN PROBE DESCRIPTION

We built Eyes-Free Art using the Microsoft Kinect for Windows V2 Software Development Kit version 2.0 and C#. The Kinect is placed approximately 4 feet above the ground and below the painting (Fig. 2). We developed the experience for five paintings that contained different objects and colors (Fig. 3).

4.1 Technical Implementation

We use Body Tracking [27] to track the person's distance from the painting and therefore the corresponding zone in which they were standing. We used the depth of the person's spine to determine their distance from the painting. By using the tactile cue on the floor, the user knows the location of each zone and is able to move between the zones. We chose to have four zones because it mirrors the four zones as described in Hall [18] and Greenberg et al. [16]. The zones of Eyes-Free Art are arranged in a 6'x15' foot space, with audio cues to ensure the user remains in front of and facing the painting. Zone 1 is closest to the painting and Zone 4 is furthest away from the painting (Zone 1: 3'-6' from painting, Zone 2: 6'-9' from painting, Zone 3: 9'-12' from painting, Zone 4: 12'-15' from painting). We chose to have each zone be the same size (3') because we wanted participants to be able to stand inside of each zone completely, and not have the probe place emphasis or preference on any zone in particular. As the user enters each zone, a verbal cue is played aloud (e.g. "You have entered Zone 1: verbal description"). With this interaction, participants can "pull" information from the system by choosing to enter each zone, and by choosing how long they spend in each zone. We provided audio through the speakers because we were imagining an immersive experience where people can explore the space together. If sound pollution is of concern, then we could include more private audio settings without posing a safety risk. For example, a user could wear a single earpiece to hear the experience privately while still being able to listen to their surroundings. For example, Sotto Voce [2] is a single earpiece guide used in museum settings, and Morris et al. 2004 found that groups of people could collaborate with one another while using private audio via a single earpiece [30]. Another alternative is bone conduction headphones that still permit blind people to listen to sounds in the environment (as reported in Rector et al. 2015 [38]).

We also tracked whether or not the person was facing the painting. If the absolute value of the difference between the z-coordinate of the right and left shoulder was above 0.1 meters, then we determined that the user was not facing the painting. If the user stops facing the Kinect, "twist your body slightly left" or "twist your body slightly right" is played aloud so the user can correct their position.

Body Tracking is also used to track the location of the user's hand. The user's hand may be used to explore the painting in Zones 2 and 3. The user raises their hand and moves it in free space within a 3'x3' square centered on their right shoulder. The region is centered on the camera in the y-axis, and ranges from ¾' to the left of the camera to 2 ¼' to the right of the camera in the x-axis. In this way, the person is able to reach all of the painting with one hand. The painting is resized while maintaining the proportions to fill the 3'x3' space as much as it can. If the painting is in portrait orientation, the height is 3' and if the painting is landscape orientation, the width is 3'. We focus on one-handed rather than bimanual interaction, since many people who are blind will need their other hand to hold a white cane or a guide dog leash. We also provide summary information in Zones 2 and 3, so the user can still get information from that zone without using their arm, in case arm movement is restricted or fatiguing. The arm movements allow participants to "pull" information from the painting, because they choose whether and where their hand moves in space.

Fig. 2. Eyes-Free Art presented in a room. The painting is projected on the wall to simulate a gallery, with the Kinect sensor below. On the floor is a white tape and cardboard "ladder" used as a tactile cue to navigate between the zones.

Fig. 3. Three of the five paintings explored in Eyes-Free Art (from left to right): 1) The Sleep of Reason Produces Monsters (Francisco Goya, 1797-1799), 2) The Blue Rider (Wassily Kandinsky, 1903), and 3) The Stone Breakers (Gustave Courbet, 1849). Two images not included: 1) Self-Portrait with Thorn Necklace and Hummingbird (Frida Kahlo, 1940): https://en.wikipedia.org/wiki/Self-Portrait_with_Thorn_Necklace_and_Hummingbird, and 2) The Red Studio (Henri Matisse, 1911): https://en.wikipedia.org/wiki/L%27Atelier_Rouge

4.2 Proxemic Audio Interface

Eyes-Free Art has multiple zones, similar to Hall's proximity theory [18] and Marquardt's application to visual interfaces and ubicomp [16]. First, we provide general background music to draw a person into the piece, followed by an interactive sonification of the colors in the painting. Third, the user interacts with the painting, learning about specific objects through sound effects. Finally, we present the most detailed and specific information, a verbal description. Note that while we chose these four zones for our design probe, the quantity and type of zones may change with the goals of a specific installation. Further, some zone types may not be applicable to all paintings; for example, abstract art may not have a sound effect zone since it lacks recognizable objects.

When the user first approaches the painting (just before entering Zone 4), some key metadata is read aloud (the painting title, artist, year, and country where it was painted). This "pseudozone" was added based on pilot-testing feedback with two sighted participants before the user study began. They indicated it was helpful for setting the stage of interpreting what occurs in subsequent zones; in a museum setting, this metadata zone would also facilitate quick navigation by allowing users to quickly scan amongst several works and then proceed with in-depth interactions only with the items of most interest. Below we will describe the zones in order from furthest to closest.

4.2.1 Zone 4: Background Music. We chose to have a background music zone based on the participants' positive reports of background music in our initial interviews (VI2 and VI6). After presenting basic facts about the painting (title, year, artist), Eyes-Free Art conveys an overall mood by playing a background music track. For each painting, we selected a genre of music based on a pairing chosen by Mechanical Turk workers. For each Turk task, we provided a picture of a painting from Fig. 3, 13 musical genres[2] including example clips and composers/artists, and asked: What genre(s) of music are most appropriate for this painting? The responses were checkboxes so the turkers could select multiple genres. Five turkers provided responses for each of the five paintings ($0.20 per response). For each painting, we chose the musical genre that had the most votes (Table 1). We anticipate that the musical genre could also be selected by an artist based on their artistic intent or by a curator based on heuristics about country, year of origin, and artistic genre.

When a user enters Zone 4, Eyes-Free Art provides a description: "People have chosen the musical genre of <genre name> to pair with this painting." The Pandora station of that genre is chosen and the first song in the queue is played aloud. We chose to pair a genre of music over a single track because it allows for the visitor to appreciate the style of music as opposed to particular components of the song.

[2] Jazz, R&B/Rap/Hip-hop, Reggae, Rock, Dance/Electronica, Classical, Baroque, Renaissance, Country, Folk, Christian/Gospel, Pop and Other.

Table 1. Music pairings for each painting as determined by the highest vote of Mechanical Turk workers.

Painting	Musical Genre
Self-Portrait with Thorn Necklace and Hummingbird	Reggae
The Sleep of Reason Produces Monsters	Classical
The Blue Rider	Baroque
The Red Studio	Rock
The Stone Breakers	Folk

4.2.2 Zone 3: Sonification. The goal of the sonification is to allow a user to have an aesthetically pleasing auditory experience while gaining a sense of the quantity and variety of colors in the painting. Another advantage to producing aesthetically pleasing sounds versus a literal translation of color is that they are more likely to seem "invisible" and blended in the museum environment, in keeping with Weiser's notion of ubiquitous computing [46]. We also chose these priorities over learning about all colors or the ability to create a full mental picture of the painting because that may require significantly more user training and degrade the casual museum visitation experience. We chose to map color to sound using sensory substitution because color-to-sound mapping has been frequently explored in technologies for people who are blind (e.g. Cavaco et al. [7], Meijer [25], and Pun et al. [37]). Other instantiations of proxemic audio could choose to use sensory substitution to convey other features of the artwork, perhaps even with mappings created at the artist's discretion.

We use Body Tracking to track the user's hand. As the user's hand is exploring the painting, an orchestral musical loop is played aloud (purchased from [4]). Musical instruments are played at different volumes to represent different colors. We chose to vary the volume of instruments to convey color to mirror other related work where researchers changed elements of music to convey changes in color. For example, Cavaco and Meijer conveyed changes in color by using aspects of sound like timbre, pitch, and volume [7, 25]. We chose to only use volume because we wanted to maintain the integrity of the sound while still communicating a change in information.

Three sets of instruments correspond to the three primary colors in RGB (orchestra – red, piano – green, harp – blue). The pixel under the x-y position of the centroid of the user's hand is played aloud. In order to make different colored regions of the painting sound noticeably distinct from one another, we choose to bin the RGB colors into nine discrete bins and choose the closest of those nine colors to that pixel. The nine colors and instrument volumes are shown in Table 2 – these mappings were chosen empirically via trial and error exploration of various mappings. If the user's hand is not hovering over the painting, no music is played.

Table 2. Map from color to instrument volume in the orchestral track [4].

Color (RGB)	Orchestra	Piano	Harp
Red (255, 0, 0)	100	10	10
Purple (255, 0, 255)	100	10	100
Blue (0, 0, 255)	10	10	100
Teal (0, 255, 255)	10	100	100
Green (0, 255, 0)	10	100	10
Yellow (255, 255, 0)	100	100	10
White (255, 255, 255)	100	100	100
Gray (128, 128, 128)	50	50	50
Black (0, 0, 0)	10	10	10

For example, if a user's hand hovers close to the top of Self Portrait with Thorn Necklace and Hummingbird (see painting at https://en.wikipedia.org/wiki/Self-Portrait_with_Thorn_Necklace_and_Hummingbird), the pixel is matched with the closest color, blue. The orchestra and piano are played at a volume of 10, and the Harp is featured at a volume of 100. When the user's hand crosses an edge in the painting (Fig. 4a), a gong that is distinct, yet cooperates with the music, is played aloud. To avoid the gong's overpowering of the sonification, it can only be played every 500 milliseconds.

When a user enters Zone 3 they hear: "You have entered Zone 3: Sonification." They hear an audio tutorial explaining how the sonification works. The user hears an initial synopsis of the painting, which is done by averaging all of the pixel colors and choosing the closest color. Finally, Eyes-Free Art instructs the user to reach their hand forward to explore.

4.2.3 Zone 2: Sound Effects. Unlike the goals of background music and sonification, which are meant to set a mood and give a general impression of the work, the goal of the sound effects is to convey literal aspects of a painting, specifically, the type and location of objects contained in the painting. The user has a chance to hear effects representing major components of the painting and to understand the spatial relation between those objects.

As the user explores the painting with their hand, the position of the hand determines which sound to play aloud. An image (Fig. 4b) is loaded with manually annotated regions representing the different objects. While manual annotation is required for the sonification zone, the task of coloring different objects is intuitive and easy to complete in a paint application. A sound effect is played aloud when the user's hand is over the corresponding region. If the sound file has finished playing or the user's hand leaves the object, then Eyes-Free Art is ready to play another sound effect. When the user's hand is not hovering over an object, a relevant background sound is played. For example, in a landscape, the user hears breeze flowing through grass. If the user's hand is not hovering over the painting, no sounds are played.

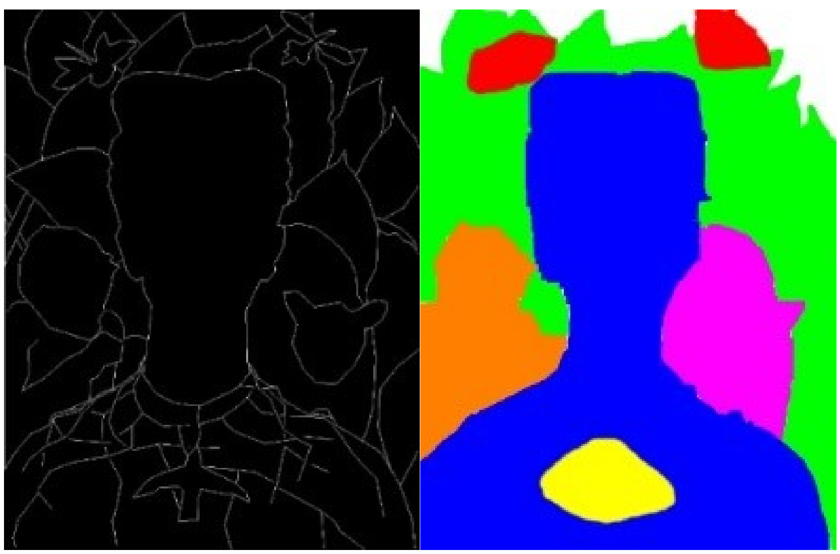

Fig. 4. From left to right: a) manually drawn edges for Self-Portrait with Thorn Necklace and Hummingbird; b) sound effect regions for Self-Portrait with Thorn Necklace and Hummingbird. The regions were manually colored (green = leaves, orange = monkey, pink = cat, red = butterflies, yellow = bird).

When the user enters Zone 2, they hear "You have entered Zone 2, Sound effects." First, Eyes-Free Art informs the user that they will hear sounds made by the objects contained in the painting. A synopsis of all of the sounds is then played aloud so the person can hear the objects represented in the painting. Finally, the user is able to explore with their hand.

4.2.4 Zone 1: Verbal Description. The user can enter Zone 1 to receive the most detailed information about the painting. A verbal description, manually curated from a combination of Wikipedia articles and an Art History textbook [22], is read aloud to the user. The description includes title, artist, year, materials used, contents of the painting, and history of the painting. We chose to manually curate the verbal description of the painting due to the lack of consistent written descriptions online; we believe with improved and consistent descriptions we could automatically scrape online text. At any point in the verbal description, the user can pause the narration by moving between the different zones to interact with the painting. The interaction on subsequent visits to each proxemic zone is faster because the instructions are not repeated unless requested by the user (instructions for each zone can be requested via the audio command "Repeat").

5 EVALUATION

To assess the ability of Eyes-Free Art to provide an enjoyable and informative audio experience for a painting, we conducted a user study with 13 participants who were blind or low vision. Our study design had every participant experience a *baseline* and an *experimental* condition. The baseline was only the detailed verbal description, similar to most museum audio guides and online resources. The experimental condition was the Eyes-Free Art design probe, which also contains the same verbal description as in the baseline as one of several proxemic zones.

5.1 Participants

We recruited 13 participants (12 females, ages: 19-71, median age: 52) who were blind or low vision to participate in our study. Seven participants were blind while six had low vision. Four of the participants identified as artists. Two of those four participants also said that being an artist was their occupation. The participants reported that they visited art museums: never (1), once a year (3), a couple of times a year (4), and once a month (5). We recruited the participants via local email lists and by connecting with a local art museum. The participants spent about an hour completing the study and were compensated with a $100 gift card.

5.2 Procedure

We interviewed participants to learn about their level of vision, artistic experience, and frequency of attending art museums. Then, the participants were told to imagine that they were in an art museum standing in front of a painting. In the experimental condition, we explained that they would slowly approach the painting, and that the audio they hear would change. Participants were given a tutorial on the tactile "ladder" (Fig. 2), where the center of the ladder allows them to move between zones, and when placing their feet on the rungs, they would know they are in zone. In this way, the participants knew how to move among zones as they were exploring the painting.

Each participant interacted with three paintings: one with the baseline condition and two with the experimental condition. We gave participants the baseline only for one painting because the verbal description was standard and we wanted to keep the study duration to about one hour; we are also not formally comparing "performance" between the two conditions – the baseline condition's presence was intended merely to remind participants of the status quo experience they would have at museums today, hence, sacrificing full counter-balancing in order to limit participants' time commitment was a conscious tradeoff. Participants had the chance to experience two paintings in the experimental condition to ensure they were able to fully explore the different zones. For each participant, we used a modified Latin Square design to choose three of the five paintings (Fig. 3), such that all five paintings were used in the baseline and experimental conditions equally. The study was counter-balanced so the participants were placed into Group 1 (baseline 1st, experimental 2nd, and experimental 3rd) or Group 2 (experimental 1st, experimental 2nd, baseline 3rd).

5.3 Data Collection

The participants used either the baseline verbal description or the experimental Eyes-Free Art condition (with the order used in Group 1 or Group 2), and we orally administered a questionnaire containing both closed-form (5-point Likert-scale from strongly disagree to strongly agree) and open-form questions, to assess their experience, learn about any perceived benefits, and receive feedback for improvement. These questionnaires were conducted after each condition. The baseline follow-up interview was about the verbal description, while the experimental interview additionally contained questions about each of the zones and the overall design probe. We did not ask participants about what they felt they saw in the painting after each condition because the detailed verbal description was a part of both conditions. The interview also contained several Likert-type statements with which participants could express agreement or disagreement. At the end of the study, we asked participants whether they preferred the baseline or experimental condition. The interviews were audio recorded and the entire session was video recorded. We did not collect data on the participant's movements.

5.4 Results

Below we discuss the differences in participant experience between the baseline and experimental conditions, followed by participant feedback for each of the zones.

5.4.1 Baseline vs. Experimental Experience. Eleven of the thirteen participants preferred Eyes-Free Art to the baseline condition because it gave the painting more dimension and the information progressed from general to specific. A Chi-squared test shows that the preference for Eyes-Free Art was significantly higher than chance ($\chi 2$ (1, N=13) = 6.23, p = 0.013). One reason for this preference was that: "[Eyes-Free Art] gave [the painting] more dimension. (P7), and "because you get a fuller appreciation of the picture (P1)." Further, participants appreciated the proxemic audio interface moving from general to specific: "[Eyes-Free Art] sort of would mirror other people just looking, because if you [a sighted person] were far away from [the painting], then you won't really know what's on [the painting], so it's just a general idea, and as you get closer to it you would actually get more out of it (P4)."

Participants also felt that they were able to gain more information about paintings in the experimental condition than in the baseline condition. For the statement, "I had a much better understanding of the contents of the painting," participants had a higher agreement in the experimental condition according to a Wilcoxon signed rank test (Z = 2.602, p = .009), where the medians were baseline = 2 and experimental = 4. P3 stated: "For a person that's not sighted... it makes you feel like you're walking away with the knowledge." Further, while the verbal descriptions were consistent between the baseline and experimental conditions, participants felt that the verbal description in the experimental condition was more informative. For the statement "After listening to the verbal description, I had a better understanding of the contents of the painting": A Wilcoxon signed rank test found there was a higher agreement with this statement in the experimental condition (Z = 2.836, p = .005) with medians baseline = 3 and experimental = 5. This may indicate that the verbal description was more informative when the person had the opportunity to gain context about the painting using the other three zones in the experimental condition, though it is possibly an artifact of novelty or social desirability bias. P13 noted that the verbal description in the proxemic audio interface "certainly pulled [the painting] all together."

Eyes-Free Art and the baseline were more comparable when conveying a sense of aesthetics or emotion; there was no significant difference (Z = 0.513, p = .608) in participants' agreement with the statement: "After using the system, I had a good sense of the general aesthetics/emotional content of the painting." For instance, P9 reported in the baseline condition that "[the verbal description] gave you a very good idea of the emotional expression and it gave a lot of detail of what was in the painting." However, participants' reactions during the study sessions seemed to indicate greater emotional involvement with the art work when using Eyes-Free Art; several participants laughed aloud with pleasure while using Eyes-Free Art, and one participant was moved to tears by the novel perspective on art that the probe provided her. P4 noted, "[Eyes-Free Art] gives you an intense feeling without spending hours [at the museum]."

Participants appreciated that, in comparison to baseline verbal descriptions, Eyes-Free Art is interactive: "[Eyes-Free Art] makes [the painting] more interactive... sometimes at museums you feel a little left out [as a blind patron] (P5)." Participants felt that Eyes-Free Art "felt more like total experience instead of just standing in front of something (P13)." In addition, P4 noted that "you can choose yourself how much time to spend on the painting." The use of the hand was useful to scan the painting: "I can move my hand back and forth, and in that way look at something more specifically (P3)." Even though participants moved themselves and their hands around a space, they did not find Eyes-Free Art to be more fatiguing than the baseline condition. For the statement "I found the system fatiguing to interact with": there was no significant difference in terms of participant agreement ($Z = -0.159$, $p = .874$).

5.4.2 Per-Zone Feedback. Each zone in Eyes-Free Art provided a unique benefit to participants, because the zones educated participants about different aspects of the painting. Participants mentioned that the background music "instantly sets a mood (P5)", while the sonification provided more information about colors: "It [the sonification] did give me a very good sense of brightness and darkness of the colors (P6)." Finally, the verbal description "helped give a complete picture (P10)."

As the participants began using Eyes-Free Art, the music zone was able to convey the emotion of the painting: "[The music] set the mood of the picture, which was really important… [Goya's] art is dark and very troubled, and the classical [music], you could get a sense of the volumes from that, what his intent was (P3)." However, the music may not be as compelling for more comprehensive painting descriptions such as genre (median = 3) or time period (median = 2) (Table 3, Music). In order for the music to provide a more educational experience, P9 suggests to "research the story about the painting to help select music that complements the art work." We anticipate that an artist or curator could choose a particular genre or song to pair with the painting more effectively than turkers. Further, a toolkit could be developed to allow for artists and designers to create their own experiences for new installations.

The sonification zone required the longest amount of training (a couple of minutes to listen to the instructions), but provided an enjoyable experience and an idea of the colors in the painting. The participants appreciated that we used composed music and only changed the volumes: "[The sonification] has got harmony to it. It's just cool. It takes away the frustration. I felt like smiling. (P9)." Participants felt that they were able to distinguish the instruments and get a sense of the colors in the painting (all medians = 4) (Table 3, Sonification): "For people who can't see colors, [the sonification] definitely gives an idea (P5)." The gong effect to demarcate edges within the paintings was generally confusing to users; they seemed more interested in having an effect to note the edges of the painting itself (in order to better understand its aspect ratio) than the edges of painted objects.

Another interesting finding is that participants did not necessarily want the music and sonification zones separate from one another. P11 and P13 suggested that the music in the sonification zone and music zone should be the same. P6 added that they would "make each [sonification] individual to the painting." We suggest that it would be better to use the same piece in both the music and sonification zones, or to combine the zones together to have a more seamless transition between zones.

Because the sound effect zone played audio contained by the objects in the painting, it provided a realistic experience: "[Eyes-Free Art] brought the picture … alive (P3)." In addition, P6 mentioned that the sound effects "created a feel you were actually in that place." Participants felt they were able to locate objects of interest (median = 4) (Table 3, Sound Effects).

Overall, participants felt they were able to navigate between the zones (median = 5), but preferences for ordering of the zones had differing responses (median = 4) (Table 3, Overall). In fact, six participants listed that they would want to hear a verbal description at the beginning (i.e., the most distant zone) to help decide if they want to visit that painting, suggesting that our "pseudozone" that announced metadata upon first approach was not as sufficiently detailed as may be desirable: "Verbal description first to have specific content especially for painting (P6)." More specifically, P11 suggested that they "would add an extra very short verbal description at the beginning." Therefore, we suggest that proxemic audio interfaces are different from proxemic visual interfaces in that they need a sufficient verbal description at the beginning to give context to the audio experience.

6 FROM DESIGN PROBE TO ART INSTALLATION

The Eyes-Free Art design probe suggests that people who are blind or low vision can independently explore paintings, and it introduced the notion of proxemic audio zones that can provide a mapping for artists to use when creating sounds and interactions for their visual paintings. Because the probe was explored in a lab setting with a single user, we needed to make design changes so it would become a more realistic art installation. To explore the artist's ability to create a proxemic audio experience, we collaborated with an artist, Keith Salmon, who is visually impaired, along with a sound artist and technician to create an art installation based on our learnings from creating and evaluating the initial design probe. We adapted Eyes-Free Art into a public installation in a larger space with multiple users. With this collaboration, as Eyes-Free Art was being considered from the piece's conception, we were able to have the artists decide the visual and auditory content and interaction to create an integrated experience. In addition, we expanded the notion of proxemic audio to address the technical challenge of supporting multiple simultaneous users.

Table 3. Medians of 13 users' agreement with statements about Eyes-Free Art (experimental condition). Separate histograms are shown for blind and low vision participants. Strongly Disagree = 1 to Strongly Agree = 5. The y range is from 0-5 participants.

Music	**Median - All**	**Histogram - Blind**	**Histogram – Low Vision**
"I was able to assess the mood of the paintings."	4		
"I was able to assess the genre of the paintings."	3		
"I was able to assess the time period of the paintings."	2		
Sonification			
"I was able to hear instruments playing at different volumes."	4		
"I was able to distinguish the instruments playing red (orchestra)."	4		
"I was able to distinguish the instrument playing green (piano)."	4		
"I was able to distinguish the instrument playing blue (harp)."	4		
"I was able to distinguish white (quieter), gray (medium), and black (louder)."	4		
"I was able to get a sense of the lightness or darkness in the paintings."	4		
"I was able to get a sense of the quantity and variety of colors."	4		
"I was able to identify specific colors in the paintings."	4		
Sound Effects			
"I was able to locate objects of interest in the paintings."	4		
Overall			
"I was able to easily navigate between the zones."	5		
"The order of the audio zones made sense to me."	4		

The artists implemented the installation in a 30'x15' room with three impressionistic drawings of a canyon landscape (Fig. 5). These drawings are presented with recorded audio composed from the landscape itself, additional sound effects, and sound recorded from the drawing process itself. We chose to record audio from the landscape itself because we had the opportunity to record the most realistic sound effects. Because our installation accounts for multiple users, each of the four corners of the room had a Microsoft Kinect pointing towards the center to mitigate the risk of occlusion. When the software detects a user, it determines which of the three drawings' audio to present based on the user's x-axis location with respect to the wall of paintings, and the user's depth will determine the corresponding proxemic audio zone, as in our original prototype. The installation is entitled “The Oregon Project”. We included a 3x5 grid of spatialized speakers so that multiple users could experience the audio installation simultaneously.

Fig. 5. The Oregon Project on display in October 2016. Three impressionistic drawings are at the front of the room with a four Kinects to track people, and a 3x5 speaker grid overhead. As a person moves and gestures in the space, composed sound tracks that correspond to location relative to the drawings are played overhead. The speaker grid provides a target, individualized sound so that there are both personalized and group experiences occurring in the space.

The zone implementations are directly inspired from the Eyes-Free Art probe and user study (Fig. 6). For instance, the piece has a “music zone” that consists of a track specifically composed for the piece (in part from the recorded landscape sounds), and played as an ambient track in the room to attract people to visit and to provide a consistent background mood track for all users. Then there are three zones in the room that present a mixture of the “sound effect” and “sonification” zones. For each zone, the primary sound components, respectively, are: 1 – most distant) presentation of recordings of the broader landscape sounds from the outdoor environment (e.g., wind and birds), 2 – middle distance) sound recordings with heightened detail of landscape sounds and targeted sounds from the drawings tuned to convey the color of where the user is pointing, as in our “sonification” zone, and 3 – closest distance) the recorded sound of the artist drawing the actual work at the front of the room.

While the zone design for our installation is different from our probe, we chose to increase the similarity of the sounds playing from each zone (after hearing the suggestions of P6, P11, and P13), so that the zones transition more smoothly. For example, in our probe transitioning from Zone 4 to Zone 3, the person would hear a song played from a Pandora station followed by an orchestral track, where these sounds differed. For “The Oregon Project,” the person would hear ambient background noise followed by the addition of sounds that provide specific details such as wind and birds. In fact, some amount of landscape sounds will be heard at a low level in the 2nd or 3rd zone (e.g. tapered blending of sounds between zones to provide a less abrupt transition). Participants liked that the zones increase in detail as the person approaches the painting, so we used the same pattern in the installation. The furthest zone of landscape sounds provides an overall picture of the scene, while the closest zone of sounds of creating the drawing presents an aural representation of the strokes and colors that dominate the visual experience when viewing the drawings up close. The involvement of an artist in designing the proxemic audio interface shows how the specific contents of the proxemic zones can change to convey artistic meaning while still adhering to the overall general-to-specific concept of the interaction.

The updated system provides a unique and seamless audio experience for multiple users while reducing noise bleed. The people in the room hear targeted sounds from a 3x5 array of 15 speakers arranged in a grid on the ceiling. Each speaker has a parabolic cone to limit the location where that speaker is audible, so that one user’s experience in one part of the room minimally bleeds into another user’s experience in a different part of the room. Depending on the user’s location, the speakers closest to her play the appropriate sound for the painting, as determined by her zone and location. Transitions between locations are eased both by careful design of the user experience and are aided by the masking sound of the “music zone” track that is played quietly in the background, as discussed previously.

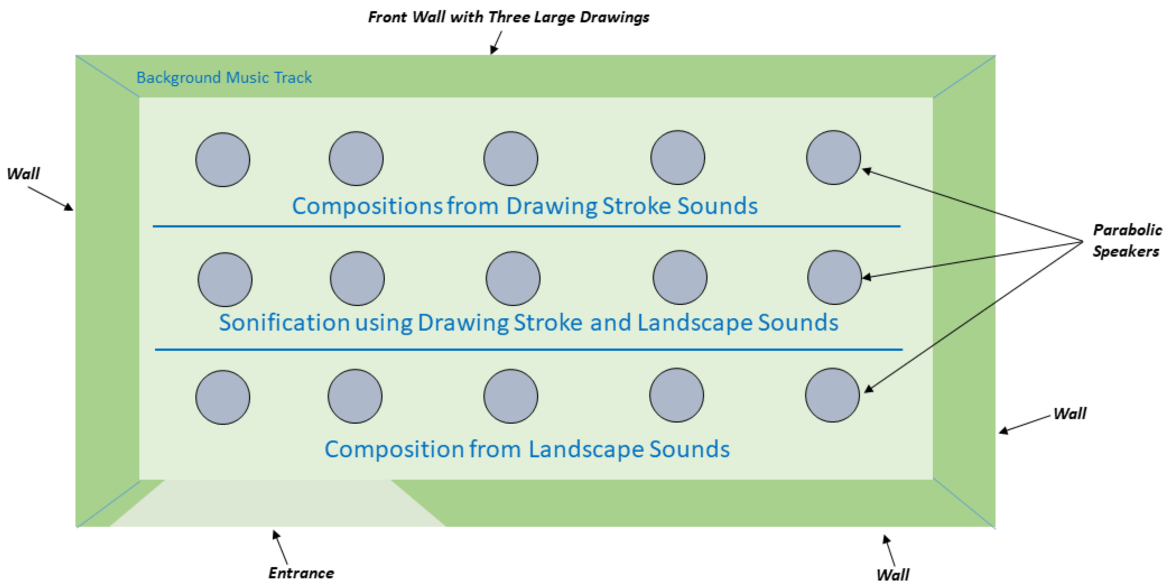

Fig. 6. Layout of the Oregon Project.

The detailed interactions with the user’s right hand as in the Eyes-Free Art design probe have been adapted to work in a larger room. When a user is further away from the painting, they do not use their hand to explore; instead, head tracking is used to determine the person’s lateral position. As the user enters the closest zone, the system smoothly transitions to tracking the person’s hand to affect the sound they are hearing. The installation was open for a nine-day long period to the public in October in Seattle [34] and a 15-day long period in Scotland, where it was viewed by many hundreds of people of all ages and levels of vision. While we did not track particular people across time, we were able to track the relative amount of time spent in each zone over the duration of the installation.

Over a six-day duration of recording people’s locations in The Oregon Project, we found that participants spent the most time in Zone 2 (40% of the time), followed by Zone 3 (35%), and the least amount of time in Zone 1 (25%) (Fig. 7). It is likely that participants clustered toward the middle of the room and the entrance, but also took the time to explore the other areas of the room (see Fig. 7).

The researchers observed the installation during one morning of the deployment in October 2016, during which four people who are blind visited (AI1-AI4). In addition, we observed The Oregon Project in a second installation during April 2017, where we interacted with one person who reported no vision and one with partial vision (AI5-AI6). Overall, we found that the participants enjoyed The Oregon Project. The participants expressed gratitude for the installation: “The existence of the project is so wonderful to me. [AI4]”

Similar to our findings for the Eyes-Free Art design probe, participants appreciated the fact that The Oregon Project is interactive. Specifically, AI5 appreciated that “the art responds in a two-way process.” AI1 noted the novelty of this experience "[The Oregon Project] enabled me to explore and appreciate the work from different perspectives in a very dynamic way... something I have never been able to do with art before." Participants found it compelling to interact with body movements in the space as opposed to learning about the artwork elsewhere: “It wouldn’t work if I was sitting and navigating with a joystick. [AI1]” AI6 said that they “liked moving to trigger the files.”

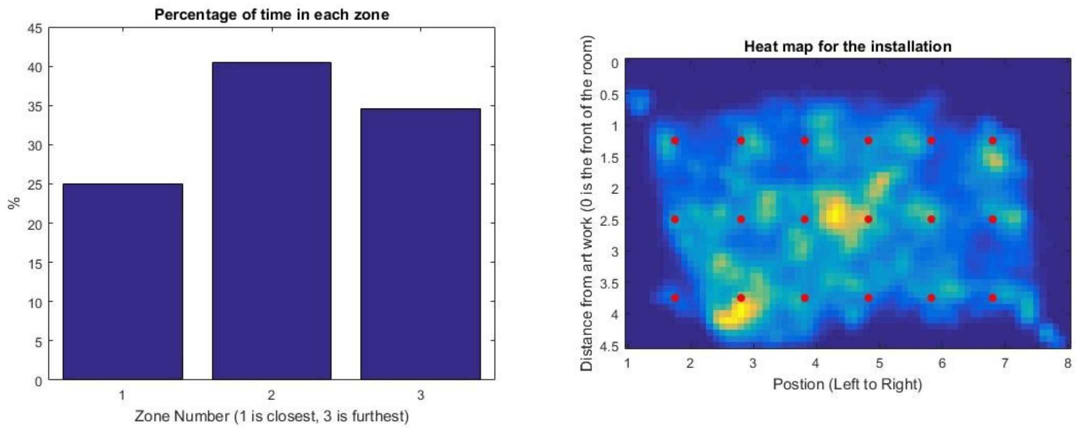

Fig. 7. Left: Percentage of time spent in each Zone. Right: Heat map of installation from the top view.

In the evaluation of our original probe, six participants preferred to hear a helpful verbal description at the beginning of the experience, and therefore we adopted this in our installation. AI1 found this helpful because "walking into [The Oregon Project] with no knowledge whatsoever would have potentially been bizarre." Additionally, AI1 interpreted the audio description as a "preparation for the experience rather than the experience itself." While the verbal description is a mandatory "zone," presenting it at the beginning gives participants context and the ability to immerse themselves in the artwork as they enter the interactive space.

Similar to our design probe results, the participants said that the experience was not fatiguing. The installation demonstrated that our modifications (including several spatialized speakers and multiple Kinects) allowed the experience to scale to multiple simultaneous viewers, and that the proxemic audio interface was valued by both sighted and blind art patrons; further, creating the proxemic audio design was valued as part of the artistic process by the work's creator.

> "I'm grateful, I haven't been to an art gallery in forever [because of her diminishing eyesight]… this [exhibit] reminds me there is art that is multimedia and I shouldn't limit my expectations. [AI3]"

7 DISCUSSION

We have shown through our design, evaluation of the Eyes-Free Art probe, and installation of The Oregon Project that we have developed a novel proxemic audio interaction, allowing people who are blind or low vision to explore paintings interactively and independently. Our work has given us valuable insights into the design of proxemic audio interfaces and into future iterations of Eyes-Free Art.

7.1 Lessons Learned

We learned that developing a proxemic audio interface, where the level of detail increases as a person moves closer to a painting, offers value for patrons who are blind or low vision. However, it is helpful for proxemic audio interfaces to begin with a more detailed verbal summary so the person has context when approaching the painting, which differs somewhat from the generic-to-detailed flow associated with visual proxemic interfaces.

Participants appreciated the interactivity of the proxemic interface, both in the ability to move among zones and in the ability to use their hand to explore more deeply within a given zone. The interaction techniques and sounds gave users a way to have a more immersive art experience. In addition, the person is able to spend as little or as much time as they would like in certain zones, which may parallel interactions of people who are sighted. These same experiences were also valued and enjoyed by sighted visitors to the art installation, suggesting that proxemic audio interactions may appeal to a broader audience beyond people with visual impairments.

From the evaluations of the design probe and installation, we have identified new research directions related to proxemic audio interfaces, and its applications to art. We discuss future research directions such as creating algorithms that automate portions of our process (e.g., mapping music genres to art genres, creating audio explanations), and creating easy-to-learn yet aesthetically-pleasing sonifications in Section 7.2.

Our design probe was an "interpretation" of existing artwork and the installation was an original art piece. While curated content can be used with this technique, creating tools and frameworks to allow artists to design proxemic interfaces as they create their artwork can provide a unique and valuable experience. These opportunities are discussed in Section 7.3.

7.2 Proxemic Audio Interfaces

We developed a proxemic audio interface where we changed the type of audio presentation (e.g. background music, changing music, sounds, and verbal) based on the person's proximity from a painting. However, it is possible to design a proxemic audio interface that utilizes only one type of audio presentation. For instance, an interface may involve a semantic zoom exclusively of verbal information, increasing the detail level of the verbal description as a person approaches the painting. For example, a verbal semantic zoom may contain: 1) the title and synopsis of the painting, 2) the genre, style, and history behind the painting, 3) the ability to explore with the hand to hear the names of the objects, and 4) the ability to explore with the hand to hear the names of the colors contained in the painting. The information presented in each zone may also change from our current approach depending on user (or artist or curator) preference. Generally, our design is that each zone of the proxemic audio interface would increase in detail as the user approaches an object (with the possible exception of a context-setting pseudozone), regardless of the type of the audio presentation.

Another consideration for proxemic audio interfaces may include whether the zones are intended for one or multiple users. The interface can be personalized, presenting a different experience based on the identity of the user, or be designed to allow multiple users to experience the interface together. If the interface is meant for individual exploration, the system may ensure the user's privacy by providing a single-piece headphone or bone conduction headphones, using directional audio, or allowing only one person at a time. While regular headphones may block out environmental sounds and conversations, a single-piece headphone as in Sotto Voce [2] has been used previously in museums settings to allow patrons to listen to their surroundings and in Morris et al. 2004 to allow groups to collaborate with one another while having private audio channels [30]. In addition, bone conduction headphones allow for people to listen to sounds in the environment (as stated by a blind participant in Rector et al. 2015 [38]). On the other hand, people might want to experience a painting or other object together, so designers would need to consider the tension between the technology and the social activity [9]. Further, designers need to determine how to present audio output based on multiple gestures from multiple users so that it is both understandable and conducive to discussion. Because proxemic audio interfaces may be public and support a "shared use" [43] of both private and collaborative group experiences, how one presents the audio information will need to be carefully designed so that the competing goals of private and public information are both possible while remaining decipherable. Our installation demonstrated one viable solution to multi-user proxemic audio, by using an array of parabolic speakers combined with gradual "bleeds" between zones rather than abrupt transitions.

Our work leaves open some questions of how to design proxemic audio interfaces. For example, the number of zones used in an interface may be determined based on the size of the space, the number of people visiting the space, or the pace that people need to walk through an exhibit. Another question may be how to design a proxemic interface for people with different levels of vision, including people who are totally blind, partially sighted, and totally sighted. People with different visual acuities may prefer the audio presentations to be of different lengths and detail. Whether to provide different interfaces based on the person's preference or provide a more universal interface is open to discussion, and may also be limited by practical concerns such as museum resources.

7.3 Eyes-Free Art in the Museum

We realize that evaluating a design probe with a lab study may not generalize to a museum experience [21]. We addressed potential issues in our installation, including multiple users encountering a painting from different trajectories in a crowded space where people may want to get close to the painting to appreciate the detail. However, we may not have captured the full dynamics of a museum experience, such as when there are multiple installations that all have proxemic audio that a patron must navigate amongst. A permanent installation is left for future work.

Another direction for future work is determining whether Eyes-Free Art should be used to interpret existing paintings or to compose audio experiences for new paintings; we worked with existing paintings when developing our design probe, as that was more practical for study. Some participants mentioned the risk of assigning audio to an existing painting without the artist's permission, particularly so with background music. The paintings selected in our study were in the public domain, so there is a decision to make of who is qualified to select sounds and music, whether it be a curator, art historian, anonymous crowd workers, or an automated algorithm. If an artist creates an experience using Eyes-Free Art as a standalone experience or to complement a visual art piece, such as in the installation created by Keith Salmon, then Eyes-Free Art can allow new mechanisms for enhancing a patron's interpretation and experience of a piece.

Because proxemic interfaces also have the goal of "immediate usability" [43], one future challenge is how to present a sonification without the necessity of training an individual. Currently, the user has to learn the mapping of red, green, and blue to instruments, as well as understanding changes in volume. Participants in the study suggested having more colors mapped to instruments to be comprehensive, but this would increase the amount of time needed for training. One possible solution may include a tutorial for sonification as the user is exploring the painting by mentioning the names of the colors for the first minute of interaction. Designing sonification techniques that balance the competing demands of learnability, expressivity, and audible aesthetics, and evaluating the efficacy of such sonifications, remains an open research problem.

8 CONCLUSION

We designed, developed, and evaluated Eyes-Free Art, a proxemic audio interface and installation that helps people who are blind or low vision explore paintings. We contributed the concept of a proxemic audio interface, demonstrated its feasibility, and discussed directions for how to integrate these interfaces in public settings. By conducting a lab-based study with 13 people who are blind or low vision, we found that Eyes-Free Art gave a good sense of the painting, in particular the colors and objects contained in the painting, and provided a rich, interactive interpretation of the art work that was appreciated by users. Our installation expanded on these ideas, and illustrated how an artist can craft proxemic audio experiences to convey artistic intent, as well as creating an installation that can be appreciated by multiple simultaneous visitors with varying levels of vision. We found that proxemic audio interfaces provided an immersive experience for both blind and sighted patrons. We hope that our work will generalize to researchers and designers who want to build immersive audio experiences that are based on tracking a user.

ACKNOWLEDGMENTS

We thank artist Keith Salmon and filmmaker Daniel Thornton for creating The Oregon Project installation. To view a video about this project visit: https://youtu.be/I--HA600Ui0

REFERENCES

[1] Cara Aitchison. 2003. From leisure and disability to disability leisure: Developing data, definitions and discourses. *Disability & Society* 18, 7 (2003): 955-969. DOI=https://doi.org/10.1080/0968759032000127353

[2] Paul M. Aoki, Rebecca E. Grinter, Amy Hurst, Margaret H. Szymanski, James D. Thornton, and Allison Woodruff. 2002. Sotto voce: exploring the interplay of conversation and mobile audio spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '02). ACM, New York, NY, USA, 431-438. DOI=http://dx.doi.org/10.1145/503376.503454

[3] Art Beyond Sight. About Art Education for the Blind. http://www.artbeyondsight.org/sidebar/aboutaeb.shtml

[4] Big Fish Audio. http://www.bigfishaudio.com/detail.html?527349
[5] The Broad. 2017. Accessibility. http://www.thebroad.org/visit/accessibility
[6] The Broad. 2017. Mobile App and Digital Tours. http://www.thebroad.org/visit/mobile-app-and-digital-tours
[7] Sofia Cavaco, J. Tomás Henriques, Michele Mengucci, Nuno Correia, Francisco Medeiros. 2013. Color sonification for the visually impaired. *Procedia Technology*, 9 (2013): 1048-1057. DOI= http://dx.doi.org/10.1016/j.protcy.2013.12.117
[8] Rene Chun. 2016. The SFMOMA's New App Will Forever Change How You Enjoy Museums. Wired. https://www.wired.com/2016/05/sfmoma-audio-tour-app/
[9] Areti Damala, Pierre Cubaud, Anne Bationo, Pascal Houlier, and Isabelle Marchal. 2008. Bridging the gap between the digital and the physical: design and evaluation of a mobile augmented reality guide for the museum visit. In *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts* (DIMEA '08). ACM, New York, NY, USA, 120-127. DOI: https://doi.org/10.1145/1413634.1413660
[10] Tilman Dingler, Markus Funk, and Florian Alt. 2015. Interaction Proxemics: Combining Physical Spaces for Seamless Gesture Interaction. In *Proceedings of the 4th International Symposium on Pervasive Displays* (PerDis '15). ACM, New York, NY, USA, 107-114. DOI=http://dx.doi.org/10.1145/2757710.2757722
[11] Lesley Fosh, Steve Benford, Stuart Reeves, Boriana Koleva, and Patrick Brundell. 2013. See me, feel me, touch me, hear me: trajectories and interpretation in a sculpture garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 149-158. DOI: http://dx.doi.org/10.1145/2470654.2470675
[12] Nicholas A. Giudice, Hari Prasath Palani, Eric Brenner, and Kevin M. Kramer. 2012. Learning non-visual graphical information using a touch-based vibro-audio interface. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility* (ASSETS '12). ACM, New York, NY, USA, 103-110. DOI=http://dx.doi.org/10.1145/2384916.2384935
[13] Lígia Gonçalves, Pedro Campos, and Margarida Sousa. 2012. M-dimensions: a framework for evaluating and comparing interactive installations in museums. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (NordiCHI '12). ACM, New York, NY, USA, 59-68. DOI=http://dx.doi.org/10.1145/2399016.2399027
[14] Cagatay Goncu, Anuradha Madugalla, Simone Marinai, and Kim Marriott. 2015. Accessible On-Line Floor Plans. In *Proceedings of the 24th International Conference on World Wide Web* (WWW '15). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 388-398. DOI: https://doi.org/10.1145/2736277.2741660
[15] Amy L. Gonzales, Thomas Finley, and Stuart Paul Duncan. 2009. (Perceived) interactivity: does interactivity increase enjoyment and creative identity in artistic spaces? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09). ACM, New York, NY, USA, 415-418. DOI=http://dx.doi.org/10.1145/1518701.1518767
[16] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. 2011. Proxemic interactions: the new ubicomp? *Interactions* 18, 1 (January 2011), 42-50. DOI=http://dx.doi.org/10.1145/1897239.1897250
[17] Rebecca Gross. 2014. The Importance of Taking Children to Museums. https://www.arts.gov/art-works/2014/importance-taking-children-museums.
[18] Edward T. Hall. 1966. The Hidden Dimension. Doubleday & Co, New York, NY, USA.
[19] Chris Harrison and Anind K. Dey. 2008. Lean and zoom: proximity-aware user interface and content magnification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08). ACM, New York, NY, USA, 507-510. DOI=http://dx.doi.org/10.1145/1357054.1357135
[20] Sergio E. Hernandez and Kenneth E. Barner. 2000. Tactile imaging using watershed-based image segmentation. In *Proceedings of the fourth international ACM conference on Assistive technologies* (Assets '00). ACM, New York, NY, USA, 26-33. DOI: http://dx.doi.org/10.1145/354324.354332
[21] Eva Hornecker and Emma Nicol. 2012. What do lab-based user studies tell us about in-the-wild behavior? Insights from a study of museum interactives. In *Proceedings of the Designing Interactive Systems Conference* (DIS '12). ACM, New York, NY, USA, 358-367. DOI=http://dx.doi.org/10.1145/2317956.2318010
[22] Horst Waldemar Janson and Anthony F. Janson. 1997. History of Art, Revised fifth edition. Prentice Hall, Inc.
[23] Eric Kabisch, Falko Kuester, and Simon Penny. 2005. Sonic panoramas: experiments with interactive landscape image sonification. In *Proceedings of the 2005 international conference on Augmented tele-existence* (ICAT '05). ACM, New York, NY, USA, 156-163. DOI=http://dx.doi.org/10.1145/1152399.1152428
[24] Nicolai Marquardt and Saul Greenberg. 2012. Informing the design of proxemic interactions. *IEEE Pervasive Computing* 2, (2012), 14-23.
[25] P.B.L. Meijer. 1992. An experimental system for auditory image representations. *Biomedical Engineering* 39, 2 (1992): 112-121.
[26] The Metropolitan Museum of Art. Audio Guide. http://www.metmuseum.org/visit/plan-your-visit/audio-guide.
[27] Microsoft. Body Tracking. https://msdn.microsoft.com/en-us/library/dn799273.aspx.
[28] Daniel Morris and Neel Joshi. 2003. Alternative "vision": a haptic and auditory assistive device. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '03). ACM, New York, NY, USA, 966-967. DOI=http://dx.doi.org/10.1145/765891.766097
[29] Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. 2006. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, USA, 1201-1210. DOI=http://dx.doi.org/10.1145/1124772.1124952
[30] Meredith Ringel Morris, Dan Morris, and Terry Winograd. 2004. Individual audio channels with single display groupware: effects on communication and task strategy. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (CSCW '04). ACM, New York, NY, USA, 242-251. DOI=http://dx.doi.org/10.1145/1031607.1031646

[31] Ann Morrison, Stephen Viller, and Peta Mitchell. 2011. Building sensitising terms to understand free-play in open-ended interactive art environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '11). ACM, New York, NY, USA, 2335-2344. DOI=http://dx.doi.org/10.1145/1978942.1979285

[32] Museum of Modern Art. Individuals who are Blind or Partially Sighted. http://www.moma.org/learn/disabilities/sight.

[33] Charles O'Neill and Kia Ng. 2008. Hearing images: interactive sonification interface for images. In *Proc. AXMEDIS 2008*, IEEE (2008), 25-31.

[34] The Oregon Project | studio99. https://msrstudio99.wordpress.com/2016/11/17/the-oregon-project/

[35] Betsy Phillips and Hongxin Zhao. 1993. Predictors of assistive technology abandonment. *Assistive technology* 5, 1 (1993): 36-45. DOI:https://doi.org/10.1080/10400435.1993.10132205

[36] Thorsten Prante, Carsten Röcker, Norbert Streitz, Richard Stenzel, Carsten Magerkurth, Daniel van Alphen, and Daniela Plewe. 2003. Hello. Wall – beyond ambient displays. In *Adjunct Proc. Ubicomp 2003*, ACM Press (2003), 277-278.

[37] Thierry Pun, Benoit Deville, and Guido Bologna. 2010. Sonification of Colour and Depth in a Mobility Aid for Blind People. In *Proc. ICAD2010*, GIT (2010), 9-15.

[38] Kyle Rector, Lauren Milne, Richard E. Ladner, Batya Friedman, and Julie A. Kientz. 2015. Exploring the Opportunities and Challenges with Exercise Technologies for People who are Blind or Low-Vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15). ACM, New York, NY, USA, 203-214. DOI: http://dx.doi.org/10.1145/2700648.2809846

[39] Andreas Reichinger, Anton Fuhrmann, Stefan Maierhofer, and Werner Purgathofer. 2016. Gesture-Based Interactive Audio Guide on Tactile Reliefs. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '16). ACM, New York, NY, USA, 91-100. DOI: https://doi.org/10.1145/2982142.2982176

[40] Andreas Reichinger, Stefan Maierhofer, and Werner Purgathofer. 2011. High-quality tactile paintings. *J. Comput. Cult. Herit.* 4, 2, Article 5 (November 2011), 13 pages. DOI=http://dx.doi.org/10.1145/2037820.2037822

[41] Seattle Art Museum. Art Beyond Sight Tours. http://www.seattleartmuseum.org/visit/accessibility#tou.

[42] Henrik Sørensen, Mathies G. Kristensen, Jesper Kjeldskov, and Mikael B. Skov. 2013. Proxemic interaction in a multi-room music system. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration* (OzCHI '13), Haifeng Shen, Ross Smith, Jeni Paay, Paul Calder, and Theodor Wyeld (Eds.). ACM, New York, NY, USA, 153-162. DOI: http://dx.doi.org/10.1145/2541016.2541046

[43] Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (UIST '04). ACM, New York, NY, USA, 137-146. DOI=http://dx.doi.org/10.1145/1029632.1029656

[44] Ron Wakkary and Marek Hatala. 2006. ec(h)o: situated play in a tangible and audio museum guide. In *Proceedings of the 6th conference on Designing Interactive systems* (DIS '06). ACM, New York, NY, USA, 281-290. DOI=http://dx.doi.org/10.1145/1142405.1142448

[45] Steven Wall and Stephen Brewster. 2006. Feeling what you hear: tactile feedback for navigation of audio graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, USA, 1123-1132. DOI=http://dx.doi.org/10.1145/1124772.1124941

[46] Mark Weiser. The computer for the 21st century. 1991. *Scientific American*, 265(3):94—104, September 1991.

[47] Tsubasa Yoshida, Kris M. Kitani, Hideki Koike, Serge Belongie, and Kevin Schlei. 2011. EdgeSonic: image feature sonification for the visually impaired. In *Proceedings of the 2nd Augmented Human International Conference* (AH '11). ACM, New York, NY, USA, Article 11, 4 pages. DOI=http://dx.doi.org/10.1145/1959826.1959837

Received February 2017; revised May 2017; accepted June 2017