2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR) | 978-1-6654-5725-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIVR56993.2022.00047

Toward inclusivity: Virtual reality museums for the visually impaired

Tycho Zaal
Information and Computing Sciences
Utrecht University
Utrecht, the Netherlands
tychozaal@gmail.com

Alkim Almila Akdag Salah
Information and Computing Sciences
Utrecht University
Utrecht, the Netherlands
a.a.akdag@uu.nl

Wolfgang Hürst
Information and Computing Sciences
Utrecht University
Utrecht, the Netherlands
huerst@uu.nl

***Abstract*—The access to virtual reality museums mostly relies on the visual sense, making it difficult if not impossible for visually impaired people to partake in the experience. We present a between-subjects study exploring if narrations and spatialized 'reference' audio combined with haptic feedback can be a sufficient replacement for the traditional use of vision in a virtual reality art museum. Our pilot study compares two implementations: A standard 'sighted' version that provides visual artifacts along with related acoustic narratives, a 'visually impaired' version with modified narratives and enhanced audio and haptics that was tested by visually impaired participants, as well as 'blindfolded' sighted individuals. Auditory and haptic feedback in the latter version were used to steer visitors towards specific virtual objects. Although the experiences of the visually impaired were obviously not statistically equivalent to the non-impaired group, results show that our method enabled them to experience the virtual reality museum adequately and find objects faster due to the additional auditory and haptic feedback.**

***Index Terms*—visually impaired, virtual reality, virtual museum, virtual exhibition**

I. Introduction

There has been increasing interest in using virtual reality (VR) as an enhancement of or even alternative to physical visits to cultural heritage sites, museums, and cultural tourism in general. It is reported that about 2000 VR museums did become available in the last few years [1]. The added value of VR in interacting with cultural and artistic artifacts and sites manifests itself in various ways. It offers opportunities to recreate authentic and immersive experiences that cannot be delivered in any other way, as in the recent Tate VR attraction where the studio of Modigliani that no longer exist is created in the VR environment [2]. Another example is the Salvador Dali Museum's VR installation that let visitors enter paintings and see the scenery from within [3]. Hürst et al. [4] turned 2D art into 3D experiences by extracting artistic style via algorithms which were further applied in VR for users to explore and be immersed in. These examples enhance the traditional way of experiencing art and culture [5]. Another advantage and opportunity VR museums offer is inclusiveness. Jung et al.'s study [6] focused on the experiences of children, elderly and disabled visitors using cultural heritage VR installations of otherwise inaccessible attractions, and reported positive results. However, empirical research on the adaptation and use of VR museums for visually impaired individuals is still scarce, and further research is needed to fully understand the challenges and opportunities in this field.

Around 285 million people worldwide have a visual impairment [7] and among them, around 23 million children and youths are estimated to be blind. Research showed that visually impaired adolescents partake in fewer leisure activities than their sighted peers, especially when it comes to visiting cinemas, theaters, and museums [8]. Furthermore, visually impaired students mainly use computers for schoolwork and to look up non-school-related information. Even though computers allow them to stay connected with sighted peers and help them cope with their impairment [9], games and virtual play often rely on ordinary graphical interfaces causing visually impaired adolescents to miss out on an important part of youth culture [10].

Enriching art exhibits with acoustic narratives in a VR setting could potentially provide a satisfying museum experience for the visually impaired. Yet, navigation in VR is often difficult for these target groups, even if their visual impairment is only partial. Via a pilot study including *sighted* (SI), *visually impaired* (VI) and *blindfolded* (BF) sighted teenagers, we demonstrate that audio narratives and a combination of auditory and haptic feedback for orientation do indeed have the potential to enable an adequate experience of an interactive VR museum for the visually impaired.

This paper is structured as follows. In Section II, we report difficulties visually impaired individuals experience in everyday life as well as in VR. We furthermore summarize VR technologies that are developed for VI in general, as well as for VR museums adapted for VI. Section III introduces our Virtual Reality Museum adaptation for VI individuals, along with the experiment setup. In Section IV, we present the results of the experiment by assessing various dimensions of the experiences reported by VI, SI, and BF participants. We conclude with a discussion in Section V.

II. Related work

Individuals with a visual impairment experience restrictions in their daily life. VI children exhibit lower levels of fitness than their sighted (SI) peers [11]. Even though their motor skills are proficient, they often lack the experience to be on the same developmental level as SI children [12], especially

2771-7453/22/$31.00 ©2022 IEEE
DOI 10.1109/AIVR56993.2022.00047

in navigating an environment. To do this successfully, VI individuals need to be aware of their surroundings and their own body. This is often referred to as *spatial awareness*, i.e. the ability to accurately locate objects in space, and position the body parts in order to make contact with these objects [13]. The same problem is observed when VI individuals navigate in VR environments. A solution to enhance spatial awareness is in using auditory cues. For example Freeman et al. used remote audio devices to improve orientation and navigation for VI, where the indicative audio would be played at fixed locations as 'reference sounds' [14]. It is furthermore observed that spatial awareness training that is done in simulated environments such as VR is directly transferable to real life [15].

In Virtual Reality, to convey spatial information through the use of audio, the system must first know where the user is in the 3D environment, and where the other objects are in relation to the user. Playing audio based on the position and orientation of the user is done using *Head-Related Transfer Functions (HRTF)* by adjusting the frequency response of the sound for each ear individually, to replicate as if it had been heard by a human in real life [16]. This can be used to create a sense of direction for the audio cues. When the user moves and rotates around, the audio stays at the same position in 3D world space, but depending on the relative position of the player it might shift from the left to the right ear. Prioritizing one ear over the other in this context is referred to as lateralization or spatialization [17]. Most game engines have built-in support for this, allowing audio to be played anywhere in the 3D environment, which is then automatically adjusted based on the player's position and orientation. These systems often only support horizontal cues, meaning differences in vertical positions are hard to distinguish [16].

Another sense that VI individuals often rely on is touch. Hence, for VR, haptic feedback is argued to be better suited for communicating information than audio [18], as it complements how VI people often use their sense of touch when exploring an environment [19, 20]. To that effect, Wagensveld and Zaal used haptic feedback to communicate upcoming events through anticipatory cues, as well as communicate immediate actions that the player needed to make. By implementing different patterns and magnitudes of vibrations for these purposes, the game was playable by both SI and VI players with similar performance. Haptic feedback has also been added to walking canes in VR environments in several studies with great success [21, 22, 23]. By adding vibration to the canes, participants experienced high levels of immersion [21].

Some games are designed with VI in mind, which use audio as the main stimuli for communicating gameplay elements [24]. These are called audio games and are split up into two different categories: *audio-only* games, which have no visual stimuli at all and *audio-based* games with some supportive visual content that is non-essential for the gameplay [25]. By having some visuals in audio-based games, SI individuals have a chance to see other people play, and low-vision players might still be able to use their vision to some extent. There are also a few VR games that were designed for VI. *Blind* is a VR puzzle game with audio cues for VI [26]. *Blind Swordsman* also makes use of 3D audio cues, in this case, to locate enemies. The interaction with the enemy is done via the press of a button which activates a sword to swing [27]. *Virtual Showdown* make use of lateralization to help locate a ball that needs to be sent to the opponent. Since the gameplay is designed in a way where the player needs to physically move to hit the ball, the resulting VR experience is more immersive.

There are only a few virtual museums that were designed with VI in mind. Robotics lab Percro designed an exoskeleton arm that provides haptic feedback upon collision with virtual artifacts [28]. A similar project was done in Prague, where haptic gloves were used in combination with VR equipment, allowing users to walk around and 'touch' the virtual objects [29]. Haptic gloves for VI were first experimented with in 2004 when CyberGlove made its first appearance [30]. Kreimeier et al. used a VR treadmill and haptic feedback to allow users to walk around in a virtual environment and feel nearby objects [31]. A VR controller would vibrate upon collision with a nearby virtual object, allowing users to 'feel' its shape and exact location. Although the study focused more on the navigation aspect, it showed that simple haptic feedback can be useful for communicating the location of nearby objects. Equipment such as the aforementioned examples allows users to 'feel' virtual objects, which is a lot more stimulating than reading or hearing about the objects.

Although these additional VR-compatible devices seem promising, they are often not available for commercial purchase, or very expensive. This makes the created experiences less accessible since the (financial) entry barrier is so high. The goal of this paper is to create an inclusive and easily accessible virtual reality museum. As such, the used equipment will be limited to equipment that is included in commercial VR sets.

III. Methodology

For this paper, an existing interactive VR Museum is expanded with additional features to accommodate for the needs of visually impaired users. In this section, we will first summarize the attributes of this original VR Museum setup, then detail the changes implemented, before moving on to the experiment setup.

A. VR Museum Design – SI and VI versions

The original VR implementation is focused on the famous painting *The Night Watch* by Rembrandt and provides an interactive virtual museum environment. At the start of the experience, the historic figures depicted in the painting come to life (see Figure 1). The players can interact with these characters to learn about their history and their importance, which is communicated using narrative voice lines. Additionally, objects that are being held by these historic figures are put on display on the flip side of the room. These can be interacted with to trigger sound effects or animations.

Several issues arise for the existing version of the museum when removing sight as the primary sense, which led to the

Fig. 1. A Snapshot from the original VR implementation of *Night's Watch*

development of a *VI version* of the original VR museum experience. Although the application has a variety of narration voice lines, none exist that clearly explain what can be seen on the painting. Besides this, some sentences are too complicated for children to understand. As such, several voice lines should be rewritten or added for visually impaired users. Another issue is how precise aiming is required for selecting the historic figures and their individual voice lines. To accommodate for this, aiming is done using the VR HMD instead, and successful hits are detected using the horizontal angle to the target. The voice lines are also combined into one, which will then be played upon interaction.

In addition, it can be difficult for VI to navigate the virtual environment when exploring different artifacts. In our set-up, in order to compensate for the lack of sight of VI, emphasis is placed on the use of audio and haptics. Although bright colors can help low-vision people with detecting objects, this will not work for fully blind people. As such, emitting indicative audio cues will be played at the location of the historic figures and the artifacts instead. This draws attention to the characters and objects and encourages the players to interact with them. Handheld controllers vibrate when the player is looking in the direction of a virtual object. Different vibration magnitudes of the two controllers can also be exploited to steer the player in certain directions, which will be activated after 20 seconds of non-interaction. This method is henceforth referred to as the 'steering' method. When the player is looking at a virtual object, they can use the trigger button on the handheld controller to interact with it. For the objects, this starts a short animation. For the historic figures, it queues a narrative story.

An overview of all the key differences between the original, SI, and the VI version of our implementation can be seen in Table I. Most of the differences affect the interaction aspect of the experience. In addition, we tested a *BF version*, which is the VI version tested by blindfolded sighted participants.

The VR experience was made entirely in Unity. SteamVR, which is a VR library made by Valve, was used also, as it has built-in support for most commercially available VR devices. For the experiments, the HTC Vive and its controllers were used. To ensure spatial audio could be properly heard, all participants wore high-quality over-ear headphones, namely the HyperX Cloud II. Unity's AudioSource component was used for spatializing and playing the audio. All diegetic audio is converted from 2D to 3D, where the audio now plays at the location of the corresponding virtual object. For example, the footsteps of the historic figures walking around now originate from the actual feet of the historic figures. Non-diegetic audio, such as the narration and background audio, has also been adjusted. An ambisonic effect has been applied to the audio files, creating a 'skybox' effect. This makes it appear as though it does not have one origin point while being less static than 2D stereo audio. This change makes it sound as though it is present in the virtual world, while still being distinguishable from the diegetic audio. Lastly, Unity's Reverb Zones were used to simulate realistic acoustics. Different settings affect the distances, pitch, and reverb of all audio in the room.

B. Pre-test

A pre-test was held with 8 VI teenagers to measure how well they were able to detect the origin point of spatialized audio in a virtual environment. In a game-like setup, players had to directly look at whichever object was making noise. The audio was spatialized, meaning it contained a sense of direction. Accuracy and speed are key variables for measuring how well the implementation worked. The accuracy is measured by taking the horizontal angle between the user's gaze direction and the direction to the target object, where a difference of 45 degrees results in an accuracy of exactly 0 percent. The median *accuracy* of the pre-test is 74.4%, and the average *accuracy* is 65.1%. This suggests that participants were able to consistently precisely locate the audio origin point. However, the few times they did make a mistake, they were off by a large margin (often by 180 degrees). The average *time between interactions* was 4.2 seconds, and the median *time between interactions* is 2.9 seconds. Considering the participants had a maximum of 15 seconds to locate the new object and perform the interaction, they were fairly fast with initiating new interactions.

Alongside the pre-test, an interview with an accessibility expert and teacher at a school for VI was conducted. The interview and results can be summarized as follows: 1) Make the experience audio-based instead of audio-only to enable VI individuals with low vision to use their sight for navigation and spatial awareness. 2) Emphasize sensory feedback, such as audio and vibrations, during the experience. 3) Communicate as much visual information as possible through auditory or haptic cues. 4) The narration should be as clear as possible since for VI individuals audio is often capable of carrying the same weight as visuals. 5) The steering method should be used as a last resort instead. By making this change, users have more control over which object or figure they want to interact with next, while also having control over the pace. Figure 2 shows the updated flow of the experience.

The pre-test also revealed that slightly aiming more to the sides of an object or a character can sometimes leads to unwanted interactions. To avoid this problem, all objects in the room should be spaced apart. Figure 3 shows the layouts of the original version (a), the pre-test (b) and the final VI version (c).

TABLE I
DIFFERENCES BETWEEN THE INITIAL, SI AND THE VI VERSION OF THE VR MUSEUM

Original version	Sighted version	Visually impaired version
Aim using controllers	Aim using controllers	Aim using HMD
All three dimensions are used when aiming	All three dimensions are used when aiming	Vertical axis is disregarded when aiming
Voicelines can be selected individually	Voicelines are combined and play upon interaction	Voicelines are combined and play upon interaction
Colliders are used to determine hits	Colliders are used to determine hits	Angle to target is used to determine hits
No additional support for helping find targets	Mild vibrations and spatial audio is added	Amplified vibrations and spatial audio is added
Exploring method	Exploring method	Steering and exploring method

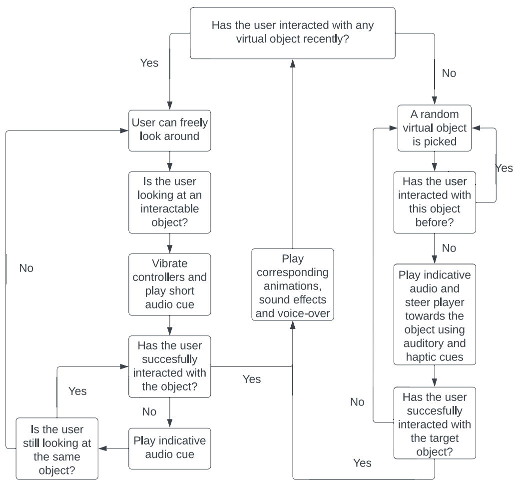


Fig. 2. Flow of the VI version

In the original version the figures stand too close to each other and some objects are obstructed by vertical pillars. During the pre-test we used an audio-only setup without visuals, and tested the sources of audio-cues and if they were placed far enough from each other to be easily discernible as individual objects and figures. For the main-test the historic figures from *The Night Watch* are distributed following a similar layout to the static objects' placements in the pre-test. Alongside the change in room layout, interacting with objects is also temporarily disabled when a user is already interacting with another virtual object to prevent accidental interactions or overlapping audio.

C. User Study

This section will elaborate on the participant groups, the experiment design, and the recorded measures.

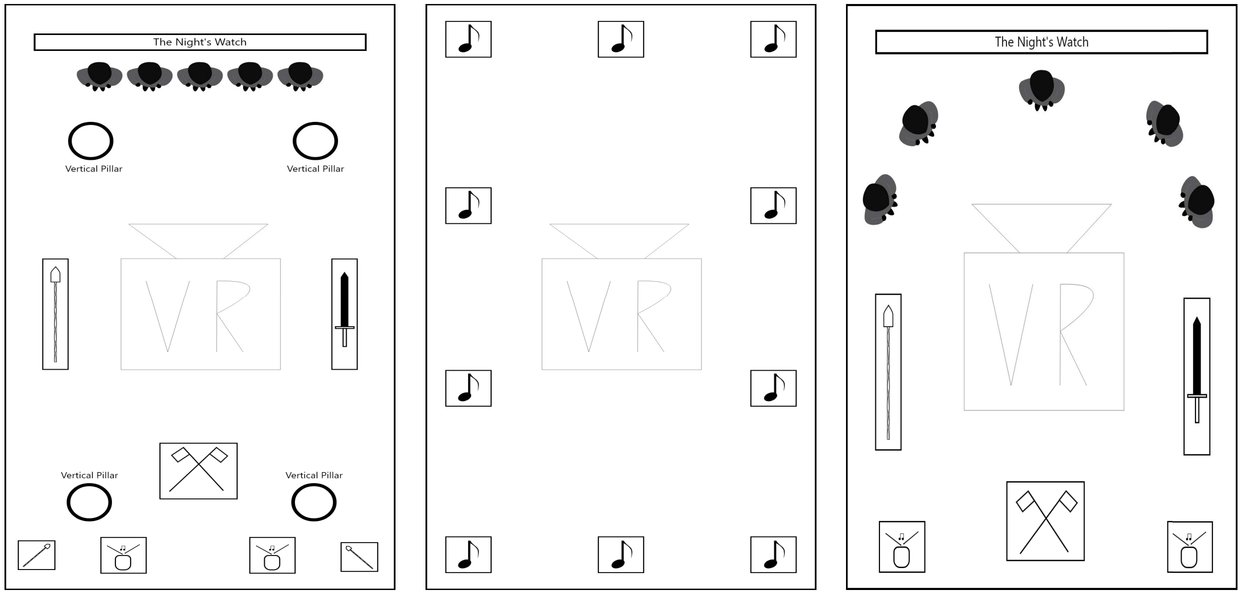
Participants. Because of the difficulty of gathering VI participants, it is fairly common to use SI individuals that are blindfolded to test implementations that are specifically designed for VI [32]. Depending on the implementation, the participants are not necessarily physically blindfolded, but, for example, visuals can also be blocked out in a VR environment. VI individuals tend to use their senses differently than those with full vision, especially their hearing, and they usually have a better sense of rhythm [33]. To test whether this affects player performance, Wagensveld and Zaal used VI, SI, and BF participants for their rhythm audio-based game [34]. The difference in performance between VI and BF participants was minimal. Hence, in general, one would expect results from a BF group to be similar to the ones from VI users.

Motivated by this, we also asked some SI individuals to experience our VI implementation while the participants are blindfolded. These participants will henceforth be referred to as the *BF group*. The other two participant groups are the *SI group*, and the *VI group*. In total 34 people participated in the experiments, of which 18 identified as male, and 16 identified as female. Of the participants, 10 were VI, 10 were BF and 14 were SI. All participants were high school students between 11 and 20 years old (M = 15.52, SD = 2.71). One 28-year-old VI employee of the high school was asked to participate as well since this person is an expert in the field of games and digital experiences for VI. The VI group consisted of participants with a variety of visual impairments, ranging from low vision to complete blindness.

Experiment design. A between-subjects design was chosen for the experiments, that is, the VI group and BF group tested the VI implementation, respectively, and the SI ones tested the original VI implementation with the conventional interaction method from the original version, and without the steering method (see Table I). Each participant first received an oral explanation about the controls and the experience. They were instructed to explore the experience in their preferred order at their own pace. Afterward, they were asked to fill in a self-report qualitative survey. For the VI participants, an SI employee of the school read the questions out loud instead.

Measures. The experiment aims to investigate how haptic cues combined with spatial audio compare to traditional vision-based interaction in a VR museum. The independent variable here is the *participating group* since the same VR experience is prepared for three different sets of participant groups (SI, VI, and BF).

Dependent variables included both quantitative and qualitative data that was recorded during the experiments. The quantitative data include which objects were interacted with, the (average) interaction accuracy, how long the entire session lasted, and how accurately the user used the controls. The qualitative data mainly focused on demographic information, and the user experience, which includes Likert-scale questions about ease of use, enjoyability, feeling of presence, understanding of the auditory cues, and understanding of the environment. The



(a) (b) c

Fig. 3. (a) Original Layout (b) Pre-test Room Layout (i.e. audio-only with haptics), (c) Main-test Room Layout (with haptics, visuals, and audio cues)

qualitative data is collected via a survey based on the Virtual Experience Text (VET). VET is a survey instrument used to measure five dimensions of Experience Design (ED) [35]: sensory, cognitive, affective, active, and relational. For our experiment, sensory, cognitive, and relational dimensions were especially important.

IV. Results

In this section, we present the results obtained from the experiment. We compare both quantitative and qualitative data from the three participant groups.

How well the users were able to interact with the environment can be analyzed by looking at a variety of variables. The *accuracy* of each interaction shows how close the player aimed to the middle of the virtual object. How long it took the player to find and precisely aim at the object is measured using the *time between interactions*. The *total playtime* shows how long it took the user to go through the entire experience. When combining these quantitative measures with self-report survey questions regarding spatial awareness, understanding of the controls, the audio and the environment, conclusions can be drawn about the effectiveness of the adaptation. The statistical results for these measures can be seen in Table II.

Tests for significant differences and statistical equivalence were done for all relevant measures. Equivalence Tests are used to look for statistical equivalence between groups. This serves as an indication of inclusiveness, where equivalence means that users have similar experiences or performances. The equivalence test was also used in order to reject the presence of the smallest effect size of interest (SESOI). The equivalence bounds were set to a standardized effect size (ES) of 0.5 [36], also known as Cohen's medium-sized effect [37]. These bounds are used to determine the range in which differences are smaller than what is considered meaningful. When statistical equivalence is found, the variable can be deemed as inclusive, since it was experienced and scored to an equivalent degree among (a combination of) the three participant groups.

A. Spatial awareness

To improve the spatial awareness of VI participants, spatial audio and haptic cues were added to help with locating objects. Narrative audio and manually-made 'reference' sounds, which complemented the features of the virtual objects, allowed users to understand what was happening around them, and which objects were in the room with them. In order to analyze the effectiveness of these implementations, the interaction accuracy, understanding of audio cues and environment, and feeling of presence are analyzed.

Interaction Accuracy. The interaction data from all three participant groups were compared in a one-way ANOVA test, showing statistically significant differences in the accuracy of each group ($p = 0.005$). The test found statistical equivalence with a 90% confidence interval (CI) between the accuracy means of the BF and the SI group (-0.34 - 0.093), but none between other combinations of participant groups. An overview of the results can be seen in Table III.

Audio Cues. The VI and the SI group have similar performances in following the audio cues, with the BF group following shortly behind. Statistical equivalence (CI = 90%, ES = 0.5) was only found between the SI and the VI group (-0.162 - 0.274). Most participants mentioned that they could properly understand the audio cues during the experiment. They also mentioned how when the game tried to steer them towards a

TABLE II
ALL RELEVANT MEASURES AND THEIR STATISTICAL RESULTS

Measure	Scale	Type	Significant differences	Statistical equivalence
Accuracy	Percentage 0-100	Quant. measure of VR experience	✓	SI & BF
Understanding of audio	Likert scale 1-5	Survey Q16	X	VI & SI
Feeling of presence	Likert scale 1-5	Survey Q13	✓	VI & BF
Time between interactions	Seconds	Quant. measure of VR experience	X	VI & SI & BF
Total playtime	Seconds	Quant. measure of VR experience	✓	SI & BF
Understanding of controls	Likert scale 1-5	Survey Q15	✓	X
Understanding of environment	Likert scale 1-5	Survey Q17	✓	X

TABLE III
ACCURACY DISTRIBUTION PER PARTICIPANT GROUP

Participant group	**Mean (± Standard deviation)**
Visually impaired	77.22% (23.97)
Blindfolded	85.20% (20.51)
Sighted	83.36% (8.79)

TABLE IV
TIME BETWEEN INTERACTIONS DISTRIBUTION PER PARTICIPANT GROUP

Participant group	**Mean (± Standard deviation)**
Visually impaired	8.24s (6.77s)
Blindfolded	11.41s (7.45s)
Sighted	9.25s (5.20s)

TABLE V
DIFFERENCES BETWEEN INTERACTION TYPES

Statistic	**Exploring**	**Steering**
VI usage	90.72%	9.28%
BF usage	97.00%	3.00%
Average accuracy	82.50%	72.25%
Average time between interaction	7.16s	4.31s

specific object, they were able to locate the spatialized audio. On average, the SI group rated their understanding the highest (4.43), closely followed by the VI (4.39), followed by the BF (3.90).

Understanding the Environment. All three groups were able to properly understand their surroundings. That being said, there were statistically significant differences ($p < 0.001$) and no statistical equivalence was found between any of the three groups. The VI still had a high mean environment understanding (4.17), although not nearly as high as the SI (4.49). Similarly to the understanding of the audio cues, the mean understanding of the environment of the BF (3.80) is a lot lower than the other two groups.

Looking at the survey question about the feeling of presence, statistically significant differences were found ($p < 0.001$), and statistical equivalence was found between the VI and BF group (-0.468 - 0.003). This is likely because the average feeling of presence of the SI is a lot higher (4.66) than those of the VI (4.07) and BF (3.8). BF participants are used to relying on vision in real life, meaning the VR experience is likely more dissimilar for them than the other groups. We expect this caused a difference in the understanding of audio and environment as well as the feeling of presence.

B. Controls

To assess the usability of the controls in detail we have prepared a series of controls-related survey questions. On top of these, the time between interactions also gives insight into how well participants were able to use the controls and interacted with the objects. In general, we observed that by changing the interactions to keep them exclusively on the horizontal axis, and using the user's head gaze direction for collision detection, the VI participants were able to successfully use the controls and interact with the environment at a statistically equivalent pace as the SI and the BF group. Table IV shows the time between interactions' distribution of each participant group.

There was no statistical difference in the mean time between interactions ($p = 0.589$) and there was statistical equivalence (CI = 90%, ES = 0.5) between all participant groups. The smallest difference was present between the VI and the BF group (-0.349 - 0.12). One reason for this difference might be in the different interaction method introduced in the VI and BF version. The VI and BF version had an additional steering method alongside the exploring method. Table V shows the differences in usage, accuracy, and time between interactions for both methods. Although the additional steering method allowed for faster interactions, the lack of feedback and clarity hindered its usability.

The understanding of the controls was rated the highest by the SI group, closely followed by the BF group. The VI rated the ease of controls and the understanding of the controls the lowest due to the lack of feedback and clarity for using the steering method. Hence, there were statistically significant differences ($p < 0.001$), and there was no statistical equivalence between the three groups. In the end, all groups were positive about being able to understand how and when to use which controls.

C. Supplementary data

We also analyzed the effect of external factors, such as age, gender, or experience for the VI, BF and SI participant groups. A one-way ANOVA test showed statistically significant differences in mean accuracy between the age groups ($p = 0.008$) as well as between the user's experience with virtual reality (p =

0.008). Older age and more VR experience tend to positively influence accuracy. On a scale of 1-5, the VI participants had the least VR experience on average (1.97), followed by the SI group (2.42). The BF group had the most experience with VR, with an average score of 4.0 out of 5. For BF group we prioritized older students, as working with a blindfold has its challenges. Hence, BF participants have a higher score of age and experience.

The participants were told to take their time to browse the VR museum. Still, the total playtime might indicate the existence of other factors. There were statistically significant differences in the mean total playtime ($p < 0.001$). A variety of external variables, such as art and history interests, experienced boredom, museum experience, and prior VR experience were investigated in the survey to see whether they affected the total playtime. In the end, no statistically significant differences were found. The total playtime was only statistically equivalent (CI = 90%, ES = 0.5) between the BF and SI group (-0.437 - -0.004). A comparison of the total playtime of all three groups can be found in Table VI.

TABLE VI
TOTAL PLAYTIME DISTRIBUTION PER PARTICIPANT GROUP

Participant group	**Mean (± Standard deviation)**
Visually impaired	355.45s (14.835s)
Blindfolded	349.80s (15.820s)
Sighted	346.18s (16.858s)

D. Observations and additional feedback

A variety of feedback was given by participants or observed during the experiments. The SI participants enjoyed the additional haptic and auditory feedback they received when looking at an object, stating that it made the experience more immersive.

Although the VI participants were able to recognize most of the haptic and auditory feedback they received, some mentioned that it could have been further amplified. A few sound effects were confusing yet intriguing, such as the audio that played during the sword fight. Participants mentioned how some additional feedback would be appreciated to let them know when a previous interaction has officially ended, or when they successfully interacted with an object or historic figure. A voice line suggesting the player to follow the sound when they are being steered also would have been appreciated.

Some participants were also confused by the steering method since they expected the vibrations to increase in power the closer they were to the object, rather than the other way around. Due to the perceived subtlety of the auditory feedback when being steered, some participants thought the audio was simply ambient noise. Once they understood that they had to look for the object that was playing the audio, they were able to quickly find and interact with it.

When it came to the narrative voice lines, some VI participants mentioned how some wording could have been more comprehensible. For example, mentioning the color of an item might not be understandable by some, especially if they were born with a visual impairment. Adding audio cues to redirect the user's vision to specific items was also a mentioned suggestion. This can be paired especially well with the narrative voice lines, which sometimes mention specific clothing or accessories of the historic figures. The VI participants applauded the realism of the experience, with several participants mentioning how they completely forgot that they were actually in virtual reality.

V. CONCLUSION

This paper developed an altered version of a virtual reality museum with the aim to verify if it can be adequately experienced by visually impaired teenagers. The virtual reality museum consists of historic figures and artifacts, which can be interacted with to prompt animations, sound effects, or narrative voice lines. A between-subjects study was designed in order to explore if narrations and spatialized 'reference' audio combined with haptic feedback is a sufficient replacement for the traditional use of vision. Alongside the VI and the SI version, a BF version was created, which blocks all vision in its entirety and was played by several SI participants to get supplementary data to that of the VI group.

Although significant differences were found in the mean accuracy of each of the three versions, all three participant groups managed to successfully interact with all virtual objects. The pace at which they did this was statistically equivalent, suggesting the adapted interaction method is just as fast to use, just slightly less accurate. An additional steering method was added to the VI version, which uses auditory and haptic feedback to steer players towards specific virtual objects. This method allowed users to find objects even faster, but due to a lack of clear feedback was found to be somewhat confusing. Even though there was no statistical equivalence in audio understanding, presence, and understanding of the environment, the VI participants positively graded these in the qualitative survey. As such, even though the VI did not have a fully statistically equivalent experience as the others, they were able to adequately experience the virtual reality museum. We can therefore conclude that the proposed idea of modifying the audio narratives and adding audio and haptics to support navigation to a traditional VR museum experience is promising and, if further developed, could lead to a more inclusive experience for the visually impaired.

ACKNOWLEDGMENT

We would like to express our gratitude to both ING and Bartiméus for their support and cooperation throughout the study.

REFERENCES

[1] Charlotte Coates. *Museum VR creates new ways for visitors to explore*. Aug. 2020. URL: https://blooloop.com/museum/in-depth/museum-vr-experiences/.

[2] Tula Giannini and Jonathan P Bowen. “Rethinking museum exhibitions: Merging physical and digital culture—past to present”. In: *Museums and Digital Culture*. Springer, 2019, pp. 163–193.

[3] *Dreams of Dalí in Virtual Reality*. en-US. URL: https://thedali.org/exhibit/dreams-of-dali-in-virtual-reality/.

[4] Wolfgang Hürst, Ferdinand de Coninck, and Xhi Jia Tan. “Complementing artworks to create immersive VR museum experiences”. In: *Advances in Computer Entertainment Technology 13th*. 2016, pp. 1–6.

[5] Timothy Jung et al. “Effects of virtual reality and augmented reality on visitor experiences in museum”. In: *Information and communication technologies in tourism 2016*. Springer, 2016, pp. 621–635.

[6] Timothy Hyungsoo Jung and M Claudia tom Dieck. “Augmented reality, virtual reality and 3D printing for the co-creation of value for the visitor experience at cultural heritage places”. In: *Place M&D* (2017).

[7] Donatella Pascolini and Silvio Paolo Mariotti. “Global estimates of visual impairment: 2010”. In: *British Journal of Ophthalmology* 96.5 (2012), pp. 614–618.

[8] Ulla Kroksmark and Kersti Nordell. “Adolescence: The age of opportunities and obstacles for students with low vision in Sweden”. In: *Journal of Visual Impairment & Blindness* 95.4 (2001), pp. 213–225.

[9] Jens P. Pfeiffer and Martin Pinquart. “Computer use of adolescents with and without visual impairment”. In: *Technology and Disability* 25.2 (July 2013), pp. 99–106.

[10] Aurélie Buaud et al. “Multimedia Games for Visually Impaired Children”. In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2002, pp. 173–180.

[11] D.L. Blessing et al. “The Effects of Regular Exercise Programs for Visually Impaired and Sighted Schoolchildren”. In: *Visual Impairment & Blindness* (1993).

[12] L.H. Schneekloth. “Play Environments for Visually Impaired Children”. In: *Journal of Visual Impairment & Blindness* 83.4 (1989), pp. 196–201.

[13] D.L. Chin. “Dance Movement Instruction: Effects on Spatial Awareness in Visually Impaired Elementary Students”. In: *Journal of Visual Impairment & Blindness* 82.5 (1988), pp. 188–192.

[14] Euan Freeman et al. “Audible beacons and wearables in schools: Helping young visually impaired children play and move independently”. In: *Human Factors in Computing Systems*. 2017, pp. 4146–4157.

[15] Paul N. Wilson, Nigel Foreman, and Danaë Stanton. “Virtual reality, disability and rehabilitation”. In: *Disability and Rehabilitation* 19.6 (1997), pp. 213–220.

[16] Camilla H. Larsen et al. “Differences in Human Audio Localization Performance between a HRTF- and a Non-HRTF Audio System”. In: *Audio Mostly Conference*.

[17] Ryan Wedoff et al. “Virtual Showdown: An Accessible Virtual Reality Game with Scaffolds for Youth with Visual Impairments”. In: Apr. 2019, pp. 1–15.

[18] Angel Walia et al. “HapTech: Exploring Haptics in Gaming for the Visually Impaired”. In: *Conference on Human Factors in Computing Systems*. 2020, pp. 1–6.

[19] Jennifer L Burke et al. “Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis”. In: *Multimodal interfaces 8th*. 2006, pp. 108–117.

[20] Bridger Scott Hodges. *The Effects of Haptics on Rhythm Dance Game Performance and Enjoyment*. Brigham Young University, 2018.

[21] Jinmo Kim. “VIVR: Presence of Immersive Interaction for Visual Impairment Virtual Reality”. In: *IEEE Access* 8 (2020), pp. 196151–196159.

[22] Andreas Kunz et al. “Virtual Navigation Environment for Blind and Low Vision People”. In: 2018.

[23] Yuhang Zhao et al. “Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation”. In: 2018, 1–14.

[24] Franco Eusébio Garcia and Vânia Paula de Almeida Neris. “Design Guidelines for Audio Games”. In: *HCI Applications and Services*. 2013, pp. 229–238.

[25] Johnny Friberg and Dan Gärdenfors. “Audio games: new perspectives on game audio”. In: *Advances in computer entertainment technology 2004*, pp. 148–154.

[26] *Blind*. en-US. URL: https://www.fellowtraveller.games/blind (visited on 10/01/2022).

[27] *Blind Swordsman*. en. URL: https://devpost.com/software/blind-swordsman (visited on 10/01/2022).

[28] Gunnar Jansson, Massimo Bergamasco, and Antonio Frisoli. “A new option for the visually impaired to experience 3D art at museums: Manual exploration of virtual copies”. In: 5 (July 2009).

[29] *Museum uses Virtual Reality to allow blind people to ‘see’ famous sculptures*. URL: www.museumnext.com.

[30] Georgios Nikolakis et al. “CyberGrasp and PHANTOM Integration: Enhanced Haptic Access for Visually Impaired Users”. In: (Jan. 2004).

[31] Julian Kreimeier and Timo Götzelmann. “First Steps Towards Walk-In-Place Locomotion and Haptic Feedback in Virtual Reality for Visually Impaired”. In: 2019.

[32] Bei Yuan and Eelke Folmer. “Blind hero: enabling guitar hero for the visually impaired”. In: *Computers and accessibility 10th ed*. 2008, pp. 169–176.

[33] Tiziana Vercillo, David Burr, and Monica Gori. “Early visual deprivation severely compromises the auditory sense of space in congenitally blind children.” en. In: *Developmental Psychology* 52.6 (2016), pp. 847–853.

[34] Lisa Wagensveld and Tycho Zaal. “Feel the Rhythm: replacing visual cues by haptics in a rhythm game for the visually impaired”. unpublished. 2021.

[35] Dustin B. Chertoff, Brian Goldiez, and Joseph J. LaViola. “Virtual Experience Test: A virtual environment evaluation questionnaire”. In: *2010 IEEE Virtual Reality Conference (VR)*. 2010, pp. 103–110.

[36] Daniël Lakens, Anne M. Scheel, and Peder M. Isager. “Equivalence Testing for Psychological Research: A

Tutorial". In: *Advances in Methods and Practices in Psychological Science* 1.2 (2018), pp. 259–269.

[37] Jacob Cohen, ed. *Statistical Power Analysis for the Behavioral Sciences*. 1988.