UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA

Accessible VR: Investigating Auditory Object Representations for Blind People

Rodrigo Henriques Nogueira

Mestrado em Engenharia Informática

Trabalho de Projeto orientado por:
Prof. Doutor João Pedro Vieira Guerreiro
Prof. Doutor André Filipe Pereira Rodrigues

2023

Acknowledgements

I would like to express my heartfelt gratitude to the individuals who have played a significant role in the completion of this thesis. First and foremost, I extend my deepest appreciation to my family. Their unwavering support, love, and understanding have been the bedrock of my journey. Their belief in my abilities and their sacrifices have enabled me to pursue this endeavor with determination and resilience. I am grateful for my friends who have been there for me every step of the way. Their encouragement, camaraderie, and willingness to lend a listening ear have provided immeasurable support throughout this thesis. Their friendship has made the challenges more manageable and the successes more joyful. I would like to extend my sincere appreciation to my thesis advisors, João Guerreiro, and André Rodrigues. Their guidance, expertise, and continuous support have been instrumental in shaping this research project. Their insightful comments, thought-provoking discussions, and meticulous review of my work have pushed me in the right direction. I am grateful for their mentorship and the opportunities provided to enhance my research skills. Lastly, I would like to express my gratitude to all the participants who took part in this study. Their willingness to share their experiences and insights has contributed significantly to the findings and outcomes of this research. In conclusion, I am deeply grateful to my family, friends, and advisor for their unwavering support, encouragement, and belief in my abilities. Without their presence in my life, this thesis would not have been possible. Their love, understanding, and guidance have been invaluable. Thank you for being a part of this journey and for enriching my academic experience.

This work was supported in part by FCT through the AccessVR project, with reference 2022.08286.PTDC (http://doi.org/10.54499/2022.08286.PTDC) and through the LASIGE Research Unit (UIDB/00408/2020, UIDP/00408/2020)

To my family.

Resumo

A realidade virtual (RV) cada vez oferece várias aplicações, no entanto, as pessoas com deficiências visuais enfrentam dificuldades de acesso. Existem ambientes de RV para pessoas com deficiências visuais, mas estes são limitados em funcionalidade e realismo. Este artigo propõe um espaço de design para adicionar representações sonoras, não visuais em ambientes de RV, permitindo que as pessoas com deficiências visuais interajam e compreendam melhor o ambiente virtual. Foi conduzido um estudo com pessoas cegas, onde experimentaram cenários de boxe virtual com diferentes tipos de feedback sonoro. Os resultados mostraram que o espaço de design permitiu a exploração de soluções que permitiram tornar a experiência de RV mais acessível para pessoas cegas. As contribuições deste trabalho são: o espaço de design proposto e a avaliação da sua utilidade na experiência do utilizador. A nossa abordagem baseia-se na formulação de um enquadramento teórico que permite aos designers de realidade virtual criar um espaço de design para a representação áudio de objetos e das suas possíveis interações. Pretendemos resolver a atual falta de padronização na definição do feedback não visual, para experiências de realidade virtual, que resulta em aplicações de realidade virtual inacessíveis (a maioria) ou na simplificação excessiva da experiência para pessoas cegas. Apresentamos um espaço de design para a representação áudio de objetos e do seu comportamento em realidade virtual como forma de apoiar o design de experiências de realidade virtual, acessíveis a pessoas cegas.

Para criar este espaço de design, seguimos um processo iterativo de definição de uma taxonomia clara das dimensões possíveis envolvidas neste domínio, beneficiando do conhecimento e experimentação. O objetivo é orientar o desenvolvimento e a investigação de futuras soluções. Esforços prévios de acessibilidade propuseram diferentes formas de representar a localização e o comportamento de objetos em ambientes virtuais, mas essa informação está dispersa em diversos artigos científicos, aplicações e jogos. Para construir este espaço de design, a equipa de investigação baseou-se em trabalhos anteriores e recorreu à sua experiência em HCI (Human-Computer Interaction), computação acessível, interação não visual, interfaces baseadas em áudio, ambientes virtuais e jogos. Durante dois meses, a equipa reuniu-se quinzenalmente para discutir e aperfeiçoar iterativamente o espaço de design. Além disso, a equipa reviu e editou o espaço de design ao implementar as aplicações de realidade virtual e preparar os estudos com utilizadores. Apresentamos

a seguir o espaço de design resultante, enfatizando que estes não são finais, devendo ser iterados com novas explorações e à medida que novas abordagens e tecnologias, sejam desenvolvidas.

Propomos uma taxonomia de propriedades que pode ser utilizada como padrão por investigadores e designers que pretendam explorar e compreender quais as escolhas de design áudio que podem criar uma experiência de realidade virtual acessível e qual é o impacto na experiência percecionada de cada escolha. São apresentados dois exemplos simulados de aplicação deste espaço de design num jogo existente. A nossa taxonomia inclui duas metacategorias (Consciência e Campo de Áudio) que definem se o feedback existe e em que condições. As sete categorias restantes definem como o áudio pode ser transmitido. Todas as representações têm um valor em cada categoria.

Neste estudo, foi realizado uma experiência com utilizadores cegos para explorar o potencial do espaço de design na criação de experiências acessíveis de realidade virtual (VR). O estudo focou-se em dois cenários: localização (S_{loc}) e movimento (S_{mov}) de objetos. Os participantes utilizaram uma aplicação de boxe em VR e experimentaram diferentes configurações de feedback sonoro para indicar a localização e o movimento do oponente virtual. Os protótipos foram desenvolvidos usando Unity3D e Oculus Quest 2, permitindo a personalização das categorias do espaço de design. Os participantes puderam testar as configurações e selecionar as suas preferências. Posteriormente, realizaram tarefas controladas para avaliar a capacidade do feedback sonoro em transmitir o significado pretendido. Após o estudo, os participantes foram entrevistados para partilharem as suas experiências e sugestões de melhorias. A análise dos dados foi efetuada através de uma análise temática com codificação dedutiva. Os resultados realçaram a importância do espaço de design na criação de experiências de VR acessíveis para pessoas com deficiência visual. O estudo contou com a participação de oito participantes legalmente cegos em cada cenário, sem experiência prévia em VR ou boxe.

Foram analisadas as preferências dos participantes e o processo de personalização ao explorar o espaço de design num ambiente de realidade virtual. Os participantes conseguiram atingir com sucesso os alvos estáticos e proteger-se dos objetos em movimento, independentemente das configurações preferidas de feedback de áudio. No entanto, a facilidade e a experiência dos participantes variaram. A Fala foi altamente preferida em relação à Sonificação, principalmente devido à certeza na localização dos objetos. No cenário S_{loc}, o áudio 3D foi preferido, enquanto no cenário S_{mov}, o áudio monofónico foi preferido. A entrega sequencial de informações foi preferida em vez de simultânea. Os participantes tiveram opiniões divididas em relação ao "trigger" de áudio, com alguns a preferir controlo e outros proximidade. A concorrência do feedback de áudio foi preferida apenas com uma fonte de cada vez. O feedback contínuo foi considerado opressivo e não intuitivo em certos contextos. Embora a eficiência e a precisão do desempenho dos participantes não tenham sido um foco específico, todos conseguiram concluir as tarefas com

sucesso, mas a experiência variou. Alguns participantes sentiram-se mais imersos e envolvidos com o feedback de áudio, enquanto outros acharam desafiador dividir a atenção entre o feedback de áudio e as ações necessárias para receber esse feedback (carregar no botão).

Ambientes imersivos de realidade virtual têm o potencial de criar experiências acessíveis, oferecendo oportunidades únicas para pessoas cegas participarem em atividades virtuais. Ao utilizar dispositivos de visualização de realidade virtual e rastreamento de mãos, os utilizadores podem controlar avatares e interagir com o ambiente virtual de forma mais intuitiva. Embora configurar esses dispositivos possa ser um desafio para pessoas cegas, os benefícios superam as dificuldades, permitindo que elas tenham uma perspectiva em primeira pessoa e controlem as suas ações com movimentos corporais.

O espaço de design é uma abordagem utilizada para criar experiências acessíveis em realidade virtual. Ele fornece um quadro que permite aos designers explorar diferentes configurações e adaptar a experiência de acordo com as preferências e necessidades dos utilizadores. No estudo apresentado, o espaço de design foi utilizado para criar configurações de feedback de áudio em dois cenários de boxe virtual. As preferências dos participantes variaram, mas foram identificadas tendências gerais que podem informar o design futuro de experiências de boxe em realidade virtual.

É importante ressaltar que não existe uma solução única para todos. As preferências dos utilizadores podem variar dependendo de suas percepções e objetivos individuais. Por isso, é essencial permitir que os utilizadores experimentem e selecionem diferentes configurações para encontrar a combinação que melhor atenda às suas necessidades. Além disso, as preferências podem mudar ao longo do tempo à medida que os utilizadores se familiarizam com a experiência.

O estudo também destaca a importância de escolhas de design, como gatilhos, tipo de som, sinal de tempo e espacialização, que afetam a experiência geral. A personalização é valorizada, mas é necessário considerar o impacto dessas escolhas na experiência do utilizador. O espaço de design oferece uma estrutura para explorar diferentes opções de design e entender suas consequências.

No final, cabe aos investigadores, designers e programadores avaliar os efeitos dessas escolhas em situações de aplicação específicas para criar experiências de realidade virtual acessíveis e envolventes.

Palavras-chave: Realidade Virtual, Inclusão, Cegos, Audio, Espaço de Design

Abstract

As virtual reality (VR) is typically designed in terms of visual experience, it poses major challenges for blind people to understand and interact with the environment. To address this, we propose a design space to explore how to augment objects and their behaviours in VR with a nonvisual audio representation. It intends to support designers in creating accessible experiences by explicitly considering alternative representations to visual feedback. To demonstrate its potential, we recruited 16 blind users (8 for each scenario) and explored the design space under two scenarios in the context of boxing: understanding the location of objects (the opponent's defensive stance) and their movement (opponent's punches). We found that the design space enables the exploration of multiple engaging approaches for the auditory representation of virtual objects. Our findings depicted shared preferences but no one-size-fits-all solution, suggesting the need to understand the consequences of each design choice and their impact on the individual user experience.

Keywords: Virtual Reality, Accessibility, Blind, Audio, Design Space

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Virtual reality (VR) is slowly becoming available to the masses at affordable prices, paving the way for a large number of applications in a variety of contexts, such as gaming and entertainment, education, and employee training. Despite the hype about immersive experiences that VR can offer, there are numerous accounts of their inaccessibility [69, 78, 86, 92]. VR applications rely heavily on visual feedback as an essential modality with limited or no tactile cues, neglecting people with visual impairments who are either being delivered with poor VR experiences or are excluded.

VR environments for people with visual impairments do exist. However, they have focused mostly on systems specifically designed for this population as a means to acquire either orientation and mobility skills [27] or knowledge about real-world locations [38]. These solutions often try to mimic real behaviours – e.g., using a white cane [58, 90] – with the main goal of transferring knowledge to the real world. As a consequence, mainstream VR which often has artificial environments that do not exist in the real world remains inaccessible. Others often over-simplify the experience (both feedback and interaction mechanisms) for the sake of access but at the expense of functionality and engagement [5, 35, 91]. For instance, many audio games restrict navigation to a grid [71], which significantly eases the ability to move in the environment but limits its exploration to a fixed set of positions instead of allowing users to freely explore the environment with their choice of walking direction and the speed. This paradigm blocks access to the more complex virtual environments and behaviours seen in mainstream VR experiences, which remain inaccessible. This is, in part, due to a lack of awareness and standardization when designing nonvisual feedback for accessible VR applications.

We propose a design space for augmenting objects with a nonvisual audio representation of location and behaviours in an immersive VR environment. The design space proposed is meant to guide the design of alternative audio feedback for objects' location, behaviours and interactions within a VR environment. We propose nine categories that can be leveraged by researchers and practitioners working on creating accessible VR content. This design space includes dimensions related to how audio feedback is provided,

what information is delivered, and how it can be triggered. To cite a few examples, one may convey information about objects using spatialized audio to hint at their current location with respect to the location of the user (e.g., left or right, near or far) or monaural sound if the location is not relevant (e.g., a warning sound or background music); audio feedback about multiple sound sources may be played to the user sequentially or concurrently; and the feedback may be given using speech or sonification.

To explore and validate the ability of our design space to support nonvisual representations of objects and their behaviours, we conducted a user study where 16 blind people experienced one of two virtual boxing scenarios: one focused on conveying the location of objects (S_{loc}), where participants were asked to understand the opponent's defensive stance (the location of their two hands); the other focused on conveying movement behaviour (S_{mov}), where they were asked to understand the attacking punches of their opponent (moving towards the user). In S_{loc}, we developed a VR environment that made it possible for the user study to be held. This environment consisted in an opponent/trainer with a head, torso and arms, with the latter being able to be moved around.

Findings show that the design space enabled the exploration of multiple solutions to support VR boxing for blind people. The user study depicted both shared and contrasting preferences for the different categories of the design space and allowed us to identify how each category affected participants' preferred design choices and their experience. Overall, and independently of their preferred configurations, participants were able to complete their tasks successfully in both scenarios.

Participants were able to complete given tasks (i.e., attacking the specific location of the opponent who are standing still in S_{loc} and guarding themselves from the opponent's punches coming from different locations in S_{mov}) at high success rate despite the variation in preferred feedback configuration across participants and scenarios. In addition, we identified how each of the different design categories affects participants' preferred design choices and their experience. Our contributions are: (1) a design space for the auditory representation of the location and behaviour of objects in virtual reality for blind people, and (2) the qualitative assessment of the potential of the proposed design space for understanding how each choice in design space categories affects the user experience.

1.1 Publications

Os resultados e contribuições apresentados nesta dissertação resultaram num artigo científico que foi aceite numa conferência international:

- João Guerreiro, Yujin Kim, Rodrigo Nogueira, SeungA Chung, André Rodrigues, and Uran Oh. 2023. *The Design Space of the Auditory Representation of Objects and Their Behaviours in Virtual Reality for Blind People. IEEE Transactions on Visualization and Computer Graphics 29, 5 (May 2023)*, 2763–2773.

https://doi.org/10.1109/TVCG.2023.3247094

1.2 Document Structure

This document is organized as follows:

- Chapter 2 – We analyzed the state of the art in order to understand and share how the work with blind people and technology is being conducted and also to have an idea of how design spaces have been used in previous research.
- Chapter 3 – We explain how we conveyed the design space we used in order to develop and complete the study
- Chapter 4 – The system design is presented in this chapter with information about how we developed the application used for the study
- Chapter 5 – This chapter shows how we conducted the user study and findings that resulted from it
- Chapter 7 – Conclusions and future work

Chapter 2

Related Work

We discuss related work along three topics: first, we describe how audio feedback has been leveraged in prior research to convey or augment information (e.g. sonification to facilitate exploring a map). Second, we discuss how blind people are interacting in virtual environments. Lastly, we highlight how design spaces have been used in previous research to promote an understanding and further explorations of certain domains.

2.1 Auditory Feedback in 2D and 3D spaces

Audio feedback has been recurrently investigated in the context of wayfinding and navigation[93, 104, 44, 41, 11, 89, 88]. For example, Heuten *et al.* [41] presented a technique to sonify a real world map that enabled blind users to build a mental map prior to their journey. The approach attempted to allow users to easily perceive relationships between objects and detect important landmarks for navigation. On touchscreens, audio feedback was investigated to improve gesture recognition [28, 74, 75, 59]. For instance, Gao *et al.* [28] designed three different types of audio feedback to support trajectory-based finger gestures. The gesture was accompanied by a discrete beep, static or gradual feedback. Their results suggest that gradual continuous feedback increases accuracy.

Other works have focused their efforts on exploring audio feedback for target acquisition and detailing target information in extended reality for 3D spaces such as VR and Augmented Reality (AR) [84, 19, 18, 29, 85, 24, 42]. For example, Ren *et al.* [84] explored audio and haptic feedback design for Mixed Reality (MR) tourism applications. They investigated the effects of spatial audio with various design factors such as volume and rhythm of audio and found that fast rhythm and adequate volume improved the user performance on target acquisition in MR applications using a head-mounted display (HMD). Similarly, Chung *et al.* [18] investigated audio and haptic feedback with spatial audio for a target acquisition task in a 3D virtual space. They found that spatial sound helps the user to understand the horizontal direction of the target, and discrete-based audio and haptic combo reduced the task completion time. The use of spatial audio is now

widely used and tries to convey information about the location of objects, often replicating sounds from the real-word [80, 90], but also trying to reproduce echolocation [6]. While there are many examples of using audio to augment interactions or to add new information about the objects that populate virtual environments, there has been little focus on standardizing the process of designing audio feedback. Thus, we propose a design space for the audio representation of objects and behaviours to support systematic procedures for making design decisions.

2.2 Sensory Substitution

Sensory substitution is a technique employed to assist individuals who are unable to utilize a particular sensory modality. This involves leveraging an alternative sensory channel to convey information that would otherwise remain inaccessible due to the limitations of their primary sensory function. Blind people make use of this technique since they cannot get information with vision, and a lot of the help they get to live their everyday lives comes from stimuli in other sensor modalities, like sound. Learning how to use this technique is a must to give the user the right information at the right time and get good results.

Árni Kristjánsson, *et al.* explained in their article [56] what were the main concerns that developers should have while designing sensory-substitution devices, making it very much applicable to this project. They mention various techniques and concerns that we should have, to make this type of adjustments, with some being more important than others. An example is how the use of such technology should not require extensive training, but it is also said that this requirement must not be a barrier to developing such devices and technologies. This is because the development of something that requires training but has shown to be of good use should still be used, as many have already done. They mention that a broad spectrum must be taken into account like comfort and ease of use, externalization, the relation between sensory substitution and perceptual illusion, etc. But the one that is really highlighted is the potential sensory overload. This happens when too much information is given to the human brain at the same time, so it cannot process everything and might cause some confusion. As an example, the use of too much environmental risk information being given to the user will most likely work against the person, making it more liable to give critical information only. Among these, there are other highlighted points regarding the design of sensory substitution, and another one that seems very relevant and probably very overlooked is that the use of sensory substitution must not interfere with other perceptual functions, as it may deteriorate performance, just like sensory overload.

The fact that blind people can't have the same sense of immersion as sighted people in video games, with gesture-based interaction, is the focus of Tony Morelli and Eelke Formel's article [67], now in the area of sensory substitution. They make it so that when-

ever a visual cue comes to the screen there's a trigger for a vibrotactile cue on a controller. A big thing that is also mentioned, is that to make sensory substitution in games for blind people, we need access to the source code, which is not usually possible. A good substitute in the "dance dance revolution" type of games is the audio cues informing what to do next, but the audio might interfere with these instructions, making it not viable, so vibrotactile is the best option. The video analysis, for a track and field game, consists in checking if the hurdle that needs to be jumped over is appearing on screen and if it is either yellow or green (green meaning it's time to jump). This is the idea for every other game that was tested here. Two user studies were conducted, one with 28 sighted players and one with seven visually impaired players. There was no significant difference in playing the standard version and the RTSS (Real Time Sensory Substitution) version for sighted players (with RTSS having a higher success rate). There was also no big difference when comparing those results with the results obtained from visually impaired players.

In another article [82] about sensory substitution, Presti, *et al.* talk about the development of WatchOut, which is a sonification technique to give information about the environment in front of them in real-time, making use of sensory substitution. Initially, they tested their approach, very roughly to assess if using mobile-driven obstacle detection was a viable option, concluding that it was. After this first test, they had some online user studies, which they considered to be a limitation, since they had no control over the participants' pool, making it impossible to control the demographic consistency of the participants. These studies helped in the development of sonification mapping, through four design iterations. With these iterations, they understood that using more than three dimensions to apply sonification in such a short time for training, was not ideal, since their accuracy dropped down from a near-perfect 0.96 to 0.71 when adding the fourth dimension. The obstacle properties chosen to be presented to the user with WatchOut were the frontal distance to the user, the horizontal position with respect to the user, and obstacle size. The latter one was ultimately dropped in this approach since it needed width and height dimensions and those types of properties were the ones that made it difficult to understand, therefore lowering the accuracy of the sonification. With all being considered, this technique looks promising as the authors seemed eager to continue exploring it in the future.

A different use for sensory substitution was proposed [48] by using sonification in moving objects so that visually impaired people could understand the characteristics of the objects captured in an event-based camera. They do this with the use of MIDI (Musical Instrumental Digital Interface), to manipulate the most comprehensive music parameters. They can make every object unique for each person since the mapping is made one-to-one, meaning one object only has one sound and vice-versa. For users to understand how to recognize each sound, they go through training, in which everything is explained through a five minute video, describing the sound, the object, and finally playing it as if

it was a real-time capture. After going through their results, they concluded that visually impaired people were more capable of recognizing the appearance and motion of the objects captured than sighted people. They also mention that when there were multiple objects present in the scene, it was more difficult to be accurate in the recognition of those objects, since the audio information was proportional to the number of objects, but people have a limited amount of information that they can perceive accurately.

In another article, Chernyshov, *et al.* [17] talk about the development of a "superhuman" sport, that resembles volleyball/tennis, since the premise of the sport is to throw a projectile into the other teams half and try to score a goal. The big difference in this sport is that the player cannot see the projectile at all and in order to know where it is, their eyes must be closed. While the eyes are closed they are using "energy" and have to direct their hand to where they think the projectile is, if they are correct there will be haptic feedback so they know it's in that direction. They used glasses with eye closure sensor and two different gloves, one for sensing and another for interaction (Sensing and Action gloves). The sensing glove had two thermoelectric modules, one to feel heat and the other to feel cold. This is because each team can only interact with the projectile if it is in the right thermal feedback for their team (i.e. Team A - heat, Team B - cold). This glove also had LRAs (Linear Resonant Actuators) to know where the ball is by vibrating and a way to keep the player aware of the game stats. The action glove only had the LRAs for aiming and as a way to keep the player aware of the stats. The sport was still in test phases in 2018.

In the topic of superhuman sports, we found a Japanese association named Superhuman Sports Society [1]. They see sports in a different way from what the usual human sees. Their whole purpose is to make it possible for everyone to enjoy sports together, regardless of age, disability, or skill. It all comes down to the way they perceive sports in the future since their idea is that we will be competing in a lot of new ways, with a lot of new technology. with the mix of robotics and our human capabilities, it is possible to have an augmented human that has superhuman abilities, making sports a little bit more entertaining for the watcher and mitigating the difference in skill between each competitor. Such abilities might include superhuman strength, flying, or teleportation for example. All this is possible with the help of virtual environments and robotics, making use of sensory substitution or even augmentation. For us, the most important part about this association/society is that they intend to make it possible for people with disabilities to play sports against someone who is perfectly healthy.

2.3 Virtual Environments for Blind People

The accessibility of virtual environments is frequently ignored. This is especially relevant for people with visual impairments, as interaction usually depends on visual stimuli and

lacks tactile cues. Research on virtual environments for visually impaired people has focused mostly on systems specifically designed for them, rather than on providing access to mainstream ones. Common approaches rely on audio and/or haptic feedback to convey information about the environment [26, 101, 52] and focus mainly on creating experiences to support mobility training and/or creating mental maps of real-world locations [23, 26, 38, 62].

Recent technological advances and the easier access to immersive VR have led to new approaches that try to improve blind people's experience through different modalities, such as more realistic audio, haptics, or locomotion. For instance, earlier approaches often used keyboards, joysticks, the smartphone, or other haptic devices to support exploring a virtual environment [57, 21, 87, 22], but more recent approaches have tried to increase immersion by supporting either walking in place (with or without a treadmill) [55] or actually walking in the real-world [45, 95] – sometimes by instrumenting a white cane for improved haptic feedback [90, 103].

The use of audio has supported multiple approaches and applications in varied contexts, such as digital games[33, 79]. For instance, NavStick [71] enables blind gamers to probe their surroundings by scanning a specific direction at a time with the controller joystick – in the context of a 3D adventure game – while The RAD [91] used sonification in racing games for an equitable experience for blind and sighted players. Other works have investigated accessible social experiences [47, 34]. For instance, VR Bubble [47] supports accessible social VR experiences by exploring peripheral awareness dividing the social space into intimate, conversation, and social bubbles, and conveying different audio feedback accordingly. Other contexts include sports and exergames where prior works have targeted, for instance, yoga [83] and tennis [66]. More recent approaches, have tried to support training or playing sports that blind people practice in the real-world – e.g., Goalball [97] and Showdown [98] – by providing an accessible VR alternative.

In order to make VR applications accessible – all, not only those specifically designed for blind people – designers have to explicitly consider how to convey alternative representations to visual feedback. While the use of audio has been fundamental to providing an accessible experience in prior work, the literature shows a panoply of approaches to convey information through this modality. Our work aims at providing a theoretical framework for designers to create accessible VR experiences.

Kreimeier *et al.*, [53] analyzed current trends in VR for blind people, focusing mostly on haptic and audio feedback, as well as the notion of perspective. They found that most VR applications for blind people use audio feedback, but not all use haptics. The use of force feedback and vibrotactile signals makes for a great way to interact with the environment, an example given by the authors is the use of this combination of haptics to convey different virtual textures when using a virtual white cane. The last point we took from this article as a primary topic for our project, is the perspective. The explanation

in this article of egocentric and exocentric interactions with the environment makes for a good understanding of the difference between both. In the first one, we can say the user perceives the virtual environment as if they were inside it, needing to understand the environment with both audio and haptic feedback with some locomotion required. With exocentric interaction, the user can just move in the virtual environment without the need for locomotion but can still use the feedback mentioned before.

As locomotion is a big part of Virtual Reality, the meaning of the article written by Kreimeier, *et al.* [54], is to try and understand what is the best option for locomotion in VR for Blind People. Four ways of locomotion were tested (Cyberith Virtualizer, Virtuix Omni, HTC Vive VR tacker based and Windows Mixed Reality Joystick). To try and get to a conclusion they ran tests with seven people who were blind/legally blind, in which the user would control an object in unity with their movement and had to try and get to the other object that was emitting sound. They concluded that the joystick was for sure the best option for locomotion between these participants since it felt safer and more practical/simple to use (Figure 2.1). They also mention that this does not change the fact that treadmills and step detection can be good in the future.

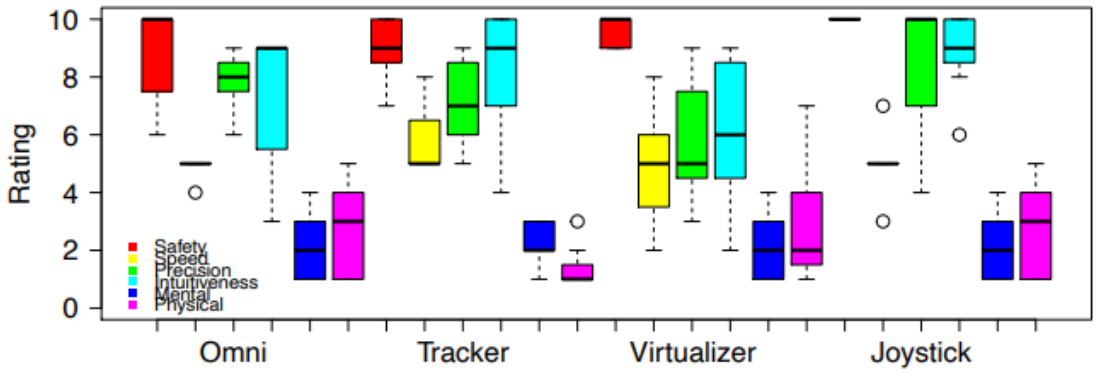

Figure 2.1: Box plots for questionnaire results regarding the four different implementations of locomotion. Red: Safety; Yellow: Speed; Green: Precision; Light Blue: Intuitiveness; Dark Blue: Mental; Pink: Physical [54]

Gluck, *et al.* wrote an article [31] about the development of a virtual reality video game, that allows blind people to drive a car on a race track. In this game players face against four AI's trying to be the first to finish a race, by completing three laps around a three and a half-mile long track (around 5.63km). For the development of the game, there was a need to deliver information that was critical to the user at the correct time, and the conclusion that they came to, was that haptic feedback was the best way to deliver that information and get a faster reaction. With this in mind, to steer, they tried to mimic the way people do it in the real world the best as possible, so the player needed to raise the right controller higher than the left one to turn to the left and the other way around to turn to the right. The bigger the difference in height from both controllers, the sharper

the turn. The haptic feedback enters in steering, making the controller vibrate more when the player needs to make a big adjustment (sharp turn) or less when the player needs a small adjustment (slight turn), so every time the player has the right placement for each controller, there will be no vibration. The HMD is also used in this game to give information to the player. The player can look up, as if they were looking at the rearview mirror, to get information about the cars behind them, or look out the windows to check for cars, which will give information too. This information that comes from head gestures is given with speech since they are not needed to be fast or are not meant for a fast reaction. The audio was not used in functionality in this game since it would make for slow reactions in the players part, so they used it for ambient only and speech. The gas and the brakes are in the controller triggers as standard.

A VR Showdown exergame was developed by Aan Hojun, *et al.* alongside an article [2] describing everything that was done and the results they got. In this exergame in Virtual Reality, they were able to make it possible for visually impaired people to play against sighted players and even get better results in games for those people with disabilities. Showdown is a paralympic sport that consists of two players that are trying to bat a ball into the opponents' goal on the other end of a table with walls that make it possible for the ball to ricochet. The player who reaches eleven goals first wins. To make this easier for visually impaired people to play, they decided to use HRTF (Head Related Transform Function), since they found it would be the best for audio, in a game that relies mostly on that type of feedback. They tested this game in three User studies conducted with two groups of people (sighted and visually impaired), in which the first one consisted in understanding how accurately the players recognized the position of the ball in the virtual table. They found that the accuracy was pretty much the same for both groups. In the second study, they tried to understand how well the two groups would perform against the AI implemented in the game. The results for this one show that 10/13 people with impairments and 9/13 people sighted, won their games. Finally, the last study was to understand which group would perform better against each other, concluding that visually impaired people got much better results, winning 67.6% of the games against sighted people. They also found, by monitoring heart rates, that playing Showdown VR, visually impaired people can experience the exercise effects of playing table tennis, making it a viable exercise tool.

2.4 Design Spaces as a Theoretical Framework

Design space is a theoretical framework that describes the different possibilities for designing a type of artifact [61]. It promotes reflecting on the attributes and choices made when designing new artifacts and is based on the domain of interest, its technological possibilities, and existing artifacts used for related purposes. It consists of a number of

dimensions deemed central to the domain, each with a set of proposed values that may be expanded after new design explorations. These dynamics make it possible to iteratively grow and adapt both to new technological innovations and to new research findings (e.g., related to human perception of auditory feedback). Design spaces are often used in Human-Computer Interaction as a framework to understand a domain and to explore innovative solutions in that domain (e.g., [14, 68, 70, 72, 105, 60]).

Researchers have been investigating sound design and VR together and separately. Some works have explored the design space of a specific type of feedback, such as conveying heart rate in VR [15], while others have tried to classify the different types of applications of a specific domain, such as AR/VR applications for assistive environments in manufacturing [12]. Garner [30] – with the goal of building a framework of VR sound – specifies multiple perspectives on how to approach sound. For instance, one may approach sound as an object (a spatial entity) that can be located through its sound [76]; or as an event, meaning the sound is related to a specific action about the object [73], among other perspectives. Jain *et al.*, [46] created a taxonomy in the context of VR sound accessibility for deaf and hard of hearing (DHH) people. Still, their goal contrasts with ours as they tried to categorize the existing sound in VR, while we aim to find alternative representations to visual feedback.

In the context of accessible computing for people with visual impairments, design spaces have been proposed as a way to promote research and innovation in domains such as nonvisual word completion interfaces [72], rich representations of visual content for screen reader users [68], and accessible visualizations [50, 106]. For creating the taxonomies, prior research has relied either on a systematic review of the domain [50], or on the authors' expertise on and analysis of the domains of interest [68, 72]. The design space created by Morris *et al.*, [68] enabled the authors to create multiple prototypes that they could then evaluate with visually impaired users. Nicolau *et al.*, [72] engaged blind participants in exploring the different categories of the design space and creating their own preferred solution. In both cases, instantiating the design space provided valuable insights into promising solutions for their own domains.

Chapter 3

Design Space

3.1 Approach

Our approach lies in the formulation of a theoretical framework that allows VR designers to instantiate a design space for the audio representation of objects and their possible interactions. We tackle the current lack of standardization when defining nonvisual feedback for VR experiences, which results in either inaccessible VR applications (the large majority) or over-simplification of the experience for blind people. In the following, we present a design space of the audio representation of objects and their behavior in VR – summarized in Table 3.1 – as a means to support the design of VR experiences that are accessible to blind people.

Category	Description	Values
Awareness	Defines if objects have an audio representation.	*None — Conditional — Full*
Audio Field	Defines users' audio spatial awareness	*Self — Perimeter — Field of View*
Trigger	Defines how audio feedback is activated	*On: Existence — Proximity — Movement — Collisions — Request*
Representation	Indicates whether the audio feedback can change	*Static — Dynamic*
Sound Type	Type of audio feedback	*Speech — Sonification*
Time Signal	Describes the type of audio signal used	*Discrete — Periodic — Continuous*
Cardinality	Number of objects considered as a audio source	*1 — 2 — ... — N*
Concurrency	Indicates sound sources play sequential or concurrent	*Sequential — Concurrent*
Spatialization	Defines the form of audio Spatialization	*Monaural — Dichotic — 3D*

Table 3.1: The design space for auditory representation of VR objects defined in our study.

3.2 Process

Defining a design space with a clear taxonomy of possible dimensions involved in this domain is an iterative process, which benefits from knowledge and experimentation. The purpose is to guide the development and research of future solutions. Prior accessibility efforts have proposed different ways to represent the location and behaviour of objects in virtual environments, but that information is scattered in different scientific articles, applications, and games. For this project we collaborated with a group of researchers

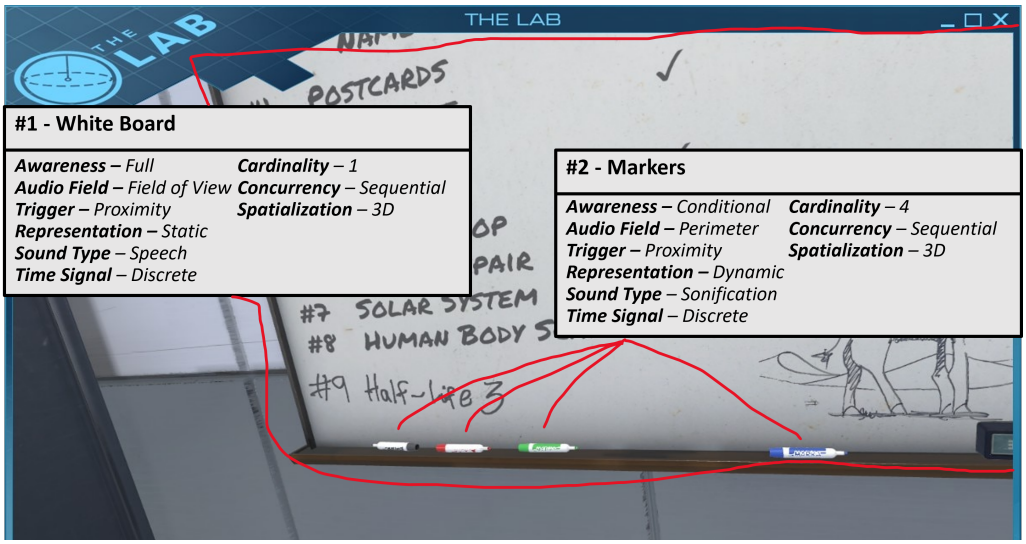

Figure 3.1: Two mock-up examples of the design space applied to two interactive objects within the scene of the game named "The Lab by Valve". #1 - *White Board* is noticeable by the player when on its field of view. The board is announced through speech with the sound coming from the object location. #2 - *Markers* are only perceivable when close to the board. Each of the markers periodically produces a sound in sequence, where the colour is associated with the pitch of the sound used, originating from their location. The closer the player hand is to one, the louder it becomes. The markers only produce sound while none is in hand.

from South Korea and in order to build this design space, the two research teams (one from Portugal and the other from South Korea) were informed by prior work and relied on their expertise in HCI, accessible computing, non-visual interaction, audio-based interfaces, virtual environments, and games. The team met bi-weekly for two months in order to discuss and iteratively refine the design space. In addition, the team revisited and edited the design space when implementing the VR applications and preparing the user studies. Below, we present the resulting design space but we reinforce that design spaces are not final. They are meant to be iterated upon with new explorations and as new approaches and technologies come to be.

3.3 Audio categories

We propose a taxonomy of properties that can be used as a standard for researchers and designers who seek to explore and understand what audio design choices can create an accessible VR experience and the impact on the perceived experience of each choice. Two mock-up examples of exercising this design space in an existing game are described in Figure 3.1. Our taxonomy includes two meta-categories (i.e., *Awareness* and *Audio Field*) which define whether or not feedback exists and under what conditions. The remaining seven categories articulate how audio can be conveyed. All representations have a value

in each category. Some of these categories were not used for this part of the research, and others were not used for the other part of the research (Korean study), but since every one of them was chosen for the two studies, they are still presented.

Awareness. The awareness category indicates if users are made aware of the existence of objects of a particular type and what are the pre-conditions to represent such objects. It can be seen as a meta-category where, ultimately, a value of *None* indicates that such an object does not have an audio representation - dismissing the remaining categories. On the other end, *Full* awareness would indicate that all instances of such object have an audio representation, being their characteristics defined by the other categories. *Conditional* means there are rules which define the awareness, for example they could be Priority (e.g., only the two closest markers) or Attribute based as in the example given (e.g., only if marker is not held). While this category pre-defines the objects that have an audio representation, the following further detail when and how it is conveyed.

Audio Field. The Audio Field category indicates when objects can have an auditory representation based on their spatial location. When outside of the audio field, the object does not produce auditory feedback. A *Field of View* audio field corresponds to the full visual field of the user, indicating that the object would provide feedback when visible (i.e. and according to its awareness). Alternatively, feedback can be based on a *Perimeter*, meaning objects would need to be within a defined area in order to produce auditory feedback. A *Self* audio field would mean that feedback only occurs on self-interactions – e.g., contact with the virtual character, self footsteps.

Trigger. The Trigger category indicates what causes the outset of auditory feedback. On one end, an object can trigger auditory feedback *On Existence*, meaning that it will start producing auditory feedback whenever it enters (or starts existing) in the user's audio field. Alternatively, feedback may be triggered *On Movement* (when the object moves) or *On Proximity* (when the object gets within a distance threshold, and where zero represents on collision with self). Other options may include audio feedback *On Collisions* or *On Request* – meaning feedback is provided when explicitly requested by the user (e.g., by clicking a button).

Representation. The Representation category indicates whether the auditory feedback has a *Static* – the feedback mapped to one object does not change – or *Dynamic* – the feedback may change according to defined rules or characteristics (where volume is adjusted given the proximity to players' hand).

Sound Type. The Sound Type category indicates whether the audio feedback is provided through *Speech* or *Sonification*. Sonification is the "use of non-speech audio to convey information", such as auditory icons – audio effects resembling real-world objects or events – earcons – artificial sounds often following musical conventions to alert users – or variations in audio characteristics to convey meaning (e.g., distance, size, or height). Specifying the possible sonification techniques has its own challenges and enters a greater

level of detail that is out of scope of this design space.

Time Signal. The Time Signal category indicates whether the audio feedback is provided with a *Discrete*, *Periodic*, or *Continuous* signal. Discrete indicates feedback is provided once; Periodic represents feedback being delivered at specific time intervals (e.g., every ten seconds); and Continuous means that feedback is provided nonstop after being triggered.

Cardinality. The Cardinality category indicates how many objects of this type are considered as a potential source of feedback at each time. This same category is used on a prior design space related to nonvisual interaction and feedback and can take values from *one to N*. In the example given all objects are rendered as potential audio sources, but there can be instances where (e.g. if there were twelve markers) it could be necessary to limit the cardinality.

Concurrency. The Concurrency category indicates whether the feedback of multiple objects is presented *Sequentially* or *Concurrently*. Similarly to Cardinality, we have borrowed this category from Nicolau *et al.*,'s design space [72]. This means that when multiple objects' feedback may overlap in time – e.g., four markers – one can either present their feedback sequentially or concurrently while trying to ensure they are distinguishable. This category is not applicable when there is a Single object as a potential source for feedback – when cardinality equals one.

Sound Spatialization. The Sound Spatialization category indicates whether the auditory feedback provided relies on any form of spatialization. *Monaural* sound means feedback is provided through a single auditory channel, which does not enable a sense of spatialization/location. *Dichotic* sound would rely on two auditory channels to convey information to the right and/or left ears (e.g., markers closer to the user's right hand would only be conveyed to the user's right ear), while *3D audio* tries to represent a specific location of a sound source relative to the user's position replicating the way people hear sound in the real world (using speech or sonification to announce its position from its location).

Chapter 4

Prototype

We show how the design space can be leveraged to explore nonvisual augmentative feedback which is showcased by applying it in the context of a VR Boxing application. First, we decided to target sports applications since physical activity has a positive effect on people's health [96, 77]. In addition, we targeted a popular sport, but that is generally inaccessible to people with visual impairments as a way to afford a new – but somehow known – experience in a safe environment. Finally, we wanted an experience that is complex enough to support a variety of possible interactions and behaviours. For example, one may want to receive feedback about the opponent's (or coach's) body, head and/or hands, but such feedback may depend on the current intention of attacking or defending.

We explored two scenarios that are related to understanding either the location of objects or their movement: 1) perceiving the location of the defensive stance of the opponent (S_{loc}), and 2) punching attacks moving toward the user (S_{mov}). The first scenario was the one we focused on during the description of this project. We explored both scenarios separately with the goal of identifying both commonalities and differences between individual explorations of the design space. Furthermore, these scenarios were explored by two different research teams – geographically distributed (S_{loc} in Lisbon, Portugal and S_{mov} in Seoul, South Korea) – aiming at showing the versatility of the design space in supporting the creation of accessible experiences. Both teams followed the same design space and procedure but explored, separately, the design space with different implementations, users, and a different but related task. The separate implementation of each team's VR application and exploration of the design space intended to create a greater variety of approaches and of possible experiences than what would be possible under a unique design exploration.

This section describes how the application came about, and the details of its development in Unity.

4.1 Implementation

A 3D project was created in order to develop this application. The first idea was to make a boxing bag, so that the user could punch it and understand where it is, getting audio feedback and vibration to the controllers whenever it hit the bag. That was used as a starting point for the application, upgrading that bag to a three-section bag Figure 4.2, with a head, chest, and stomach, with the intent of each section giving different types of feedback making them unique to each other. Finally, it was decided that for the purpose of this project, the bag had to simulate an offense-only boxing coach, with its mitts telling the user where they were.

Everything in this project was made in Unity3D, with free assets from the Unity Asset Store, so that the use of VR would be compatible with the application (Oculus Integration, ie.).

4.2 The opponent/Bag

The first bag was just a cylinder, with some mass and an anchor point to the top Figure 4.1, making it work like a physical boxing bag. There was no feedback on where to hit, and the only feedback the user would receive was a vibration to the controllers and a 3D spatialized audio telling that the bag was hit.

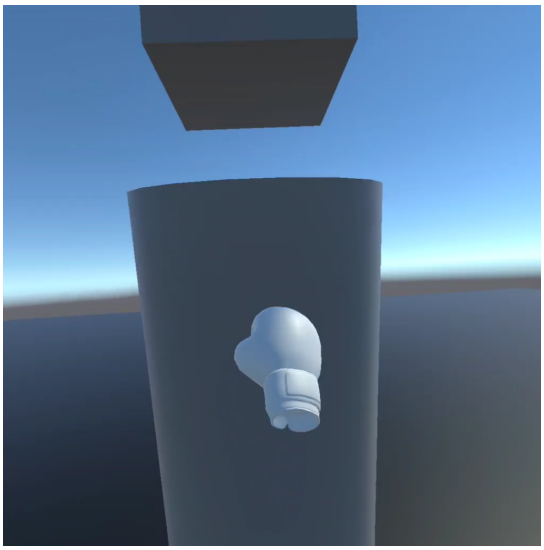

Figure 4.1: The first bag design

The second bag, as said before, was divided into three sections (head, chest, and stomach) Figure 4.2 so that it would feel different to punch each section, and start to tell the user where to punch (although this was not implemented during this phase). Each of the sections that the bag had would give different audio feedback and would have different vibration patterns so that the user could understand which of the sections was hit.

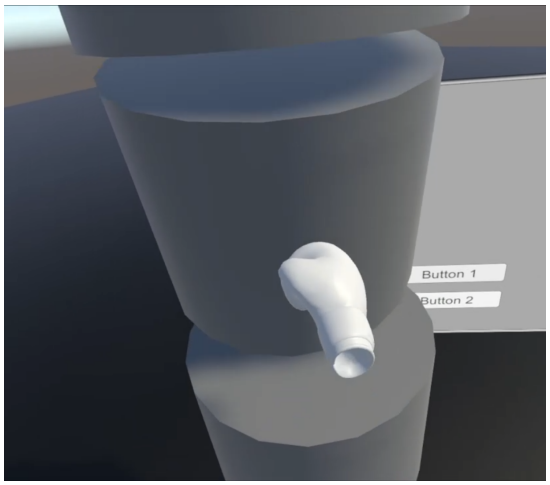

Figure 4.2: three-Section bag

The final design of the bag would make it feel like a boxing coach/opponent. This was because it was composed of a head, a torso, and two arms/mitts Figure 4.3. Every part of this whole object had different on-hit sounds and vibration patterns so that the user could perceive which was which more easily. The mitts started by having random position patterns inside a 2x3 grid for simplicity; [up(head), down(torso)] × [left, center, right], so each mitt would switch position after a window of four to eight seconds. Soon after the possibility to control the position of the mitts was added, making that the primary feature for changing the mitt position during testing.

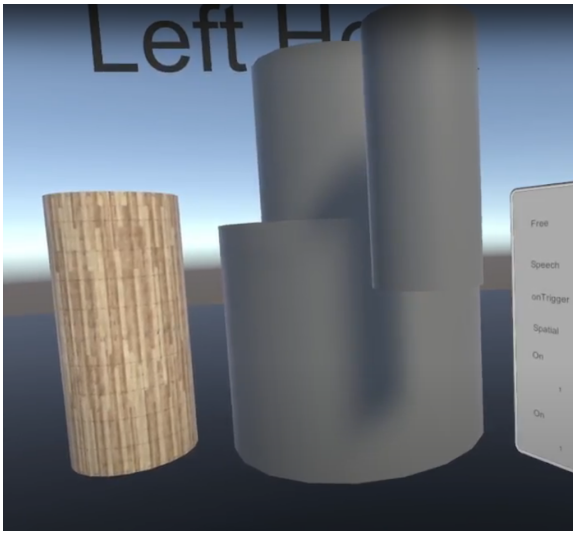

Figure 4.3: Final bag design

4.3 Controls

We used the Oculus Quest 2 VR set in this project and the Oculus Quest 2 controllers. These consist of two ergonomically designed controllers, one for the left hand and one for the right hand, identical, only differing in button nomination, and the placement of the buttons is mirrored. The joysticks were usable but irrelevant to this project, so they would only serve a purpose during developmentFigure 4.4, with the left joystick being used to walk around and the right joystick to look around the scene.

Figure 4.4: Unused joysticks

The ”X” button on the left controller was used for the Discrete category in the design space, letting the user get audio feedback referring to the position of the right mitt, and the ”A” button on the right controller served the same purpose but for the left mittFigure 4.5.

Figure 4.5: Button X and A.

Lastly, the right ”Grip” button was used to recenter the opponent/coach and put it at arm’s distance from the userFigure 4.6. The two ”Trigger” (left and right), ”Y”, ”B”, and the left ”Grip” buttons were not used. The Quest 2 controllers are supposed to be the reference in the application for the user’s hands. So in order to mimic the use of the hands, the audio, and haptic feedback for punching the mitts/opponent, were dependent

on how fast the user punched, meaning that if the punch was harder, then the controller vibration would be stronger and the audio would be louder.

Figure 4.6: Right grip button

4.4 Exploration

The two scenarios allowed us to explore the design space of VR boxing both with attacking and defending in mind. We convey auditory feedback from the hands of the opponent – framed as the participant's coach in the study. In S_{loc}, we indicate the defensive stance of the opponent, by conveying feedback about the location of their hands, while in S_{mov} we indicate the attacking attempts of the opponent, by trying to convey both location and movement. In both scenarios, we fixed **Awareness** (as Full) and **Audio Field** categories (as Field of View), and **Cardinality** as two, since we wanted to always be able to convey feedback about both hands. Note that we have fixed the opponent's location to be in front of the participant's avatar. The nature of both scenarios led us to fix the **representation** category, but differently among scenarios. In S_{loc}, representation is static, meaning the location of the hands is conveyed but not its movement. On the other hand, in S_{mov}, representation is dynamic as the audio feedback changes depending on speed and proximity/distance.

All the other categories were explored. **Time Signal** (*discrete*, *periodic*, and *continuous*), **Concurrency** (*concurrent* or *sequential*), and **Sound Type** (*speech* or *sonification*) were explored in full. Although **Sound Spatialization** was explored too, we did not explore the monaural condition given the known advantages of sound spatialization when using multiple sound sources [10]. **Trigger** category could either provide feedback on existence (meaning we would always have feedback about the hands' location), on movement (only when the hands change position), or on request (when the user explicitly requests it). Note that there were differences in the way each category was explored for the two scenarios, and those derived from an independent exploration of the design space by the two distributed research teams, and from two related, but different scenarios.

We used two different types of audio cues in S_{loc} for different chosen categories. These cue types were speech and sonification. As the name says, speech was used to tell exactly where the opponent's arms were positioned through words (i.e. "Up right", "Down front",...), with all of the voices being generated by a text-to-speech program, making it clear for understanding to the user. Besides speech, we also used two different sounds for sonification, as mentioned just above, one being a sonar-type sound, and the other a continuous C4 musical note. The sonar sound was used in any sonification category that didn't require continuous feedback, and the continuous C4 note was used in those instances where continuous feedback was required. These two audio types made it possible for the user to use non-continuous sonification and still feel comfortable with the sounds being emitted and also the continuous feedback could actually be continuous, without any interruption of the sound.

We also highlight that combinations among these categories resulted in different-sounding approaches. For instance, in S_{loc}, combining sonification with a periodic time signal corresponded to a sonar-like sound, and a continuous beep when combined with a continuous time signal. In both scenarios, we used the sound pitch to indicate if the target was at the head (higher pitch) or body (lower pitch) height. In addition, in S_{mov}, volume indicated distance.

Figure 4.7: PC user's screen

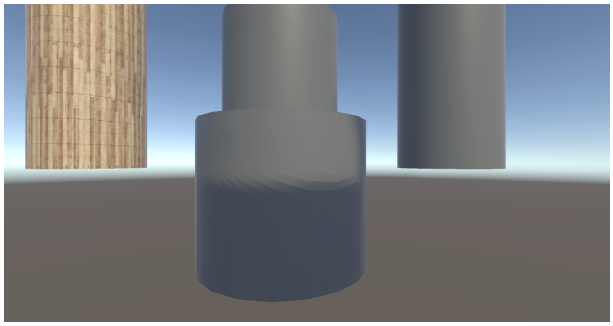

Figure 4.8: VR user's screen

Chapter 5

User Study

The design space is the means – and not the end – to build interactive VR experiences that are accessible to people with visual impairments. For that reason, we conducted a user study where blind users explored different instantiations of the design space, aiming to assess its potential to support the design (and experimentation) of nonvisual, audio representations of objects and their dynamics in VR environments. In particular, we addressed two scenarios:

- *Scenario 1 - Location (S_{loc})* : understanding the *location of stationary objects* in the virtual environment

- *Scenario 2 - Movement (S_{mov})* : understanding the *behaviour of moving objects* in the virtual environment

5.1 Apparatus

The prototype was implemented using Unity3D running on an Oculus Quest 2 VR system (including a headset and two controllers). The prototype enabled the researchers to customize the different categories and values of the design space. We used two laptops: one capable of running Virtual Reality (with an NVIDIA Geforce RTX 2060) and another to select the configurations and present visual feedback to the researchers. Audio customization (e.g., spatialization) was implemented with Unity3D native assets with default configurations (e.g., regarding reverb values). During the study, the researchers calibrated the VR Headset and the position of the avatar in the VR space and guided the participant to the intended location. Participants were asked to use headphones as a way to improve spatialized audio. Information about configuration preferences and task success/failure was logged by the system. This chapter will also take account what was done on the other side of the investigation (South Korea work), since the findings are related for each case and those conclusions were thought about in accordance with both sides.

5.2 Procedure

Each session started with a demographics and VR, Gaming, and Boxing experience questionnaire. Next, we presented a brief introduction to the goal of the study. Participants were informed they would be exploring a variety of audio feedback configurations to convey information about a virtual boxing opponent, their coach, in an immersive virtual environment and choose their preferred configuration.

Participants were then assisted to wear an Oculus Quest 2 and its respective handheld controllers. Then they were given a brief overview of what they were wearing and how it enabled them to interact with an immersive environment by hitting their coach's hand/arm protections. This was followed by a practice session where participants started immediately inside the boxing environment, with a punching bag in front of them. We then told them they were allowed to move their arms and hit a punching bag to get familiar with the environment. During practice, there was no audio feedback other than the sound of collisions. This was so the user would understand the physicality of using immersive VR and experience the environment without any location/behaviour sound other than hit collision of their hands.

Then we started the exploration of the design space. We explored each category, one by one, and let participants try all the possible options within that category and select their preferred one. When moving to the next category, participants would keep their prior preferences but were always allowed to revisit previous selections and update their preferences. Participants would then start receiving the first audio configuration and the environment changed to now feature an opponent on a fixed location in front of the participant. We could control the environment on the PC to allow to quickly change audio feedback configurations and adjust opponents behaviour (e.g., moving the arm mitts). During this process, an experimenter guided participants to explore the design space one category at the time allowing them to both experience and customize the experience. Participants guided the experience and in each category, they were free to ask (and encouraged) to experience or change previously explored categories (e.g. changing Sound Type from speech to sonification when exploring Audio Source). Akin to the work by Nicolau *et al.* [72], the procedure mimicked an optometry appointment where interface variations were experienced until the participant and researcher were confident in the selection of the participant's overall preferred configuration. When trying each configuration, participants could try out the application freely, by punching the mitts of the opponent. Sound feedback informed participants if they have missed or hit the target – e.g., the sound of hitting the hand/arm protections, head, or body was different.

After selecting their preferred configuration, participants were asked to perform a short, controlled task to assess its ability to convey the intended meaning. Participants were asked to punch the coach's hand protections, by understanding their location. Participants had five trials, where they had a maximum of five attempts (per trial) to hit one of

the coach's hands, without time constraints. The hands' positions would change between trials. We limited the possible hand locations to a 2×3 grid for simplicity; [up(head), down(torso)] × [left, center, right]. With up or down (meaning head or torso) and left, center, or right.

After completing the task, participants were asked to walk the researcher through their thought and selection process. Finally, we conducted a ten minute semi-structured interview focused on their thoughts about the experience, and possible changes they would like to see or be able to configure/change on the audio feedback provided. All participants were compensated for their time.

5.3 Participants

For S_{loc}, we recruited eight legally blind participants, seven of them with light perception at most and one with residual sight in one eye. One participant's age ranged between 18-30, five between 30-50, two were over 50. As for S_{mov}, we recruited eight participants who are legally blind. Six of them were totally blind and two had residual vision. P16 had a minor hearing impairment in the right ear. Three participants' ages ranged between 18-30, four between 30-50 and one was older than 50. Overall, two participants (P8, P10), reported having tried VR once. None of the participants had experience with boxing.

Table 5.1: Participants' audio configuration preference for S_{loc} (P1-P8), and S_{mov} (P9-P16).

Scenario	PID	Trigger	Sound Type	Time Signal	Concurrency	Spatialization
S_{loc}	**P1**	On Movement	Sonification	Discrete	Sequential	3D
	P2	On Movement	Speech	Discrete	Sequential	3D
	P3	On Movement	Speech	Discrete	Sequential	3D
	P4	On Request	Speech	Discrete	Sequential	3D
	P5	On Request	Speech	Discrete	Concurrent*	Dichotic
	P6	On Existence	Speech	Periodic	Sequential	3D
	P7	On Request	Speech	Discrete	Sequential	3D
	P8	On Existence	Speech	Periodic	Sequential	3D
S_{mov}	**P9**	On proximity	Speech	Discrete	Concurrent	Monaural
	P10	On proximity	Speech	Discrete	Sequential	Monaural
	P11	On proximity	Speech	Discrete	Concurrent	Monaural
	P12	On proximity	Speech	Discrete	Sequential	Monaural
	P13	On proximity	Speech	Periodic	Sequential	Dichotic
	P14	On movement	Speech	Continuous	Sequential	3D
	P15	On movement	Speech	Continuous	Concurrent	Dichotic
	P16	On movement	Speech	Periodic	Sequential	Monaural

5.4 Data Analysis

We performed a thematic analysis with primarily deductive coding [94]. We created a codebook based on our concepts of interest, which included the categories and values for the design space and 17 other codes associated with how the configuration was perceived (e.g., Novelty, Comfort, Overload). The codebook was created after the user studies were conducted when researchers had familiarity with the data but before the coding process. The codebook was discussed and iterated by the whole research team and used to code all audio transcriptions for both scenarios. The thematic analysis result presented below encapsulates the discussions that ensued based on the coding, observations, and notes taken during the user studies. Note that researchers translated participants' quotes to English as sessions were conducted in non-English speaking countries.

5.5 Findings

We examine participants' preferences (Table 5.1) and the personalization process when engaged in exploring the design space. We first present the success rate with their preferred method in each case study. Next, we discuss the established themes from reflecting, discussing and iterating over the findings by collaborative sense-checking the data resulting from both scenarios.

5.5.1 Success Rate

As a quantitative assessment of the effectiveness of the participants' preferred configuration, we asked participants to hit five different locations of the opponent standing still (S_{loc}) or guard themselves against moving objects coming from six different locations (S_{mov}) at the end of the design space exploration. As a result, most participants completed the task successfully on their first try regardless of their configuration choices (Table 5.2. For S_{loc}, all participants were able to hit the target with their preferred configuration in all tasks, being that in 82.5% (33 out of 40) of the trials the target was hit on the first attempt. Similarly, for S_{mov}, six participants were able to guard themselves from the target in all trials and the other two guarded themselves successfully in half of the trials (87.5%, 42 out of 48).

To be specific, for S_{loc}, all participants were able to hit the target with their preferred configuration in all tasks. Of the 40 attempts (five trials per participant), 33 targets were hit on the first try, five on the second, and lastly two on the third try.

Similarly, for S_{mov}, six participants were able to guard themselves from the target in all tasks. Of the 48 attempts (six trials per participant), in 42 targets were guarded on the first try. Two participants (including one with a hearing problem), guarded against three punches out of six on the first try.

Table 5.2: Tries it took each participant of S_{loc} to hit the mitts out of five tries.

PID	Test 1	Test 2	Test 3	Test 4	Test 5	
P1	1	3	1	1	1	
P2	1	1	1	1	2	
P3	2	1	1	1	1	
P4	1	1	1	1	2	
P5	1	1	1	2	1	
P6	1	1	1	1	1	
P7	2	1	1	1	3	
P8	1	1	1	1	1	**Total Average**
Average	1.25	1.25	1	1.125	1.5	1.225

5.5.2 The Overall Preference

The preferred configuration while exploring the design space is shown in Table 5.1. Regarding **Sound Type** – in both scenarios – *Speech* was highly preferred (N = 15 out of 16) when compared with *Sonification* because it conveyed greater certainty about the position of elements. As for **Concurrency**, the majority of the participants (N = 12) preferred the information to be delivered sequentially rather than concurrently even when there were multiple changes happening simultaneously[1] mostly due to cognitive load. For the same reason, ten participants preferred **Feedback Mode** to be *Discrete*.

On the other hand, configurations in the remaining dimensions varied depending on the scenario. For instance, in **Spatialization**, 3D audio was highly preferred in the S_{loc} scenario, while monaural audio was more popular for the S_{mov} scenario. It is relevant to note, however, that the monaural condition was not explored in the S_{loc} scenario. Likewise, preferences for the **Trigger** category were different between scenarios. The preferred trigger type varied between participants in the S_{loc} scenario while participants in the S_{mov} scenario slightly preferred feedback triggered on proximity. We further detail the participants' preferences and their rationale in the upcoming sections.

5.5.3 Speech Feedback for Precise Information

All but one participant chose to use *Speech* over *Sonification* as **Sound Type**, mainly because it provided the exact location (e.g., "upper left") of the objects of interest. P8 mentioned:

[1]Note that P5 chose *Concurrent* combined with an *On Request* trigger, meaning he could experience concurrent feedback if pressing both controllers simultaneously. Still, he never pressed both buttons simultaneously and therefore did not experience concurrent feedback

> *"I think it is the assertiveness in the message because the voice tells us clearly, where it is... where the target is. Top left, down right, whatever."*

This preference contrasted with our expectations, especially when conveying information about moving objects (i.e., S_{mov}), given the ability of sonification techniques to convey feedback about dynamic changes [75, 8, 43]. One possible reason is the limited number of discrete options (a set of six predefined locations such as "upper right"), which make it easy and quick to describe using speech. A less restrictive environment with continuous values (e.g., allowing the location to be anywhere in the environment) could be more difficult to describe via speech and therefore put sonification as an appropriate candidate. In this scenario, the need to interpret the audio cues and convert them to a spatial location in the environment made it harder – and more cognitively demanding – to identify the exact location of objects compared to a simple speech message.

Another explanation could be the participants' little exposure to sonification. Unlike speech feedback, sonification may require some training to comprehend its meaning [75, 102]. Conversely, participants found in speech a familiar modality that they are exposed to daily on their devices. P5 stated,

> *"The voices are good, we already know them [i.e. familiar TTS voices]. The voices you got are our friends. Therefore it is very good, and they are very perceptible".*

In addition, the use of different voices (male and female) to convey information about each of the hands, helped participants to better identify the feedback source, as described by P8:

> *"The voice is the more effective of all, as it also has the detail of having a female voice on the left and a male voice on the right (...) It helps us identify much better and separate much better what is left and what is right."*

5.5.4 Spatialization was preferred, but not always!

The S_{loc} scenario has shown a major preference (seven out of eight) for *3D audio* in comparison to *Dichotic* presentation of audio, mainly because it transmits the real position of the target, making it more realistic than presenting sound in only one of the ears. When used together with speech, *3D audio* worked as a complement and helped participants gain confidence about the location of objects. For instance, P7 mentioned:

> *"The sound, which sounds different from left to right, informed me about the direction anyway, so I could grasp the location a little more clearly."*

In contrast to S_{loc}, the S_{mov} scenario has also explored a *monaural* condition that ended up being preferred by the majority of participants (five out of eight). Note that

in S_{mov} we used the HMD speakers, which may result in a worse perception of location when compared to the use of headphones. Still, the main reason for preferring *Monaural* was related to transmitting distance (which is not explored in S_{loc}), which made it more difficult for participants to detect the sound when the object was further away. In this scenario, participants wanted to be aware that one punch was coming and therefore preferred the condition where that information was available – and more clearly – as early as possible. Thus, despite the advantages of *3D audio* in conveying more realistic feedback and information about direction, representing distance needs careful consideration to make sure that relevant information is clearly transmitted to the users even when further away.

In addition, *Monaural* sound was preferred by P16, who has a hearing impairment in one ear, because it allowed him to rely on one ear without missing useful information, unlike *3D or Dichotic audio* which requires both ears to interpret the feedback. It suggests the need to further consider the impact of design decisions when users have multiple impairments:

> *"There is only one reason. Your right and left ears are not the same. So, I think the accuracy will be higher if listening to the right and left sides together."* (P16)

5.5.5 The Effect of Perceived Mental or Physical Loads on Preferred Trigger Options

Participants were split in their preference for the audio trigger. Three preferred to control when to get feedback (*On Request*) as they can expect when they will get feedback just by pressing a button on the hand-held controller and act upon it. P7 stated:

> *"I prefer to press the button (...) because this way I control the game."*

On the other hand, some participants found that pressing a button to receive feedback was an additional step that made the experience more demanding, particularly given the novelty of the experience. P2 specified that,

> *"Note that it distracts you a bit, having to press a button. it is another mechanical thing that you have to do."*

As for S_{mov}, slightly more than half of the participants chose *On proximity* over *On Movement*, suggesting that they would like to receive feedback only when the moving objects are close enough to demand action. Otherwise, it may be overwhelming (or unnecessary) to keep track of changes in the movement, which in this case was likely caused by relatively slow punches and longer times to react. P11 said,

> *"With the first option [On movement], I have to detect the sound beforehand and prepare for it. But, with this one [On proximity], it is enough to react as soon as I hear the sound. There is no need to hold my hands in advance."*

5.5.6 Concurrent and Continuous Feedback to Convey Urgency

Concurrency of audio feedback was important to explore in both scenarios, because the opponent may move their two hands simultaneously. Conveying feedback one hand at a time takes more time for conveying the same amount of information. However, while past research has shown people's ability to identify relevant information levering multiple audio sources [10, 37], this overwhelmed most participants when they tried to understand both audio sources concurrently, leading them to choose the other option. P10 stated that,

> *"The concurrent one was frantic as it told all the sounds of the punches even before I guard them."*

Still, the decision to convey multiple sounds concurrently or sequentially may be influenced by time constraints. For instance, in S_{loc} it is not time-sensitive to know the location of both hands, since the participant could hit the location of either of the coach/opponent's hands. As a consequence, participants did not see the advantage of getting that information more quickly with concurrent feedback.

Continuous feedback was also linked to a higher cognitive load due to the amount of information transmitted. We expected more participants in S_{mov} to choose *Continuous* feedback than the ones in S_{loc} as it can represent the dynamic changes during the continuous movement. The result came out as expected for S_{loc}; none of the participants preferred *Continuous* feedback. However, only two out of eight participants chose the continuous feedback in S_{mov} because others found it overwhelming and unintuitive when combined with speech feedback – note that this may be different in contexts where participants prefer sonification instead of speech. However, some participants commented that continuous feedback can be useful when giving users a sense of urgency. which might be useful if that is the experience designers are striving for their users to have. For example, P14 said that,

> *"I think this sound [Continuous] has the effect of being able to feel the pressuring situation and making it more immersive because it keeps telling me [like an alert]."*

5.5.7 The Desire for Challenging Configurations for Game-Like Experiences

Our main objective was to explore the participants' perceptions of object location and movement. However, participants' preferences were highly influenced by a tradeoff between perceiving these elements and what they thought was an interesting challenge as in a game-like environment (e.g., when hitting their coach's hands).

> *"Rather, it would be advantageous to simplify the sound in the process of running the game."* (P12) *"I differentiated between gameplay and auditory*

> *perception. In terms of gameplay, it is more challenging the more obstacles we have, right?"* (P8)

Participants' approach to this experience has clearly influenced their preferred configuration. For instance, participants who were more focused on perceiving the environment would prefer to trigger feedback *On Request* in order to have more control over the experience. In contrast, those approaching it as a game-like experience associated the lack of control over the feedback as a more challenging experience and preferred *On Movement*,*On Proximity* or *On Existence* triggers.

Additionally, P8 explicitly mentioned that he selected sequential feedback because he was able to better perceive the location of both hands, but that he would feel more challenged and engaged if receiving concurrent feedback. He felt that these configurations could fit different game levels, where players could progress based on expertise and success:

> *"What I would do is have levels, for the easiest level have it discrete, for example, for the intermediate difficulty having the periodic, and then continuous (...) And when you are used to it, you move to the next level."* (P8)

These comments support that the design is highly influenced by the context and main objectives of the VR application and that designers should consider how each feedback decision contributes to their intended experience.

5.5.8 Uncontrollable feedback associated with challenge and game-like behaviour.

While we did not intend to design the scenarios as games, some participants felt like they were playing a game (e.g., hitting the opponent), and the more participants like to have game-like experience, the tendency for choosing *Periodic*, *On Movement*, or *On Proximity* over *On Request* was stronger. For instance, when participants were in control over when feedback was provided *On Request*, we observed participants taking a more passive approach to the experience, waiting for the researcher to prompt them to move even if they had just triggered feedback. On the other hand, when the feedback trigger was set to either *Periodic*, *On Movement*, or *On Proximity*, participants were immediately more engaged and started to try hitting the opponent or defending themselves without waiting for any prompt from the researchers. They associated the lack of control over the feedback as a more challenging experience.

> *"What I would do is have levels, for the easiest level have it discrete, for example, for the intermediate difficulty having the periodic"* (P8)

Participants were keen to suggest on how the experience could include more unprompted feedback to create a more immersive experience, such as adding movement to the adversary and footsteps, and having the feedback being unpredictable.

"No no, I do not mean to speed up, it is for example - Down right punch sound, Top left punch sound*, Top right punch sound* [P7 vocalized using irregular intervals between commands]"* P7

There was a clear trade-off between control and reliably when understanding the object's location or the behaviour over game challenge and engagement. However, for each participant despite the choices they had to make were the same, the consequences for their experience vastly differ, resulting in the spread of preference shown in **??**.

5.5.9 Feedback Realism for Immersive Experiences

Both scenarios provided feedback about the coach/opponents' hand position in addition to feedback on collision. Still, participants commented that this felt short of a fully immersive experience, which could be augmented by realistic sounds that could either provide useful information – the opponent breathing to convey their location – or just background sound (e.g., the crowd cheering). Also, many participants suggested having additional sound cues for landmarks of the surrounding space and adversary. For instance,

> *"I mean, if the opponent was talking, we could tell if they were on the left or right." (P6)*

Others suggested the "arena" could have different crowd sounds to reference when they get away from the opponent. Another frequent request was to have human voices – specifically of a coach – more excited, shouting the location the user should hit or even shout from the side during a competitive match. For instance,

> *"I think it will be more immersive if you include the sound of the coach shouting from the side or the sound of the audience." (P13)*

5.5.10 The Potential of VR for Embodiment

The two boxing scenarios were unembellished, having only the bare necessity for the experience. While the experience could be more immersive, participants perceived it as a reliable proxy of what boxing is actually like:

> *"I'll tell you something funny, I never payed any attention to boxing, but I came to see how it would be. I once saw a boxing bag and gloves, but I had never punched the bag. I found it super interesting, for those who can't see, using this they have a perfect notion of boxing."* (P5)

One participant suggested augmenting their own body movement in addition to the representations of other objects to better understand their position in the world. This is not surprising given that blind people, particularly congenital blind people, are at risk of having posture deficiencies and tend to have lower posture awareness [4]. P10 provided the following example:

> *"A person stretches their arm up and the sound goes up, stretches [the arm] down it goes down, to the right to the right, or left.. In other words, we could be hearing the sound corresponding to my arm movement"*

Participants were keen to suggest on how the experience could include more unprompted feedback to create a more immersive experience, such as adding movement to the adversary and footsteps, and having the feedback being unpredictable.

5.5.11 Experiencing the Unexplored with VR

While the experience could be similar to that of a boxing coach, it would hardly be fully autonomous for blind users in the real world. This sort of approach to inaccessible sports (and other activities) can enable blind people to autonomously experience a variety of contexts that would otherwise be out of – or difficult to – reach.

> *"Most blind people love baseball very much. But even though there are a lot of baseball games, we cannot do it because we do not know where the ball is coming from. But if this research develops, it would be really helpful if blind people could play games like this with sighted people on an equal basis."* (P14)

5.6 Discussion

In this section, we present lessons learned after exploring a design space of the auditory representation of virtual objects' location and of their behaviours. In addition, we demonstrate how our design space exploration may help inform the design of accessible VR boxing experiences.

5.6.1 Immersive VR Environments for Supporting Accessible Experiences for Blind People

Immersive VR environments have the potential to create accessible experiences, opening new avenues both for creating holistic experiences that are far more accessible than any real-world location, as everything is digitized and tracking of virtual elements is accurate and immediate and for providing access to mainstream VR applications.

In immersive VR environments, users' heads and hands are tracked with HMDs and hand-tracking devices (e.g., hand-held controllers). This can be cumbersome for blind people in terms of setup, compared to traditional video games and other applications with 2D input and output devices. However, it comes with significant benefits as not only users can take a first-person perspective, but also control their avatar with their own body movements – e.g., changing the viewpoint by moving their head orientation or punching by moving their hands. This contrasts with the major challenges faced by blind users when dealing with camera controls and joystick/keyboard input in traditional 3D environments. This opportunity for embodiment has put our VR Boxing applications as safe proxies for this sport, as "*for those who cannot see, using this they have a perfect notion of boxing*" (P5). Moreover, because virtual environments are not bound by real-world limitations (e.g., object detection, tracking), developers can use precise information about every object in the scene (label, position, orientation) as well as user events, and have full control of the environment (e.g., changing the object properties or the feedback design) as demonstrated in prior studies [19, 20, 65, 49, 3, 7, 16].

In this study, we used such knowledge (in particular, hand location) to experiment with various configurations with blind users and understand the impact of each design decision. Still, our application was fairly basic and participants felt the need for a more realistic and immersive experience (e.g., the audience or the coach shouting). Designers and developers should consider how to combine both realistic audio feedback (e.g., the opponent breathing or the audience cheering), and the audio representation of visual objects (supported by our design space) that do not necessarily produce sound (e.g., the opponent's hands). We encourage designers and developers to facilitate immersive VR environments for blind people for designing and supporting accessible interaction.

5.6.2 The Design Space as a Framework to Create Accessible VR Experiences

The design space supported the creation of multiple configurations to represent objects and their movement in two different scenarios. Participants' variance in preferences alongside the high success rates when performing the study tasks show the viability of the different configurations implemented. However, the main goal of the design space is to support designers and developers in exploring different solutions, and not to perform generalist comparisons to reach a preferred, one-size-fits-all solution.

5.6.3 No One-Size-Fits-All Solution

While we often strive for the "one" best solution, when it comes to augmenting objects with audio feedback our findings show a great variety of preferences among blind users, both within and across scenarios. This suggests that designers should consider providing

users (and designers themselves) the ability to personalize their configurations.

As evident in **??**, the preferred configuration varied depending on how each participant perceives a task and what they value. For example, participants who wished to be challenged as in game-like environments preferred the audio feedback to be given automatically. For instance, P8 envisioned a sequence of game levels where configurations get increasingly more challenging to keep the users engaged. On the other hand, participants who appreciate the sense of user control preferred feedback played on request. Yet, all of them were able to achieve high success rates implying that the effective configurations can differ depending on the user and the context. We could see that participants preferred different spatialization and control over the feedback mode and trigger and Sound Type, Concurrency and Spatialization were the only three categories where there appeared to be a clear preference by most participants.

We highlight the importance of allowing: 1) designers to explore the design space to create experiences that are in line with their context and objectives – e.g., is the main goal to convey precise information about the object location or to challenge users with more demanding tasks? 2) users to try out and select different configurations to understand what are their preferred combinations. Participants suggested that the experience enabled them to have walk through an interactive tutorial experimenting and adapting to best fit their behaviours and performance. Meanwhile, findings also imply that the preferred configuration may change over time as users would become familiar with the feedback design. For instance, novice users may prefer speech feedback for its low learning cost but they may switch to sonification for efficiency after long-term use. In addition, one should note that even when participants chose the same configuration, the reasons for their decision may vary. Thus, it is essential for designers to focus more on the participants' thought process, in particular on how each feedback decision was made rather than the decisions themselves in order to provide a more personalized experience.

> *"I think these are layers, they are an adaptive phase of the game, meaning, in the first phase a less experienced person, perhaps, would benefit . . . from only having information when pressing the button. But when you are used to it, you go to the next level."* (P8)

5.6.4 Leveraging the Design Space to Create a Virtual Boxing Experience

We presented a scenario in the context of virtual boxing to understand and exemplify how our proposed design space can be leveraged to design augmentative audio feedback for blind people. By exercising the design space, we found no one size fits all solution, with participants having a wide selection of configurations. Still, we were able to identify trends, specific to our boxing scenarios, that inform the future creation of VR boxing experiences:

- *Trigger*: When familiarizing players with the demands of the game, trigger on request may lower the challenge and allow players to solely focus on understanding the relationship between their actions and the feedback received. *On Request*, can also accommodate interaction where players are in control of the pacing of the task but this also demands action from users, which may increase their load.
- *Sound Type*: *Speech* is to be selected due to its ability to provide precise information about location, given a limited number of (short) options (e.g., top left, bottom right).
- *Time Signal*: Since actions are relatively quick, and are a sequence of one-time events that are mostly relevant on occurrence, *Discrete* time signals are the most suitable option.
- *Concurrency*: if the time delay is minimal in sequential feedback, the disadvantages of concurrent feedback (mental load) outweigh its benefits. Still, it can be used to give a sense of urgency and challenge.
- Spatialization: use *3D audio* when conveying location where the distance of an object does not change.

Although we advocate for the ability to customize, it is equally relevant to highlight that these choices significantly affect the experience someone has with virtual environments (e.g., make it more or less perceivable or engaging). We believe the design space can provide a framework from which to explore the viable design options and let users show the consequences of the designs while simultaneously uncovering latent feedback desires. It is then up to researchers and designers to select what we want the experience to be.

Chapter 6

Conclusions & future work

In this study, we investigated the potential of virtual reality (VR) as a tool to provide feedback to blind individuals. Our findings demonstrate that VR holds great promise in enhancing the sensory experiences of blind users and enabling them to perceive and interact with virtual environments. Through a comprehensive examination of different visual feedback techniques, we gained valuable insights into the most effective approaches for conveying visual information to blind people. These findings can have significant implications for the development of assistive technologies and the advancement of inclusive design principles for visually impaired individuals.

We proposed a design space with nine categories to support designers in exploring various audio representations of virtual objects and their behaviours in order to create accessible VR experiences for blind people. We demonstrated how the design space can be used with two VR boxing example scenarios. We performed two independent design space explorations, by separate research teams, showcasing its flexibility as a theoretical framework capable of supporting the exploration of multiple viable solutions for augmenting virtual objects with a nonvisual, audio representation. In addition, it enables designers to reflect on the impact of different design decisions, supporting them in creating accessible experiences.

While our study might contribute with some valuable insights into the use of VR for blind individuals, there are several avenues for further exploration and improvement in this field. The development of standardized guidelines for designing accessible virtual environments is crucial. Currently, there is a lack of consensus on the best practices for creating inclusive VR experiences for blind people. Establishing a set of guidelines and principles, based on user feedback and expert evaluations, would greatly assist developers and designers in creating more accessible and engaging VR applications.

Furthermore, it would be ideal to expand the scope of VR applications beyond mere entertainment. Virtual reality has the potential to revolutionize education, employment training, and social inclusion for blind individuals. Future work should explore the integration of VR technology in these domains, addressing specific needs and challenges

faced by blind users. For example, VR could be utilized to simulate real-world scenarios, providing immersive training experiences for job skills, mobility, and independent living.

Lastly, the usability and accessibility of VR hardware should be improved to cater to the diverse needs of blind users. Efforts should be made to develop lightweight and portable VR devices with built-in accessibility features, such as speech recognition and gesture control, to ensure ease of use and seamless integration into daily life. In conclusion, our study highlights the immense potential of virtual reality in providing visual feedback to blind individuals. By leveraging auditory and haptic cues, VR can offer a rich and immersive sensory experience, enhancing spatial awareness, navigation abilities, and object recognition. However, further research and development are needed to refine the existing techniques, establish design guidelines, expand application domains, and improve hardware accessibility. With continued advancements in technology and a multidisciplinary approach, virtual reality could possibly become a transformative tool for empowering blind individuals and fostering inclusivity in society.

Appendix A

Study Guide

Let's test various configurations for a boxing VR application.

Throw a punch and leave your arm straight; (Distance)

We have a predefined configuration, which we will go through 1 by 1 to find out what you prefer;

1st Spatialization (3D, dichotic)

At the moment, feedback on the position of the arms is given by pressing the bottom button, operated by the right thumb.

Test yourself;

Task 1;

Task 2;

Want to try the other one?

2nd speech (speech, sonification)

Test yourself;

Task 1;

Task 2;

Want to try the other one?

3rd Trigger (OnRequest, OnMovement)

For OnRequest you need to press the button, for OnMovement you don't need to the press any button;

Test yourself;

Task 1;

Task 2;

Want to try the other one?

4th Feedback Mode (Discrete, Periodic, Continuous)

Continuous will only be sonified;

Only as Discrete you'll need to use the buttons;

Test yourself;

Task 1;

Task 2;
Want to try the other one?
5th Concurrency (sequential or concurrent)
Test yourself;
Task 1;
Task 2;
Want to try the other one?
Is there a setting you haven't tried that you want to try/do you think is better?

Appendix B

Interview Guide

B.1 Before Testing

VRBoxing

Muito obrigado por terem aceitado participar neste estudo. O meu nome é Rodrigo Nogueira e sou estudante de Mestrado em Engenharia Informática na FCUL.

O objetivo da tese é perceber de que forma poderemos tornar realidade virtual mais acessível a pessoas cegas. Para isso, estamos a testar diferentes formas de passar a localização de diferentes objetos no ambiente virtual e perceber de que forma é mais fácil para um cego processar essa informação de forma correta.

* Indicates required question

ID

Your answer

Nome

Your answer

Idade *

- <18
- 18 - 30
- 30 - 50
- >50

Género *

- ◯ Masculino
- ◯ Feminino
- ◯ Outro

Tipo de Cegueira

- ◯ Total
- ◯ Sensibilidade a luz
- ◯ Other:

Experiência com Realidade Virtual *

- ◯ Sim
- ◯ Não

Experiência com boxe *

- ◯ Sim
- ◯ Não

Como classificaria o seu conhecimento de boxe? (De 1 a 5 onde 1 é Pouco ou nada e 5 especialista) *

- 1
- 2
- 3
- 4
- 5

Como classificaria a sua experiência com jogos? (De 1 a 5 onde 1 é Pouco Experiente e 5 Muito Experiente) *

- 1
- 2
- 3
- 4
- 5

Submit Clear form

B.2 After Testing

Question 1: What did you think of this experience?

Question 2: I would like you to explain to me again your preference and why you preferred these settings

Question 3: What would you change or add to the audio you heard during the experience?

Question 4: Can you imagine other ways to know the opponent's position, other than the ones we showed?

Bibliography

[1] Superhuman sports society. https://superhuman-sports.org/, 1 2022.

[2] Hojun Aan, Sangsun Han, Hyeonkyu Kim, and Kibum Kim. Remote virtual showdown. In *SIGGRAPH Asia 2020 XR*, SA '20, New York, NY, USA, 2020. Association for Computing Machinery.

[3] Kevin Allain, Bas Dado, Mick Van Gelderen, Olivier Hokke, Miguel Oliveira, Rafael Bidarra, Nikolay D Gaubitch, Richard C Hendriks, and Ben Kybartas. An audio game for training navigation skills of blind children. In *2015 IEEE 2nd VR workshop on sonic interactions for virtual environments (SIVE)*, pages 1–4. IEEE, 2015.

[4] Abdullah Z Alotaibi, Ahmad Alghadir, Zaheen A Iqbal, and Shahnawaz Anwer. Effect of absence of vision on posture. *Journal of physical therapy science*, 28(4):1374–1377, 2016.

[5] Ronny Andrade, Melissa J. Rogerson, Jenny Waycott, Steven Baker, and Frank Vetere. Playing blind: Revealing the world of gamers with visual impairment. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–14, New York, NY, USA, 2019. Association for Computing Machinery.

[6] Ronny Andrade, Jenny Waycott, Steven Baker, and Frank Vetere. Echolocation as a means for people with visual impairment (pvi) to acquire spatial knowledge of virtual space. *ACM Transactions on Accessible Computing (TACCESS)*, 14(1):1–25, 2021.

[7] Oscar Ariza, Gerd Bruder, Nicholas Katzakis, and Frank Steinicke. Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 327–334. IEEE, 2018.

[8] Stephen Barrass and Gregory Kramer. Using sonification. *Multimedia systems*, 7(1):23–31, 1999.

[9] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.

[10] Adelbert W Bronkhorst. The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica*, 86(1):117–128, 2000.

[11] Ayana Burkins and Regis Kopper. Wayfinding by auditory cues in virtual environments. In *2015 IEEE Virtual Reality (VR)*, pages 155–156. IEEE, 2015.

[12] Sebastian Büttner, Henrik Mucha, Markus Funk, Thomas Kosch, Mario Aehnelt, Sebastian Robert, and Carsten Röcker. The design space of augmented and virtual reality applications for assistive environments in manufacturing: a visual approach. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*, pages 433–440, 2017.

[13] Michele Capella-McDonnall. The need for health promotion for adults who are visually impaired. *Journal of Visual Impairment & Blindness*, 101(3):133–145, 2007.

[14] Stuart K Card, Jock D Mackinlay, and George G Robertson. The design space of input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 117–124, 1990.

[15] Hao Chen, Arindam Dey, Mark Billinghurst, and Robert W. Lindeman. Exploring the design space for multi-sensory heart rate feedback in immersive virtual reality. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, OZCHI '17, page 108–116, New York, NY, USA, 2017. Association for Computing Machinery.

[16] Yi Fei Cheng, Hang Yin, Yukang Yan, Jan Gugenheimer, and David Lindlbauer. Towards understanding diminished reality. In *CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2022.

[17] George Chernyshov, Kirill Ragozin, Jiajun Chen, and Kai Kunze. Dubhap: A sensory substitution based superhuman sport. In *Proceedings of the First Superhuman Sports Design Challenge: First International Symposium on Amplifying Capabilities and Competing in Mixed Realities*, SHS '18, New York, NY, USA, 2018. Association for Computing Machinery.

[18] SeungA Chung, Kyungyeon Lee, and Uran Oh. Investigating three-dimensional directional guidance with nonvisual feedback for target pointing task. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 206–210. IEEE, 2020.

[19] SeungA Chung, Kyungyeon Lee, and Uran Oh. Understanding the two-step non-visual omnidirectional guidance for target acquisition in 3d spaces. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 339–346. IEEE, 2021.

[20] SeungA Chung, Soobin Park, Sohyeon Park, Kyungyeon Lee, and Uran Oh. Improving mealtime experiences of people with visual impairments. In *Proceedings of the 18th International Web for All Conference*, pages 1–11, 2021.

[21] Antonio Cobo, Nancy E Guerrón, Carlos Martín, Francisco del Pozo, and José Javier Serrano. Differences between blind people's cognitive maps after proximity and distant exploration of virtual environments. *Computers in Human Behavior*, 77:294–308, 2017.

[22] Chetz Colwell, Helen Petrie, Diana Kornbrot, Andrew Hardwick, and Stephen Furner. Haptic virtual reality for blind computer users. In *Proceedings of the third international ACM conference on Assistive technologies*, pages 92–99, 1998.

[23] Erin C Connors, Elizabeth R Chrastil, Jaime Sánchez, and Lotfi B Merabet. Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches. *Frontiers in human neuroscience*, 8:223, 2014.

[24] Konstantinos Dadamis, John Williamson, and Roderick Murray-Smith. 3d cues for human control of target acquisition in auditory augmented reality. 2022.

[25] Gaël Dubus and Roberto Bresin. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one*, 8(12):e82491, 2013.

[26] Agebson Rocha Façanha, Ticianne Darin, Windson Viana, and Jaime Sánchez. O&m indoor virtual environments for people who are blind: A systematic literature review. *ACM Transactions on Accessible Computing (TACCESS)*, 13(2):1–42, 2020.

[27] Agebson Rocha Façanha, Ticianne Darin, Windson Viana, and Jaime Sánchez. O&m indoor virtual environments for people who are blind: A systematic literature review. *ACM Trans. Access. Comput.*, 13(2), aug 2020.

[28] BoYu Gao, HyungSeok Kim, Hasup Lee, Jooyoung Lee, and Jee-In Kim. Effects of continuous auditory feedback on drawing trajectory-based finger gestures. *IEEE Transactions on Human-Machine Systems*, 48(6):658–669, 2018.

[29] BoYu Gao, Yujun Lu, HyungSeok Kim, Byungmoon Kim, and Jinyi Long. Spherical layout with proximity-based multimodal feedback for eyes-free target acquisition in virtual reality. In *International Conference on Human-Computer Interaction*, pages 44–58. Springer, 2019.

[30] Tom A Garner. *Echoes of Other Worlds: Sound in Virtual Reality: Past, Present and Future*. Springer, 2017.

[31] Aaron Gluck, Kwajo Boateng, and Julian Brinkley. Racing in the dark: Exploring accessible virtual reality by developing a racing game for people who are blind. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1):1114–1118, 2021.

[32] Aaron Gluck, Kwajo Boateng, and Julian Brinkley. Racing in the dark: Exploring accessible virtual reality by developing a racing game for people who are blind. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 65, pages 1114–1118. SAGE Publications Sage CA: Los Angeles, CA, 2021.

[33] David Gonçalves, Manuel Piçarra, Pedro Pais, João Guerreiro, and André Rodrigues. " my zelda cane": Strategies used by blind players to play visual-centric digital games. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, pages 1–15, 2023.

[34] Inês Gonçalves, André Rodrigues, Tiago Guerreiro, and João Guerreiro. Inclusive social virtual environments: Exploring the acceptability of different navigation and awareness techniques. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–7, 2023.

[35] David Gonçalves, André Rodrigues, and Tiago Guerreiro. Playing with others: Depicting multiplayer gaming experiences of people with visual impairments. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20, New York, NY, USA, 2020. Association for Computing Machinery.

[36] João Guerreiro. Towards screen readers with concurrent speech: where to go next? *ACM SIGACCESS Accessibility and Computing*, (115):12–19, 2016.

[37] João Guerreiro and Daniel Gonçalves. Scanning for digital content: How blind and sighted people perceive concurrent speech. *ACM Transactions on Accessible Computing (TACCESS)*, 8(1):1–28, 2016.

[38] João Guerreiro, Daisuke Sato, Dragan Ahmetovic, Eshed Ohn-Bar, Kris M Kitani, and Chieko Asakawa. Virtual navigation for blind people: Transferring route

knowledge to the real-world. *International Journal of Human-Computer Studies*, 135:102369, 2020.

[39] João Guerreiro, Dragan Ahmetovic, Kris M Kitani, and Chieko Asakawa. Virtual navigation for blind people: Building sequential representations of the real-world. In *The proceedings of the 19th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2017.

[40] Thomas Hermann, Andy Hunt, and John G Neuhoff. *The sonification handbook*. Logos Verlag Berlin, 2011.

[41] Wilko Heuten, Daniel Wichmann, and Susanne Boll. Interactive 3d sonification for the exploration of city maps. In *Proceedings of the 4th Nordic conference on Human-Computer Interaction: Changing roles*, pages 155–164, 2006.

[42] Emil R Hoeg, Lynda J Gerry, Lui Thomsen, Niels C Nilsson, and Stefania Serafin. Binaural sound reduces reaction time in a virtual reality search task. In *2017 IEEE 3rd VR workshop on sonic interactions for virtual environments (SIVE)*, pages 1–4. IEEE, 2017.

[43] Oliver Höner, Andy Hunt, Sandra Pauletto, Niklas Röber, Thomas Hermann, and Alfred O Effenberg. Aiding movement with sonification in "exercise, play and sport". In *The sonification handbook*. 2011.

[44] Ibrar Hussain, Ling Chen, Hamid Turab Mirza, Kong Xing, and Gencai Chen. A comparative study of sonification methods to represent distance and forward-direction in pedestrian navigation. *International Journal of Human-Computer Interaction*, 30(9):740–751, 2014.

[45] Gesu India, Mohit Jain, Pallav Karya, Nirmalendu Diwakar, and Manohar Swaminathan. Vstroll: An audio-based virtual exploration to encourage walking among people with vision impairments. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–13, 2021.

[46] Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. A taxonomy of sounds in virtual reality. In *Designing Interactive Systems Conference 2021*, pages 160–170, 2021.

[47] Tiger Ji, Brianna R Cochran, and Yuhang Zhao. Vrbubble: Enhancing peripheral awareness of avatars for people with visual impairments in social virtual reality. In *The International ACM SIGACCESS Conference on Computers and Accessibility*, New York, NY, USA, 2022. Association for Computing Machinery.

[48] Zihao Ji, Weijian Hu, Ze Wang, Kailun Yang, and Kaiwei Wang. Seeing through events: Real-time moving object sonification for visually impaired people using event-based camera. *Sensors*, 21(10), 2021.

[49] Oliver Beren Kaul and Michael Rohs. Haptichead: 3d guidance and target acquisition through a vibrotactile grid. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2533–2539, 2016.

[50] Nam Wook Kim, Shakila Cherise Joyner, Amalia Riegelhuth, and Y Kim. Accessible visualization: Design space, opportunities, and challenges. In *Computer Graphics Forum*, volume 40, pages 173–188. Wiley Online Library, 2021.

[51] Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John H Flowers, Nadine Miner, and John Neuhoff. Sonification report: Status of the field and research agenda. 2010.

[52] Julian Kreimeier and Timo Götzelmann. Two decades of touchable and walkable virtual reality for blind and visually impaired people: A high-level taxonomy. *Multimodal Technologies and Interaction*, 4(4), 2020.

[53] Julian Kreimeier and Timo Götzelmann. Two decades of touchable and walkable virtual reality for blind and visually impaired people: A high-level taxonomy. *Multimodal Technologies and Interaction*, 4(4), 2020.

[54] Julian Kreimeier, Pascal Karg, and Timo Götzelmann. Blindwalkvr: Formative insights into blind and visually impaired people's vr locomotion using commercially available approaches. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '20, New York, NY, USA, 2020. Association for Computing Machinery.

[55] Julian Kreimeier, Pascal Karg, and Timo Götzelmann. Blindwalkvr: formative insights into blind and visually impaired people's vr locomotion using commercially available approaches. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pages 1–8, 2020.

[56] Árni Kristjánsson, Alin Moldoveanu, Ómar I Jóhannesson, Oana Balan, Simone Spagnol, Vigdís Vala Valgeirsdóttir, and Rúnar Unnthorsson. Designing sensory-substitution devices: Principles, pitfalls and potential1. *Restorative neurology and neuroscience*, 34(5):769—787, September 2016.

[57] Orly Lahav and David Mioduser. Haptic-feedback support for cognitive mapping of unknown spaces by people who are blind. *International Journal of Human-Computer Studies*, 66(1):23–35, 2008.

[58] Anatole Lécuyer, Pascal Mobuchon, Christine Mégard, Jérôme Perret, Claude Andriot, and J-P Colinot. Homere: a multimodal system for visually impaired people to explore virtual environments. In *IEEE Virtual Reality, 2003. Proceedings.*, pages 251–258. IEEE, 2003.

[59] Grégory Leplâtre and Stephen A Brewster. Designing non-speech sounds to support navigation in mobile phone menus. 2000.

[60] Danielle Lottridge, Rebecca Weber, Eva-Rae McLean, Hazel Williams, Joanna Cook, and Huidong Bai. Exploring the design space for immersive embodiment in dance. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 93–102. IEEE, 2022.

[61] Allan MacLean, Richard Young, Victoria Bellotti, and Thomas Moran. Design space analysis: Bridging from theory to practice via design rationale. *Proceedings of Esprit*, 91:720–730, 1991.

[62] Shachar Maidenbaum and Amir Amedi. Blind in a virtual world: Mobility-training virtual reality games for users who are blind. In *2015 IEEE Virtual Reality (VR)*, pages 341–342. IEEE, 2015.

[63] Alexander Marquardt, Christina Trepkowski, Tom David Eibich, Jens Maiero, and Ernst Kruijff. Non-visual cues for view management in narrow field of view augmented reality displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 190–201. IEEE, 2019.

[64] Alexander Marquardt, Christina Trepkowski, Tom David Eibich, Jens Maiero, Ernst Kruijff, and Johannes Schöning. Comparing non-visual and visual guidance methods for narrow field of view augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3389–3401, 2020.

[65] Stefan Carlo Michalski, Ancret Szpak, Dimitrios Saredakis, Tyler James Ross, Mark Billinghurst, and Tobias Loetscher. Getting your game on: Using virtual reality to improve real table tennis skills. *PloS one*, 14(9):e0222351, 2019.

[66] Tony Morelli, John Foley, Luis Columna, Lauren Lieberman, and Eelke Folmer. Vi-tennis: a vibrotactile/audio exergame for players who are visually impaired. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, pages 147–154, 2010.

[67] Tony Morelli and Eelke Folmer. Real-time sensory substitution to enable players who are blind to play video games using whole body gestures. *Entertainment Computing*, 5(1):83–90, 2014.

[68] Meredith Ringel Morris, Jazette Johnson, Cynthia L Bennett, and Edward Cutrell. Rich representations of visual content for screen reader users. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–11, 2018.

[69] Martez Mott, John Tang, Shaun Kane, Edward Cutrell, and Meredith Ringel Morris. "i just went into it assuming that i wouldn't be able to have the full experience": Understanding the accessibility of virtual reality for people with limited mobility. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20, New York, NY, USA, 2020. Association for Computing Machinery.

[70] Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. Requirements and design space for interactive public displays. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1285–1294, 2010.

[71] Vishnu Nair, Jay L Karp, Samuel Silverman, Mohar Kalra, Hollis Lehv, Faizan Jamil, and Brian A. Smith. Navstick: Making video games blind-accessible via the ability to look around. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, page 538–551, New York, NY, USA, 2021. Association for Computing Machinery.

[72] Hugo Nicolau, André Rodrigues, André Santos, Tiago Guerreiro, Kyle Montague, and João Guerreiro. The design space of nonvisual word completion. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '19, page 249–261, New York, NY, USA, 2019. Association for Computing Machinery.

[73] Casey O'Callaghan. Sounds and events. 2009.

[74] Uran Oh, Stacy Branham, Leah Findlater, and Shaun K Kane. Audio-based feedback techniques for teaching touchscreen gestures. *ACM Transactions on Accessible Computing (TACCESS)*, 7(3):1–29, 2015.

[75] Uran Oh, Shaun K Kane, and Leah Findlater. Follow that sound: using sonification and corrective verbal feedback to teach touchscreen gestures. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–8, 2013.

[76] Robert Pasnau. What is sound? *The Philosophical Quarterly*, 49(196):309–324, 1999.

[77] Frank J Penedo and Jason R Dahn. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current opinion in psychiatry*, 18(2):189–193, 2005.

[78] Kaitlin U. Phillips. Virtual reality has an accessibility problem. Scientific American, January 2020.

[79] Manuel Piçarra, André Rodrigues, and João Guerreiro. Evaluating accessible navigation for blind people in virtual environments. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–7, 2023.

[80] Lorenzo Picinali, Amandine Afonso, Michel Denis, and Brian FG Katz. Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. *International Journal of Human-Computer Studies*, 72(4):393–407, 2014.

[81] Irwin Pollack and James M Pickett. Cocktail party effect. *The Journal of the Acoustical Society of America*, 29(11):1262–1262, 1957.

[82] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca A. Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. Iterative design of sonification techniques to support people with visual impairments in obstacle avoidance. *ACM Trans. Access. Comput.*, 14(4), oct 2021.

[83] Kyle Rector, Cynthia L Bennett, and Julie A Kientz. Eyes-free yoga: an exergame using depth cameras for blind & low vision exercise. In *Proceedings of the 15th international acm sigaccess conference on computers and accessibility*, pages 1–8, 2013.

[84] Gang Ren, Side Wei, Eamonn O'Neill, and Fenfang Chen. Towards the design of effective haptic and audio displays for augmented reality and mixed reality applications. *Advances in Multimedia*, 2018, 2018.

[85] Vincent Reynaert, Florent Berthaut, Yosra Rekik, and Laurent Grisoni. The effect of rhythm in mid-air gestures on the user experience in virtual reality. In *IFIP Conference on Human-Computer Interaction*, pages 182–191. Springer, 2021.

[86] AJ Ryan. Thoughts on accessibility issues with vr. Ablegamers, 2020.

[87] Jaime Sánchez, Angelo Tadres, Alvaro Pascual-Leone, and Lotfi Merabet. Blind children navigation through gaming and associated brain plasticity. In *2009 virtual rehabilitation international conference*, pages 29–36. IEEE, 2009.

[88] Toros Senan, Bart Hengeveld, and Berry Eggen. Sounding obstacles for social distance sonification. In *Proceedings of the 17th International Audio Mostly Conference*, AM '22, page 187–194, New York, NY, USA, 2022. Association for Computing Machinery.

[89] Stefania Serafin, Niels C Nilsson, Erik Sikstrom, Amalia De Goetzen, and Rolf Nordahl. Estimation of detection thresholds for acoustic based redirected walking techniques. In *2013 IEEE Virtual Reality (VR)*, pages 161–162. IEEE, 2013.

[90] Alexa F Siu, Mike Sinclair, Robert Kovacs, Eyal Ofek, Christian Holz, and Edward Cutrell. Virtual reality without vision: A haptic and auditory white cane to navigate complex virtual worlds. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–13, 2020.

[91] Brian A Smith and Shree K Nayar. The rad: Making racing games equivalently accessible to people who are blind. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12, 2018.

[92] Grant Stoner. Vr is here to stay. it's time to make it accessible. Wired, March 2022.

[93] Jing Su, Alyssa Rosenzweig, Ashvin Goel, Eyal de Lara, and Khai N Truong. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, pages 17–26, 2010.

[94] Gareth Terry, Nikki Hayfield, Victoria Clarke, and Virginia Braun. The sage handbook of qualitative research in psychology, Sep 2017. Thematic Analysis.

[95] Lauren Thevin, Carine Briant, and Anke M Brock. X-road: virtual reality glasses for orientation and mobility training of people with visual impairments. *ACM Transactions on Accessible Computing (TACCESS)*, 13(2):1–47, 2020.

[96] Darren ER Warburton, Crystal Whitney Nicol, and Shannon SD Bredin. Health benefits of physical activity: the evidence. *Cmaj*, 174(6):801–809, 2006.

[97] Michiharu Watanabe, Takahiro Miura, Masaki Matsuo, Masatsugu Sakajiri, and Junji Onishi. Goalbaural-ii: An acoustic virtual reality training application fornbsp;goalball players tonbsp;recognize various game conditions. In *Computers Helping People with Special Needs: 18th International Conference, ICCHP-AAATE 2022, Lecco, Italy, July 11–15, 2022, Proceedings, Part II*, page 79–88, Berlin, Heidelberg, 2022. Springer-Verlag.

[98] Ryan Wedoff, Lindsay Ball, Amelia Wang, Yi Xuan Khoo, Lauren Lieberman, and Kyle Rector. Virtual showdown: An accessible virtual reality game with scaffolds for youth with visual impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–15, New York, NY, USA, 2019. Association for Computing Machinery.

[99] Evette Weil, Melissa Wachterman, Ellen P McCarthy, Roger B Davis, Bonnie O'Day, Lisa I Iezzoni, and Christina C Wee. Obesity among adults with disabling conditions. *Jama*, 288(10):1265–1268, 2002.

[100] Graham Wilson and Stephen A Brewster. Using dynamic audio feedback to support peripersonal reaching in young visually impaired people. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 209–218, 2016.

[101] Shamima Yasmin. Virtual reality and assistive technologies: A survey. *International Journal of Virtual Reality*, 18(2), 2018.

[102] Tsubasa Yoshida, Kris M Kitani, Hideki Koike, Serge Belongie, and Kevin Schlei. Edgesonic: image feature sonification for the visually impaired. In *Proceedings of the 2nd Augmented Human International Conference*, pages 1–4, 2011.

[103] Yuhang Zhao, Cynthia L Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2018.

[104] Yuhang Zhao, Elizabeth Kupferstein, Hathaitorn Rojnirun, Leah Findlater, and Shiri Azenkot. The effectiveness of visual and audio wayfinding guidance on smartglasses for people with low vision. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–14, 2020.

[105] Daniel Zielasko, Jonas Heib, and Benjamin Weyers. Systematic design space exploration of discrete virtual rotations in vr. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 693–702. IEEE, 2022.

[106] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum (Proc. EuroVis)*, 2022.

[107] Franz Zotter and Matthias Frank. *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature, 2019.