Audio Personalisation for Accessible Augmented Reality Narratives

Joshua Gregg

MSc by Research

University of York
Electronic Engineering

April 2022

Abstract

Augmented Reality (AR) is rapidly growing across a diverse range of uses. Along with other Extended Reality (XR) technologies, it has the potential to not only make the world more accessible but to provide new and unique opportunities for impaired users. This will not happen accidentally but requires greater thought and development with impaired users central to decision making.

This work explores whether creative and personalisable audio practices can be used to develop mobile-based AR content which is accessible to Visually Impaired (VI) users and includes them in enjoyable, immersive ways.

It begins by assessing the accessibility of current AR applications (apps). Despite developments in the use of AR to make the world more accessible, most mobile AR content does not take access needs into account. 37 AR apps were tested against a rubric of features with few found to be accessible. In Study 1, 8 VI participants validated these results by evaluating a subset of these apps.

This work then tests the efficacy of accessibility strategies centred upon two concepts. First, using well-designed audio creatively to include and immerse users in meaningful and enjoyable ways. Second, allowing users to personalise experiences to suit their needs. To this end, a bespoke narrative AR app was developed, with accessibility features initially equivalent to those in Study 1. In Study 2, 6 VI participants and 5 content creators displayed positivity towards the strategies and suggested ideas for their implementation. Based on this, the app was updated to include enhanced audio accessibility features. These were evaluated in Study 3 by 12 VI participants and 3 content creators. The results demonstrate positivity towards the strategies' potential to provide an enjoyable, immersive experience. However, they underline how difficult AR is for VI users, highlighting fundamental challenges which need addressing before such strategies can be applied.

Contents

List of Tables

List of Figures

Acknowledgements

Completing research like this is much more than just an individual effort. Over the course of this project, many have contributed in all sorts of ways.

I would like to express my gratitude to XR Stories for funding this studentship and to BBC R&D for providing support and resources. Without you, this research would not be possible. This is also extended to the Royal National Institute of Blind People (RNIB), London Vision, the Technology Association of Visually Impaired People (TAVIP), VI networks within the BBC and the University of York Disability Support and Programme Design and Learning Technology teams. Thank you for going above and beyond in helping me to get this research out there and find willing volunteers to participate in it. To those willing volunteers in many cases, the tasks asked of you were difficult but you responded with honesty and enthusiasm and were willing to try something new and unusual. Thank you.

Particular thanks goes to my supervisors, Gavin Kearney and Lauren Ward. At every stage of this work your support, encouragement and guidance have been inexhaustible, spurring me on and providing a perfect environment for academic growth. It has been a privilege to work with you both. This recognition is also extended to my thesis advisor, Helena Daffern who provided feedback that was not only valuable but that shaped some of the contributions of this work that I am proudest of.

To everyone else at AudioLab, it's been a privilege to be a part of the community, despite much of the time during this project being spent in lockdowns. Getting to see you all in person again, and share lunchtime conversations (ranging from the ridiculous to the existential) helped provide fresh motivation.

To my son, who was born not long after this commenced. The sleep deprivation caused may have made this more difficult at times, but I would not have it any other way. The joy you have brought to my life is immeasurable. To my wife, Eva Gregg. It's not easy to express just how much one person can champion another, remaining patient and sane in the midst of it all. Thank you.

Last, but by no means least to both sets of parents and siblings, you have provided board game-based breaks, days out, food, alcohol, helped with childcare and household jobs, and the list goes on. Without your love and support, this would have been both much less fun and much more difficult.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1

Introduction

1.1 Motivations

An estimated 2.2 billion people in the world live with some degree of vision impairment, a number which is expected to increase [1, 2]. Support for blind and partially sighted people is frequently lacking and they often experience negative attitudes and difficulties gaining employment and with everyday activities such as travel [3]. One area which often falls short in its inclusivity is technology, especially when innovations are made without thought for accessibility. This can lead to 'best-fit' access solutions which are added on as an afterthought and may miss unique opportunities afforded by new innovations for ease of use and enjoyment. Disability advocates are concerned that these trends are continuing into Extended Reality (XR) technologies, which although relatively new, are growing rapidly and have limited accessibility solutions [4, 5]. This is a missed opportunity, as the immersive and adaptable nature of XR technologies has the potential to include and even provide unique opportunities to impaired users [6, 7]. There is an opportunity in this present moment to address these issues with input from Visually Impaired (VI) users central to discussions, decisions and developments.

Of XR technologies, Augmented Reality (AR) is growing the fastest and is outpacing Virtual Reality (VR) [8, 9]. In AR, an enhanced version of reality is created by overlaying digitally created content such as information, objects, sounds or characters onto the user's real-world environment [10]. This is done through a device, such as a smartphone or Head Mounted Display (HMD). A further definition of AR and its context within AR technologies is given in Section 2.5.1. The success of Pokémon GO can be attributed with making AR a familiar term to the general public [11]. However, the uses of AR are incredibly, and even surprisingly, diverse. They range from gaming, entertainment and social media to education, healthcare, manufacturing, retail, art and beyond [12, 13]. The growth of AR is driven by the ubiquity of smartphones [8, 14] and it has the potential to transform the ways we interact, explore and learn [13]. Some of these benefits are filtering through to VI users, with novel uses of AR in image enhancement, navigation systems and providing audio labels to objects [7]. Although these developments are positive, the majority of AR content and in particular mobile AR content does not take VI access needs into consideration. This is interesting as current smartphones have a plethora of accessibility features built into their Operating System (OS) [15–17], but little thought or research has been put into the accessibility of mobile AR for VI users [18]. This work aims to highlight this problem and suggest some solutions to include VI users in mobile AR, by building on previous accessibility research.

1.2 Project Overview

1.2.1 Research Context

Personalisation in Accessibility As Chapter 2 will describe, accessibility research is increasingly moving from 'one size fits all' strategies to personalisable services. This involves adapting to individual preferences and needs, with users able to have control over how media is delivered to and made accessible for them. This is driven in part by a greater acknowledgement of the individuality of a user's experience of an impairment and their strategies and preferences in how they therefore approach technology. For example, two people who on paper experience the same visual impairment may have vastly different experiences and ways they approach and adapt technology to their needs [3]. Therefore, this diversity of preferences is linked not just to impairment but to aesthetic predilections [19]. This research trend is also driven by technological developments and availability (e.g. faster than ever internet speeds, delivered to mobile devices with more processing power than ever [20]) and by new tools for the creation and presentation of media (e.g. object-based audio [21, 22], non-visual interaction modalities [23–25]). This research trend also emphasises accessibility as part of creative processes, not just as an afterthought [26].

Adapting methods from TV and Film This project builds on successful accessibility research in film and TV. The Enhancing Audio Description (EAD) project [19, 27–30] explored soundtrack enhancement, spatial audio, and first-person narration as alternatives to traditional audio description. The Narrative Importance (NI) concept [31, 32] developed methods for end-users to personalise soundtrack complexity, improving both speech intelligibility and content comprehension. Both sets of research emphasise user individuality and provide personalisable solutions. However, their potential to work together and application to XR is unexplored. This project is therefore an exploratory first step, assessing whether these strategies are applicable in mobile AR and have potential for other XR applications.

Research Ethos This project is driven by the priorities expressed by VI users. It aims to understand how VI users' needs can be met in meaningful, enjoyable, and immersive ways. It focuses on accessibility as part of creative practice and emphasises the potential of audio to create virtual worlds VI users can autonomously explore rather than just having it described to them. This project also aims to create usable tools and repeatable workflows and involves content creators, VI or sighted, to achieve this. There is little other research into the accessibility of mobile AR which considers both VI and content creator opinions. It is also hoped that the successes and lessons learned from the pilot app developed will provide a starting point, both for research and for the development of more accessible mobile AR available to the general public.

1.2.2 Project Objectives

The strategies proposed in this project centre upon using the creative potential of audio to include and immerse users in meaningful and enjoyable ways. The central hypothesis of this project is:

It is possible to use creative and personalisable audio practices to develop mobile AR content which is accessible to VI users and includes them in enjoyable, immersive ways.

Objectives

The development of this hypothesis is further explored in Chapter 2 and from it, the following objectives of this work are established:

1. Establish the current state of accessibility research in AR, especially as applied to mobile-based AR.
2. Establish the accessibility of currently available AR apps.
3. Develop a mobile-based, narrative AR app as a testing platform.
4. Assess the efficacy of the new enhanced and personalisable audio accessibility strategies developed.

The first two objectives can be thought of as laying the foundations in understanding the needs of VI users and how they are currently being met, or not, in AR. They allow for the development of the research questions this work asks, which are then used to meet the third objective.

Research Plan

Given the project motivations and research context, several sub-objectives can be identified which support the investigations of this work:

1. Obtain first-hand insight into the accessibility of current AR apps from VI users themselves. Therefore, a user experience study of AR apps must be conducted.
2. Confirm that the accessibility level of the testing AR app is initially comparable to currently available AR apps, as previously established, through an initial user experience study.
3. Obtain initial feedback on VI and content creator perceptions of the accessibility strategies proposed, as well as giving opportunity for discussion of other improvements and ideas. Consult content creators on how the strategies could be implemented and what possible workflow tools they may necessitate. This can be part of the initial user experience study. The testing app will be used to facilitate these discussions.
4. Based on this feedback, determine the changes needed to create an updated version of the app with dedicated accessibility features. Update the app with these features.
5. Evaluate the updated app, gaining feedback on VI and content creator perceptions on the success of the accessibility features through a final user experience study.
6. Based on this feedback, establish whether the accessibility strategies proposed are feasible. If obstacles remain, use feedback from the final user experience study to present potential solutions and directions for further research.

1.3 Thesis Structure

This chapter provides an introduction to the body of work presented in this thesis. This includes the project motivations, objectives and original contributions made.

Chapter 2 establishes the research questions foundational to this work by exploring the needs of VI users, the characteristics the of AR medium and by reviewing developments in accessibility practice and research that are either already used in AR, or that provide transferable approaches. This includes the EAD and NI concepts.

Chapter 3 uses this review to establish the accessibility of current mobile AR applications. First, AR apps are tested against standard accessibility features and advanced audio features. Second, VI participants themselves validate these results by using and providing feedback on a subset of the apps in Study 1. Through this they provide insight on their experiences using AR apps, share challenges faced and begin to suggest solutions including initial thoughts on the feasibility of the access strategies proposed in this work.

Chapter 4 presents the bespoke app designed for this project. It describes Study Two, in which a two-scene version of the app with accessibility equivalent to those in Study 1 was deployed to VI users and XR content creators to gain a baseline assessment. Results of surveys and interviews show the issues faced, possible improvements, and the value of personalising the quantity and loudness of audio layers. A comparison of consumer and content creators' perceptions is also made.

Chapter 5 summarises how the app was updated to include dedicated accessibility features. These features are based on NI and EAD approaches and suggestions from Study 2.

Chapter 6 presents results of the final deployment of the app in Study 3. Here VI users and content creators provide their feedback on the updated app, allowing the efficacy of these accessibility strategies to be evaluated.

Chapter 7 summarises and evaluates the completed work. Based on this evaluation, practical suggestions are made for future use of the research strategies proposed and further research avenues are discussed.

1.4 Original Contributions

The following original contributions are made by this work:

- An overview of accessibility research for AR and research in other mediums which may provide transferable approaches.
- A comprehensive set of user studies demonstrating that the majority of mobile AR apps are inaccessible and that fundamental challenges must be addressed for AR to become accessible to VI users.
- The presentation of VI opinions on how mobile AR could be made accessible to them.
- An initial evaluation of whether EAD and NI based accessibility strategies work in an AR context.

Chapter 2

A Review of Accessibility in Augmented Reality

2.1 Introduction

The foundations for the rest of this work are laid in this chapter. It first highlights the prevalence of visual impairments and then seeks to clarify the needs of Visually Impaired (VI) users themselves. Relevant practices, concepts and research into accessibility are then highlighted, including the Enhancing Audio Description (EAD) and Narrative Importance (NI) concepts. It then turns to the medium of Augmented Reality (AR) itself, exploring the unique opportunities and challenges it presents. Developments in accessibility research and industrial practices that are directly related to AR are then reviewed. This allows the first objective of this work to be met, which is establishing the current state of accessibility research in AR, especially as applied to mobile-based AR. From this, we can begin to develop suggestions for how AR could be made accessible using creative and personalisable audio practices.

This chapter aims to provide a holistic approach, considering equally the needs of VI users, the challenges of the AR medium and the potential solutions presented by relevant research and practices.

2.2 The Need - Visual Impairment

2.2.1 Prevalence of Visual Impairment

The number of those living with a condition affecting their eyes is significant. Literature by the World Health Organisation (WHO) states that 'everyone, if they live long enough, will experience at least one eye condition in their lifetime' [2]. They do however acknowledge that trying to even put an estimate to how many of these cases translate to people living with visual impairment is nearly impossible, as people may have multiple conditions or may have conditions not yet reported. Their latest estimate in 2019 was 2.2 billion people, with up to 1 billion of these cases either being preventable or awaiting treatment. In the UK, work commissioned by the Royal National Institute of Blind People (RNIB) estimates that up to 2 million people are living with significant sight loss [1, 33]. This figure includes those who are not formally registered either because they are waiting to be, or whose level of sight loss prevents them from being allowed to register.

Levels of sight loss increase with age [2] and age-related macular degeneration is the leading cause of sight loss in the UK [1, 3]. Ageing populations, coupled with the increase of other confounding medical factors (e.g. diabetes) and lifestyle factors (e.g. sedentary lifestyles and unhealthy eating) mean that levels of sight loss are predicted to rise both in the UK and globally [1, 2]. It is estimated that in 2030 there will be 2.7 million people living with sight loss in the UK, and up to 5 million in 2050 [1, 33].

2.2.2 Characterisation of Visual Impairment

When talking about sight loss there is a wide range of terminology used, which can be reviewed in [34]. Not everyone who experiences a condition affecting their eyes will experience long-lasting sight loss [2]. In the UK, being certified as severely sight impaired (blind) or sight impaired (partially sighted) depends upon two criteria [3, 35]

1. Visual acuity – your central vision, the vision you use to see detail. This criterion is the most commonly used to assess vision worldwide [2]. It is assessed using a chart with letters or numbers that get smaller the further down they are. The results are given as a fraction, with the distance at which the chart is viewed as the numerator and the distance at which a healthy eye would be able to read that line as the denominator. 6/6 is referred to as normal visual acuity.[1] In general, a visual acuity of <3/60 is classed as severely sight impaired (blind) and a visual acuity of <6/60 is classed as sight impaired (partially sighted) [2, 35]. [2, 35]

2. Visual field – how much you can see around the edge of your vision while looking straight ahead.

Categorising Participants

In the rest of this work, the following categories are therefore used when discussing the level of sight-loss, particularly in categorising study participants:

- Partially sighted
- Blind with residual vision
- Blind
- Other - to allow participants freedom of expression if they feel their visual impairment doesn't fit into a previous category.

In all the studies presented in this work, participants self-report which category they belong to and are given the chance to provide further clarifying information. This follows a similar format to previous accessibility research [28, 29]. These categories provide the appropriate level of detail for work of this scope that is non-medical and exploratory.

[1] This is also referred to as 20/20 vision, which is the North American terminology [36].

Table 2.1: A summary of the RNIB's findings in the 2015 report 'My Voice' [3], which sought to understand the views and experiences of VI individuals. The headline findings from the report are shown and are categorised to aid readability.

Categorisation	Finding
Sight conditions and sight loss	There was huge variation in the experiences and views of blind and partially sighted people
	Residual sight was unstable, and people experience changes in sight over a long period of time
	Some people wait many years for their sight loss to be registered
	Sight loss was the biggest barrier that people face
Wellbeing	Feelings of wellbeing were lower among blind and partially sighted people when compared to the rest of the population
	The majority of people did not receive any emotional support in relation to their sight loss
	Blind and partially sighted people feel cut off from the people and things around them
Suppport	Only a small proportion received key elements of practical support
	Most blind and partially sighted people relied on some form of care, and often this was informal and unpaid
	Public awareness and attitudes towards sight loss was poor
Practical impacts	Travel and transport remains a major issue for blind and partially sighted people, especially obstacles on the pavement
	Many blind and partially sighted people have limited choice about they spend their free time
	Access to key information was an ongoing problem
	There was a generational divide in the use of technology
Financial impacts	The proportion of people in employment has decreased
	Blind and partially sighted people of working age were struggling financially

2.2.3 The Impact of Visual Impairment

It is important to highlight that behind all the numbers are real people with unique experiences. Work by Cattaneo stresses the importance of the age at which sight loss begins and individual differences such as the nature of sight loss, locomotion abilities and compensatory training [34]. Research by the RNIB also highlights "huge variation in the experiences and views" and the importance of these factors, as well as also acknowledging the effects of additional disabilities [3]. However, Cattaneo does also remind us that the "world of blind individuals is as rich as that of the sighted, or even more so" [34]. This motivates accessibility research to tap into this rich world by acknowledging and providing for their individuality. This is consistent with the increasingly accepted [37, 38] 'social model' of disability. This model places the onus on society to make itself accessible for all individuals [26]. In this model, it is society and surroundings that disable people [39], rather than their impairment, as the antithetical 'medical model' contends [40]. It is this shift in thinking which is a motivating factor for the move to personalisable accessibility services described in Section 2.4.

Despite this shift in thinking there are many ways in which society remains inaccessible to VI individuals. Work commissioned by the RNIB in 2015 attempted to understand the experiences of blind and partially sighted people across all areas of their life [3]. Its findings are summarised in Table 2.1. These findings show that those with sight loss remain a group that can be vulnerable, isolated and experience significant impacts on their quality of life [33].

The Impact of Technology

Digital technology has the potential to provide new opportunities to VI users [41] and when new technology is accessible, it can help them to become independent [3]. It is possible for new technologies to unintentionally provide accessibility improvements as they are developed [42]. However, it can't be assumed that because a technology is new, it is more accessible. Especially as those developing it may not aware of accessibility needs and solutions. For example, a study of 197 developers and digital technology students showed gaps in their accessibility knowledge despite standards, tool kits and efforts to promote accessibility [40]. The RNIB's findings on technology show that two-thirds of VI users were unable to effectively use new technology [3]. When technology isn't accessible, it can become another barrier to inclusion and participation in society.

Many solutions exist for making technology such as computers or mobile smartphones accessible to VI users. They fall within the following categories [15, 16, 37, 43]:

1. Presenting visual content in ways that allow the best use of any remaining vision. This can include:
    - Design considerations - such as the type of elements used, their layout, size, colour, spacing or the contrast between them (i.e. text and icons). This may be coded in for all users or ideally, users may have the ability to change any or all of these to best suit themselves.
    - Magnification.
2. Providing alternative (non-visual) methods for users to access the content. This can include:
    - Audio - such as audio description, screen readers, Text to Speech (TTS), voice recognition technologies.
    - Braille - such as braille displays, braille embossers.
    - Other tactile methods such as haptics or tactile overlays.

Whilst many of these solutions are already used in the Operating System (OS) of computers and smartphones, this does not mean that all content on these devices is compatible with them. Developers have to ensure their work is compatible [17, 44]. Chapter 3 highlights this problem for mobile AR applications (apps). Some of these solutions are used or extended in the accessibility research that this project builds on, as explored in Section 2.4. Many of them also form the basis for suggestions made by industrial standards and research into creating accessible AR, as explored in Section 2.5.

2.3 Existing Audio Technologies

This section explores existing audio technologies which have the potential to make content more accessible to VI individuals.

2.3.1 Sonification

Sonification is a technique where audio is used to present information to the user, which may be 'information about data under analysis, or about the interaction itself, which is useful for refining the activity' [45]. There are several motivations for this including making the most of using all our senses and harnessing the ability of our listening system to understand complex information with multiple layers of understanding [46].[2].

Of particular interest are sonification techniques that use sounds to signal an event [45]. They can be used to alert the user that something has happened or needs to happen and provide information about what it is. There are several approaches. Auditory icons [47] attempt to create an audio-based 'caricature' of what is being represented, using sounds that the user should be able to associate with the given meaning. For example, in OSX the action of deleting a file is accompanied by the sound of paper being crumpled up [46]. Earcons [48] are more abstract, often using musical tones to create a representational system that the user can learn. This means that it does not assume that the user will already have certain associations with sounds and meaning like auditory icons, so therefore training is required to understand earcons which does give them the potential to communicate more complex information [46]. One common example of earcons is the use of ringtones in mobile phones, where the user can set different ringtones for different callers, thus recognising the caller without having to look at their device [46]. The third method, spearcons [49] speeds up descriptive speech, even to the point of it not being recognisable. As no associations are assumed and no training is required, there is some evidence that users can understand spearcons more quickly and accurately [49]. A good summary of studies exploring the relative merits of each method can be found in [50].

There has been exploration of these methods as an accessibility tool for VI users. For example, Mascetti et al [51] showed that after a short training, various sonification modes can aid non-visual exploration of shapes, Drossos et al [52] used both auditory icons and earcons to create an well-received accessible game for blind children and Merabet et al [53] showed that using iconic sounds to help train users about the layout of a previously unfamiliar building was both successful and engaging. Although their precise implementation of sonification varies, the studies all give credence to the potential utility of sonficiation for VI users. As we shall see, the concept of sonification underpins a number of the research solutions projects which use AR as an accessibility tool (Section 2.6.4). For example, sonification has been shown to be useful in helping VI users navigate zebra crossings [54]. Its ability to provide information on interactions proves particularly useful in the AR medium which centres upon user interaction with real-time input.

[2] A good starting point for exploration into sonification concepts and research is 'The Sonification Handbook' [46]

2.3.2 Spatial Audio

At its most basic, spatial audio is used to describe audio that tries to create a realistic audio experience similar to the three dimensional way that we hear the world [55]. It tries to give us information about the location of sounds and the acoustic properties of the environmental context [10, 55]. There are many methods for its creation and reproduction.[3] With developments such as spatial audio being brought into television setups [56] and added to streaming services, such as Apple bringing spatial audio to their music library [57], spatial audio technology is becoming increasingly available to the end-user in their daily life.

A reason for the increased use of spatial audio is its ability to create experiences that are more realistic and immersive. It is especially used with Extended Reality (XR) technologies so that the location of an object is matched in the visual and auditory domains as the user explores [10]. There is an increasing realisation that spatial audio can be just as immersive as 3D visual environments [58] and is important for a sense of presence [59]. Indeed, some evidence shows that it can have a bigger impact than the visuals in an experience on enjoyment and feeling a sense of presence [60].

In this work, the primary spatial audio technology discussed and used is binaural audio, as it only requires headphones. This means it can be easily accessed using mobile devices [61]. Binaural audio uses complex functions which are defined for each ear, taking into account the effect of the various cues that we use in decoding the sound information in real life [62]. These include the difference in the loudness that a particular source presents to each ear (Interaural Level Difference, ILD), the difference of arrival time for sound from a particular source to each ear (Interaural Time Difference, ITD), and the effect (or filtering) of the human head, shoulders and outer ear (pinna) upon the sound [63]. As the sound at each ear will be affected separately by these cues, our brains use the differences between the sound each ear receives to deduce the location of the sound source. The Head Related Transfer Function (HRTF) encapsulates these cues [63, 64]. When imposed on clean audio signals it deceives the brain into locating sounds at virtual locations. The realism of the auditory experience can also be increased by using head tracking technology, which allows the binaural renderer to move between different HRTFs as the head moves [64].

As well as helping to make experiences more believable, enjoyable and immersive, spatial audio also has the potential to present more information to the user. It is commonly believed that VI individuals have a more accurate sense of hearing than sighted individuals. By reviewing numerous pieces of research which investigate this claim, Katz and Picinali conclude that where VI individuals excel is in the way they process, or exploit auditory information, especially when using that information to navigate [65]. They then go on to present several studies in which the potential of spatial audio to help VI individuals explore virtual environments is explored. As spatial audio becomes more widespread there is excitement about the potential uses of spatial audio as an accessibility tool [66]. It can provide more novel ways in which location information is presented, such as helping guide VI users in real life [66] or virtual [65] environments. Drossos et al [52] assert that it could be particularly powerful at presenting more information concurrently, especially when combined with sonification techniques. Its potential as an accessibility tool is one of the suggestions made by the EAD project (Section 2.4.2) and it has already been used in several AR research projects (Section 2.6.4) and apps that have been designed with VI accessibility in mind (Section 2.6.5).

[3]The interested can find an introduction to the development of these technologies and concepts in Francis Rumsey's book 'Spatial Audio' [55].

2.3.3 Object-based Audio

Most mediums, like TV, traditionally deliver a final product to the end-user, in which all constituent parts (such as individual video or audio tracks) have already been combined in ways decided by the creators. In an object-based paradigm, these constituent parts are delivered to the end-users as independent objects alongside metadata describing how each object should be used [67, 68]. Based on this information, the objects are then combined and rendered on the users device. To this end, there are several different formats in use including ADM, Dolby Atmos, MPEG-H and DTS:X [32, 69]. Figure 2.1 gives a visual representation of the differences between traditional and object-based paradigms.

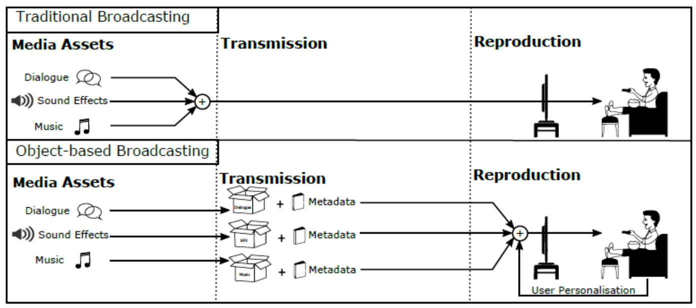


Figure 2.1: Visualisation of object-based broadcast in comparison to channel-based, noting the potential end-user input which can influence reproduction of the media assets in the home. Reproduced from Lauren Ward [32].

The shift to this paradigm has two main advantages:

Flexibility - with traditional workflows, the creator may have to develop several different representations depending on the loudspeaker configuration of the device(s) being targeted [68]. This can mean several mixes are produced, each for a different loudspeaker configuration [70]. However, object-based paradigms are 'format agnostic' [70] meaning that production can be done once for a variety of devices or listening formats [67, 68]. This is of particular benefit as spatial audio (see Section 2.3.2) becomes more widespread. The accompanying metadata can encode a precise 3D coordinate location for each audio object [70], which can then be rendered for headphone based (i.e. binaural) or speaker based (5.1, 7.1 and beyond) spatial formats [22, 71]. Or if the user does not have such a setup available, an optimised 2D mix can still be defined. Object-based audio therefore makes it easier to achieve universal delivery, where reproduction is fully optimised for all possible playback devices [22].

End user interaction - as the end-user is not being given a finished product, object-based audio opens up the possibility of user interaction, with them being able to adapt the mix to best suit their needs and tastes [22]. This has many potential applications, for example, dialogue enhancement where viewers are able to control the relative levels of commentary and court sound in a tennis broadcast [67], or greater exploration such as being able to freely navigate a football game [70]. End users being able to personalise audio mixes also holds great possibilities for accessibility. Fraunhofer, who helped to develop MPEG-H audio have provided a summary of the potential solutions object-based audio can present for different end-user needs [22]. These range from controlling relative levels to altering the complexity of language used. Ward and Shirley [21] presented the most comprehensive systematic review to date of uses of audio personalisation using object-based audio to improve broadcast audio for the hearing impaired.

However, at this point there is little to no discussion of how this same technology may be used to benefit VI users. What discussion there is mostly centres on simple level or positioning adjustments for Audio Description (AD) [22]. There is much discussion around how AD itself could be improved in more advanced ways, especially in the light of both object-based and spatial audio (see Sections 2.4.1, 2.4.2). As we have seen, VI needs are widely varied and individual. Further consideration may allow object-based audio to become a key technology to offer personalisable accessibility for VI end users.

2.4 Established Accessibility Solutions and Research

This section explores existing solutions and research for making content accessible to VI individuals. Many of these solutions are applied to established, linear media such as film and Television (TV). These include the two research concepts that this work builds on, the EAD project (Section 2.4.2) and the NI concept (Section 2.4.4). This section limits itself to solutions that could be relevant to the AR medium, meaning most are applied within the world of media.[4]

2.4.1 Audio Description

Audio Description (AD) provides a way in which VI individuals can hear content that they cannot see. It is used all around the world [72], and provides access to a wide range of mediums including film and TV, theatre, museum exhibitions and more [73]. The idea has been around since the 1960's when Chet Avery was inspired by the closed captioning of films for the hearing impaired, although the concept and workings of it weren't more fully developed until the 1970's by Gregory Frazier [73].

Characterisation of AD

In most cases, AD takes the form of additional descriptive narration which is slotted into pauses in the existing dialogue [74, 75]. The content of the additional narration and the precise way in which it is interwoven into the content can vary between different mediums and standards. Despite fragmented world-wide development, AD is crucial as the main method VI consumers can be included in many mediums [72].

In terms of the content of AD, most guidelines or standards break their advice down into the components of describing when, where, who and what [76]. The OFCOM in the UK puts it as "characters, locations, time and circumstances, any sounds that are not readily identifiable, on-screen action, and on-screen information" and is also strict about describers not adding any information, but limiting themselves to what is on the screen only [75]. In terms of when AD can be used most guidance suggests that AD shouldn't encroach upon the rest of the soundtrack, VI users can hear and interpret it for themselves [76]. There is also guidance around the delivery, for example, OFCOM suggests it should be present tense and uncoloured by the describer, through it being unobtrusive and impersonal [75]. Although, some variety and emotional range is allowed, as long as it suits the content [75]. This is similar to other worldwide standards [76].

[4]For readers wishing to gain more information about other solutions used throughout society to help VI individuals in their day to day lives, the RNIB website is a good place to start: `https://www.rnib.org.uk/`

Prevalence of AD

The prevalence of AD depends upon the medium. In the UK, the Communications Act of 2003 means that all television broadcasters have to provide AD for at least 10% of their content [77]. Although many of the larger broadcasters including the BBC, Channel 4, ITV and Sky have committed to increasing this to at least 20% [72]. In cinema, most of the big Hollywood distributors provide AD on the majority of their UK releases, with about 40% of UK cinemas equipped to deliver it to the viewer via separate headsets [72]. Accessibility is more limited in increasingly popular streaming services [27]. However, it is improving as the 'Digital Economy Act' [78] of 2017 allows OFCOM greater regulation of streaming services. The BBC has committed to ensuring that all content originally broadcast with access services will have the same services available on BBC iPlayer [79] and Netflix began adding AD to their content in 2015 [27].

Criticisms of AD and Research Suggestions

It is argued that AD is able to conjure up vivid images when well-chosen words are used [73], which may allow listeners to construct a version of the story similar to that experienced by a sighted person [80]. However, criticism of its current format is increasing within the research community including that it is often an afterthought, that it reduces the users' sense of presence and that it hasn't developed with technology.

Universal design theory involves ensuring all potential uses and users (including those with impairments) are considered, from the very beginning of the product or service being created [81]. However researchers such as Fryer [82–88] and Romero-Fresco [26, 89, 90] have shown that AD is very much an afterthought, which is added on after the fact without collaboration from the original artistic team. This means it is possible to suggest that AD actually violates universal design principles, as it is not considered from the beginning of creation, but is done separately, by separate people [81]. Furthermore, AD can actually exclude VI users, as it may have some effect on their social inclusion, either due to exclusivity as they listen to it on a separate headset [86], or due to feeling sighted family or friends don't want to have to put up with it, especially as it may overlap with music or sound effects [28]. Furthermore, VI users are only given the option of this one solution with no recognition of their individual preferences and needs [30]. Fryer summarises it in this way "Traditional AD is typified by five characteristics: it is exclusive; neutral; non-auteur; third-party and post hoc" [87]. However, the very fact that AD is a vital service [74], which does make all the difference in VI users being able to access and understand content [72] means that solutions must and are being found. Cavallo and Fryer have suggested 'Integrated AD', which involves the creative team and is part of the creative process [91]. This is similar to other propositions [26, 27]. It also has an impact on the format of AD, as we shall see below. It is already believed that the questions such research has raised has had an impact and the role of the audio describer is indeed changing [87], providing more authentic experiences [86].

A second criticism is that AD is not as engaging to the listener as it has the potential to be. We have seen how the delivery of AD is tightly regulated, to prevent it colouring the listener's experience. This means it has been shown to be informative, but not entertaining [81] and that its neutrality can be dull and lack emotional connection [91]. There has been a particular focus in research on how this affects the sense of presence that the listener has. Defining precisely what presence is has proved difficult, although Lombard and Ditton's 1997 work perhaps provides the best approach [59]. They break it down into concepts of social richness, realism, transportation, immersion, users acting like they could interact with the medium or treating the medium as a social entity. These threads are then pulled together into presence being “the perceptual illusion of nonmediation”, with several factors affecting it which range from attributes of the media, its impact on different senses, its interactivity and more cognitive factors such as its social realism or the nature of the task/activity. With AD, work has been done to challenge the neutrality of the describer’s voice [92] and that human voices can elicit emotion in a way TTS is unable to [85]. Cinematic AD is a variant of Integrated AD which includes information on camera shots and editing, and has been proved to increase the sense of presence for VI users [84, 89, 90]. This has even been applied in other mediums, such as Virtual Reality (VR) [82] and theatre [93].

A third criticism of AD is that it has changed very little since its inception, despite technological advances. These advances can be as simple as fixing issues AD users have with the existing format, such as complaints about the intelligibility of the AD [80]. Or they could explore a whole new, creative approach to AD. This is one of the central tenets of the EAD project [19, 27–30], whose suggestions we shall explore further in Section 2.4.2. Object-based audio technology also opens up the possibility of end-users personalising AD, such as altering its relative level or spatial positioning within a mix [22]. Another approach is to make AD multi-modal, such as in [94] where haptics is suggested, so that information can still be given even during important parts of the soundtrack, and to help prevent an overload of auditory information. This links to the second criticism, as making use of technological advancements in both audio and other modalities may provide more engaging AD.

2.4.2 Enhanced Audio Description

Central to the approach of the EAD project is its acknowledgement of the variety of experiences and needs of VI users [19, 27–30]. It aims to create new user-centred methods for AD, not replacing traditional AD, but giving VI users a greater choice of the accessibility tools available to them [27]. These methods are characterised by their attempt to make up for how “AD guidelines have failed to acknowledge how advancements in digital audio could be game-changers” [27]. They assert that using these advancements provides an experience that is both informative and entertaining [29]. It also sought to demonstrate the importance of involving content creators, and the creative possibilities of doing so when technological advancements are utilised [19].

In [28], the EAD project first sought to obtain background information from VI individuals on their use of AD and their opinions on its advantages and disadvantages. As well as confirming how essential such services are, the results provide further evidence for the need to acknowledge the diversity of VI individuals. The EAD methods were then developed in [29] by consulting focus groups to obtain feedback on each method individually, with additional listening tests used to investigate the effectiveness of the spatial audio. The methods are as follows:

- **Soundtrack enhancement.** By using well-designed sound effects audiences can be shown what is happening through what they hear, meaning there is less need for verbal descriptions. Sound effects may have several purposes, such as conveying actions, establishing shots, conveying abstract scenes, establishing the presence of characters and creating auditory logos to help listeners recognise an important element or event. [29, 30]. There are some similarities to sonification approaches here in the use of audio as a representational and informational tool (Section 2.3.1).

- **Spatial audio.** By using binaural audio, listeners can determine the location of characters and objects, as well as a more detailed understanding of the layout and characteristics of the space that action is taking part in [30].

- **First-person narration.** The first two methods seek to minimise the need for audio description so that VI individuals can experience the media in the same way as sighted individuals, and to minimise masking of original soundtracks. However, further description and clarification can still be needed. Chions concept of i-voice [95], where a characters internal voice is used to give descriptions is proposed [29]. The intention of this is that it allows the description to feel more like it is an intentional part of the piece, and encourages the listener to identify with the describer. Particularly key to this method is the idea that the original scriptwriter should be involved, to ensure cohesiveness.

After proposing these methods, [29] then assesses their efficacy, finding them a fully feasible alternative to AD and capable of providing information, enjoyment and accessibility. This was further confirmed in [30], which also demonstrates sighted individuals accepting and even enjoying material that incorporates EAD methods. It therefore argues that EAD strategies have the potential to create experiences that both sighted and VI individuals can enjoy together, making it an example of successful universal design. The project also considered the need to educate future content creators in [19]. By getting students to implement EAD methods in a short film, they were able to gain insight into how successfully people outside the research team could implement their methods. Through some of the shortcomings in the student implementation, the project was able to determine methods for educating content creators in the use of EAD methods, such as inviting them to consider how accessibility could be integrated into their creative processes rather than regarding it as something outside their remit that is added on afterwards.

The EAD project has been successful at showing it is possible to provide an alternative to traditional AD which uses technological advancements in exciting, creative ways. The progression of EAD research and its integration into production workflows informs this research as the use of sound effects and spatial audio is not uncommon in XR practices. However, thinking of them as accessibility tools is a fairly new concept.

2.4.3 Audio Games

An inspiration for the EAD project is the field of Audio Games [27]. Although games can utilise well designed audio content, most games still rely upon visuals to play the crucial role in conveying necessary information [58]. Audio games reverse these roles, using 'auditory displays' to convey all the necessary information, with visuals that are very limited and may not even be necessary to understanding and enjoying the game [52, 58]. They were originally developed by VI programmers or amateurs as a way to create blind-accessible games [96] but have garnered interest even from sighted users [97]. In fact, there is a belief that they could provide interesting experiences for users with any level of sight [58, 96].

Audio games often use narration, sound effects, atmospheric sounds, music and even clever use of sound processing, such as altering reverberations or filtering the audio [27]. These are often used creatively to build a believable world and to present relevant information or guide users through the process of accessing information or undertaking interactions. This is similar to some of the techniques used in sonification. Binaural audio is also used extensively, as it can effectively increase the amount of information the audio is providing by giving listeners a sense of the location of objects and the space they are in [27, 52].

In a study with blind children, Drossos et al [52] concluded that for an audio game to be accessible, they must provide constant audio feedback for any events, that thought must be put into how distinguishable these sounds are especially if multiple layers of information are being communicated and that binaural audio is particularly key for navigation. Friberg and Gärdenfors [58] demonstrated the potential of audio games to communicate complex, multi-layered information in an educational context.[5]

Audio games further underline the potential of creative audio methods like EAD to provide accessibility in XR solutions, as there are many overlaps between XR and gaming. For example, the common need for users to react and interact with information in real-time. There are already some instances of Audio only AR (AAR) and even AR Audio games (Section 2.5.6).

2.4.4 Narrative Importance

The concept of Narrative Importance (NI) was first put forward by Lauren Ward in her PhD thesis [32]. It is centred around using object-based audio to make broadcast services more accessible for hard of hearing individuals. Although there has been work into using object-based audio in this way [21], this has not always been holistic in its approach, as content creators are rarely involved and attempts to meet hearing impaired needs are either oversimplified to just making speech intelligible [31] or result in interfaces that are overly complex and thus difficult to use [98].

Standard approaches for making content accessible to the hard of hearing often rely on some form of adjusting the levels or clarity of dialogue relative to other sounds [21]. The NI approach allows for much more subtlety and for a closer representation of how a content creator may intend the sound design to convey the story to the end-user [31]. It can thus be thought of as prioritising comprehension of the story over simply ensuring intelligibility of dialogue. This approach was developed by considering the opinions and needs of hard of hearing individuals and the opinions and workflow needs of content creators. Its success was evaluated through studies with these same two groups.[6] Ward started from basic storytelling principles, rethinking how content creators go about designing sound. Rather than mixing sounds by commonly used categories such as 'sound effects', 'dialogue' or 'ambience', Ward's approach considers the importance of every sound object to the story being told. Each sound object is given a piece of metadata which assigns it a NI categorisation. The end-user has one simple control, which will alter the relative level of all sound objects depending on their category [31]. These categories were kept to 4 simple to understand levels, with a dB multiplier attached to each as follows: essential (+3)/high (0)/medium (-12) and low (-48) importance [31]. These multipliers, M, are defined by the users choice, N, to be between 0 (standard mix) to 1 (fully NI defined mix), thus adjusting the original level ($Level_{orig}$) of the sound object to set its adjusted output level ($Level_{adj}$)[32]:

$$Level_{adj}(dB) = Level_{orig}(dB) + [M_1 * N](dB). \tag{2.1}$$

Although it works in very different ways, NI has many conceptual similarities to EAD and other methods explored thus far. It acknowledges the capacities of audio to creatively

[5] For the interested, a good place to find out more about audio games, and even a comprehensive list of games to give a go, `https://audiogames.net/` is an excellent resource and Rovithis [96] can provide definitions of the sub-categories that exist within the field.

[6] Further reading of [32] will provide the interested more depth on the development and testing of the NI concept.

story-tell, through its consideration of the potential importance of each sound object. The solutions it proposes acknowledge the individuality of access needs by giving users meaningful choices over how they experience content. It also seeks to involve content creators in making accessibility decisions. One of the key takeaways from NI is that the approach both manages to give the end-user a significant amount of power and choice over how they experience content, whilst keeping the manner in which they do so simple and intuitive.

Although a NI framework has thus far only been developed and tested in the context of accessibility for hard of hearing individuals, its underlying concepts are highly transferable. The needs of VI individuals can be very different to the hard of hearing. However, we have seen the importance of audio-based solutions to creating VI accessible content through the successes of EAD and other approaches. Therefore allowing VI users control of how precisely these solutions work in a manner similar to NI could have great potential in providing personalisable accessibility solutions.

2.5 The Medium - Augmented Reality

In the following sections, accessibility solutions for AR are explored, beginning with a description of the medium itself.

2.5.1 Context within XR technologies

Augmented Reality (AR) is part of a larger suite of technologies called Extended Reality (XR), which include Mixed Reality (MR) and Virtual Reality (VR). These technologies hold in common in their ability to alter the user's reality, by adding to it or taking it over completely, allowing them to interact with it in new ways [10, 99]. Some believe that these technologies may increasingly converge, especially as more powerful computing allows them to be delivered by the same device [100]. However, at present they have several distinctions:

- **AR** adds layers of virtual elements on top of the users real world [101] to enhance it [10]. This can be information, digital objects or media [102].
- **VR** replaces the users reality with a new, digital environment through the use of surrounding displays, or more commonly a headset [99, 102]. The user is fully immersed in this environment, losing their sense of the real world [10].
- **MR** sits between AR and VR. In it, the real and virtual environments coexist and can interact with each other [10].

There is often an implicit assumption that XR technologies are visual, or at least vision-centric with "non-visual interaction modalities underused and undervalued" [23]. Although this may often be the case, it is possible for the virtual elements that are augmenting or replacing the users environment to be audio-based, or even use other senses.

Many of the ideas behind XR can be traced back to as early as 1838, but it is only as recently as the 2010's that these technologies started to grow and become of interest [10].[7] Many commentators believe that XR is at a tipping point, now moving past the stage where 'early adopters' are the only people on board, to being adopted more rapidly in the mainstream [101–103]. The products and experiences being created have moved from

[7] As well as providing standards for the development of XR, the Cyber-XR coalition's documentation also give an interesting historical timeline of XR technologies [10].

proof of concept-type solutions to something that is effective and scalable for businesses and consumer use [14]. This is due to the dropping costs of both hardware and software and the increasing impact of XR in non-entertainment industries [8].

2.5.2 The delivery of AR

Given that AR augments or enhances the real world, the devices it is delivered on need a range of sensors to capture the real-world environment. These can include a camera, depth sensor, accelerometer, gyroscope, GPS for location-based content and even high-speed internet for content that updates in real-time [12]. It also needs to have a method of displaying the augmented layers alongside the real-world input. When sensors and displays are combined, users can explore virtual elements from all angles by reorienting their device or even walking around them and may be able to interact with them using gestures or movement [99]. At present, there are two common methods for delivering AR:

- **Head-Mounted Displays (HMD)** use small displays or projection technologies placed directly in front of the users eyes [10]. This may take the form of a headset (such as Microsoft HoloLens [104] or Magic Leap [105]) or try to appear more like traditional glasses (such as Google Glass [106]). As technologies improve and become more compact, it will become possible for headsets to be as compact as traditional glasses, with some predicting that the same glasses could be used for AR and VR [100].

- **Mobile Smartphones** have an increasing number of the sensors required, and an interactive display already built into them [23]. Given that they are already widely used, they provide an obvious way for AR to be delivered to the general public and have driven much of the growth of AR to date [8, 14]. They have already altered the face of the gaming market, becoming the main gaming device [11]. Although some expect them to eventually be replaced by HMD's [8, 9, 100], they remain the easiest and most widespread methods that most people can access AR. They are also increasingly used by VI individuals [3, 107]. This is why this work focuses on mobile AR.

- **Other methods**, such as Heads Up Displays (HUD) for vehicles are beginning to become available, with manufacturers such as Mercedez-Benz [108] and Panasonic [109] leading the way.

Most development guides break AR experiences down into three sizes of 'playspace'; tabletop, room or world-size [99, 110]. This can influence how exactly the user can be expected to explore and interact with the virtual elements. For example, a user is usually expected to move around and 'look' at virtual elements much more in a world sized experience, whereas in a tabletop experience, they may expect to remain stationary and explore virtual elements more through manipulating them.

These considerations make AR an especially challenging medium for accessibility. The user needs to understand both what is being presented to them (i.e. information, objects, a narrative) and how to simultaneously navigate and interact with both the real world and the virtual world [18].

2.5.3 Uses of AR

The potential uses of AR are only just beginning to be explored. It has potential application in art, automotives, construction, education, entertainment, gaming, healthcare, logistics, manufacturing, marketing, museums, retail, tourism and training [8, 12, 102]. Released in 2016, Pokémon GO brought AR into the mainstream and achieved huge commercial success [11]. In it, users search the real world for animated Pokémon characters. It is perhaps the most widely known example of AR and is often used when explaining what AR technology is. Today, AR has become widely used in mobile smartphone gaming, opening up whole new gaming experiences [11]. However, it is in applications outside of entertainment where AR is particularly expected to rise [8]. Central to many of its other uses are information sharing with the ability to simulate scenarios effectively in real-time. Manufacturers such as Lockheed Martin and Mercedez-Benz have used AR to streamline technician training and assembly processes [104]. Logistics companies are using AR to streamline stock-picking processes [106]. AR has allowed health care professionals to become more accurate at intravenous insertions by showing a map of veins on the surface of the patients skin [101]. NASA has even simulated Mars rover repairs using AR [8]. These examples are by no means exhaustive. As the contextual base for AR is the real world, its potential uses are limited only really by the imagination of content creators.[8]

2.5.4 The Future of AR

Reports by the XR Association [8] and PWC [102] provide interesting modelling and expert opinions on the future of XR. PWC represents the uptake of XR in an 's-shaped' adoption curve, with the world at the crucial point in its adoption right now [102]. They estimate that VR and AR will give global GDP a $1.5 trillion boost by 2030. However, their estimates show that AR is already providing the bigger global GDP boost ($33 vs $13.5 billion in 2019) and will continue to do so until 2030 ($1092.4 vs $450.5 billion). The AR market is expected to grow to $31 billion by 2023 [111] and in terms of market revenue, 76% of industry leaders think AR will overtake VR, with most (49%) thinking this will be in 3-5 years [8]. Many of these leaders (52%) thought that ease of access to AR will be the factor most responsible for it overtaking VR [8]. As already stated, much of this growth to date is due to the pre-existing ubiquity of smartphones [8, 14]. A central factor behind these positive outlooks is a convergence of advances, such as the advent of 5G [10], technologies such as eye-tracking and gesture recognition [14] and more effective powering systems [100].

Given the incredible growth of the AR industry and the increasing exploration of its uses, we can only assume that this and the other XR technologies are going to become an increasing part of all areas of our lives.

2.5.5 Factors Affecting user experience in XR

Several factors affect user experience in XR media. We have already explored presence in Section 2.4.1. The other that is often discussed is immersion. Similarly to the definition of presence, the exact definition of immersion is debated. However, a model of immersion directly applicable to XR was recently developed by Hyunkook Lee [112]. This is shown in Figure 2.2. In it, the attributes physical presence, social/self-presence and involvement work together to make an experience immersive. The ways that a user is engaged in these overlaps and may be based on sensory, narrative or task/motor input. For these attributes

[8]The two-part book 'Augmented Reality Games' [12, 13] seeks to understand why AR is so powerful, starting from the "phenomenon of the Pokémon GO game" and then exploring its wider use. It does so through the lens of gamification, providing interesting perspectives on both the current and future potentials of AR.

to be present or engaging, the immersive system must be plausible, interesting and interactive, with each of these properties influenced by the factors subjective to the user; internal reference, personal preference and skills/knowledge. The way this model breaks down the immersivity of an experience down provides an extremely useful framework for evaluating a users engagement with XR.

As dicussed, two of the main challenges for VI users in AR is their ability to understand what is being presented and how to navigate and interact with the AR. When matched up with this conceptual model, this shows that we are most interested in their narrative and task/motor engagement. Therefore, the interestingness and interactivity of AR systems developed will be of particular interest in assessing the effectiveness of this work. Interestingness is mostly assessed in this work by the extent to which VI individuals enjoyed using AR and understood the information or narratives being presented. Interactivity is mostly assessed by the extent to which they succeed in using AR and the difficulties they face in doing so.

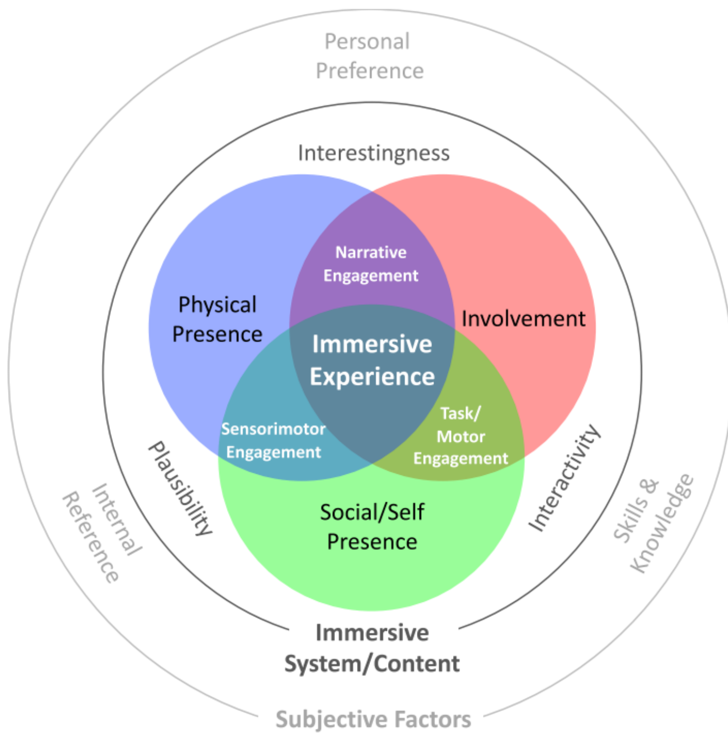


Figure 2.2: The conceptual model of immersive experience (IE) in XR. Reproduced from Hyunkook Lee [112].

2.5.6 Audio AR

Although AR is not limited to enhancing the visual world of the user, visuals have been the focus of much of its development and uses to date. However, work has been done into Audio AR (AAR), which specifically focuses on enhancing what people can hear.

Creating Enhanced Auditory Worlds

The first use of AAR is more close to ‘mainstream’ AR, with AAR being used to add to the audio ‘world’ around the user. To do this, these sounds needs to be so believable that they are indistinguishable from reality [113]. This is often done using spatial audio [23]. To hear the real world, users may use open headphones, bone conduction devices or head-mounted devices with built in speakers, such as the Bose Frames [114]. When closed headphones are used hear-though (or pass-through) technologies are used which capture the real sound environment and play it back with the virtual sounds added [115]. This has been used in a variety of settings. Ren et al [23] investigated AAR (and haptic methods) to provide information and guidance in tourism situations where visual displays are not always appropriate as the user may be walking, wanting to look around, or screens may have poor visibility due to direct sunlight. Nagele et al [24] used Bose Sunglasses to facilitate participatory theatre performances, with participants able to navigate around and interact with both the experience and each other. It highlights the high levels of presence they felt in the experience as they were still able to see, hear and react to each other, whilst being fully immersed in the virtual world at the same time.[9] The BBC further pushed the boundaries of interactive AAR experiences with ‘Looking for Nigel’ which adds GPS, voice recognition and physical gestures to take you on an adventure where you interact with alternate realities in your local park [116]. Rovithis et al [20] have also investigated the combination of AAR with Audio Games (see Section 2.4.3), which is a relatively new area. By creating and testing the game “Audio Legends”, they demonstrated how easily a user can become familiar with AAR games and they can be challenging, enjoyable and immersive.

Enhancing the Users Hearing

The second use of AAR is to enhance the hearing of the listener. The real world audio objects can be subjected to noise suppression, signal enhancement or even moved to different locations, with this processing adapting to context and user preference [117]. As in the previous use, additional virtual sounds may also be used to provide extra information or highlight real world stimuli. This could have particular impact for the hard of hearing [117]. Hearing aids which are adaptable to user preference and environment are already available from Bose [118]. Facebook’s Project ARIA [113] is looking to take this even further, giving the user ‘perceptual superpowers’ [119].

Given its reliance on audio to present information, AAR has capacity to be inherently accessible to VI individuals. However, only making AAR content for VI individuals is not the answer in itself, as this would mean creating separate content for VI individuals, which does not provide an equal experience [18]. Instead, AAR has the potential to provide answers to how ‘mainstream’ AR could become more accessible, if AR audio were to be enhanced alongside visuals. It has already been shown that enhancing visual AR with audio using such methods improves the users localisation accuracy and understanding of the virtual space [23]. In many ways, these methods are very similar to those suggested by the EAD project and those used in Audio Games. They all hold use of spatial audio, well-designed sound effects (even sonification) and in some cases, use of spoken guides.

[9] Nagele et al [24] also provides a review of other uses of AAR in interactive theatre.

2.6 Accessibility Solutions for AR

In this section, current accessibility standards, methods and research for AR are reviewed. Doing so allows an assessment of the current use and future applicability of the accessibility methods suggested for traditional media in previous sections. It is important to distinguish between two categories of work discussed.

- Assistive AR - this is work that uses AR technology or content to make the world more accessible by addressing a specific access need [18]. It is often bespoke, especially when used in a research context. Therefore, it is less likely to be easily available or adaptable to the general public but give us a good idea of 'state of the art' uses of AR for VI individuals. The mobile AR apps explored in this section, as well as many of the research projects, fall into this category.

- The accessibility of AR - this is work which seeks to establish how accessible AR technology or content is and/or provide suggestions about how it could be made accessible. Industrial standards and research exploring how accessible AR is to VI individuals fall into this category.

At this point, the question is how could AR be made accessible to VI individuals, rather than whether or not it is currently accessible. Although a brief overview of the accessibility of XR is provided, a more specific review of the accessibility of mobile AR content is given in Chapter 3.

2.6.1 XR and Accessibility

The Potential of XR to Include

The same is true of XR technologies as other new technologies (see Section 2.2.3). They have huge potential to not only be themselves accessible but to perhaps provide new ways of including and giving opportunities to visually impaired users [120]. The Information Technology and Innovation Foundation (ITIF) has highlighted this in a 3-part series of reports [6, 121]. This begins by highlighting the possibilities AR and VR have for creating new assistive technologies and of being an agent for social inclusion by making services, experiences and work more accessible through providing new ways for VI individuals to participate [6]. They highlight that it is the location-independent and multi-sensory nature of XR which gives it such potential. Froissard et al [122] highlight AR's ability to help with surface edge detection as the most important contribution it can give VI individuals. As we shall see in Sections 2.6.4 - 2.6.5, helping users to navigate is the focus of much of the currently existing AR for VI accessibility research and content.

Conversely, XR also has the potential to exclude VI users. The second part of the ITIF series focuses on how XR could pose health and safety risks, bias and discrimination risks and financial, physical, technical, and societal barriers [4]. They highlight several areas needing consideration for VI users to be included in XR technologies including user privacy, bystander privacy, in-world safety, physical safety and comfort, accessible design, technical requirements, non-technical barriers, representation and bias and discrimination. For those interested in designing accessible AR, this provides a useful starting point.

Overall Assessment of the state of XR Accessibility

Before we move on to *how* AR may be made accessible to VI individuals, it is worth assessing how accessible XR is perceived to be overall. There is concern in both the research and industrial communities that not enough has been done to date to make XR accessible. XR has grown so rapidly that accessibility solutions are still limited [5]. 'Temporary impairments' can and have occurred when such new technologies are created, because the needs of impaired individuals have not been considered in hardware design [40]. Many disability advocates are concerned that previous trends of not considering accessibility needs until after a technology has been developed are simply continuing into XR technologies [4].

However, there is hope. Existing organisations are stepping in and new organisations have been created to provide suggestions and standards for the development of accessible XR (Section 2.6.2). Engagement with impaired communities will be central to progress [4].

2.6.2 Standards and Guidance for the Accessibility of AR

Over the last few years, several efforts have begun working towards accessible design in XR. Table 2.2 highlights these efforts, providing some background information on their origins and the contributions they are making. Some have come from established organisations, such as the Institute of Electrical and Electronics Engineers (IEEE) or the World Wide Web Consortium (W3C). Others come from new organisations, set up by concerned stakeholders working together, such as the XR Association. These efforts have several aims including the creation of standards and guidance, collaboration with stakeholders and researchers and the communication and promotion of accessible practices.[10]

From Table 2.2, we can see that these efforts have produced 3 main resources that provide guidance on creating accessible XR; the Cyber-XR coalition's standards [10], the W3C's XRAU working draft [123] and the XR Association's developers guide [99, 124, 125]. There are of course other guidelines provided by specific XR manufacturers, such as Oculus [126] or Magic Leap [127]. At present, none of these focus solely on AR, but provide guidance for all XR technologies. Table 2.3 attempts to summarise recommendations made that could apply to AR and VI individuals. The W3C guidelines especially go into much more detail and the exact ways that each document frames solutions are not consistent. For this sort of guidance to be effective in creating more accessible XR, developers must consider accessibility from the start [121]. All those investigated in Table 2.3 do advocate accessibility being part of creative processes from the ground up. Much of this guidance bears some similarities to existing solutions for making technology accessible to VI users (see Section 2.2.3). Whether this in an issue or not is still to be determined as little thought has yet been put into the specific needs of VI users in mobile AR [18]. Not all the suggestions in Table 2.3 work for all VI individuals as visual considerations only apply to those with some vision.

It is interesting to note that many of the audio methods in Table 2.3 correspond with those used in AAR and in the traditional media approaches investigated; spatial audio, use of sound effects and use of speech (TTS or AD). It is also encouraging to see that there is some acknowledgement of personalisation in these suggestions. Almost all the visual considerations are framed as allowing the user choice in altering visual characteristics. It is possible to see direct similarities between NI methods and the suggestion of users having the ability to remove or reduce audio elements.

[10] For the interested, the third part of the ITIF series provides a more comprehensive overview [121].

Table 2.2: Efforts working towards accessible XR

Name	Origins	Contributions	Refs
Cyber-XR Coalition	Launched in 2020 by the XR Safety Initiative (not-for profit)	Focus on creating standards, policy guidance and providing oversight. Created 'Immersive Technology Standards for Accessibility, Inclusion, Ethics, and Safety'.	[10]
Global Initiative on Ethics of Extended Reality	Industry connections group launched in 2020 by the IEEE	Focus on ensuring the ethical development of XR technologies. White papers, policy recommendations and workshops.	[128]
Partnership on Employment & Accessible Technology (PEAT)	Organisation funded by U.S. Department of Labor	Develop policies and practices for technologies emerging in the future of work. White papers, created 'The Accessibility Playbook for Emerging Technology Initiatives'.	[129]
XR Accessibility User Requirements (XRAU)	Public draft by W3C's Accessible Platform Architectures (APA) Working Group, first published in March 2020	Create standards, currently in progress. It is expected that the Web Content Accessibility Guidelines (WCAG) 3.0 standards will soon be completed and will include finalised recommendations for XR	[123]
XR Access Initiative	A Research Consortium at Cornell Tech, in collaboration with Yahoo. Formed in 2019	Research network and symposiums, collaboration into guidance and standards (e.g. W3C, PEAT), create and sign-post resources.	[130]
XR Association Developers Guide	The XR Association includes Google, HTC, VIVE, Microsoft, Oculus from Facebook, and Sony Interactive Entertainment.	Developers guide released in 2020.	[101]

Table 2.3: Guidance applicable to making AR content accessible to VI users. Elements refer to anything within the AR experience that a user may want to understand or interact with - text, objects, etc.

Suggestion	Cyber-XR Coalition [10]	W3C's XRAU working draft [123]	XR Association's developers guide [125]
Visual considerations			
Altering the size of elements	×	✓	✓
Altering fonts	×	×	✓
Altering foreground/background colours of elements	✓	✓	✓
Altering brightness levels	×	×	✓
Adding contrasts or edge enhancements when highlighting elements	×	×	✓
Recolouring the interface	×	×	✓
Providing symbols, shapes or textures as alternative ways of recognsising elements	×	×	✓
Scrim/overlays for text readability	×	×	✓
Removing or reducing unnecessary details	×	✓	✓
Audio considerations			
TTS and/or AD	✓	×	✓
Audibly identifying elements as they are explored	✓	✓	✓
Removing or reducing unnecessary audio elements	×	✓	✓
Spatially accurate sound design	×	✓	×
General considerations			
Ability to undo/redo	×	×	✓
Allowing users more time to perform interactions	✓	✓	✓
Filtering or sorting elements	×	✓	×
Bypassing challenging/difficult experiences	×	×	✓
Compatible with existing assistive technologies	✓	✓	×
Alternate methods of input/interaction	✓	✓	×
Save progress	×	×	✓

It is excellent that many efforts now exist to make XR accessible, and that these are beginning to lead to increasingly comprehensive standards and guidance. Unfortunately, this does not necessarily translate to content being accessible to impaired users. As Crabb et al have shown in [40], the existence of standards, guidance and tool kits do not always lead to developers being aware of and implementing accessible design. They also point out that such resources do not encourage development with accessibility in mind from the beginning, but that instead if developers use them, they assess accessibility retrospectively. Furthermore “accessible design in AR/VR is far too complex for guidance to be limited to baseline technical standards” [121]. Developing with the involvement of VI individuals is always necessary even when standards are fully implemented [40, 44, 121]. There are solutions being developed to make such input more available to developers. For example, Brunel University and the University of Cambridge have announced that they are beginning the ‘Inclusive Immersion project’ [131] inspired by the University of Cambridge’s Inclusive Design Toolkit [132]. They aim to add additional tools and software plugins for immersive environments to help developers simulate impairments and find solutions.

2.6.3 Research into the Accessibility of AR

Mobile AR

There has been very little work into how mobile AR can be made accessible to VI users. We thus have a limited understanding of the challenges that are faced and of potential solutions. However, Herskovitz et al [18] have made valuable contributions in this area. They attempted to formulate a meaningful understanding of how we use AR, so that challenges for VI users can be identified. They reviewed 105 apps from the Apple app store. From these, a taxonomy of tasks needing to be completed when using mobile AR was identified. Figure 2.3 shows a reproduction of this. They break down the tasks into 5 main blocks, with lower-level items describing the different ways a user may go about completing them.

Using the taxonomy shown in Figure 2.3, Herskovitz et al [18] then created prototype accessible AR apps. To do so, they made the structure of AR scenes work with VoiceOver (the built-in screen reader for iOS) [16] so that 3D AR objects could be interacted with in the same way that VoiceOver handles targets in other apps. They also added a feature allowing users to freeze the AR scene, so they can grasp what is going on without the content moving around.

Five prototype apps were matched up with building-block tasks. One was involved scanning surfaces for *Establishing Physical/Virtual Correspondence*. Two involved placing virtual objects on a surface for *Creating Virtual Content*. The final two involved on locating virtual objects for *Observing Virtual Content*. For the tasks that had two apps, one of these apps used more standard methods which explained what was happening as the user explored (termed ‘camera-based placement’ and ‘camera search’) whilst the second version guided them, giving instructions on possible placements of objects or guidance to find them (termed ‘guided placement’ and ‘guided search’). Based on these 5 apps, two full AR experiences were made to evaluate how successfully participants could use more complete AR apps with multiple tasks involved. A retail app allowed users to place furniture in their physical space, using ‘camera-based placement’ and ‘camera search’ functionality. An educational app allows users to place a virtual solar system in their space and then explore informational panels. This used ‘guided placement’ with ‘camera search’. These apps thus required users to successfully scan, place virtual objects, locate virtual objects, select them and move them. Therefore elements of all the building-block tasks are evaluated in the full AR experiences.

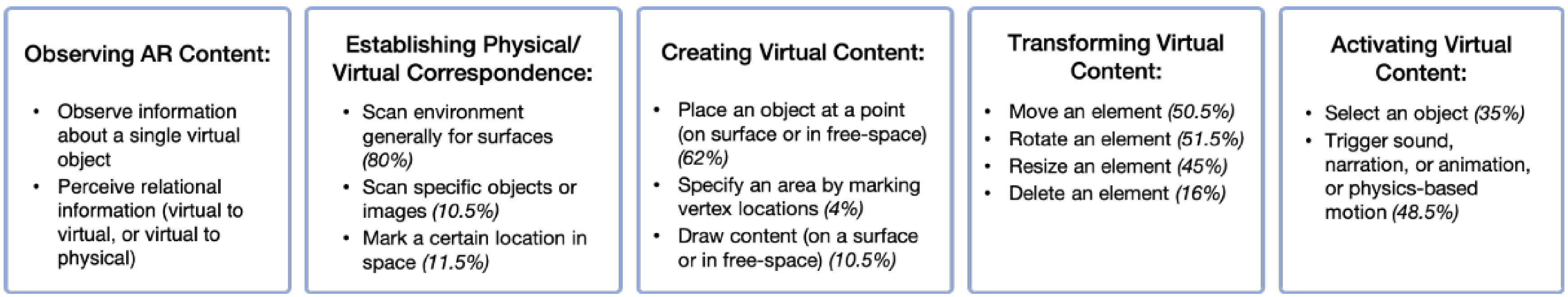


Figure 2.3: A breakdown of ‘building-block’ tasks needed to interact with AR apps, broken down into higher level categories. Percentages indicate how many of the 105 AR apps reviewed contained that task. Reproduced from Herskovitz et al [18].

Evaluation of these apps with 10 participants demonstrated positivity towards their accessibility methods. The participants were able to use the apps, felt confident doing so and even had a fairly strong grasp of the virtual objects around them. They felt that completing the tasks was not too physically or mentally demanding. Some interesting further challenges were also identified. First, the different ways VI individuals may have interpreted virtual elements. For some, the virtual elements became just something else in their 'mental map' of the space, whereas other struggled to understand them, in part due to their non-physicality. Second, they identified the need for richer contextual descriptions. This included more information about both the physical environment such as warnings to prevent collisions and about the virtual objects. For example, rather than just telling the user that a surface has been detected, giving them information about the distance to it.

Head-Mounted Display AR

Outside of mobile AR, there has been some work done to determine what low-vision people can actually see in AR glasses by Zhao et al [133]. They suggested that developers need to think about the type, colour and size of virtual elements. If these are taken into account, they suggest that AR glasses could be powerful assistive tools, aiding in the reading of text, visual searches and navigation.

Research into the accessibility of VR

There has been much more research into methods for making VR accessible to impaired users. This is partially because it is easier to conceive of methods for making a whole virtual world accessible, rather than making an overlapping real and virtual world accessible [18].

A particularly impactful project in this area is the Immersive Accessibility (ImAc) project [5, 25, 38, 134, 135]. Through the creation of an end-to-end platform, they have integrated accessibility services into 360° video VR content, allowing users to both produce and play accessible content [25]. They have also considered the unique needs of AD in VR [38, 134, 135]. User testing has shown that using spatial audio made the experience more immersive and that if AD is placed where the action is happening, it helps users to orient themselves within a scene [134].

Other work has been done into how VI users can effectively explore VR environments Picinali et al [136] have shown that VI users can use only auditory cues to learn the configuration of a new environment. Picinali has further tested the use of VR environments to help VI individuals create a cognitive map of new environments, as doing so can be a long process that may be difficult or socially uncomfortable to complete in situ [137]. Results demonstrate the plausibility of using spatial audio in this way. Siu and a Microsoft research team [138] have also explored the use of a modified to help VI users navigate virtual environments. Their blind participants commented positively on the power of the spatial audio; that it immersed them and provided information on orientation, distance and occlusion.

2.6.4 Research into Assistive AR

Within research into VI individuals using AR, there is a focus on using AR as an assistive technology, a tool to make the world more accessible [18]. Most projects attempt to use AR to overcome one of two challenges:

1. Navigation, which can be further split up into challenges of finding and getting to points of interest, and avoiding potential dangers or obstacles.

2. Interpreting the world around the user [120]. Scanning to identify things of interest (i.e. objects, text, signs, faces).

Table 2.4 reviews 11 such research projects categorised by the above challenges and the end product made. A summary of the aims of each project is given. 6 of the projects (55%) provide navigational solutions and the remaining 5 (45%) help users with interpreting the world around them by scanning and providing feedback as users explore. These challenges are linked as interpreting the world around the user can aid in navigation solutions, as well as being an aim unto itself. Therefore where projects cross this boundary, they are categorised by their primary aim. Given these aims, all of these AR solutions can be thought of as existing in a world-sized 'playspace'. The most recent project included, PeopleLens [139] explores AR's potential to help with social sensemaking by helping a blind child with conversational cues. The proposed system gives the user information about who is in their vicinity and where they are, then attempts to help the user orient themselves towards the persons face. It also provides an external visual cue to people around the user so they know if they've been detected. This is particularly interesting as it takes interpreting the world around the user to a whole new level.

A more comprehensive version of Table 2.4 with further clarifying details can be found in Appendix C.

Research projects like these exhibit the 'state of the art', bespoke, solutions. In most cases smartphones (4, 36%) or HMD devices (4, 36%) are used with custom software. However, in some cases, custom hardware has also been created. NAVIG [50] uses a bespoke system of head-mounted cameras and sensors with a body-mounted GPS system, and VISTE [140] uses maps augmented with audio and tactile cues.

The methods that the research solutions use to communicate information to the user are also shown in Table 2.4. Figure 2.4 provides a further visual aid for counts of features that line up with the accessibility strategies explored for linear media (EAD and NI). This allows for easy comparison with similar evaluations of assistive mobile AR apps in Section 2.6.5 and most other mobile AR content in Chapter 3. Although EAD includes i-Voice, spoken instructions are not included in these counts as its equivalent, screen reader technology is an encouraged standard practice in mobile applications. Therefore, 'audio description' is considered a standard accessibility feature for these investigations.

Figure 2.4: Summary of enhanced audio features used in AR research projects

Table 2.4: Research projects using AR as an assistive technology. Features assessed include Spoken instructions (Spok.), Enhanced Visual Instructions (Vis.), Spatial (Binaural) Audio (Spat.), SFX to convey actions or help with navigation (SFX.), Personalisable audio options (Pers.), Haptic Feedback (Hap.) and references. *This project compares mono audio to a stripped-back version of spatial audio (stereo with ILD and ITD adjustments). The projects are also given a categorisation (Cat.) based on whether they attempt to assist users with navigation (N) or interpreting the world around (I).

Name	Aim	Cat.	Year	Delivery	Spok.	Vis.	Spat.	SFX.	Pers.	Hap.	Refs
NAVIG	Using GPS and realtime sensors to aid navigation	N	2012	Bespoke	✓	×	✓	✓	×	×	[50]
CueSee	Visual cues to aid product search	I	2016	HMD	×	✓	×	×	×	×	[141]
VizLens	Adds screen reader-functionality to real world interfaces	I	2016	Mobile smartphone	✓	✓	×	✓	×	×	[142]
ZebraX	Auditory interface to aid crossing the road.	N	2016	Mobile Smartphone	✓	×	×*	✓	×	×	[54]
NavCog3	Indoor navigation assistant	N	2017	Mobile smartphone	✓	×	×	✓	×	✓	[143]
VISTE	AR maps teaching spatial thinking	I	2017	Bespoke - augmented map	✓	✓	×	✓	×	✓	[140]
An augmented reality sign-reading assistant for users with reduced vision		I	2019	HMD	✓	✓	×	×	×	×	[144]
Designing AR Visualizations to Facilitate Stair Navigation for People with Low Vision		N	2019	Projection based AR, Smartglasses	✓	✓	×	✓	×	×	[145]
Walking Support for Visually Impaired Using AR/MR and Virtual Braille Block		N	2020	HMD and mobile smartphone	✓	✓	×	×	×	✓	[107]
CamIO	Audio interface providing guidance to target	I	2020	Mobile Smartphone	✓	×	×	✓	×	×	[146]
PeopleLens	Identifies and tracks people, providing cues to help the user make eye contact	I	2021	HMD	✓	×	✓	✓	×	×	[139]
				Total	10	6	2	8	0	3	11

Review of Assistive AR Research Projects

In these projects, there is a reliance on spoken cues, with all projects using them except CueSee [141], which provides only visual cues. Perhaps surprisingly, many of the projects (6, 55%) do use some form of visuals. These are designed with low, rather than no vision in mind. All of them except CueSee [141] are usable if the user has no residual vision, although Zhoa et al [145] and Hommaru and Tanaka [107] may be particularly challenging, but unfortunately do not provide details on how effective usage for those without any residual vision would be. A high number (8, 73%) of them do also use sound effects, with some mention of using them as earcons or auditory icons [50, 142].

There is very little use of either spatial audio or haptics. Unfortunately, the 2 (18%) that use spatial audio do not comment upon how effective it is. However, based on user feedback, CAMIO[146] recognises that using spatialised sound would improve it. None of the projects allows the user to personalise the audio to suit their access needs and there is no discussion about the potential of this as a feature.

The extensive use of audio, in particular spoken cues and sound effects, strengthens the argument made for the potential of audio to include VI individuals in AR. However, limited use of spatial audio and personalisable methods demonstrate that just like traditional media, these methods are still underutilised even in cutting edge research.

2.6.5 Assistive Mobile AR Apps

Using AR to make the world more accessible is not limited to a research context. Similar solutions are now available to the average user via their mobile smartphone, which remains the most attainable way to use AR. It is estimated that 90% of visually impaired people have a mobile smartphone [107]. There are also a number of existing assistive wearable solutions which use AR.[11]

Apps selected

Apps that use AR as an assistive technology seek to overcome the same two challenges of navigating and interpreting the world around a user (see Section 2.6.4).

It is attempted here to present and discuss those which are most established and known. Meaning they can be easily discovered by searching words such as ‘blind AR’ on the Google and Apple app stores. When doing so, it was important to filter out apps that don’t actually use AR technology. For example, Be My Eyes [156] shows up in such a search. Whilst it has won widespread recognition, it relies upon volunteers providing visual assistance through a video call, so is not included. Table 2.5 presents the 8 apps selected. The majority of these (6, 75%) help the user with interpreting the world around them. Most of these apps give real-time spoken feedback as the user scans their environment. Although, they all require the user to switch between different modes such as text, currency, barcodes, colours and free exploration. However, Sullivan Plus [154] and Tap Tap See [155] do not scan in real-time, but provide a description based on a photo or video taken. This gives users the additional challenge of first finding the item. All apps selected are available free of charge.

[11] For an overview of assistive wearable solutions, see [147].

Table 2.5: Descriptions of the assistive mobile AR apps investigated. The apps are given a categorisation (Cat.) based on whether they assist users with navigation (N) or interpreting the world around (I).

App Name	Cat.	Description	Refs
Soundscape	N	Uses binaural audio cues (spoken and sound effects) to help exploration, navigation and finding points of interest	[148]
Envision	I	Describes the world around by reading out text, describing objects and scanning to find specific objects	[149]
Lookout	I	Describes the world around by reading out text and describing objects	[150]
Seeing AI	I	Describes the world around by reading out text and describing objects	[151]
Supersense	I	Describes the world around by reading out text and describing objects	[152]
Lazarillo	N	Guidance through voice messages, indoors and outdoors for exploration, navigation and finding points of interest	[153]
Sullivan Plus	I	Describes a photo of the world around by reading out text and describing objects	[154]
Tap Tap See	I	Describes a photo or video of the world around by reading out text and describing objects	[155]

Criteria for Reviewing Mobile AR Apps

These apps were assessed against against two sets of criteria, shown in Table 2.6. These criteria are also used to assess mobile AR apps not designed for an assistive purpose in Chapter 3.

If apps have minimal support for a feature, they are categorised as having 'some' support. Where necessary Table 2.6 provides further descriptions to clarify what a criteria is and what counts as 'full' or 'some' support. Apps are classed as having full support for a feature if it works comprehensively throughout the app. For standard accessibility features, 'some' support often means that the feature only works for User Interface (UI) elements, such as menus, and not in the actual AR parts of the app .

Details of how to access the full assessment table of these apps can be found in Appendix C. This also provides detailed explanations, especially where the features used don't correspond precisely with the assessed categories.

Table 2.6: Assessment criteria used for investigating mobile AR apps.

Criteria	Description
Standard accessibility features - these match with existing solutions given in Section 2.2.3. They are part of both the Android and iOS operating systems [15, 16]. Apps can work with them by ensuring compatibility. Not all apps do this [17, 44], so the assessment checked if developers implemented bespoke solutions	
Audio description	Features which use speech to convey information, such as TTS or screen readers (i.e. VoiceOver [16] or Talkback [15]). Apps have full support for this if it works throughout. A number of apps only provide support in menus, so are classed as having 'some' support.
Subtitles	At an OS level, any spoken content can also be converted to visible text. If implemented within the app, spoken audio content is also displayed using text.
Magnification	Apps are categorised as 'some' if not every element can be magnified (i.e. exclusively UI/AR).
Colour changing settings	Settings to invert the colours, or correct individual colours to suit the users needs. Colour Inversion
Enhanced audio features - these line up with EAD and NI features, except from the exclusion of i-voice due to screen readers being standard on mobile smartphones.	
Spatial audio	Binaural audio over headphones
Sound effects	Must be used to convey actions/interactions or help with navigation. Apps have full support if sound effects are used throughout. 'Some' refers to apps that only use sound effects for a subset of elements (i.e. exclusively UI/AR) or only one specific action
Personalisable audio	Apps have full support if there is refined control over the audio options. If users can only turn on/off categories of sound, they are classed as having 'some' support.

Review of Assistive Mobile AR Apps

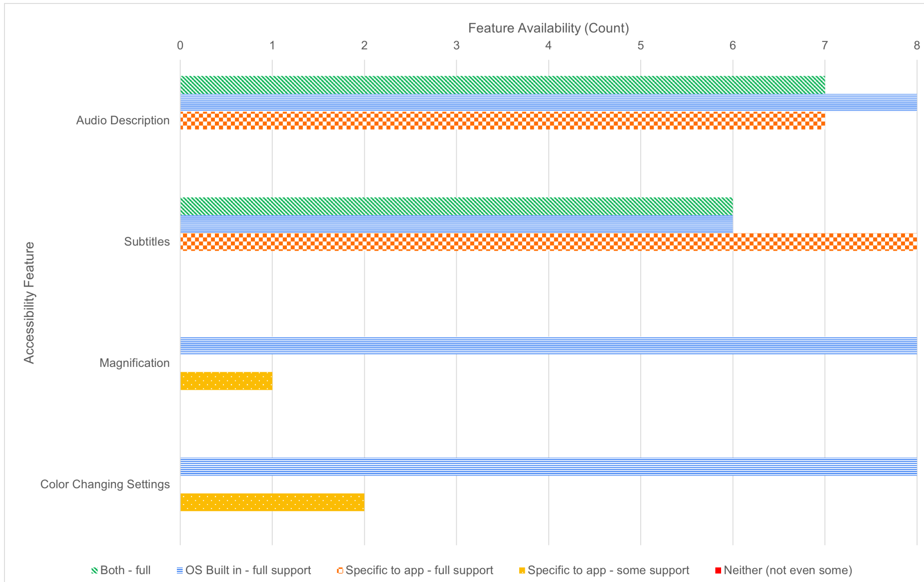


Figure 2.5: Summary of accessibility features used in assistive mobile AR apps.

Figure 2.5 shows how much support the apps have for standard accessibility features. It gives counts of whether apps have full support for OS built in tools, bespoke solutions or both of these. It also gives counts for whether features have some support for features using OS tools or specific features. Finally, a count is given of apps that have no support of the feature, neither using OS tools nor implementing bespoke solutions. Unsurprisingly, all the apps investigated have a high level of compatibility with standard accessibility features, with almost complete support for OS built in accessibility features.

Audio description. Screen readers worked with UI and AR elements in all apps. This is surprising, as most VR & AR (60%) content is made using Unity [14]. Apps made with Unity do not easily integrate with OS screen readers on iOS or Android without additional coding outside the Unity environment or third-party toolkits. All apps except Tap Tap See [155] also gave spoken descriptions when using their main functionality (i.e. describing objects or navigational hints) regardless of whether the OS screen reader was enabled.

Subtitles. This was the only feature not supported by all apps at an OS level. However, all apps ensured that any spoken information was also displayed as text.

Magnification. Only Sullivan Plus [154] implemented additional magnification features to OS by allowing the user to zoom in on the real world whilst doing AR scanning.

Colour changing settings. Envision [149] and Lazarillo [153] added their own dark and light modes in addition to OS options.
Figure 2.6 shows how much support the apps have for enhanced audio features. Details are not needed for whether apps support OS or bespoke implementations of these, as they are not standard practice. Therefore, ‘yes’, ‘no’ or ‘some’ is used to indicate if apps support these features. These can be compared to assistive AR research projects in Figure 2.4.

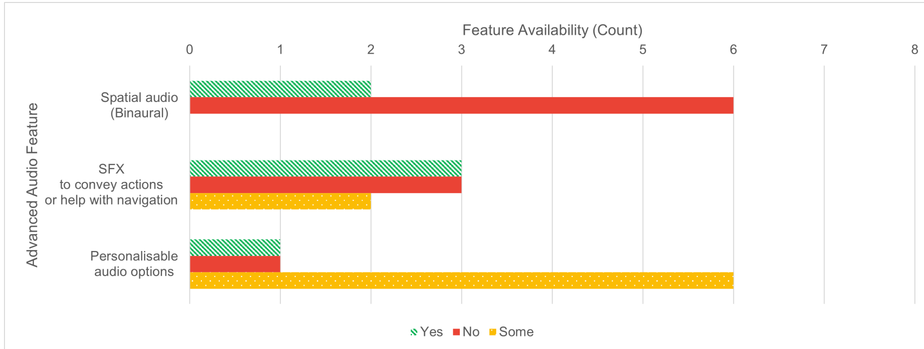


Figure 2.6: Summary of enhanced audio features used in assistive mobile AR apps.

Spatial Audio. Only the apps from Microsoft, Seeing AI [151] and Soundscape [148] used spatial audio. As for the assistive AR research projects investigated, this demonstrates the underutilisation of this powerful tool.

Sound Effects. Usage of sound effects was mixed. There was an even split between those with full support and those with no support (3 each).The 2 remaining apps only used sound effects to indicate one specific action, which in both cases was the primary task required. This is disparate to the assistive AR research projects, which mostly used sound effects.

Personalisable audio options. Most of the apps allowed users some control over the audio. In most cases, this meant controlling the speed of the screen reader, sometimes with the additional options to turn off specific sound effects. Only Lazarillo [153] gave finer control, with options to alter the speed of the screen reader, its pitch and relative volume. None of the assistive AR research projects allowed the user an control. It is promising that despite this, these publicly available apps implement some level of personalisation.

These apps convincingly demonstrate that in certain contexts mobile AR can be fully accessible to VI users. They illustrate just how powerful it can be as an assistive tool, with many industry awards between them. They indicate the need for compatibility with standard OS accessibility features, whilst also implementing creative solutions of their own. For example, Soundscape [148] creatively uses audio 'beacons', which the user then navigates to based on distance (loudness) and location (spatial audio) cues.

The extensive use of spoken instructions especially compared to sound effects and spatial audio cues could indicate one of two things. Either that spoken instructions are essential for making mobile AR accessible compared to other methods. Or that other methods are under-explored and therefore underutilised. Both of these may be true, given that the same is true in traditional media and assistive AR research; spoken instructions are relied on whilst sound effects and spatial audio remain underutilised.

However, using these apps to establish methods for the accessibility of mobile AR is limited. If we return to the building block tasks identified by Herskovitz et al [18], only the *Establishing Physical/Virtual Correspondence* task is ever needed in these apps. Whilst the existence of such apps is excellent, it is not enough for VI access to AR to be limited only to distinct content types. Thus further work into how all mobile AR can be made accessible is desperately needed, as demonstrated by the lack of work in this area (see Section 2.6.4).

2.7 Conclusions

2.7.1 Summary

This chapter has begun to address the first objective of this work, establishing the current state of accessibility research in AR especially as applied to mobile AR.

Foundational to all other areas is a consideration of the needs of VI individuals. Section 2.2 outlined the following:

- The prevalence and characteristics of visual impairments, with the criteria for certification and past research informing the categorisation of participants employed throughout the rest of this work (partially sighted, blind with residual vision, blind or other).
- Unfortunately, there is still much work to be done including VI individuals, who are often vulnerable and isolated with much of society not being fully accessible to them.
- The experience, needs and views of VI individuals are highly diverse. From this, a fundamental principle of this work is derived, that any accessibility attempts should take this diversity into account by providing personalisable services.
- Technology has huge potential to either provide leaps in the inclusion or barriers in the exclusion of VI individuals. Which of these applies often depends upon whether accessibility has been considered from the ground up when designing new technology.
- An overview of existing accessibility solutions.

Technologies and solutions that have been successfully applied to established, traditional media were then examined in Section 2.4:

- Sonification and spatial audio are powerful techniques, which can allow multiple layers of information to be provided to the user simultaneously.
- Object-based audio is of great benefit for accessible systems. Its flexibility allows one production to be optimised for any listening format or device. It opens up the possibility of end-user interaction, making it a compelling method to deliver personalisable accessibility services.
- AD is a vital service for the inclusion of VI individuals, However, several criticisms have led to suggestions that AD should be designed from the start of creative processes with input from original creative teams, that guidelines should allow the spoken content itself to be more engaging, and that greater use should be made of advances in audio technology.
- The EAD project created a successful alternative to AD. The methods involve the use of soundtrack enhancement, spatial audio and first-person narrative. Thinking of such methods as accessibility tools may be vital to unlocking answers to accessibility in XR.
- Audio Games further demonstrate the potential of methods suggested by the EAD project. They also revealed that using such methods can create experiences that both sighted and VI individuals enjoy equally.

- The NI concept illustrates that personalisable audio solutions can be simple whilst remaining highly effective for the end-user. It further underlines the creative potential of audio to story-tell and include by considering the importance of every sound object in a piece of content. Further credibility is also given to the notion of involving personalisable services with content creators involved from the start of creative processes.

The medium of AR itself was then considered in Section 2.5, to establish both the opportunities it affords and also what unique requirements it demands of accessibility services:

- AR was defined with an overview of its context within XR technologies.
- Methods for AR delivery and its wide-ranging uses were explored.
- AR is growing incredibly rapidly, even outpacing VR.
- The unique challenges of AR were established. Users need to both understand what is being presented to them and how simultaneously navigate and interact with both the real world and the created virtual world.
- Factors affecting user experience in XR were explored, with particular attention to immersion. Based on past research, the interestingness and interactivity of AR systems were identified as key features of interest when assessing the effectiveness of accessibility methods developed in this work.
- AAR holds potential answers to how 'mainstream' AR could be made more accessible. Its reliance on spatial audio, sound effects and spoken guides prove that EAD-like methods make sense within an AR context.

Consideration was then given to current AR accessibility solutions in industrial standards, research and the use of AR as an assistive technology in Section 2.6. From this came the following insights:

- XR technologies hold huge potential to include and even provide new opportunities for VI individuals.
- The flexible and multi-sensory nature of XR technologies means that they could be accessible in interesting ways.
- There is rising concern that impaired individuals are being left behind in XR technologies, with accessibility solutions once again being designed after the fact.
- Recent efforts have begun to seek understanding of and promote the visibility of accessibility needs in XR. From this, guidance is beginning to be created. This bears many similarities to methods in established methods. Use of spatial audio, sound effects, spoken guides and personalisation methods are all acknowledged. However, guidance does not always translate to developers thinking about accessibility from the start of a project. Further, impaired individuals must always be involved.
- Research into how mobile AR can be made accessible is lacking. Nonetheless, a beginning has been made into understanding the challenges VI users face. A taxonomy of the tasks needing to be completed within AR provides a starting point for solutions. Consideration of how VI individuals interpret virtual elements and the need for rich contextual descriptions will be crucial for solutions to be successful.

- Research into how VR can be made accessible has shown that VI users can effectively explore audio-based environments using spatial audio.
- Research has demonstrated the efficacy of AR as an assistive tool. It can help users to navigate and to interpret the world around them.
- Assistive AR research makes extensive use of audio, and in particular spoken cues and sound effects. However spatial audio and personalisable methods are under-explored.
- There are excellent assistive mobile AR apps available, which can help VI users in their everyday lives. They are compatible with standard accessibility features that are built-in at an OS level. Spoken cues play a particularly important role. Spatial audio has not been explored much in this context. However, the use of sound effects and allowing users some personalisation is fairly common.
- Assistive mobile AR apps only demonstrate limited solutions for the wider accessibility of AR to VI users. This is because they are only one content type within AR, where VI users only interact with one type of AR task.

2.7.2 Making AR accessible for VI users

It is possible to see several threads running throughout the above summary that may provide answers to how AR can be made accessible to VI individuals.

Firstly, that all technology, including AR, holds the potential to either include or exclude VI individuals. What makes the difference is whether accessibility solutions are considered from the ground up or after the fact and whether impaired individuals are involved so that solutions fully cater for their diverse needs and preferences.

Secondly, that creative use of audio is a powerful tool for the inclusion of VI individuals. The EAD project, NI concept, Audio Games, AAR, accessibility guidance for AR, mobile AR accessibility research, assistive AR research and assistive mobile AR apps hold this in common. All use spoken guidance and in most cases, this is particularly relied upon. To the point where in some cases, it is the primary way that information is communicated. Sound effects are also widely used as a tool for communicating information. Spatial audio has not been explored in mobile AR accessibility research, assistive AR research and assistive mobile AR apps. However, it should be explored further given its use in all other areas and proven ability to communicate multiple layers of information effectively. Further to this, Herskovitz's study [18] into the accessibility of AR established the need for rich contextual descriptions to help VI users fully understand what is happening in mobile AR apps. Perhaps solutions could be found to this through the use of sonification and spatial audio techniques. For example, spatial audio may be capable of providing contextual information about the location of physical or virtual objects, without the need for lengthy spoken explanations. Personalisation techniques are similarly unexplored in either research or assistive AR, but their successful use in traditional media and the fact that they are suggested in some AR guidance means that they merit exploration.

2.7.3 Research Questions

From the suggestions above, and the central hypothesis of this work (outlined in Section 1.2.2), it is possible to refine two research questions. These will be investigated during the completion of the third objective of this work, which is to assess the efficacy of the accessibility strategies propose:

1. What challenges do the dual purposes of the soundscapes in XR content, narrative and navigation, present for this approach to accessibility
2. How can the concept of Narrative Importance for personalisable accessible audio be extended into the audio description domain for blind and partially sighted viewers?

The first question links to the part of the hypothesis which posits that ‘it is possible to use creative audio practices’, and the second to the part that posits ‘it is possible to use personalisable audio practices’.

2.7.4 Scope of this work

In testing of the research questions outlined, it is important to be precise about the scope of this work:

- This work is limited to AR for mobile smartphones only, as at present these are the most widely available way for VI, and indeed any, users to access AR content. As the studies presented took part during the COVID-19 pandemic, this also allowed for direct delivery of AR content to participants own devices.
- The approaches foundational to this work (EAD, NI) have been trialled in media with a narrative. Therefore, the AR content explored will be narrative-based. As this is the first step in assessing their applicability to AR, this allows for the most direct transferal of methods these approaches suggest, and thus the most valid evaluation of their success at this stage.

Chapter 3

Accessibility of Mobile Augmented Reality Applications

3.1 Introduction

This chapter is focused on achieving the second objective of this work, which is to establish the accessibility of currently available Augmented Reality (AR) apps. In the previous chapter, the potential of AR to include and provide opportunities for Visually Impaired (VI) users was showcased. Current standards and potential solutions were discussed, providing a focus for the research questions of this work and features that could be used to improve the accessibility of AR for VI users. In this chapter, we assess currently available mobile AR apps against these features. First, 36 apps are assessed in Sections 3.2 - 3.3. Results from this are then validated by Study One, which takes into account the experiences and suggestions of VI users themselves in Sections 3.4 - 3.7.

3.2 Assessment of Current Mobile AR Applications

Modern smartphones have many built-in accessibility features [15, 16]. However, this does not stop them from being challenging for VI users [44], with the majority not fully able to make the most of new technologies [3]. An interesting approach by Ross et al [17] used an epidemiological approach, considering different accessibility barriers in mobile apps as 'diseases' that should be approached, measured and addressed across the whole population. Using this approach they analysed 100 Android apps, finding that all of them had accessibility errors, with 72% having five or six errors. In Chapter 2 several assistive mobile AR apps were investigated. These were designed specifically to make the world more accessible to VI individuals. They were found to be almost entirely compatible with standard accessibility features, with some creative solutions of their own including some instances of enhanced audio features being utilised. However, as these apps are designed specifically with VI users in mind, they do not necessarily give any indication of how accessible mobile AR is more generally. The following sections assess whether the wider category of mobile AR apps are also accessible, or if they fall short, as it would seem that most mobile apps do.

Presented here is a summary of the work so that conclusions can be drawn. Details on how to find the full spreadsheet of mobile AR apps investigated can be found in Appendix C. This also provides detailed explanations, especially where the features used don't match up precisely with the assessed categories.

3.2.1 Assessment Criteria

The apps assessed in this section were compared to the same set of features as the assistive mobile AR apps examined previously (see Section 2.6.5). They are split up into two categories: First, standard accessibility features including audio description, subtitles, magnification and colour changing settings. Second, enhanced audio features including spatial audio, sound effects and personalisable audio. The full list of these features, with clarifications where needed are shown in Table 2.6.

Tools for determining the accessibility of an app do exist, for example Google Accessibility Scanner [157], which was used in [17]. At present there is no such tool that exists across both the Android and iOS platforms. Furthermore, such tools do not always capture the nuances of AR apps. In particular, they are poor at recognising that the virtual elements of AR also need to be made accessible to users. This means that they may tell you an app is fully compatible with screen-reader technology, when in fact virtual AR elements are not recognised or described. Therefore, as with the review of assistive apps in Chapter 2, each app was used by the author and assessed against the criteria given. Android apps were tested using a Google Pixel 3A running Android 12. iOS apps were tested on an iPhone 7 running iOS 14.8.1.

3.2.2 Apps Selected

This assessment was conducted on 37 mobile AR apps. These were selected by searching 'Augmented Reality' on both the Android and iOS app stores and selecting the top results. As much as possible, apps that were available on both Android and iOS were chosen. However, 5 were not available on iOS and 1 was not available on Android. Of these, two were different takes on using AR to measure distances, one by Apple and one by Google, so were therefore only available on their own platforms.

Table 3.1 provides information on the distribution of different app categories. Although there is a range of categories included, games and educational are more prevalent. This is to be expected as starting with Pokémon GO, AR has grown hugely in gaming, which continues to be a driver for its popularity [11, 12].

Table 3.1: Categories of AR apps investigated for various accessibility and advanced audio features.

App Category	Count	%
Art & Design	1	2.7
Education	9	24.3
Entertainment	3	8.1
Game	16	43.2
Health & fitness	1	2.7
Social	3	8.1
Tools	3	8.1
Travel & Local	1	2.7

It can be difficult to precisely categorise the playspace [110] that an AR app uses as many apps adapt depending on the space that the user allows them. When using the apps, an assessment was made of which playspace size was the most intuitive when using it, based on contextual information such as the size of elements used, or how the user is expected to explore them. The majority (19, 51.4%) of the apps utilised a world sized playspace, with the remaining apps fairly evenly split between tabletop (8, 21.6%) and room (10, 27.0%).

3.2.3 Assessment of Current Mobile Apps

Standard Accessibility features

Standard accessibility features are shown in Figure 3.1. These can be compared to the assistive mobile AR apps in Figure 2.5. In the same way, it gives counts of whether apps have full support for OS built in tools, bespoke solutions or both, whether features have some support for features using OS tools or specific features or whether apps have no support of the feature in either OS or bespoke implementations.

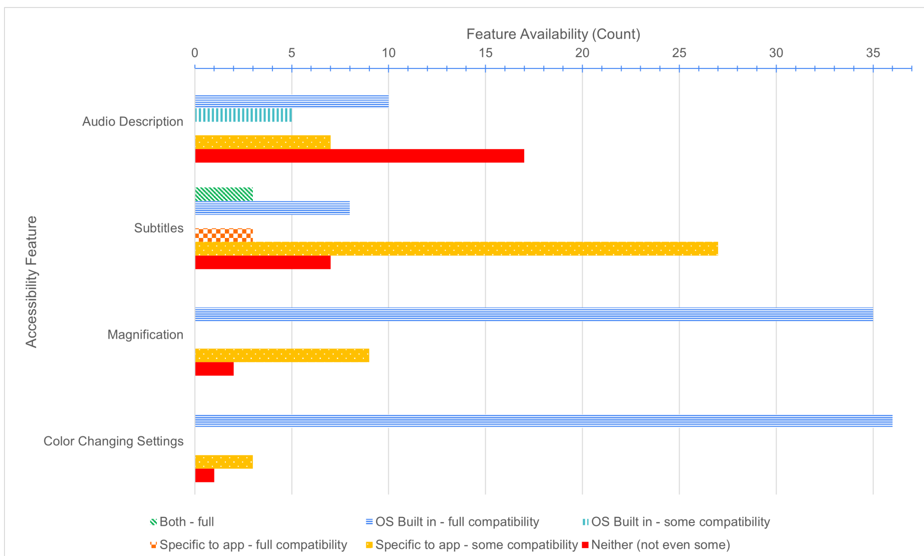


Figure 3.1: Summary of accessibility features used in current AR apps. Full details can be found via Appendix C.

Audio description. Compatibility with screen readers was mixed, unlike the assistive AR apps which all had full compatibility with screen readers at an OS level and mostly implemented built-in support too. Here, the majority (17, 36%) had no support for screen readers whatsoever. When this is the case, users who rely on OS screen readers are unable to even get past the first screen of the app, as nothing on the screen is recognised by them and thus nothing can be interacted with. On Android, Talkback just says "Game view" for all screens, whilst on iOS VoiceOver says "Direct Touch only". Of the 15 (41%) that were compatible with OS built-in screen readers, a third only had some support, meaning the built-in screen reader worked for User Interface (UI) elements, but not with the AR itself, which in practice means that the app's function is not screen-reader compatible. Of the 7 (19%) that implemented their own screen reader, all only had some support, mostly giving information about what was happening in the content but not information on UI or how to navigate and explore AR elements. All but 2 of the apps with their own solutions were not compatible with built-in screen readers. This means that users who fully rely on screen reader technology would encounter difficulties with the bespoke solutions. To first access them, they have to turn off their OS-level screen reader or nothing would be recognised, with only warnings spoken as described above. This would perhaps be justified if the bespoke support then received was appropriate. However, the incomplete support of the bespoke solution makes actual use of the apps extremely difficult. There's no way of telling how many of these apps were made using Unity, but given its ubiquity, it's likely to be many of them [14]. Therefore it's not surprising that there is low compatibility with screen readers, given Unity's lack of easy support for them without additional coding outside the Unity environment or third-party tool-kits.

Subtitles. Support for this was much more mixed than with the assistive mobile AR apps. Most of the apps did not have any spoken content (8, 22%). However, the majority (27, 74%) had some custom support, showing some text-based instructions (mostly for initial set-up) or character speech as text. A further 3 (8%) had full custom support, showing all information and character speech as text within the app. These 3 were all also fully compatible with OS built-in subtitles.

Magnification. Almost all apps (35, 95%) were fully compatible with OS built-in magnification features. For the two that weren't, magnification did occur but it rendered core content unviewable and unusable due to how they reacted and they also did not implement their own custom solution. The 9 (24%) that implemented custom magnification all did so only for the AR content, and not the UI elements. This level of support is similar to the assistive mobile AR apps.

Colour changing settings. All but one app (36, 97%) had full compatibility with OS built-in colour changing settings. A small number of apps (3, 8.1%) added their own 'dark mode' settings as well. Again, this is similar to the assistive mobile AR apps.

Enhanced Audio Features

Enhanced audio features are shown in Figure 3.2. These can be compared to assistive mobile AR apps in Figure 2.6 and assistive AR research projects in Figure 2.4. As previously, only 'yes', 'no' or 'some' is needed here, as use of these features is always an additional choice by the developer.

Spatial Audio. Only 1 app (3%) had support for spatial audio. As with the assistive AR research projects and assistive mobile AR apps, this once again demonstrates its underutilisation.

Sound Effects. The majority (21, 57%) did not make use of sound effects and those that did only made limited use of this feature. For most, this meant sound feedback when specific UI elements (and sometimes AR content) were interacted with. This is the opposite of the assistive AR research projects, most of which used sound effects. The assistive mobile AR apps had more mixed usage, but some did have full support and as many had some support as had no support. Meaning overall adoption of sound effects was greater in assistive AR than general mobile AR.

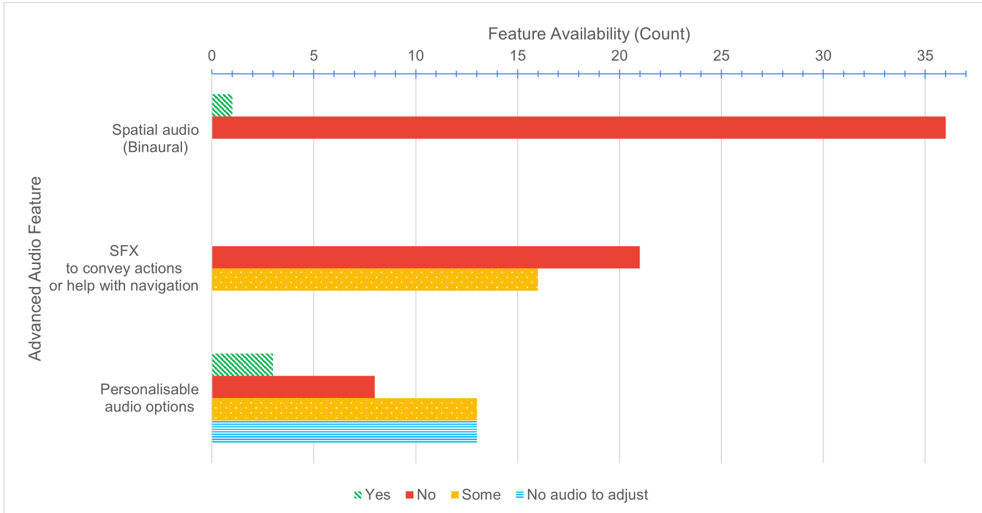


Figure 3.2: Summary of enhanced audio features used in current AR apps. Full details can be found via Appendix C.

Personalisable audio options. A large proportion (13, 35%) of the apps didn't have any audio content. A further 8 (22%) did have audio, but with no personalisable options. Most of those who did provide options only had some support (13, 35%), mostly the option to turn off sound effects and/or music. 3 apps (8%) did allow the user fine control of the levels of different categories of audio. Setting aside apps with no audio to personalise, this is somewhat similar to the assistive mobile AR apps, most of which had some support for personalisable audio options. This is despite assistive AR research projects having no support for this feature whatsoever.

3.3 Discussion of Current Mobile AR Applications

The current mobile AR apps assessed demonstrate that there are shortcomings in the accessibility of mobile AR apps. They exhibit limited support even for OS built-in and recommended accessibility features. Crucial compatibility with built-in screen reader features is low, with alternate solutions not going far enough to provide meaningful support. However, support for magnification and colour changing features is more positive.

The fact that not all apps had any audio content impacted the use of subtitling and personalisable audio options. For both these features, most apps with audio had some support. This perhaps indicates that personalisable audio options can be viably used, but are not pushed far enough by developers at present. The almost non-existent use of spatial audio and minimal use of sound effects demonstrates that creative use of audio, both in general and as an accessibility tool is simply not a consideration in most mobile AR app development.

3.4 Study One Material : Selecting AR Applications

To validate the results of the assessment of current mobile AR apps, Study One asked VI participants to use and provide feedback on a subset of the apps assessed. Those that were selected attempt to account for as much diversity as possible. Particularly in size of playspace

and which of Herskovitz's AR building block tasks are utilised [18]. Whether or not they have a narrative was also considered, as narrative content allows for the most valid initial exploration of EAD and NI strategies in AR (see Section 2.7.4). This does mean that the categories of mobile AR apps selected may be more homogeneous than is ideal.
Table 3.2 summarises the apps selected including their accessibility and enhanced audio features. Not having screen reader support may prove frustrating for VI users, especially those who rely more completely on this sort of technology to use their devices. To get past initial screens users have to disable TalkBacK (Android) and VoiceOver (iOS), which may prevent users getting very far into the apps. However, this does reflect the current state of mobile AR apps (as Section 3.2 demonstrated). Likewise, none of them have spatial audio or personalisable audio options. A more detailed discussion of the content of each app follows.

Table 3.2: Summary of apps selected for Study One. Full details can be seen in Appendix D

App Feature	**Angry Birds AR**	**Pokémon GO**	**WWF Free Rivers**
Category	Game	Game	Educational
Playspace size	Tabletop	World	Room
AR Building block tasks [18]	'transforming virtual content', 'activating virtual content' and 'observing AR content'	'establishing physical/virtual correspondence'	'creating AR content', 'creating virtual content' and 'activating virtual content'
Audio description	-	-	No OS built in, but some specific to app support
Subtitles	No OS built in, but some specific to app support	No OS, but some specific to app support	Yes, OS and own text based instructions
Magnification	Yes, OS	Yes, OS	Yes, OS and can zoom in on AR content
Colour changing	Yes, OS	Yes, OS	Yes, OS
Spatial Audio	-	-	-
SFX to convey actions/help with navigation	Some	Some	-
Personalisable audio options	Some	Some	-

3.4.1 Angry Birds AR

Angry Birds AR provides a nice example of a simple AR game. The concept of the game may also be one that users are familiar with, as the non-AR version(s) of Angry Birds are popular mobile games.

Figure 3.3 gives an idea of the gameplay of this app, which revolves around placing a play space where virtual towers are created, which the user then knocks down by firing birds at using a slingshot.[1] Angry Birds AR was also deemed as containing content appropriate for any participant to use, as it is rated 'PEGI 3 (suitable for all age groups)' on the Play Store [158] and '9+ (Infrequent/Mild Cartoon or Fantasy Violence)' on the App Store [159].

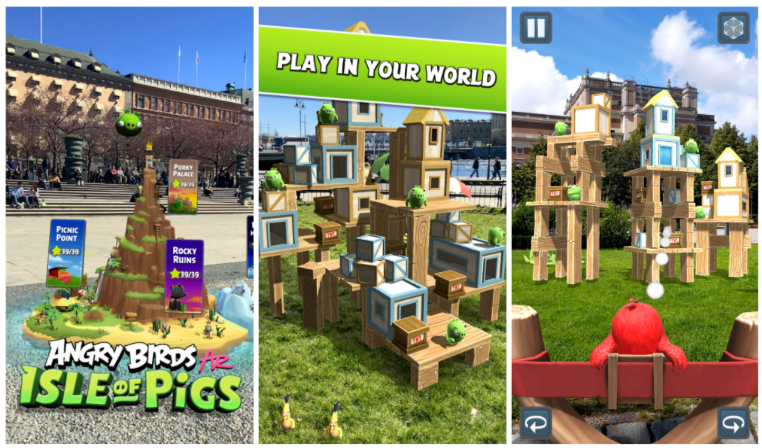


Figure 3.3: Game play screenshots of Angry Birds AR, the first app participants used in Study One. Taken from the Google Play Store listing [158]

Angry Birds AR has limited accessibility features (see Table 3.2), with no support for screen readers and only some text-based instructions on initial set up. As with all the apps, magnification and colour changing settings are supported at an OS level. It also has limited support for enhanced audio options, with only some sound effects on button presses and actions in AR. Users can turn sound effects and music on and off individually, but have no control over their levels.

3.4.2 Pokémon Go

The popularity of Pokémon Go and the impact it has had on mobile AR content made it an obvious choice. Figure 3.4 gives an idea of the gameplay of this app, which revolves around the user searching the real world for virtual Pokémon characters, which they can then interact with. [2] It was deemed appropriate for all participants to use as it is rated 'PEGI 7 (Mild Violence)' on the Play Store [160] and '9+ (Infrequent/Mild Cartoon or Fantasy Violence)' on the App Store [161].

Pokémon Go has limited accessibility features (see Table 3.2), with no support for screen readers and only some text-based instructions on initial set up. As with all the apps, magnification and colour changing settings are supported at an OS level. It also has limited

[1]The interested can find more a more comprehensive summary of the narrative and gampleplay of Angry Birds AR on its Play Store [158] or App Store [159] listings.

[2]The interested can find more a more comprehensive summary of the narrative and gampleplay of Pokémon Go on its Play Store [160] or App Store [161] listings.

support for enhanced audio options, with only some sound effects on button presses and actions in AR. Users can turn sound effects and music on and off individually, but have no control over their levels.

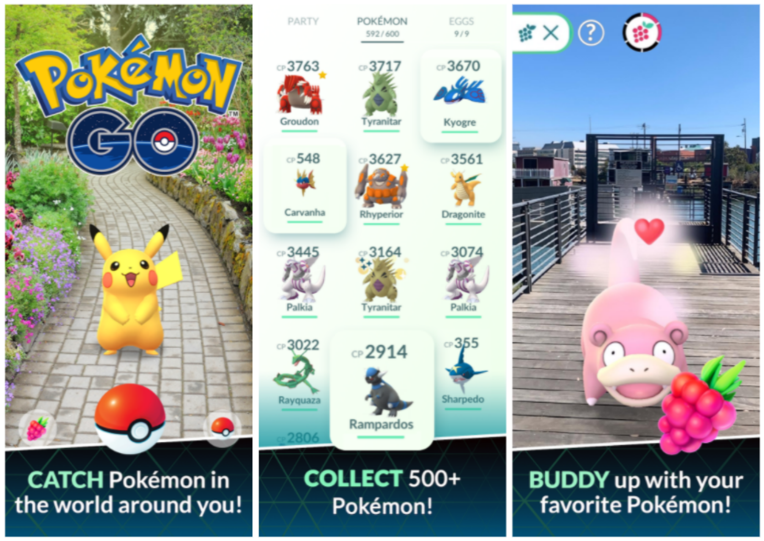


Figure 3.4: Game play screenshots of Pokémon Go, the second app participants used in Study One. Taken from the Google Play Store listing [160]

3.4.3 WWF Free Rivers

Out of the three apps, WWF Free Rivers gets the user to explore around a single virtual AR object the most. It has a clear narrative for the user to follow, and also provides some ambient sounds which give cues as the user gets near certain parts of the object whilst exploring.

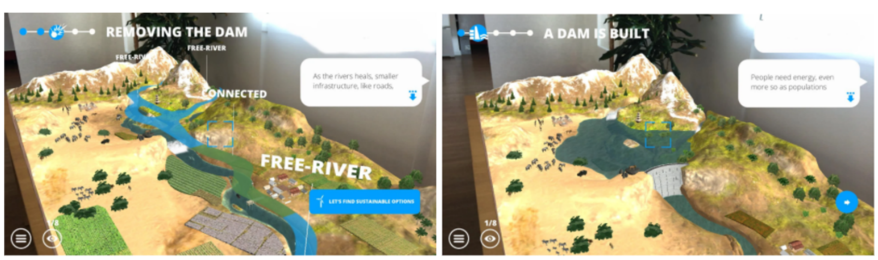


Figure 3.5: Game play screenshots of WWF Free Rivers, the third app participants used in Study One. Taken from the Google Play Store listing [162]

Figure 3.5 gives an idea of the gameplay of this app, which revolves around the user creating and then exploring a virtual river.[3] It was deemed appropriate for all participants to use as it is rated 'Everyone' in the Play Store [162] and '4+' on the App Store [163].

WWF Free Rivers has the most accessibility features of the apps (see Table 3.2). It has no support for OS built-in screen readers, but built-in narration reads out some information

[3]The interested can find more a more comprehensive summary of the narrative and gampleplay of WWF Free Rivers on its Play Store [162] or App Store [163] listings.

(the introduction and main narrative) to you, but not instructions or extra information, such as about the extra characters you can explore. It doesn't read out information on interface buttons and information. All audio is converted to text by OS built-in subtitling and text displaying all spoken instructions is also displayed. As with all the apps, magnification and colour changing settings are supported at an OS level. It has no support for enhanced audio options.

3.5 Methodology

This section outlines how Study One was designed, including instructions given to participants and what information was collected from them. Full ethical approval was obtained from the University of York Physical Sciences Ethics Committee. This can be found in Appendix A. Details of how to access the full survey instrument can be found in Appendix D.

3.5.1 Recruiting Participants

Participants were recruited via a variety of networks including the RNIB, London Vision, the Technology Association of Visually Impaired People (TAVIP), VI networks within the BBC and the University of York Disability Support and Programme Design and Learning Technology teams. The opportunity of being interviewed for BBC Radio 4's 'In Touch' also promoted the study to an even wider audience (see `https://www.bbc.co.uk/programmes/m000w4t6`). Participants were offered Amazon vouchers as a recompense for their time.

As this and both other studies presented took part during the COVID-19 pandemic, participants had to be able to use the apps and provide feedback using their mobile smartphone, in their own home. Their smartphone had to meet the criteria of having a camera and:

- Apple - any iPhone with iOS 12.0 or later.
- Android - must have access to the Google Play store and be compatible with Google Play Services for AR. To see the full criteria and list compatible devices, follow this link (in general, you must be running Android 7.0 or newer): `https://developers.google.com/ar/devices`.

A further requirement was that individuals with severe-profound hearing loss could not take part. Given the aims of this work, participants needed to be able to effectively assess audio-based accessibility methods. These requirements were stated clearly when advertising for participants.

3.5.2 Study Design

As this work contains three studies asking VI participants to use and evaluate AR apps in similar ways, many questions presented here are used across all studies. This is an intentional decision to allow for easy comparison of results.

In all studies, surveys were used to collect participants information and their feedback on the apps. They were created and delivered using Qualtrics (`https://www.qualtrics.com/uk/`. This allows the survey creator to check that the survey is accessible and complies fully with Web Content Accessibility Guidelines (WCAG) 2.0. Making sure that all supporting material, and especially the surveys, were fully accessible was extremely important. As

Table 3.3: Participant instructions for the apps in Study One.

Instructions for Study One
Once you have downloaded the app: 1. Make sure you have some way of making notes available before you get started. Then you can keep a note of your experience, and problems you may face. 2. Complete the tutorial/introduction - the app will automatically take you through this the first time you use it. 3. Spend some time playing/experiencing the application. We would recommend you do this for at least 15 minutes. Try to use as many features as you can. However, if you wish to use the apps for longer, then the more the better! 4. Much of the content can be explored whilst seated. However, for some elements, you may wish to look around or move to look at your playspace from different perspectives. Make sure your immediate surroundings are clear of obstructions and trip hazards. In particular, avoid walking backwards. To experience the audio fully, we recommend that you wear headphones whilst using the applications.

participants could potentially be frustrated if they were unable to use the apps, everything else in the studies needed to be as easy to use as possible. Across all studies any time participants gave information, they were asked to provide unique identifiers in the form of the last three letters or digits of their postcode or zip code and the last three digits of their telephone number. This allowed for results to be linked across studies.

In Study One, participants were given a single Qualtrics link. This first contained an information sheet on the purpose of the study and what they were expected to do. Once participants had read the full information sheet and provided their consent, they were then asked demographic questions and about their past experiences with Extended Reality (XR) (Section 3.5.3). They were then given a link to each app in turn, with instructions for use. They had to fill out the feedback questions (Section 3.5.4 and 3.5.5) for each app before moving onto the next, so their experiences were not influenced by the other apps. Table 3.3 shows the instructions given to participants before they used each app.

3.5.3 Demographic and Past Experience with AR

The set of questions in Table 3.4 were used in all studies to find out some basic information on the background of the participants. Some questions aren't asked until later studies (B5-7). Questions B1-B3.1 established basic information about the participants. This includes asking them to self-report details of their visual impairment using the categories established in Section 2.2.2. Participants were asked if they had any known hearing loss, to ensure they could effectively assess audio-based accessibility strategies. Questions B4 and B4.1 established what sort of previous experience participants had with XR, as familiarity could influence both their ability to use mobile AR apps and their perceptions of its accessibility.

Table 3.4: Survey items addressing participants background. B1-4.1 were used in all studies, B5 was used in Studies Two and Three (denoted by a ˆ) and B6-7 were only used in Study Three (denoted by a *). As all surveys except those for content creators were only open to those with a visual impairment, the option 'none' for questions B3 was only used in the content creator surveys (Study Two and Three).

ID	Demographic question	Type of response
B1	What is your age?	Free text
B2	Do you have any known hearing loss	Yes/No/Unsure
B3	What sort of visual impairment do you have?	Partially Sighted/Blind with residual vision/Blind/Other/None
B3.1	If you selected Other, or wish to provide more details, please specify	Free text
B4	How familiar are you with Extended Reality? This includes Virtual Reality, Augmented Reality and Mixed Reality.	I've never heard of it before/I'm familiar with it, but have never experienced it myself/I'm familiar with it and have used it a couple of times/I'm familiar with it and use it regularly
B4.1 was only shown if participants answered B4 saying that they had used XR.		
B4.1	What sort of Extended Reality have you experienced?	Free text
B5ˆ	Have you heard 'The Turning Forest' prior to commencing this survey?	Yes, once/Yes, more than once/No
B6*	Did you take part in our first study, in which you used and provided feedback on the previous version of the application?	Yes/No
B7*	Did you take part in our second study, in which you used and provided feedback on the 3 existing AR applications?	Yes/No

Perception of XR

Table 3.5 shows questions that were used to establish how accessible participants perceived XR to be. In all studies, these were asked before they had used the apps/AR testing app. The aim was to identify what preconceptions VI individuals hold about XR. It is designed to give insight into how they approached the studies. Question XR3 was only used in Study Three, to determine whether the redeveloped AR testing app had altered their perceptions of the accessibility of XR.

Table 3.5: Survey items addressing participants perceptions of XR. XR1-2 were used in all studies and XR3 was only used in Study Three (denoted by a *).

ID	Questions about previous experience with XR	Type of response
XR1	Do you perceive Extended Reality as being easy for you to access? Please provide a ranking between 1 and 5, where 1 is easy to access and 5 is difficult for you to access.	$1 \rightarrow 5$
XR2	Please elaborate on your answer	Free text
XR3*	Having used our app, has your perception of how accessible Extended Reality is for you changed? You can select multiple.	Easier to understand/More enjoyable/Did not change/Less Enjoyable/Harder to understand/Not sure

Accessibility Features Used

In Study One only, questions were asked about what type of accessibility features participants would normally use on their mobile smartphones. These are shown in Table 3.6.

Table 3.6: Survey items for Study One addressing accessibility features required by participants, questions AC1-2.

ID	Accessibility features question	Type of response
AC1	What, if any, accessibility features would you normally use on your mobile smartphone? You can choose multiple.	Screen Reader facility such Talkback (Android) or VoiceOver (iOS)/Braille reader facility, such as BrailleBack (Android) or VoiceOver Braille option (iOS)/Screen magnification facility, such as Magnification (Android) or Zoom (iOS)/Subtitles or Captioning/Display Adjustments such as High Contrast Text, Larger text or Inverted Colours/Other
AC2	If you selected Other, or wish to provide more details, please specify	Free text

3.5.4 Experience of the Apps

Table 3.7 shows questions asked to understand participant's experience of the AR apps. They repeated these questions three times, once for each of the apps used. Questions A1-3 ascertain details of how participants have used the app, ensuring that they used headphones to fully enjoy audio features and spent a reasonable length of time using it. Questions A4-6 establish participant's enjoyment of the app. Questions A7 and A8 then allow participants to discuss the accessibility of the app, using a 1-5 Likert scale to determine how easy it was for them to use, with free text allowing participants to provide reasons that they may have found it difficult. For comparison, questions identical to A1-A8 were used in Studies Two and Three (see Tables 4.4 and 6.1).

For study one, questions A9 and A10 focused specifically on what standard accessibility features would have been helpful for VI participants to use the apps. The options given match up with the work already presented assessing the accessibility of currently available AR apps. Feedback from VI users themselves helps to pinpoint exactly which features are necessary for them.

3.5.5 App Sounds

Table 3.8 shows questions which asked participants more specifically about the sounds in the app. As discussed in Chapter 2, soundscapes in AR can convey both narrative and navigational information, with the proposed solutions seeking to do so in a way that enhances the accessibility of the medium. These questions thus ask the participants to consider separately 'sounds that provide you with cues for navigating the app' and 'sounds that provide you with cues to help you with understand the narrative, or story, of the app'. To ensure participants understood what was being asked, clarifying statements were used. These are shown in italics in Table 3.8. After that point, the questions are identical except for the swapping of the phrases 'navigating' or 'understanding the narrative'.

Questions NV1-3/ST1-3 established whether sounds in the app were helpful and allowed participants to pick which categories of sound were helpful for them. Question NV4/ST4 give participants the chance to give their feedback on whether additional sound would help them, which gives a basic idea of whether the strategies proposed may be welcomed by VI users. Questions NV5-6/ST5-6 then asked which categories of additional sounds could be used to help with navigating/understanding the narrative. These help us to assess what specific parts of the strategies proposed may be most helpful to VI users.

In Study Two, the same questions are asked, but about the bespoke AR testing app. This allows for assessment of whether currently helpful and potentially helpful sounds match up between existing AR apps and the app used for testing the research questions of this work. In Study Three similar questions are used, but only those asking which sounds were helpful, to assess the success of the updated AR testing app. There are also additional questions to assess the additional audio features that were added to the app (see Table 6.2).

Table 3.7: Survey items in Study One addressing key concepts about the users experience of each app including enjoyment, ease of use and accessibility features needed. These questions were repeated for each app.

ID	Questions about app experience	Type of response
The following questions are about your experience whilst using (app name).		
A1	Have you downloaded and installed the app?	Yes/No
A2	How long did you spend using this app?	Free text
A3	What sort of headphones did you use for listening to the app? Please provide make and model.	Free text
A4	Did you enjoy using the AR app?	Yes/No/Unsure
A5	What did you enjoy?	Free text
A6	What didn't you enjoy	Free text
A7	Was the AR app easy for you to use? Please provide a ranking between 1 and 5, where 1 is easy and 5 is difficult.	$1 \rightarrow 5$
A8	What difficulties did you encounter while using the app?	Free text
A9	Would any specific accessibility features have helped you to use the app?	None/Screen Reader compatibility/Braille reader compatibility/Screen magnification compatibility/Subtitles or Captioning compatibility/Display Adjustment compatibility (such as High Contrast Text, Larger text or Inverted Colours)/Other
A10	Do you have any further suggestions about how the app could be made easier for you to use?	Free text

Table 3.8: Survey items in Study One addressing how helpful sounds where for navigation and narrative comprehension and the use additional sounds to enhance the apps. These questions were repeated for each app.

ID	Questions about Navigational Sounds	Type of response
The following questions will ask you about sounds within (app name) that provide you with cues for navigating through the AR app. Navigation may include finding your way around the app, exploring AR experiences and interacting with AR objects.		
NV1	Did any sounds within the app help you with navigating?	Yes/No
NV2	Which sounds were particularly helpful for navigating? You can select multiple.	Sound Effects/Voice Over/Music/Ambient Sounds/Other
NV3	Please elaborate on your answer	Free text
NV4	Do you think that additional sounds would make navigating easier?	Yes/No
NV5	What types of additional sounds would particularly help with navigating? You can select multiple	Sound Effects/Voice Over/Music/Ambient Sounds/Other
NV6	Please elaborate on your answer	Free Text
ID	Questions about Narrative Sounds	Type of response
The following questions will ask you about sounds within (app name) that provide you with cues to help you understand the narrative, or story, that the application follows.		
ST1	Did any sounds within the app help you with understanding its narrative?	Yes/No
ST2	Which sounds were particularly helpful for understanding its narrative? You can select multiple.	Sound Effects/Voice Over/Music/Ambient Sounds/Other
ST3	Please elaborate on your answer	Free text
ST4	Do you think that additional sounds would make understanding the narrative easier?	Yes/No
ST5	What types of additional sounds would particularly help with understanding the narrative? You can select multiple	Sound Effects/Voice Over/Music/Ambient Sounds/Other
ST6	Please elaborate on your answer	Free Text

3.6 Study One Results

The following sections present the results from Study One. Appendix D gives details of how to access the spreadsheets with all participant responses. Analysis of the results was primarily undertaken in R, making use of ggplot2 for visualisations, with the RColourBrewer 'Set2' palette, which is colour-blind friendly. The reporting functionality within Qualtrics was also used, mostly for creating word clouds. For visualisations made in Qualtrics, the hexadecimal colour codes from the 'Set2' palette were copied across for consistency.

Despite recruitment efforts across a wide variety of networks (see Section 3.5.1), it was difficult to find a large number of participants for any of the studies. Therefore, the statistical power of the studies is limited. However, given the exploratory nature of this work, they can be thought of as focus-group studies which give a strong indication of real experiences and opinions of VI individuals, with some depth.

3.6.1 Demographic and Past Experience with AR

Some participants recruited took part in multiple or even all three studies. Full details of the demographic information for participants, aggregated across all studies is shown in the Appendices, in Tables B.1 and B.2.

A total of 8 participants were recruited for this Study, P1-7 and CC1. A summary of their demographic information is shown in 3.9. Their ages ranged from 20-63, with a mean of 37 (median 35.5). None reported having hearing loss. The study was weighted towards those who self-reported as being 'blind with residual vision', which is the middle category given for level of sight. The participant (P7) who reported themselves as 'other' had strabismus. Further discussion in the interviews for Study Two determined that they have residual vision and are towards the more sighted end of the participants who took part. This means that there is a fairly even spread around 'blind with residual vision' of those with less sight ('blind', 25%) and those with a greater level of sight ('partially sighted', 12.5% and 'other' 12.5%).

Table 3.9: Participant demographic information for Study One.

Question (ID in brackets)	Responses	Count	%
Visual Impairment (B3)	Partially sighted	1	12.5
	Blind with residual vision	4	50
	Blind	2	25
	Other	1	12.5
Familiarity with XR (B4)	I'm familiar with it and have used it a couple of times	3	37.5
	I'm familiar with it, but have never experienced it myself	4	50
	I've never heard of it before	1	12.5

Whilst most participants had some familiarity with XR (7, 87.5%), few had ever used it (4, 50%) had never used it before). 5 (62.5%) had either never used it or never even heard of it. For the 3 (37.5%) who had used XR before, all stated that they had used AR, with one referring specifically to Pokémon Go. One also had some experience of Virtual Reality (VR) in the form of 'Oculus headset [and] PS4 VR'. Overall, participants can be thought of as inexperienced with AR.

Analysis was done to determine if there was a significant relationship between participants' level of sight (B3) and their familiarity with XR (B4). Unfortunately, given the small sample size, the criteria for a Chi-Squared test for association are not met. Therefore Fisher's exact test of independence was used. There was not a statistically significant association between participants' level of sight (B3) and their familiarity with XR (B4) ($p = 0.743$).

Perception of XR

Figure 3.6: Participant ratings of how easy to use the AR apps were and XR as a whole. Question XR1 from Table 3.5 and A7 from Table 3.7. Includes Angry Birds AR (A7B), Pokémon Go (A7P) and WWF Free Rivers (A7W) as well as their prior perception of XR (XR1). As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

Participants were asked what their perception of XR was before using the apps (XR1). Figure 3.6 shows a visualisation of their answers alongside how easy they rated the apps to be once they had used them. They rated XR a mean of 3.63 out of 5 difficulty. However, there seemed to be a lot of uncertainty on this point, as can be visibly be seen in the box plot (Figure 3.6) and through the standard deviation of 1.69. However, given that the median rating is 4.5, the data is negatively skewed towards participants having a perception of XR as being difficult to use.

Further analysis was done using Kruskal-Wallis tests to determine two different relationships. Due to the small sample size, no significant difference were seen relating level of sight (B3) or XR familiarity (XR1) with difficulty ratings. It is interesting to note that of the participants who had actually used XR before, ratings of 5,4 and 2 were given.

Figure 3.7 shows a word cloud of participant responses when they were asked to explain why they rated XR this way (XR2). This was created in Qualtrics, which automatically applies lemmatization and filters out articles and other common but not meaningful words.

The 30 most used words are shown, with size indicating those that are the most common. Interpreting the words in context shows that participants think of XR as a largely visual medium, which is difficult for them to access. This is backed up by phrases such as *'it is largely visual...so I assumed that I wouldn't get much out of it'* (P2), *'I can't do it because you have to be able to see'* (P1). Further, participants discussed how it is not designed with them in mind. They cited incompatibility with accessibility solutions *'depend[s] on holding a phone in the right direction while managing mobility aids at the same time'* (P3) and *'apps are not accessible with some accessible features'* (P1), which could cause *'potential safety issues'* (P3) and them to *'fatigue quickly and/or migraines'* (P6).

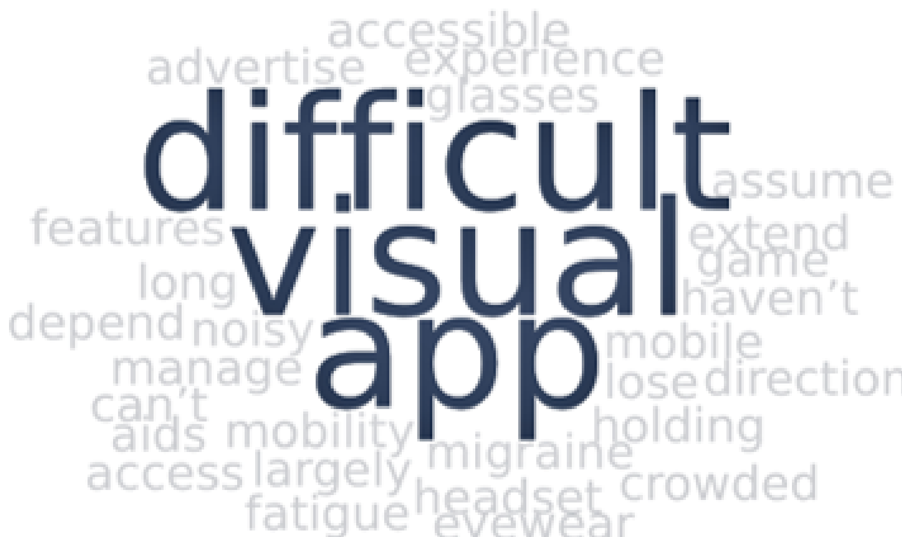


Figure 3.7: Word cloud showing participants reasoning for how easy to use they perceive XR to be, Question XR2 from Table 3.5. Created in Qualtrics.

Accessibility Features Used

Table 3.10 shows which accessibility feature(s) participants would usually use. Screen readers are by far the most common (6, 75%), with no use of braille readers. Most of those who used screen readers didn't make use of other accessibility features. There was a clear relationship between participant's level of sight (B3) and the accessibility features they use (AC1). All those who self-reported as blind or blind with residual vision stated they would usually use screen-readers. Those with greater levels of sight (partially sighted or other) didn't make use of screen readers (2, 25%) but made use of display adjustments and subtitles/captioning. No participants use braille readers, but all other features were evenly used with 2 participants (25%) each.

Table 3.10: Accessibility features required by participants in Study One. These are questions AC1 and AC2, as shown in Table 3.6. Free text extra information from AC2 is in the 'other' column).

ID	Level of sight	None	Screen Reader	Braille reader	Screen magnification	Subtitles or Captioning	Display Adjustments	Other
P1	Blind with residual vision	-	Yes	-	Yes	-	-	Enlarged font size, bold text
P2	Blind	-	Yes	-	-	-	-	-
P3	Blind with residual vision	-	Yes	-	-	-	-	-
P4	Blind with residual vision	-	Yes	-	-	-	-	-
P5	Blind with residual vision	-	Yes	-	-	-	-	-
P6	Partially sighted	-	-	-	-	Yes	Yes	-
P7	Other	-	-	-	Yes	Yes	Yes	-
P8	Other	-	Yes	-	-	-	-	-
Count	0	6	0	2	2	2	1	
%	0	75	0	25	25	25	12.5	

3.6.2 App experience

Several questions were asked of participants to determine their experience using the apps, shown in Table 3.7. Here, the most significant results are examined in turn; whether participants enjoyed the apps (A4-6), how difficult the apps were to use (A7-8) and whether specific accessibility features would have helped (A9-10).

All participants but one wore headphones whilst using the apps (question A3). Most (4, 50%) used Apple Airpods, with the rest using a variety of other brands (AKG EO-IG955, Bose E700 and an unknown model of Beats).

The amount of time that participants spent using the apps varied widely (question A2). Angry Birds AR was used for an average of 16.2 minutes, but use ranged from 5-40 minutes, giving a large standard distribution (10.91). Pokémon Go was used for an average of 16 minutes, but use had the largest range of 2-60 minutes and thus the largest standard distribution (18.29). WWF Free Rivers was used for an average of 12.38 minutes, with a range of 4-40 minutes giving a large standard distribution (12.05). Most participants remained fairly consistent in how long they spent using the apps except for two who significantly changed their time spent for Pokémon Go. One went from 40 minutes on both other apps to 15 minutes on Pokémon Go, whilst the other spent 20 minutes on Angry Birds AR and 15 minutes on WWF Free Rivers, but 1 hour on Pokémon Go. The wide range of usage times can perhaps be explained by the difficulties participants faced either meaning that they persevered for a long time to try and get the app working for them, or they simply gave up.

App Enjoyment

The below presents participant responses to whether they enjoyed the apps (A4) and why they did or didn't enjoy the apps (A5-6). Figure 3.8 shows participants' enjoyment ratings for all three apps. It is clear that the majority of participants did not enjoy the apps, although there is some variance as to the extent to which they were enjoyed or not. Across all apps, one participant rated themselves as unsure whether they enjoyed it. This was the same participant for both Angry Birds (A4_B) and Pokémon Go (A4_P). Clarification of their ratings is found below.

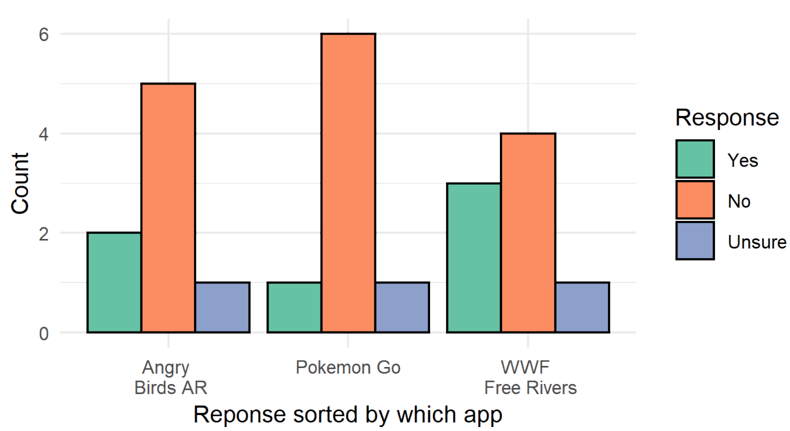


Figure 3.8: Participant ratings of whether they enjoyed using the apps. Question A4 from Table 3.7. Angry Birds AR (A4_B), Pokémon Go (A4_P) and WWF Free Rivers (A4_W).

Angry Birds AR. As with all the apps, the majority of participants (5, 62.5%) stated that they didn't enjoy this app. It is rated in the middle out of the apps (see Figure 3.8, A4_B), with both the second most participants who didn't enjoy it (5, 62.5%) and the second most participants who did enjoy it (2, 25%). Fisher's exact test of independence was used to determine that there was not a statistically significant association between a participant's level of sight (B3) and whether they enjoyed the app (A4) ($p = 0.1667$).

Based on their responses, it would seem that both those who did enjoy Angry Birds AR had previously used the original, non-AR version of the app. Further to this, one talked about using their own magnifiers and glasses to be able to overcome the lack of screen reader support. When asked what they enjoyed about the app (A5) most of those who didn't enjoy it gave no answer. However, P5 did say *'it was easier than I thought it would be'*, explaining their ranking with *'it wasn't my type of enjoyment'* when asked what they didn't enjoy (A6). The rest described the AR as making it *'more interesting'* (P6) and *'more interactive'* (P7) than the original game. When asked what they didn't enjoy about the app (A6), most of those who didn't enjoy it described it as being incompatible with screen readers. P3 described great difficulties due to VoiceOver not working, stating that for them there was *'no explanation of game play, topography of screen, what gestures did what, how to win'*, rendering the game *'unplayable'* for them. Even those who did enjoy it also described it as *'not the most accessible'* (P1) and found it *'difficult to see any of the instructions and guidance'* (P1), further describing struggles such as with how they could *'place the play environment'* (P1). The participant who was unsure about the app seemed to be it means that they *'have to hold your phone in a specific position'* (P7).

Pokémon Go. This was the least liked app, with 6 (75%) participants stating they didn't enjoy it (see Figure 3.8, A4_P). Only 1 did enjoy it (12.5%). Fisher's exact test of independence was used to determine that there was a statistically significant association between a participant's level of sight (B3) and whether they enjoyed the app (A4) ($p = 0.03571$). In this case, those who were blind or blind with residual vision all stated that they did not enjoy using the app, whereas the one participant who did was partially sighted, and the one who was unsure self-identified as other (see Section 3.6.1 for more detail on their visual impairment). When asked what they enjoyed about the app (A5), all those who didn't enjoy it gave no response, except one who did like the *'bright vivid colours'* (P1). The ones who didn't say no to A4 described it as a *'nice new way'* (P7) that *'bring[s] Pokémon to life'* (P6). When asked what they didn't enjoy about the app, the main complaint was the lack of compatibility with screen readers. It was also mentioned that the fact *'I had to go outside'* (P7) was an issue for multiple participants, with one concluding because of this that *'this is not a safe game to play'* (P1). The one participant (P1) that didn't mention screen readers did say they struggled with *'very little instructions on how to play'* and that content was not large enough on the screen for them. Even the one participant who did enjoy the app stated *'it's easier to play without AR'* (P6).

WWF Free Rivers. Of all the apps, this was the most enjoyed with 3 (37.5%) participants stating they did enjoy it (see Figure 3.8, A4_W). But it was still poorly received overall, with half (4, 50%) saying that they didn't enjoy the app. Fisher's exact test of independence was used to determine that there was not a statistically significant association between a participant's level of sight (B3) and whether they enjoyed the app (A4) ($p = 0.7429$).

When asked what they enjoyed about the app (A5), those who didn't enjoy it didn't identify anything they liked about it. Those who did enjoy it liked that *'the main points* *were read by a character/expert to explain what was happening on the map'* (P1). They also described good navigation, graphics and liking its horizontal orientation. When asked what they didn't enjoy about the app (A6), Those who didn't enjoy it all mentioned lack of compatibility with their screen reader, describing it as *'completely inaccessible'* (P2), *'frustrating'* (P4) and a *'fail'* (P3). Those who did enjoy it described buttons and details *'too small to see or appreciate'* (P6) and it being *'a bit repetitive'* (P7). The one who was unsure, despite liking the navigation said that it was simply *'not of interest for me'* (P5).

App Difficulty

The below presents participant responses to how difficult to use the apps were (A7) and what difficulties they faced when using them (A8). Figure 3.6 displays participant ratings for how difficult the apps were. Their prior perception of how difficult XR would be to use (XR1) is also included. Participants clearly found the apps difficult to use, with all rating a 4 or higher (both mean and median). The greatest agreement on the difficulty can be seen for Pokémon Go (A7P) and the most uncertainty for WWF Free Rivers (A7W). They are further discussed individually below.

Figure 3.9 shows word clouds of what difficulties participants faced (A7) in all the apps. There was some overlap here with questions A10 (what did you enjoy), A6 (what didn't you enjoy), and question A10 (any other suggestions on how could be made accessible). One participant simply referred back to question A6 when answering A7, so their responses were copied across for analysis. The word cloud was again created in Qualtrics. The 30 most used words are shown, with size indicating those that are the most common. Some other cleaning had to be done. For example, some iOS users shortened VoiceOver to 'VO' and Android users referred to 'Talkback'. In both cases, voiceover was substituted, as this allows all responses to be counted and gives the clearest representation to the reader.

(a) Angry Birds AR

(b) Pokémon Go

(c) WWF Free Rivers

Figure 3.9: Word-clouds showing the difficulties participants faced when using the apps, question A8 from Table 3.7. Responses shown for all 3 apps together. Created in Qualtrics.

Angry Birds AR. As with enjoyment, it would seem that this in the middle of the apps for how difficult it was rated to be (mean 4.13, median 4.5), and how much variance there was in the participant ratings (standard deviation 1.13, see Figure 3.6 A7B). Nobody rated it 1 (easy). Fisher's exact test of independence was used to determine that there was a statistically significant association between participants' difficulty rating (A7) of the app and whether they enjoyed it (A4) ($p = 0.0119$). Looking further into the data, all those who rated it a 5 difficulty also said they did not enjoy using it. However, the relationship isn't a clear one. Both of those who rated it a 4 difficulty stated that they did enjoy it, but those who rated it a 2 or 3 difficulty both stated they were unsure or didn't enjoy it.

A Kruskal-Wallis test determined that there were no significant differences (chi-squared = 4.0873, df = 3, p-value = 0.2522) between participants of different levels of sight's (B3) rating of how difficult they found the app to use (A7).

From the word cloud (Figure 3.9a) it is clear that the biggest difficulty participants faced (A8) was not being able to access screen reader functionality. This was best summed up by one participant who said that without it, it is *'impossible to read any parts of the app, know what you're doing, check settings of the app or simply navigate around the app'* (P4). For those that did enable direct touch with VoiceOver on iOS or turn off Talkback on Android, they describe *'randomly tapping the screen'* (P2) and making *'no discernible progress'* (P2), with one saying they *'eventually discovered some haptic feedback that did something; but I know not what. No indication by voice of buttons, options or actions'* (P3). Those who didn't talk about screen readers talked about issues with placing the environment, depth perception and trying to *'juggle lots of equipment'* (P1) whilst holding the phone in the correct place to make it accessible to them. They also described an overall lack of understanding of what was going on because there weren't enough instructions for *'a novice'* (P5).

Pokémon Go. As well as being the least enjoyed app, this was the app that participants found the most difficult (mean 4.5, median 5). Participants were fairly unanimous in this, with nobody rating it 1 or 2 and a relatively low standard deviation of 0.93 (see Figure 3.6 A7P). Fisher's exact test of independence was used to further determine that there was a statistically significant association between participants' difficulty rating (A7) of the app and whether they enjoyed it (A4) (p = 0.0357). Looking further into the data, all those who rated it a 5 difficulty also said they did not enjoy using it. However, out of the remaining participants who rated it a 3, half stated they enjoyed it and half were unsure.

A Kruskal-Wallis test determined that there were no significant differences (chi-squared = 7, df = 3, p-value = 0.0719) between participants of different levels of sight's (B3) rating of how difficult they found the app to use (A7).

From the word cloud (Figure 3.9b) it is once again possible to see that lack of compatibility with screen readers was again the greatest cause of difficulties (A8). This made it so that nothing on the screen was accessible, and if they did enable direct touch on VoiceOver for iOS or disable Talkback on Android, *'randomly tapping gave no positive results'* (P2). Those who didn't talk about screen readers had the most other points for this app, mentioning factors such as *'lots of visual clutter'* (P7), *'the writing was too small to see'* (P1), that *'sounds made navigation difficult'* (P5) and that *'trying to throw pokeballs accurately is more difficult with the AR as it makes it more wobbly'* (P6).

WWF Free Rivers. As well as being the most enjoyed app, this was also rated the least difficult to use (mean 4, median 4.5), although again nobody rated it 1. Though ease of use was far from unanimous with the highest standard deviation, 1.20 (see Figure 3.6 A7W). Despite this, we can firmly say that it was not an easy app for participants to use. A Fisher's exact test of independence was used to shows that there was a statistically significant association between participants' difficulty rating (A7) of the app and whether they enjoyed it (A4) (p = 0.0143). All those who rated it a 5 difficulty also said they did not enjoy using it. The only participants who stated they enjoyed it rated it a 2 or 3 difficulty.

A Kruskal-Wallis test determined that there were no significant differences (chi-squared = 3.9315, df = 3, p-value = 0.269) between participants of different levels of sight's (B3) rating of how difficult they found the app to use (A7).

Once again, lack of compatibility with screen readers was the biggest difficulty faced (A8), with users describing *'randomly tapping'* (P2) and *'even if your touches were activating anything you have no idea what that response could be'* (P4). Those who didn't mention screen readers talked about it *'being confusing'* (P5) and difficulties with it being too small to see *'some details'* (P6) and *'to read'* (P1).

Requested App Accessibility Features

The accessibility features participants stated they would want to help them use the apps are shown in Table 3.11. Figure 3.10 provides a visualisation for this data and compares it with the features participants would usually use (see Table 3.10).

Table 3.11: Participant responses to Question A9 (Table 3.7), which accessibility features would makes the apps easier to use in Study One.

App	Screen Reader	Braille reader	Screen magnification	Subtitles or Captioning	Display Adjustments	Other
Angry Birds AR	6	0	2	0	3	1
Pokémon Go	6	0	0	0	3	0
WWF Free River	7	1	2	0	3	0

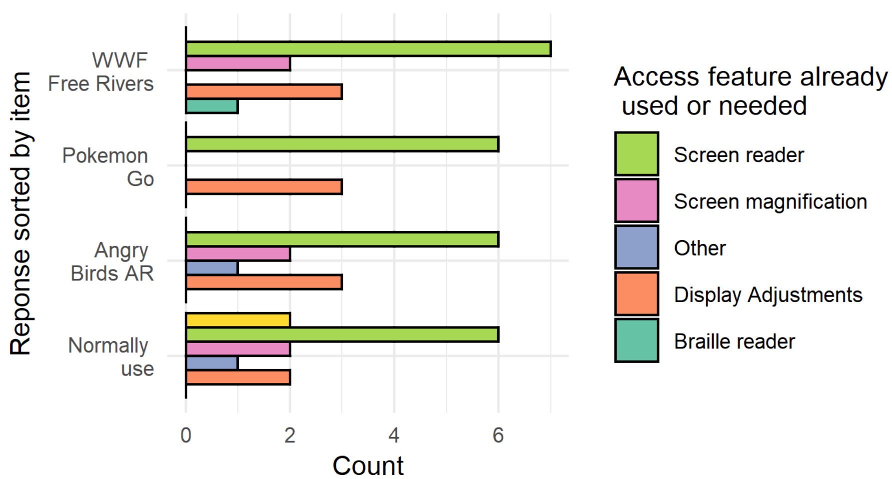


Figure 3.10: Participant ratings of which access features would be helpful, question AC1 from Table 3.6 and A9 from Table 3.7. Shows ones they use already (AC_1), Angry Birds AR (A9_B), Pokémon Go (A9_P) and WWF Free Rivers (A9_W).

Screen reader compatibility was once more the most suggested feature for all apps, wanted by 6 (75%) to 7 (87.5%) of participants. Screen magnification and display adjustments were both suggested by multiple participants. The one suggestion of an 'other' feature for Angry Birds AR was followed by a description of full support for a screen reader, although they did suggest 'auto ducking of game volume when voiceover is talking'.

The choices that people gave stayed consistent with the features they said they usually use (Table 3.10). However, those who used subtitles and screen magnification didn't ask for those features across the apps used at all. As the features participants usually used were consistent with their level of sight (see Section 3.6.1) and their choice of features that would be helpful in the apps were consistent with the features they usually use, their choice of features that would be helpful in the apps was also consistent with their level of sight.

Further Suggestions

The main result of question A10 (any further suggestions) was further repeated affirmation of the need for screen reader compatibility. For Angry Birds AR it was also suggested that *'clearer instructions'* (P5) or a *'better tutorial'* (P6) and bigger, more boldly coloured elements within the app would help. For WWF free rivers, two participants suggested a zoom feature, either *'to zoom in after the initial placement'* (P6) or to *'zoom in to certain areas to explore more'* (P1). P1 further mentioned that they thought this would *'add interest'* and *'be great for everyone not just people with vision impairments'*. However for Pokémon Go, the least enjoyed and most difficult to use app, P4 expressed their frustration by saying *'even a little accessibility support would be a start'*. All those who self-reported as blind or blind with residual vision stated that screen-readers would be helpful. Those with greater levels of sight (partially sighted or other) didn't ask for screen readers (with one exception for WWF Free Rivers) but requested display adjustments.

3.6.3 App Sounds

Several questions were asked to determine relationships between sounds and use of AR apps, shown in Table 3.8. Participants were asked if sounds were helpful overall and which categories of sounds were helpful (NV1-3/ST1-3). They were then asked if additional sounds would help and which categories of additional sounds would help (NV4-6/ST4-6). The results are divided into these two areas. These questions were repeated for both navigational sounds (NVx) and sounds for understanding the narrative of the app (STx).

Did sounds help?

The below presents participant responses to whether sounds in the apps were helpful (NV1/ST1) and which sounds in the apps were helpful (NV2-3/ST2-3). Figure 3.11 shows counts of responses for whether participants thought sounds were helpful. Overall, it would seem that sounds were not always helpful. This is the clearest for Pokémon Go (NV1_P, ST1_P). However, participants were split as to whether sounds were helpful for WWF Free Rivers (NV1_W, ST1_W). and whether they helped with understanding the narrative of Angry Birds AR (ST1_B).

For each app, Fisher's exact test of independence was used to determine several associations. First, between level of sight (B3) and whether participants thought sounds helped with navigating (NV1) or whether they thought sounds helped with understanding the story (ST1). Second, between whether participants enjoyed the app (A4) and whether they thought sounds helped with navigating (NV1) or whether they thought sounds helped with understanding the story (ST1). Results are only reported where significant associations were found.

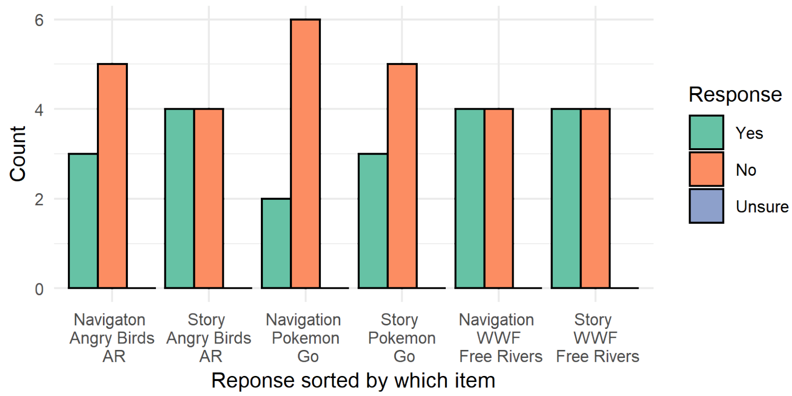


Figure 3.11: Participant ratings of whether sounds were helpful in the apps, questions NV1 and ST1 from Table 3.8. Each app has 2 categories of responses, for whether sounds helped them to navigate (NV1) or to understand the narrative (ST1). Ordered by app, then in each category; Angry Birds AR (NV1_B, ST1_B), Pokémon Go (NV1_P, ST1_P) and WWF Free Rivers (NV1_W, ST1_W).

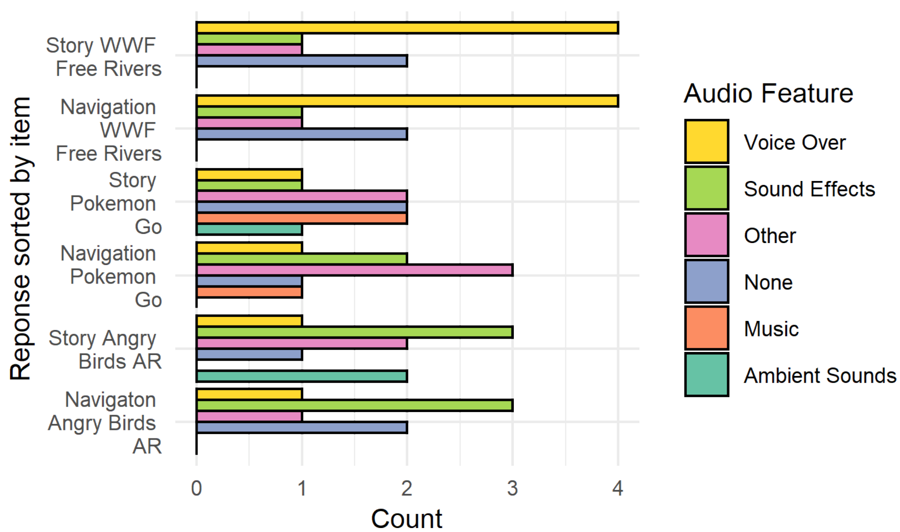


Figure 3.12: Participant ratings of which sounds were helpful in the apps, questions NV2 and ST2 from Table 3.8. Each app has 2 categories of responses, for whether they thought additional sounds would help them to navigate (NV2) or to understand the narrative (ST2). Ordered by app, then in each category; Angry Birds AR (NV2_B, ST2_B), Pokémon Go (NV2_P, ST2_P) and WWF Free Rivers (NV2_W, ST2_W).

Figure 3.12 shows what types of sounds participants thought were helpful in the apps. Overall voice-over, sound effects and the other category seem to be the most helpful. There doesn't seem to be much difference between participants perceptions of which sounds helped with navigation or understanding the narrative. Participants didn't find music or ambient sounds helpful in most cases, except for music in understanding the story of Pokémon Go (ST2_P) and ambient sounds for understanding the story of Angry Birds AR (ST2_B). Participants who selected the other category often talked about having any other sounds in an app being pointless to them unless the app was first compatible with screen reader functionality. There did seem to be some confusion around the term 'voiceover' with some participants choosing it, but then talking about how screen reader functionality was needed, not something they had.

Angry Birds AR. The majority of participants (5, 62.5%) said that sounds didn't help them to navigate the app (NV1_B, Figure 3.11). However, participants were split in half over whether they thought sounds helped them to understand the story (ST1_B, Figure 3.11).

Sound effects were chosen as the most helpful sounds for both navigation (NV2_B, Figure 3.12) and understanding the story (NV2_B, Figure 3.12). However, several participants rated no sounds as helping navigate (2, 25%), but sounds were overall more helpful for understanding the story as only 1 participant stated no sounds helped (12.5%). Interestingly, ambient sounds were mentioned as helping understand the story by 2 participants (25%). Given the app's lack of support for screen readers, and that there was no in-built narration, it is not surprising that only 1 participant rated voice-over as helpful for either navigating or understanding the story (12.5%).

For navigating (NV3), one of those who rated sound effects as helpful said that *'the sound of the slingshot pulling back when you drag the bird back'* (P1) was the most helpful, and the other two talked about hearing noises to indicate what was going on and when buttons were pressed. Those who didn't respond or responded other/voiceover all talked about voiceover being the most necessary, as it is *'our eyes'* (P5), and that without it explaining *'what sounds meant or game play in general it was pointless'* (P3). One of the ones who selected voiceover did this as a feature they would need, not as something they did have.

For understanding the story, two participants selected sound effects and ambient sounds because *'these add to the atmosphere'* (P7) and *'help you understand what was happening'* (P6). There was further talk of needing screen reader functionality for anything else to make sense. One of the ones who selected voiceover did this as a feature they would need, not as something they did have.

Pokémon Go. Participants were clear that sounds didn't help to either navigate the app (NV1_P, Figure 3.11) or understand the story (ST1_P, Figure 3.11). This is not surprising, as it was rated the most difficult and least enjoyed app (see Section 3.6.2). Fisher's exact test of independence determined that there was a statistically significant association whether participants enjoyed the app (A4) and whether they thought sounds helped with navigating (NV1) ($p = 0.03571$).

Several participants chose other for sounds that helped them with navigation (NV2_P, Figure 3.12) and understanding the story (NV2_P, Figure 3.12). For navigation (NV3), those who selected other talked about no sounds being helpful if no voiceover to describe. For understanding the story (ST3), they used other to say nothing was helpful. Sound effects were chosen by 2 (25%) as being helpful navigating (NV2_P), with those who selected this stating they *'helped me understand where to go'* (P7) and helped with knowing buttons had been activated. Music was mentioned by 2 (25%) as helping understand the story, that it helped them to *'understand the vibe'* (P7) and made it *'more intense'* (P6). P6 stated that the music and sound effects together made it *'more immersive'*. The participant who selected voice-over for both sound purposes did this as a feature they would need, not as something they did have.

WWF Free Rivers. Participants were split in half on whether sounds helped them navigate (NV1_W, Figure 3.11) or understand the story (ST1_W, Figure 3.11). This does mean that it has the highest approval rating of the sounds, with 4 (50%) saying sound helped with either purpose. This is not surprising, as it was rated the least difficult and most enjoyed app (see Section 3.6.2). However, it's important to remember that the app was still rated difficult overall, which perhaps explains why half thought sounds were not helpful for either purpose.

Participants gave the clearest idea of a single sound category that did help with both navigation (NV2_W, Figure 3.12) and understanding the story (NV2_W, Figure 3.12). Voice-over was helpful for 4 (50%) participants for both sound purposes. Although this app is still not compatible with a screen reader, it is the only one to include some built-in audio description. For navigating, participants described it being something they were *'able to learn from'* (P1), that it was *'clear'* (P7) and how *'when the speaker stopped I knew I was able to find and press the next cue'* (P6). For understanding the story, participants talked about how *'during the tutorial [it] was great'* (P1), that it *'told the story'* (P1), it *'fully explained what was happening'* (P6) and *'made it clear what to do'* (P7). However, for understanding the story one participant did this as a feature they would need, not as something they did have, referring to screen reader functionality. Further, the participant that selected other for navigating talked about how they couldn't access any of the audio without a screen reader. 2 participants (25%) did state that none of the sound categories helped them with using the app, both of them had said that no sounds helped in response to NV1/ST1.

Additional Sounds

The below presents participant responses to whether they thought additional would be helpful in the apps (NV4/ST4) and which categories of additional sounds would be helpful (NV5-6, ST5-6). Figure 3.13 shows that across the board, participants were positive about the use of extra sounds in the apps.

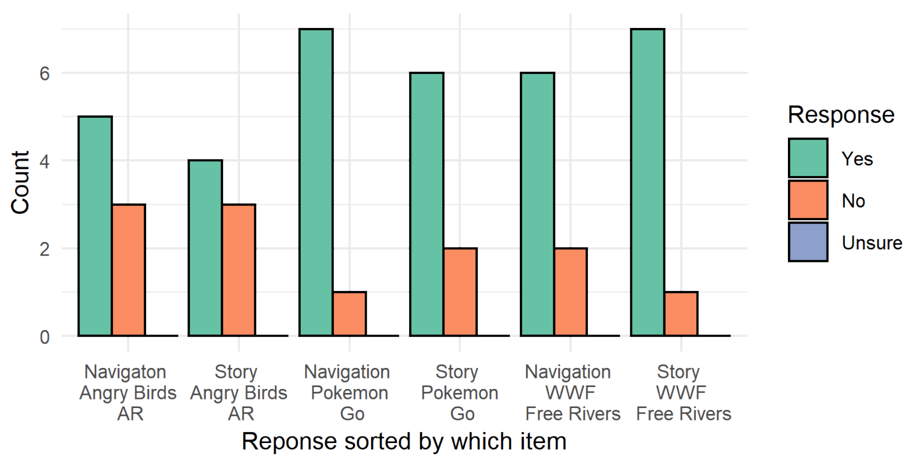


Figure 3.13: Participant ratings of whether additional sounds would be helpful in the apps, questions NV4 and ST4 from Table 3.8. Each app has 2 categories of responses, for whether they thought sounds helped them to navigate (NV4) or to understand the narrative (ST4). Ordered by app, then in each category; Angry Birds AR (NV4_B, ST4_B), Pokémon Go (NV4_P, ST4_) and WWF Free Rivers (NV4_W, ST4_W).

For each app, Fisher's exact test of independence was used to determine several associations. First, between level of sight (B3) and whether participants thought additional sounds would help with navigating (NV4) or whether participants thought additional sounds would help them to understand the story (ST4). Second, between whether participants enjoyed the app (A4) and whether participants thought additional sounds would help with navigating (NV4) or whether participants thought additional sounds would help them to understand the story (ST4). Results are only reported where significant associations were found.

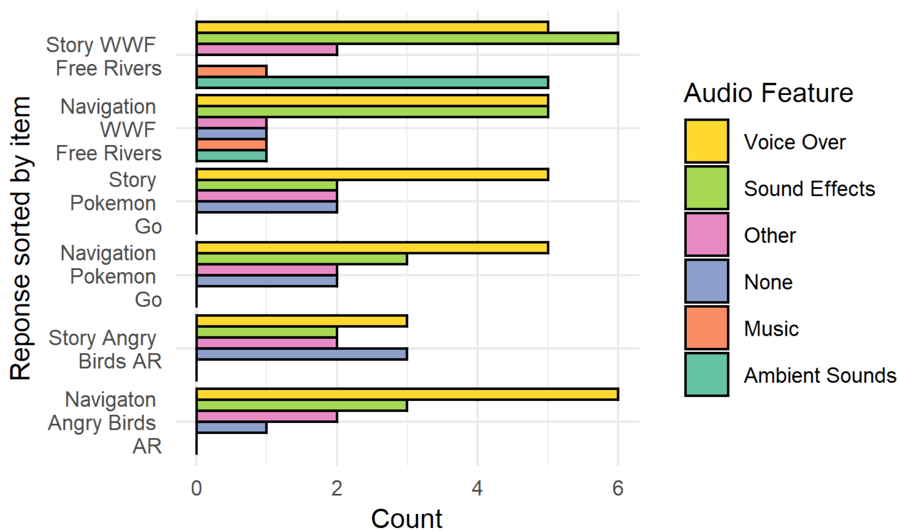


Figure 3.14: Participant ratings of which additional sounds would be helpful in the apps, questions NV5 and ST5 from Table 3.8. Each app has 2 categories of responses, for whether they thought additional sounds would help them to navigate (NV5) or to understand the narrative (ST5). Ordered by app, then in each category; Angry Birds AR (NV5_B, ST5_B), Pokémon Go (NV5_P, ST5_P) and WWF Free Rivers (NV5_W, ST5_W).

Figure 3.14 shows what sounds participants thought would be helpful if they were added (NV5, ST5). Voice-over was by far the most popular feature, followed by sound effects. Voice-over was chosen by at least 4 (50%) in all cases but one. There doesn't seem to be much difference between participants perceptions of which sounds helped with navigation or understanding the narrative. There were two notable exceptions to these preferences, that 3 (37.5%) thought no sound categories would help them understand the story of Angry Birds AR (ST5_B) and that 5 (62.5%) thought additional ambient sounds would help them to understand the narrative of WWF Free Rivers.

Angry Birds AR. Participants were generally the least sure about whether additional sounds would help with this app. 5 (62.5%) said that additional sounds would help with navigating (NV4_B, Figure 3.13) and 4 (50%) said that additional sounds would help with understanding the story (ST4_B, Figure 3.13). One participant didn't respond to ST4_B.

Voice-over was chosen by the majority of participants (6, 75%) for helping with navigating the app (NV5_B, Figure 3.14). They said it would be *'helpful for [the] level page to readout the level numbers as you click on them'* (P1), that it could give *'menu options or when clicking something brings up text'* (P6), and that using it means *'less reading and straining your eyes'* (P7). Sound effects were chosen by 3 (37.5%) for helping with the navigating, with one who chose both sound effects and voice-over could *'be used to indicate the direction that you need to go or need to do a action in that direction'* (P4). The 2 (25%) who selected other for sounds to help with navigating talked further about the use of screen readers, with one giving the further suggestion that *'without a tutorial even voiceover could not help'* (P3).

Unlike all other apps and sound purposes, voice-over was not rated highly for helping users understand the narrative of the app (ST5_B, Figure 3.14). One of those who did select voice-over stated it could give *'a hint as to how close you are on the aim'* (P1). Several participants (3, 37.7%) stated that none of the additional sound categories would help, either because felt they had *'already said this'* (P3) or because without screen reader functionality, everything else was pointless because *'I did not even know what I was supposed to do or how to do it'* (P3).

Pokémon Go. Participants were clear that additional sounds would help in this app, with 7 (87.5%) saying they would help with navigating it (NV4_P Figure 3.13).

Voice-over was chosen by 6 (75%) of participants for helping with both purposes of sound in the app (NV5_P, ST5_P 3.14). For navigating, participants focused on having screen reader functionality as it would mean *'less reading and less visual clutter'* (P7). P2, who said none for navigating stated that '*any sonic feedback would be useful'*, as they were a screen reader user who was unable to get into the app (see Section 3.6.2). P5, who selected other for additional navigation sounds stated that *'additional sounds make it too confusing'*. Participants stated that voice-over would help with understanding the story by helping them *'to understand what I needed to do'* (P1), which would also *'make it more immersive'* (P6). The participant who said that no additional sounds would be helpful for understanding the story said there *'would be too much noise'* (P7).

WWF Free Rivers. Participants were clear that additional sounds would help in this app, with 7 (87.5%) saying they would help with understanding the story (ST4_W Figure 3.13). This is interesting as this app was rated as having the most helpful sounds already in the app.

Voice-over was again strongly requested, by 6 (75%) of participants for both sound purposes (NV5_W, ST5_W Figure 3.14). Interestingly, for understanding the story ambient sounds were ranked just as highly and sound effects higher (7, 87.5%). For navigating, a number of participants chose both sound voice-over and sound effects, stating they could provide *'sounds for the progression buttons'* (P1) and that *'together with accessibility support so you know the actions you need to take to interact with it, audio cues could then be used for direction finding and directional interaction*' (P4). The participant who selected other for navigating said that *'anything else would confuse things'* (P5). Participants had a number of suggestions for how sound effects and ambient sounds could help tell the story of the app such as *'river sounds to make it more immersive and have sounds related to what was being said'* (P6), *'in the animal layer world map part specifically'* (P7) and *'include building sounds when building the roads or dams, a crash noise when destroying or removing the road and dam, more sounds on the raft section of the tutorial as this was impossible to know what was happening'* (P1). As with navigation, one participant selected both sound effects and voice-over stating that *'with accessibility support together with voice guidance you would know what you need to do and how you can interact excessively, then with added audio Fx you could know which direction you proceed or which direction you need to interact'* (P4).

3.7 Discussion of Study One

Study One validates the results of the assessment of current mobile AR apps (see Section 3.2). Significant associations were found between difficulty using the apps and enjoyment levels, meaning that the difficulties VI participants experienced made them unenjoyable. This matched up with participants' prior perceptions of XR as a difficult to use, visual medium which is incompatible with their accessibility needs and solutions. Interestingly, there seems to be no relationship between a participant's level of sight and how difficult and enjoyable they found the apps to use. Study One therefore furthers the argument that the needs, experiences and views of VI individuals are highly diverse. Even two users of the same level of sight may have totally different difficulties with an AR app, further demonstrating that accessibility solutions should be personalisable.

Throughout all the feedback given, the importance of screen readers was highlighted. Particularly for those who are blind or blind with residual vision, who may not even be able to get into apps without screen reader support, or will just end up *'randomly tapping'* the screen. The need for spoken cues explains why WWF Free Rivers was rated the easiest to use and most enjoyable app. Although it does not have screen reader support, it does feature some spoken narration which guides the user both through the story and using the app. However, participants demonstrated how they can make the most of what is there, finding help in other categories of sounds (particularly sound effects).

Participants were positive about the use of additional sounds to help make AR apps more accessible to them. Voice-over was the most requested feature, which most participants seemed to take as meaning screen reader functionality. It was repeatedly stated that without such support, any other sounds would be pointless. Sound effects followed voice-over as the second most requested additional sound. There was some interest in the use of other categories, such as music or ambient sounds.

Finally, it is interesting that in their choices of both what existing sounds were helpful and what additional sounds could be helpful, participants made little distinction between how sounds could help with navigation or understanding the story. Again, this can perhaps be explained by the fact that without screen reader compatibility, all other use of the apps was impossible for many participants. As P2 stated, *'any sonic feedback would be useful'*.

3.8 Chapter Summary

This chapter has addressed the second objective of this work, establishing that currently available mobile AR apps are not accessible for VI users. They are neither compatible with standard accessibility features nor do they make use of enhanced audio features.

First, in Sections 3.2 - 3.3, 37 current mobile AR apps were assessed against standard accessibility and enhanced audio features, with the following results:

- The apps all supported OS built-in magnification and colour changing settings, but importantly, had little support for OS built-in audio description (screen readers) or subtitles.
- Alternate screen reader solutions specific to the apps did not go far enough to provide meaningful support.
- Most apps had some support for text-based instructions or character speech as text.
- Spatial audio is an under-utilised tool (1 app used it) in current mobile AR apps.

- Sound effects were not used by most apps. Those that did use them, didn't use them extensively enough to convey actions or help users with navigation.
- The almost non-existent use of spatial audio and minimal use of sound effects demonstrates that creative use of audio, both in general and as an accessibility tool is simply not a consideration in most mobile AR app development.
- Many of the apps that had audio did allow users some level of personalisation. However, this was mostly very limited and not designed with accessibility in mind.

Second, in Sections 3.4 - 3.7, Study One further validated the inaccessibility of mobile AR apps by asking the opinions of VI users themselves. This gave the following results:

- VI individuals see XR as a difficult, visual medium that is not designed with their accessibility needs in mind.
- Participants found the apps difficult to use, which correlated with them being unenjoyable.
- Audio description (in the form of screen reader compatibility) is the most important feature used by the VI participants and therefore needed to make AR accessible to VI users.
- Participants were positive about the use of additional sounds to help make AR apps more accessible to them. Spoken cues (in the form of screen reader compatibility) are most needed, but participants were optimistic about use of other sounds, in particular sound effects.
- There is little distinction between whether categories of sounds are helpful for navigation or understanding the story. This may be because of their inter-related nature in an AR app, meaning that helping with one helps with the other. However, participants made it very clear that without an app being accessible to them, they are unable to understand either of these purposes of soundscapes in AR.

Chapter 4

Study Two - Prototyping an AR Testing App

4.1 Introduction

The remaining chapters present the work done to meet the fourth objective of this work, which is to assess the efficacy of the enhanced and personalisable audio accessibility strategies proposed (Section 1.2.2). Meeting the first two objectives has provided a solid foundation that demonstrates the gap between standards and state-of-the-art solutions and the actual accessibility of current mobile Augmented Reality (AR) to Visually Impaired (VI) users. The remaining work seeks to bridge that gap, testing the central hypothesis of this work ((Section 1.2.2) and therefore the key research questions (Section 2.7.3) which all revolve around it.

To do so required a mobile AR application (app) which could be a testing platform. Developing this is the third objective of this work. This chapter first presents the AR testing app created. The accessibility of this app is then validated through Study Two. As well as taking into account the opinions of VI participants, as Study One did (Chapter 3), this study also involved Extended Reality (XR) content creators. As with Study One, the accessibility of the app, user experience and sounds within it which have, or could, be helpful are assessed. However, further questions are asked about what features would make the app more accessible. Participants were also asked to perform a ranking task on the sounds within the app so that they could be coded in a way similar to Ward's narrative importance work [32], as described in Section 2.4.4. Content creators were asked further questions about both their current workflows and what extra tools they may need to use audio as an accessibility tool. As well as being asked survey questions, both groups of participants were also interviewed to allow greater freedom to express their opinions and suggest solutions in a less structured format.

4.2 The AR Testing App

AR content with a narrative is necessary in this first step to establishing the validity of practices previously used with narrative content in traditional media (see Section 2.7.4). As a project partner, BBC R&D provided assets to facilitate app development from 'The Turning Forest' [164], which was previously released as an interactive Virtual Reality (VR) experience. Screenshots from this are shown in Figure 4.1. The following plot description is reproduced from [164]; *'The Turning Forest, is a fantasy scenario set in a forest. The scene opens with two children playing in an autumnal forest. After one of the children runs away, the remaining child encounters a large friendly monster. The child walks with the creature through the forest, and rides on the creature's back as it swims across a river. When the creature and the child reach the other side of the river, the season has changed from autumn to winter.'* Both audio assets, in the form of ADM files [69, 165] and visual assets were provided.[1]

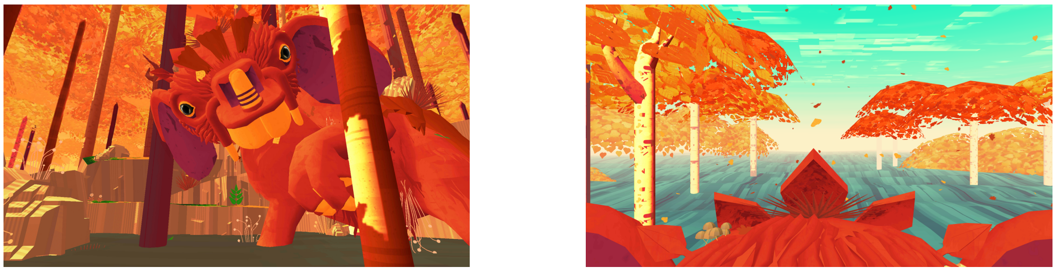


Figure 4.1: Game play screenshots of the original Turning Forest VR experience. Taken from [166].

App development tools

Unity [167] was used to develop the app, using a combination of the Unity 3D editor and C# scripting. Apps made with Unity do not easily integrate with OS screen readers on iOS or Android (Section 2.6.5). However, across the various platforms of PC, Console and Mobile, more than 50% of games are made on Unity, with 5 billion app downloads a month and 2.5 billion regular active users a month [168]. What's more, 60% of AR and VR content is made in Unity [14]. Therefore working with Unity has the potential to allow for maximum reach of any solutions developed to other AR content. Details of how to find screenshots of the Unity development of the app can be found in Appendix G.

Unity's AR Foundation framework was used to provide all AR functionality to the app. The framework supports AR features (such as plane detection, point clouds, environment probes, etc...) for multiple platforms, including Apple's ARKit, Google's ARCore, Magic Leap and Hololens. This meant that the Unity project of the app could be built for both Android and iOS smartphones, meaning participants using either OS could be recruited for studies.

App Structure

Unlike the original VR version, the AR app was split into several 'Scenes', which are outlined in Section 4.2.1. There were several reasons for this. First, different interactions were expected of users in each scene, so this allows feedback to be correlated with Herskovitz's building-block tasks for interacting with AR [18] (see Section 2.6.3). Second, splitting it up allows users to more easily talk about what they did/didn't understand or have trouble navigating in each scene. Third, splitting it up allowed for Scenes 3 and 4 to not be released in the initial version for Study Two, so that participants who took part in Studies Two and

[1] The Audio Definition Model (ADM) is an object-based audio delivery format, providing metadata to describe audio objects [165].

Three had some content they were unfamiliar with. Finally, by splitting the assets between scenes, the whole experience ran more smoothly.

Audio Assets

The original piece contains a complex and interesting audio world, with layers including narration, sound effects and ambient sounds. These were kept as close as possible to the original. However, in this initial version of the app, spatial audio was not used to allow comparison of whether it aided accessibility when added in the final version. The audio in this version is stereo down-mixed, although objects do get louder and quieter with distance. Furthermore, although voice-over (the narrator) and sound effects existed in the piece, these were not designed specifically to aid accessibility. The narrator tells the story of the piece from the perspective of the main character. The sound effects help tell the story, for example the creatures footsteps as it approaches.

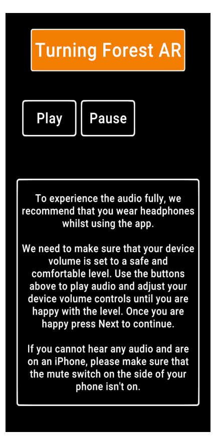
Visual Assets

In addition to the assets from the VR experience, the Low Poly Forest pack provided all the foliage used in the app [169], Simplistic Low Poly Nature provided extra animals [170] and Low Poly water was used to create animated water for Scenes 3 and 4 [171].

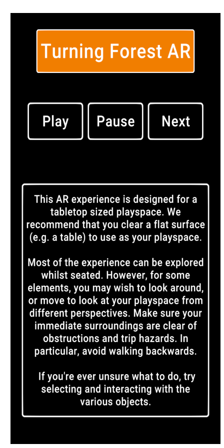
Assets had to be optimised for mobile. Especially those from the original 'Turning Forest' as they were designed for use in VR headsets. Given that participants would need to use the app on their own devices, it needed to be as backwards compatible as possible for lower specification or older devices. For visual assets, polygon counts had to be reduced, materials converted to using mobile shaders, texture sizes decreased and compressed. Animations were made more efficient by simplifying them and reducing the overall amount of animated assets. Where possible use of functions that run every frame, such as *Update()*, was avoided. When such functions were necessary, (where possible) they were made to run their functions every few updates rather than every single frame.

4.2.1 Testing App Narrative Summary

Appendix G details how to find high-quality app screenshots and a walk through video G.



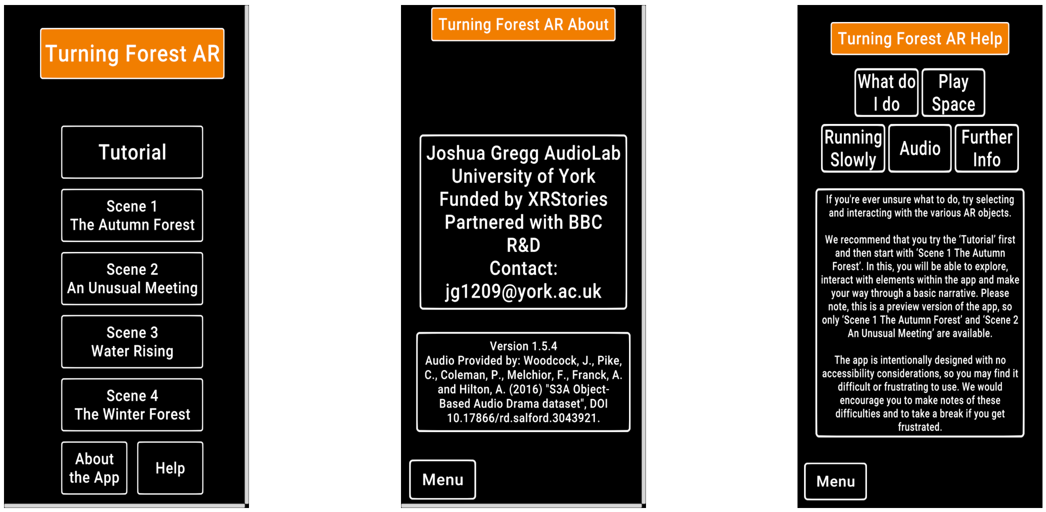
(a) Screen guiding users to set a safe audio level



(b) Screen showing further safety information

Figure 4.2: Start page screenshots from 'The Turning Forest' AR test app.

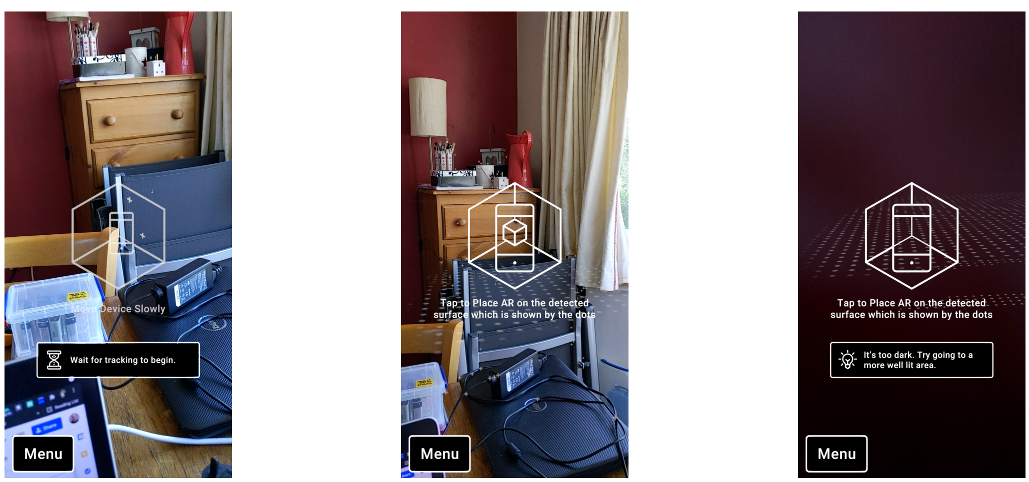
As shown in Figure 4.2, when the application is first opened, the user is guided through the process of setting their headphone volume to a safe and comfortable listening level (Figure 4.2a). Once they have pressed 'Play', further safety information is given and the 'Next' button appears, to allow them to move onto the menu (Figure 4.2b). From here, they can access the tutorial and explore the scenes of the AR experience.



(a) App main menu (b) App about screen (c) App help screen

Figure 4.3: Menu, About and Help page screenshots from 'The Turning Forest' AR test app.

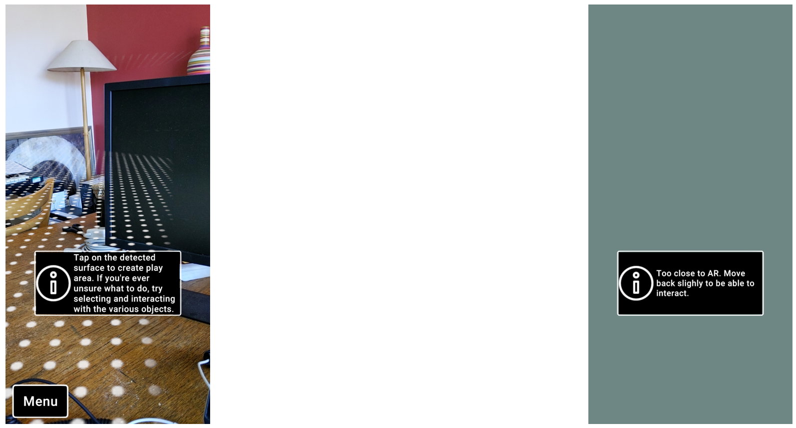
The menu (Figure 4.3a) allows the user to access the tutorial and explore the scenes of the AR experience. There are also pages with information about the app (Figure 4.3b) and help for issues users may come against (Figure 4.3c).



(a) Instruction to move device around to find surface (b) Instruction to place first AR object (c) Issue detection, not enough light shown

Figure 4.4: Tutorial screenshots from 'The Turning Forest' AR test app.

The tutorial guides the user through the basics of using AR. The screenshots below show the simplest version of this process. The user is guided to move their device around slowly, whilst surfaces are located on which AR objects can be placed (Figure 4.4a). Once surfaces have been found, they are then able to tap the screen to place an AR object (Figure 4.4b). However, if the device is having trouble finding surfaces on which AR objects can be placed, it will provide the user with suggestions. For example, “Try moving at a slower pace” or “It’s too dark. Try going to a more well lit area” (Figure 4.4c).

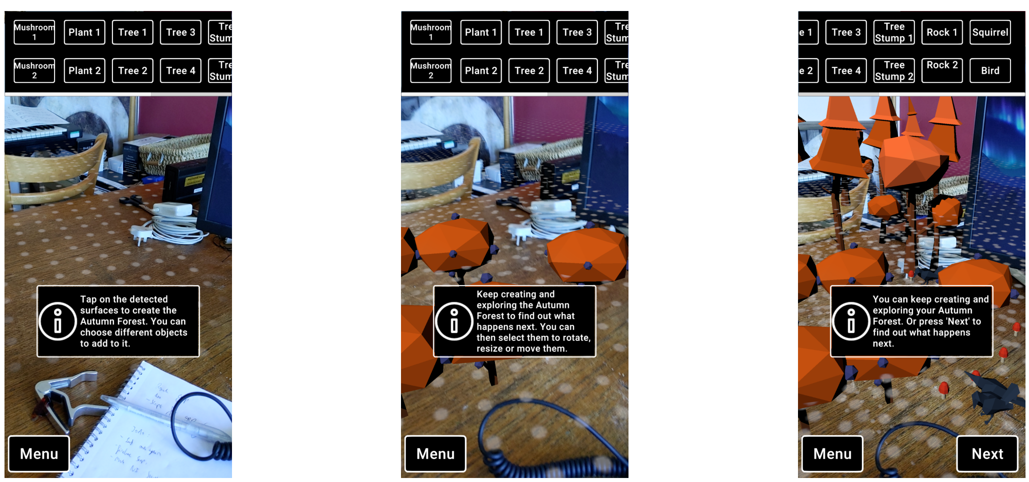


(a) Instruction for placing AR playspace

(b) Hint when a user is too close to an AR object.

Figure 4.5: Screenshots of features used in all scenes of ‘The Turning Forest’ AR test app.

The narrative unfolds as the user explores and interacts with the experience. This is complemented by various layers of audio, including narration, music, sound effects and the ambient/atmospheric sounds. All scenes except Scene 1 start by asking the player to place the AR play space (Figure 4.5b). In all scenes, if the player gets too close to an AR object, a hint will appear and the visuals will be blurred (Figure 4.5a).



(a) Instruction to place forest objects

(b) Instruction to rotate, move or resize forest objects

(c) End of scene hint

Figure 4.6: Scene 1 screenshots from ‘The Turning Forest’ AR test app.

Scene 1: The Autumn Forest opens within the magical autumn forest. The player is invited to create ‘The Autumn Forest’ in their environment, by placing objects such as trees, plants, rocks and even animals around themselves. Some of the hints that players are given as they go about exploring are shown in Figure 4.6. As the forest is built, we discover that two children are playing together there, letting their imaginations run wild. But then one day, one of the children runs away, leaving the other behind.

(a) Instruction to clear foliage (b) Instruction to touch the creature (c) Instruction to create music using the creatures teeth

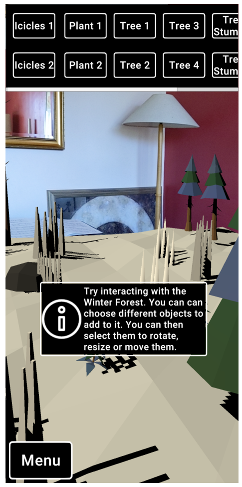
Figure 4.7: Scene 2 screenshots from 'The Turning Forest' AR test app.

Scene 2: An Unusual Meeting. The remaining child is now alone in the forest. The player clears some of the foliage, so they can see more clearly (Figure 4.7a). A huge, fantastical creature appears from nowhere, staring at the remaining child. But it turns out that the creature is friendly and lets the child touch it (Figure 4.7b). It opens its mouth to reveal teeth that can create beautiful music (Figure 4.7c).

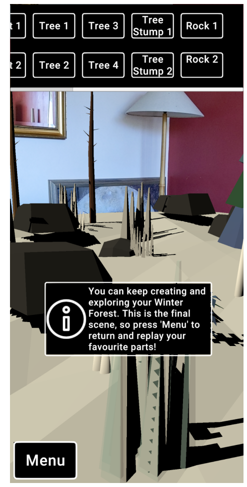
(a) Instruction to move foliage around and attract the creatures attention (b) Instruction to grab hold of the creature (c) Instruction to create music by interacting with the birds

Figure 4.8: Scene 3 screenshots from 'The Turning Forest' AR test app.

Scene 3: Water Rising. Water begins to fill the forest, worrying the child. However, the friendly creature lowers itself and allows them to ride on its back. This is achieved by first moving some foliage to attract the creatures attention (Figure 4.8a) and then grab hold of the creature (Figure 4.8b). They swim together across a river, where the player can create music by interacting with the birds and fish (Figure 4.8c).



(a) Intruction to place forest objects and interact with the autumn forest



(b) End of game hint

Figure 4.9: Scene 4 screenshots from 'The Turning Forest' AR test app.

Scene 4: The Winter Forest. The creature and child reach the other side of the river. As they do so, everything freezes over and the season changes to winter. The player can add to the 'The Winter Forest' around themselves and create music by interacting with objects. Some of the hints that players are given as they go about exploring are shown in Figure 4.9.

4.2.2 Testing App Accessibility

Table 4.1 shows the level of support the initial testing AR app has for both standard accessibility features and enhanced audio features. These can be directly compared to the apps used in Study One in Table 3.2. This list of features was also used to assess assistive mobile AR apps in Section 2.6.5 and current mainstream AR apps in Section 3.2.
In line with current AR apps in Section 3.2, the testing app has limited support for standard accessibility features. It has full support for OS built-in magnification and colour changing settings. Its audio is compatible with OS built-in subtitles and some built-in instructions are also provided. It does not support OS built-in screen reader technology. Given that the content comes with narration, it is classed as having 'some' support for audio description built into the app, as this provides contextual information that certainly helps to understand the story of the app and could be used to infer what interactions are necessary. However, it does not provide any directions on interacting with either User Interface (UI) or AR elements.

As seen in Study One (Sections 3.6-3.7) the lack of support for screen readers may prove frustrating for VI users, especially those who rely more completely on them. As with the apps in Study One, to get past initial screens users have to disable TalkBack (Android) and VoiceOver (iOS). This may again mean that some users do not get any further than the initial screens, but it reflects the current state of mobile AR apps (Section 3.2).

Table 4.1: Summary of accessibility features used in the initial test AR app

App Feature	Type	Answer
Audio Description	Specific to app	Some - narrative provides some contextual information about the AR scenes, not UI
	OS built in screen reader (TalkBack on Android/VoiceOver on iOS)	No
Subtitles	Specific to app	Some - text based instructions
	OS built in (Android/iOS Subtitles)	Yes
Magnification	Specific to app	No
	OS built in (Android Magnification/iOS Zoom)	Yes
Colour changing settings: Colour Correction, Colour Inversion	Specific to app	No
	Built into OS (Android/iOS)	Yes
Advanced Audio Feature		**Answer**
Spatial Audio (Binaural)		No
SFX to convey actions or help with navigation		Some - audio when AR objects are selected, not UI
Personalisable audio options		No

4.3 Study Two Methodology

This section outlines the design of Study Two, including what information was given to and collected from participants. Full ethical approval was obtained from the University of York Physical Sciences Ethics Committee. This can be found in Appendix A. Details of how to access the full survey instrument can be found in Appendix E.

4.4 Study Design

This study is structured similarly to Study One (3.5) with participants first using the app and then providing feedback on it. Throughout the survey and interviews 'sounds that provide you with cues for navigating' and 'sounds that provide you with cues to help you with understand the narrative, or story' are considered separately. This is to help delineate where difficulties occur and how effective solutions may be, given the previously established purpose of soundscapes in AR (see Chapter 2).

This study allows the development of the final app to be one of 'participatory' design, where the ideas of stakeholders are taken into account. This is particularly important given that the key stakeholders are VI users and the issues to be resolved concern accessibility. It would have been preferable to have several rounds of iterative feedback-development loops, so that changes made could be evaluated and improved upon. However, the time constraints of the project made this impractical.

Distributing the app

Participants were first given a Qualtrics link to an information sheet on the purpose of the study and what participation involved. The main instructions given to them on how to approach this study are shown in Table 4.2. This Qualtrics link also collected their unique identifiers, what OS their mobile smartphone uses and the email address they use on their app store. It also asked *'Do you create content for, or conduct research into media or gaming, which includes TV, Radio and Film, or Extended Reality media or gaming, which includes Virtual Reality, Augmented Reality and Mixed Reality?'.* The email address and OS details allowed for the app to be distributed to them correctly. This initial version of the app was named 'Turning Forest AR Preview' to help participants understand why Scenes 3 and 4 were not included, and that their feedback would guide the further development of the app.

For Android users, an 'Internal Testing' track on the Google Play store was used, which was managed using Google Play Console [172]. Once their email was added to the list of testers, they could be sent a link that guided them through installing the app via the Google Play store. Although this is more complicated to set up than simply sending the users an .APK file, it has the advantage that most users will already be familiar with installing apps via the Google Play Store, without having to alter additional settings. Further, due to the amount of audio content the compiled app was larger than the file limits (100MB for .APK or 150MB for .ABB files). Distribution using the Google Play store allowed easy utilisation of .OBB expansion files.

For iOS users, 'Test Flight' was used [173]. Once their emails are added to the list of testers, participants receive an automatic email with details of how to download the app. However, they were also sent an email with a joining link, so that more specific instructions on using this app for the study could be communicated.

When contacted with the download links, participants were reminded of the instructions shown in Table 4.2. For users who rely upon Screen Reader functionality, they were provided with more specific instructions so that they could get past the load up screen of the app. On Android this meant turning TalkBack off temporarily. On iOS, users had to activate 'Direct Touch'. For any user who has played games, or used apps with their own built-in screen reader these processes would likely be familiar. As not all participants may have had this kind of experience, detailed instructions were provided. Given that turning off their screen reader may be unnerving for users who rely on it, and may they may have had difficulty using the app, the following paragraph reassured them that their feedback would be fully valid and useful, even if they didn't manage to get very far:

> *The app is intentionally designed with no accessibility considerations, so you may find it difficult or frustrating to use. Specifically, the app doesn't work with TalkBack/VoiceOver, which unfortunately is the same as most Augmented Reality apps. We would encourage you*

to make notes of these difficulties and to take a break if you get frustrated. You can also get in touch with jg1209@york.ac.uk. You may not be able to use all of the app, but that doesn't mean you can't still fill in the survey. All feedback is valuable. However, if you can experience some of the AR scenes, even with sighted help, that would be amazing!

Table 4.2: Participant Instructions for Study Two and Three. These were repeated with clarifying information at relevant points. The phrase in italics was ony used in Study Two.

Instructions for Study Two
This study has five main parts which are: 1. Reading the following information 2. Downloading an application and installing it. 3. Spending some time using the app. We recommend that you try the 'Tutorial' first and then start with 'Scene 1 The Autumn Forest'. In this, you will be able to explore, interact with elements within the app and make your way through a narrative. *Please note, this is a preview version of the app, so only Scenes 1 & 2 are available.* 4. Completing a survey based on your experiences when using the app. You will have to provide consent to take part in this. The survey is fully WCAG 2.0 compliant. 5. Once you have completed the survey, you will be invited to attend an informal interview to further discuss your experience and suggestions. You will have to provide consent to take part in this. It is anticipated that the whole process will take approximately 2 hours. If you do not wish to participate in an interview/study group, you can use the app and then complete the survey, which would take approximately 1 hour.

Survey

Two separate Qualtrics surveys were created for content creators and VI participants. Participants were sent the relevant one depending on their answers to the initial joining Qualtrics link. Sections 4.4.1 to 4.4.5 detail the questions asked. Many of the questions are similar to those used in Study One, except those in Sections 4.4.4 and 4.4.5. Content creators answered all the same questions as VI participants, but with additional questions pertaining to their professional experience, shown in Section 4.4.1 and questions about how they might integrate personalisable accessibility strategies into their workflows, shown in Section 4.4.5.

Interviews

Once participants had filled out the survey, they were sent a final email inviting them to provide further feedback in an informal interview. A final Qualtrics link was used to obtain consent for data collected in the interviews and find what times participants were available to interview. Details on the format of the interviews can be found in Section 4.4.6.

4.4.1 Demographic and Past Experience with AR

As with Study One, questions B1-B4.1, shown in Table 3.4, were asked to establish some basic information on the demographic of all participants including their age, hearing loss, visual impairment (if applicable for content creators) and if they had any previous experience of XR. Question B5 was also used to assess whether participants had any previous familiarity with 'The Turning Forest' as it might bias their experience.

Questions XR1 and XR2, shown in Table 3.5 were also used to determine participants perceptions of XR. Question XR1 was altered for content creators to ask 'Do you perceive Extended Reality as being easy for visually impaired users to access?'.

Additional questions for Content Creators

Table 4.3 shows the additional demographic questions that content creators were asked. Except for PRO6-8, these questions were also used in Study Three (Chapter 6). These questions were adapted from those used by Ward in their work with content creators when investigating the Narrative Importance concept[32]. As a wide variety of content creators were invited to take part in the study, these questions helped pinpoint exactly how relevant their feedback is to either the medium of AR, work with narrative content or accessibility work.

Questions PRO1-5 established more general information about their professional experience including what medium and genre they usually work with, what organisation they work in and how long they have worked in content creation. Participants were given lists of options, with the chance to provide further information if no option corresponded closely enough to their actual background. Questions PRO8-9 further determined whether they had specific experience with either object-based or XR productions. Question PRO6-7 attempted to understand how they approach mixing and whether their approach would be compatible with an object-based personalisable accessibility approach (such as Narrative Importance). They are asked whether they consider the importance of sounds, both to the narrative and separately to navigating an experience.

Table 4.3: Survey items addressing content creators professional background PRO6-8 were not used in Study Three (denoted with a *).

ID	Questions about background	Response
PRO1	What medium do you most commonly work in?	Television/Radio/Film/Gaming/XR (including VR, AR and MR)/Other
PRO1.1	If you selected Other, please specify	Free text
PRO2	What genres(s) of content do you most commonly work in?	Documentary/Drama/Entertainment/News/Sports/Lifestyle/Comedy /Music/Action/Adventure/RPG/ Simulation/Strategy/Other
PRO2.1	See PRO1.1	Free text
PRO3	Which of the following best describes the organisation you currently work for?	Freelancer/Small or Community Broadcasting/National or Large Broadcaster/Small Independent Production house/Large Independent Production House/University/Other
PRO3.1	See PRO1.1	Free text
PRO4	Which of the following best describes the majority of the work you do?	Track Laying/Dubbing Mixer/Producer- /Teacher/Lecturer/Researcher/Student - including PhD and MSc/Assistant Producer/Program Editor/Journalist Reporter/Mastering Engineer/Sound Recordist/ Sound Mixer/Audio Designer/Audio Programmer/Audio Tech Designer/Voice Designer/Music Tech Designer/Music Implementer/Other
PRO4.1	See PRO1.1	Free text
PRO5	How many years have you worked in content creation?	Free text
PRO6*	Is the importance of a sound to navigating an experience something you consider when you mix?	Yes/No/Unsure
PRO7*	Is the importance of a sound to the narrative something you consider when you mix?	Yes/No/Unsure
PRO8*	Have you ever worked on an object-based audio production before?	Yes/No/Unsure
PRO9	Have you ever worked on an Extended Reality production before? This includes Virtual Reality, Augmented Reality and Mixed Reality.	Yes/No

Table 4.4: Survey items in Study Two addressing the participants experience of the AR app including enjoyment and ease of use. The same questions with additions are used in Study Three, shown in Table 6.1. Questions marked with a * were also used in Study One.

ID	Questions about app experience	Response
EX1*	How long did you spend using the AR app?	Free text
EX2*	What sort of headphones did you use for listening to the app? Please provide make and model.	Free text
EX3	What sort of environment did you use the AR app in? For example, outside or sat down at a table.	Free text
EX4*	Did you enjoy using the AR app?	Yes/No/Unsure
EX5*	Was the AR App difficult for you to use? Please provide a ranking between 1 and 5, where 1 is Easy and 5 is difficult.	$1 \rightarrow 5$
EX6*	What difficulties did you encounter while using the app?	Free text
EX7	Do you have any suggestions about how the app could be made easier for you to use?	Free text

4.4.2 AR App Experience

Table 4.4 shows the questions participants were asked about their experience using the AR app. These are similar to those used in Study One, as shown in Table 3.7. Questions EX1-2 ensure that they spent a reasonable length of time using the app and wore headphones. EX3 established what sort of playspace they used for the app as this may alter how they interacted with it, such as if they walked around and explored from different angles. This may also alter how useful different navigation cues were. EX4 attempts to understand participants' enjoyment of the app, using three categories (yes/no/unsure) comparable to question A4 in Study One. Questions EX5-7 began to allow participants to discuss the accessibility of the app, using a 1-5 Likert scale to determine how easy it was for them to use, with free text allowing them to provide reasons that they may have found it difficult and suggestions for how it could be made easier. Less free-text questions were used than Study One, as participants repeated similar responses when asked what they did/didn't enjoy, what difficulties they faced and how apps could be made easier. Along with some additions, these exact questions were used in Study Three. Table 6.1 shows the questions and additions.

4.4.3 AR App Sounds

Participants were asked about the sounds in the app and if additional sounds could enhance it. These were the same as those used in Study One, as shown in Table 3.8. A small change was made to the introduction of the questions for those related to the narrative (ST1-ST6). Participants were given access to a document containing a summary of the app's story, as shown in Section 4.2.1. This meant the introduction to these questions was as follows: *The following questions will ask you about sounds within the application that provide you with cues to help you understand the narrative, or story, that the application follows. For a short* *summary of the story in the scenes you experienced, click here (link to summary).*

Table 4.5: Sounds ranked in Study Two. Each sound was ranked twice, for navigational importance and for narrative importance. All sounds were from Scene 2: An Unusual Meeting.

ID	Sound
NI1	Creature Breathing
NI2	Forest Ambience
NI3	Narrator
NI4	Music
NI5	Creature Selected Sound - Roar - created by user interaction
NI6	Creature Deselected Sound - Growl - created by user interaction
NI7	Musical Teeth Sounds - created by user interaction
NI8	Creature Moving Sounds
NI9	Tree Selected Sound - created by user interaction

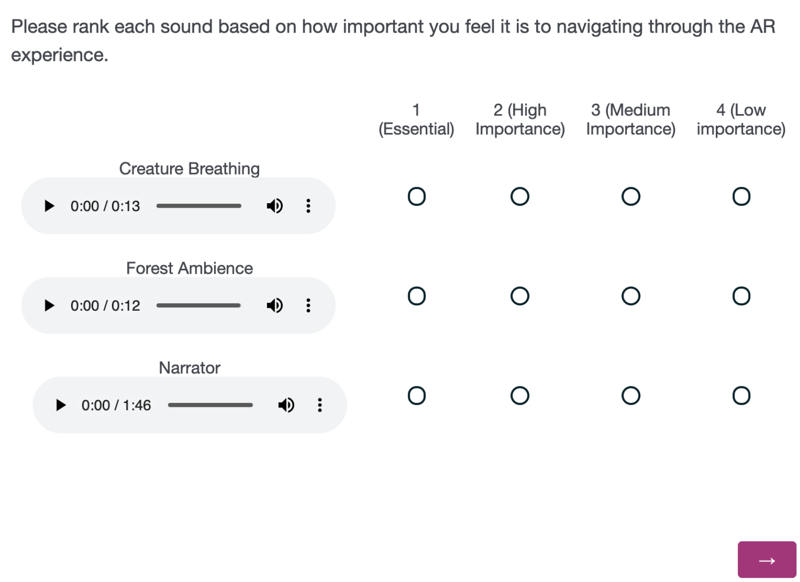


Figure 4.10: An example of the audio ranking interface used in Study Two.

4.4.4 Ranking Task

The questions asked in this section endeavour to understand how a Narrative Importance like personalisation approach could be used in an AR context. The format of this task and the questions used were adapted from work done by Ward [32]. The task involved users ranking each of the chosen sounds on a scale of 1-5, which matched with the narrative importance levels defined by Ward (see Section 2.4.4). Participants were asked to repeat the task twice, asking them to first consider how important the sounds were to navigating through the app, then consider how important they were for understanding the narrative of the app. Table

4.6 shows the instructions participants were given, including the preamble used to help them establish what criteria they were using to rank the sounds.

Table 4.6: Participant Instructions for ranking audio in Study Two.

<table>
<tr><th>Instructions for ranking audio</th></tr>
<tr><td>The following questions will ask you to categorise sounds from the AR application you just used, based on the sound's importance to navigating through the AR app. Navigation may include finding your way around the app, exploring AR experiences and interacting with AR objects. The sounds are taken from 'Scene 2 An Unusual Meeting'. You can watch this video (link to walk through video on YouTube) for a reminder of what happens. For each sound, you should select one of the following categories using the radio buttons on the right:
• Essential (1),
• high importance (2),
• medium importance (3),
• or low importance (4),
based on how important you feel it is to navigating through the AR experience.</td></tr>
<tr><td>The following questions will ask you to categorise sounds from the AR application you just used, based on the sound's importance to understanding the narrative, or story, of the AR app. The sounds are taken from 'Scene 2 An Unusual Meeting'. You can watch this video (link to walk through video on YouTube) for a reminder of what happens. For a short summary of the story in the scenes you experienced, click here (link to summary of story). For each sound, you should select one of the following categories using the radio buttons on the right:
• Essential (1),
• high importance (2),
• medium importance (3),
• or low importance (4),
based on how important you feel it is to following the content's narrative.</td></tr>
</table>

Table 4.5 shows the sounds selected for the task. All sounds were taken from 'Scene 2: An Unusual Meeting', as the sounds in that scene equate most closely to those Ward used in their Narrative Importance ranking task [32]. Figures E.1 and E.2 can be used to give an idea of the sounds Ward used.

The interface participants used is shown in Figure 4.10. They weren't provided with further information on what the categories meant or how much each sound category would be boosted or attenuated. They could preview each sound, but not how it would sound if it was set to the narrative importance level chosen.

After each time of completing the task, participants were asked questions about their rankings, shown in Table 4.7. These attempt to understand the rationale behind their rankings. Question R2 can also provide further information to help establish what sounds are and aren't important to them. Question R3 also seeks to establish whether having some control over sounds based on their importance would be useful to participants.

Table 4.7: Survey items in Study Two addressing how the difficulty of the ranking task, criteria the participants used and how helpful they thought it would be. These were repeated for navigational (coded as R$-NR) and narrative sounds (coded as R$-NAV)

ID	Questions about ranking task	Response
R1	Generally, how easy or hard did you find it to assign an importance category to each sound? Please provide a ranking between 1 and 5, where 1 is easy and 5 is difficult.	$1 \rightarrow 5$
R2	What criteria did you use to decide the importance level of each sound object?	1 Free text
R3	Do you think it would be helpful if you were able to change the levels of the sounds that helped you with *navigating the app/understanding the story*? For example, you would be able to turn up those you ranked as more important and turn down those you marked as less important. Please provide a ranking between 1 and 5, where 1 is helpful and 5 is not helpful.	$1 \rightarrow 5$

4.4.5 Workflow Questions

Content creators were asked 3 additional questions about how they would integrate personalisation into their own workflows. These are shown in Table 4.8. WF1-2 try to understand how different using a ranking based system for how important sounds are to narrative/navigation is to their work using 1-5 ratings. WF3 allows content creators a chance to suggest tools for users to personalise their audio experience in their content.

Table 4.8: Survey items in Study Two addressing workflow tools for content creators.

ID	Questions about workflow	Response
WF1	How much does the idea of ranking sounds due to their navigational importance differ from your current workflows? Please provide a ranking between 1 and 5, where 1 is not different and 5 is very different.	$1 \rightarrow 5$
WF2	How much does the idea of ranking sounds due to their importance for understanding the narrative differ from your current workflows? Please provide a ranking between 1 and 5, where 1 is not different and 5 is very different.	$1 \rightarrow 5$
WF3	What tools would you require to allow users to personalise their audio experience so that your content was more accessible to them?	Free text

4.4.6 Interview Format

All interviews took place over Zoom [174]. They were recorded using Zoom's built-in functionality. They were one on one and undertaken by the author. Participants were told to set aside an hour for the interview. Before the recording was started for each interview, participants were given some background information, assurances of their confidentiality and the chance to ask any questions. This information is shown in Table 4.9

Interviews were informal and semi-structured to encourage participants to share their own ideas and suggestions. Participants were encouraged to discuss the topics they wanted to, with their ideas mostly guiding the directions the interviews took. The survey gave in-depth insights on participant's experience of the app, and what extra sounds could be helpful. Therefore, the interviews focused more on getting higher-level feedback from participants on the concept of using audio to make AR more accessible for VI users. It was also much easier in an interview setting to work through their experience of the app step by step and try to focus on particular sticking points and what potential solutions might exist. A small number of questions were prepared in advance to keep the conversation flowing in the right direction and facilitate consistency across interviews, ensuring that important themes were covered. These are shown in Table 4.10.

Table 4.9: Introductory statement for interviews, given before recording was started.

Interviews Introduction
Thank you for joining me to talk about your experiences using Augmented Reality, and how it can be made more accessible for you. I would like to know both what you did, and didn't like and how improvements can be made. There are no wrong answers, only differing opinions. Keep in mind that I'm just as interested in negative comments as positive comments, and at times, the negative comments are the most helpful. I'll record the session because we don't want to miss any of your comments. People often say very helpful things in these discussions and I can't write fast enough to get them all down. We will be on a first-name basis today, but I won't use any names in my reports. You may be assured of complete confidentiality. Do you have any questions? I'll begin recording now.

Table 4.10: Questions used in Study Two to guide the interviews. Where there are multiple questions under the same number, those after the first are follow up questions.

Interview Questions
1. Tell me a bit about yourself, your visual impairment and how you use technology and make technology accessible for you in your everyday life? 2. Have you used AR before? 3. Do you think extended reality is something that is hard for someone with visual impairment to access? 4. What did you think of the AR application? 5. Work through Scene by Scene (a) How far did you get in this Scene? (b) What caused problems? (c) Can you think of any particular solutions to the problems you encountered? 6. Do you think using audio creatively could help it be more accessible? How would you go about doing it? *Repeat the questions, more specifically focusing on spatial audio, sound effects and a guide character.* 7. Do you think that having control over the sounds that you can hear within the app would have helped you? (a) How would you decide what was important? (b) What would be the easiest way for you to control it? (c) Do you think the number of categories that we got you to put sounds into was appropriate? 8. Do you have any other suggestions for how the app could be made easier for you to access, that you haven't been able to discuss already?

4.5 Survey Results

The following sections present the survey results from Study Two, for both VI participants and content creators. Appendix E gives details of how to access spreadsheets with full participants responses, as well as analytical work and originals image files of visualisations. Results analysis was done in R, making use of ggplot2 for visualisations, with the RColourBrewer 'Set2' palette, which is colour-blind friendly. The reporting functionality within Qualtrics was also used, mostly for creating word clouds. For visualisations made in Qualtrics, the hexadecimal colour codes from the 'Set2' palette were copied across for consistency.

Despite efforts to recruit across a wide variety of networks (see Section 3.5.1), it was only possible to find a small cohort of 6 participants. Therefore, the statistical power of this study is limited. However, it gives rich qualitative data about the real experiences and opinions of VI individuals and content creators.

4.5.1 Demographic and Past Experience with AR

Some participants took part in multiple or even all three studies. Therefore, their anonymised demographic information has been aggregated and displayed in Appendix B. Full details of VI participants can be seen in Table B.1 and full details of content creators can be seen in Table B.2.

Table 4.11: VI participant demographic information for Study Two.

Question (ID in brackets)	Responses	Count	%
Visual Impairment (B3)	Partially sighted	1	16.7
	Blind with residual vision	2	33.3
	Blind	1	16.7
	Other	2	33.3
Familiarity with XR (B4)	I'm familiar with it and have used it a couple of times	3	50
	I'm familiar with it, but have never experienced it myself	3	50

VI participants. A total of 6 VI participants took part; P1,4 & 6-9. Table 4.11, shows an overview of their demographic information. Out of these participants, 5 went also went on to take part in interviews (see Section 4.6). Their ages ranged from 19-63, with a mean of 32 (median 26.5). None reported having hearing loss (B2) and none had ever heard 'The Turning Forest' before (XR2). This study was weighted more towards those with a greater level of sight. Only 1 participant (16.7%) self-reported as fully blind. Through giving further detail in the survey (B3.1) and further discussion in the interviews, it was determined that the two who self-reported as 'other' are more towards the sighted end. One (P7) had strabismus and the other (P8) is 'short sighted in one eye and long sighted in the other, registered sight impaired in one eye and severely sight impaired in the other (I think). My eyes also shake making text and depth difficult'.

All VI participants reported being familiar with XR (B4). None had used it regularly, but half (3, 50%) had experienced it themselves a couple of times. They reported having used a variety of VR headsets and some AR games, for example, 'Oculus Quest VR, Hive VR, Pokemon Go AR, PS4 VR'. However, one of them was unsure what they had used as it was some time ago. Overall, participants can be thought of as having a beginner level of experience with AR.

Content Creators. A total of 5 content creators completed the survey for this Study; CC1-5. Table 4.12, shows an overview of their demographic information. Out of these content creators, 2 went on to take part in interviews (see Section 4.6). They had an average age of 33.6 (median 37), with a range of 23-45 years old. None reported any known hearing loss (B2). One (CC1) reported as being blind (B3). Most (60%) had heard 'The Turning Forest AR' before taking part in this study.

The content creators had quite a wide range in their years of experience (PRO5), between 0.5 - 15 years. On average, they had worked as a content creator for 4.8 years (median 3.0). This means that they can be thought of as mostly in the early stages of their content creating careers, with only one of them having more than 5 (15) years of experience (CC2). They also worked for a variety of organisations, with the largest contingent being those who worked for a national or large broadcaster (2, 40%). Those who selected 'other' worked for a 'game development studio' and a 'not for profit'. However, the majority of them mostly worked as a researcher (3, 60%), with both the others holding audio-based roles. They also worked with a variety of genres of content (PRO2), with the majority working across multiple genres. The results in Table 4.12 are aggregated so that if only one content creator worked with multiple unique genres, these are combined. The most popular were entertainment and music with 2 (40%) each. One only selected other, stating 'don't work with content, more so experiences' whilst another used other to provide more detail; 'Health-related (mainly audiology), i.e. implementation in AR/VR of clinical audiological assessments'.

All the content creators were familiar with XR. Most (4, 80%) had used it before, although only 1 (20%) uses it regularly. The only one who had not used XR was the VI content creator (CC1). The majority of the content creators (3, 60%) worked mostly with the XR medium (PRO1). However, it is interesting that out of these, only 1 (20%) uses XR regularly. Further to this, none had ever worked on an XR production before (PRO9).

The content creators had a variety of experience, with most being researchers and most working with XR. However, their minimal experience of XR and the fact they are at an early stage in their careers should be considered when interpreting their responses.

Table 4.12: Content Creator Demographic information for Study Two.

Question (ID in brackets)	Responses	Count	%
Familiarity with XR (B4)	I'm familiar with it and use it regularly	1	20
	I'm familiar with it and have used it a couple of times	3	60
	I'm familiar with it, but have never experienced it myself	1	20
Medium (PRO1)	XR (including VR, AR and MR)	3	60
	Gaming	1	20
	Television	1	20
Genre(s) (PRO2)	Action/Adventure	2	20
	Drama/News/Sports/Comedy	1	20
	Entertainment	2	40
	Music	2	40
	Simulation	1	20
	Other	2	40
Organisation (PRO3)	National or Large Broadcaster	2	40
	University	1	20
	Other	2	40
Majority of work (PRO4)	Researcher	3	60
	Audio Programmer	1	20
	Sound Recordist	1	20
Years in content creation (PRO 5)	less than 1	2	40
	1-9	2	40
	10+	1	20
Worked on an object-based audio production (PRO8)	Yes	2	40
	No, but I am familiar with the concept of object-based audio	2	40
	No, and I am unfamiliar with the concept of object-based audio	1	20
Worked on an XR production (PRO9)	No	4	80
	Did not answer	1	20
Heard The Turning Forest before (XR2)	Yes, more than once	1	20
	Yes, once	2	40
	No	2	40

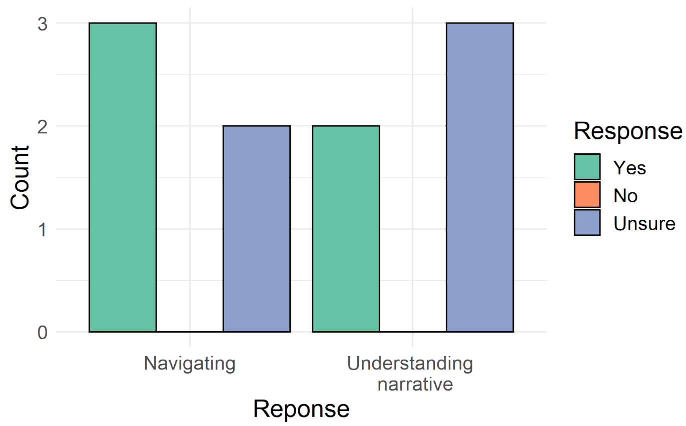
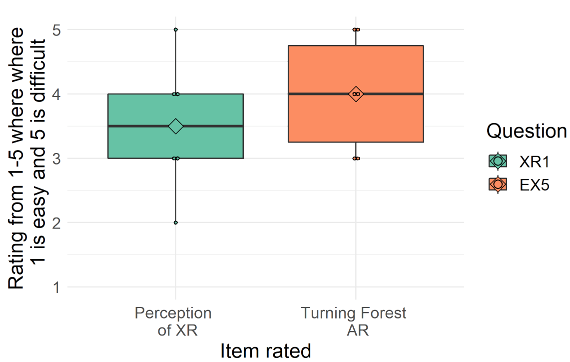


Figure 4.11: Content creator responses for whether they consider the importance of sounds when when mixing. Split into 2 categories of responses, for whether they think of the sounds importance to navigation (PRO_6) or to the narrative (PRO_7). Questions from Table 4.3.

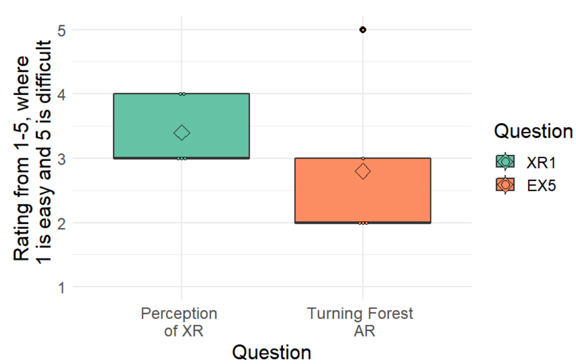
Content creators had some familiarity with object-based audio (PRO8). The majority (4, 80%) had at least heard of the concept, but only 2 (40%) had used it before - one gaming as an audio programmer and one in XR as a researcher. They were then asked whether they consider the importance of a sound to navigating or understanding the narrative when mixing. Their responses to this are shown in Figure 4.11. There was a lot of uncertainty about whether they would or not. This is not surprising given that most had never actually worked with object-based audio (3, 60%). The 2 content creators (40%) who had actually worked with object-based audio responded yes to both questions. Therefore, it suggests that once experienced with object-based audio, making decisions on sounds based on their importance to navigation or narrative does become a natural consideration.

Perception of XR

Participants were asked how difficult to use they perceived XR as before they used the app.



(a) VI Participants



(b) Content Creators

Figure 4.12: Participant ratings of how easy to use XR as a whole (XR1) and the testing AR app (EX5) are.). No content creators responded to XR1 based on how they perceived the difficulty of XR for VI users. Questions XR1 from Table 3.5 and EX5 from Table 4.4. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants. Figure 4.12a shows a visualisation of their rating alongside how difficult to use the rated the AR testing app once they had used it. They rated XR as a mean of 3.5 difficulty. There was fair agreement on this, with a median also of 3.5 and a standard deviation of 1.04. Meaning that overall they perceived XR as fairly difficult to use.

Content creators. Figure 4.12b shows a visualisation of their rating alongside how difficult to use the rated the AR testing app once they had used it. Question XR1 was slightly altered for content creators, to ask how difficult they perceived XR to be for VI individuals, rather than themselves. They gave it a similar rating to the VI participants, with a mean of 3.4 out of 5 difficulty (median 3). They were more unanimous, with a comparatively small standard deviation of 0.55 and no ratings above 4 or below 3.

As none had ever worked in XR, this will not have influenced their judgement. The VI content creator was one of the two participants who rated it most difficult (4 out of 5).

4.5.2 AR App Experience

Several questions were asked to determine participant's experience using the app, shown in Table 4.4. They are examined in turn; whether participants enjoyed the app (EX4) and how difficult the app was to use (EX5-7).

The amount of time that VI participants spent using the app varied widely (EX1). It was used for an average of 46.67 minutes (median 35), but use ranged from 10 minutes to 2 hours, giving a very large standard distribution of 40.21 minutes. The wide range of usage times can perhaps be explained by the difficulties participants faced either meaning that they persevered for a long time to try and get the app working for them, or they simply gave up. All used headphones, with the majority (3, 50%) using Apple in-ear headphones and all the rest using other in-ear models (AKG EO-IG955, Creative Aurvana Air, Bose QC Earbuds). All used the app sat down, with only 1 (16.7%) not using it at a table.

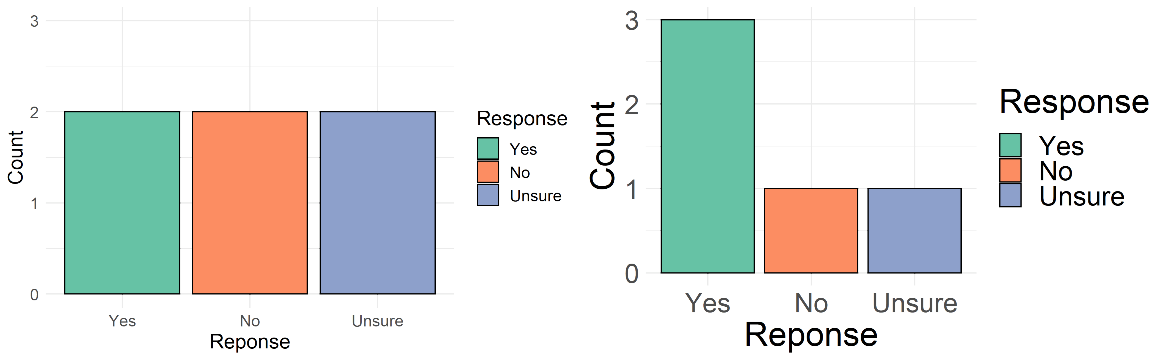
The amount of time content creators spent using the app varied less widely (EX1). They spent an average of 20 minutes (median 19.2 minutes), with a range of 8-35 minutes and a standard deviation of 10.1 minutes. Apple in-ear headphones were once again the most popular choice of headphones, with one not reporting using any and other use of over-ear headphones (Bomaker , AKG K240 Studio). They all used the app indoors, mostly (3, 60%) sat a table, with a minority (2, 40%) making use of the floor.

App Enjoyment

The below presents participant responses to whether they enjoyed the app (EX5).

VI participants. Figure 4.13a shows VI participants' enjoyment ratings for the AR testing app. It would seem that overall, VI participants were uncertain over whether they enjoyed the app. Only 2 (33.3%) stated that they actually enjoyed it, with an even split between the rest stating they didn't enjoy it or being unsure (2, 33% each). Fisher's exact test of independence determined that there were no significant associations between ratings of app enjoyment (EX4) and sight level (B3) or familiarity with XR (B4).

Content creators. Figure 4.13a shows content creators' enjoyment ratings for the AR testing app. The majority (3, 60%) stated that they did enjoy the app. The rest were evenly split between being unsure and not enjoying the app (1, 20% each). The blind content creator was the only one to state that they did not enjoy the AR testing app.



(a) VI Participants (b) Content Creators

Figure 4.13: Participant ratings of whether they enjoyed using the initial AR testing app. Question EX4 from Table 4.4.

App Difficulty

The below presents participant responses to how difficult to use the app was (EX5), what difficulties they faced when using it (EX6) and their suggestions for how it could be made easier (EX7).

VI participants. Figure 4.12a shows VI participants' ratings of how difficult they found the app to use. They found it difficult to use (EX5) and more difficult than they had previously perceived XR to be (XR1). They rated it a mean of 4 out of 5 difficulty (median also 4). There was a fair spread of ratings, with 2 each rating it a 3, 4 or 5 difficulty (standard deviation 0.89). This means that none of them rated it as easy to use (1 or 2 out of 5). The average rating of 4 out of 5 is similar to the apps investigated in Study One (see Section 3.6.2), which had means ranging 4-4.5. This supports the validity of this AR testing app as a useful platform for the studies done in this work because as a baseline with no extra accessibility features, it is as accessible as the apps in Study One.

Fisher's exact test of independence was used to determine that there was not a statistically significant association between participants' difficulty rating (EX5) of the app and whether they enjoyed it (EX4) ($p = 0.4667$). However, both of those who rated it a 5 did not enjoy the app.

Figure 4.14a shows a word cloud of participant responses when they were asked what difficulties they faced when using the app (EX6). Half of them (3, 50%) cited incompatibility with their screen-readers. All 3 were participants who self-reported as blind or blind with residual vision (see Section 4.11). This made it impossible for them to progress past the start of the app as *'all parts of the screen say "Direct touch area" when touched'* (P4) and *'Voiceover didn't read anything no matter where you tapped or swipe'* (P4). Two of them managed to enable direct touch on VoiceOver for iOS or disable Talkback on Android. P1 then used their *'strong lenses for very close up work'* but '*could[n]'t see the instructions for what I needed to do or the different objects available to place'*. P4 used sighted assistance via an AIRA [175] call, which meant they were able to locate some buttons, but all this resulted in was *'the blur on [the] screen changed to another blur'*, and *'I could tell that when I moved* *the phone the screen had shapes and shadows which changed so I know the camera was being used'*, but this was useless to them without spoken feedback. Those that didn't mention screen-readers talked about how the information displayed in the app was too small or not high enough contrast for them; *'the text did not enlarge'* (P1) so *'playing the actual scene was difficult because I couldn't see the instructions for what I needed to do or the different objects available to place'* (P1) and *'environment dots difficult to see as very faint, especially on white/lighter backgrounds'* (P6). Those who were able to see instructions thought that they weren't comprehensive enough, stating *'it was difficult to understand what to do in the scenes sometimes'* (P7) and *'instructions in the tutorial were vague and unclear'* (P6).

Figure 4.14a shows a word cloud of participant responses when they were asked what suggestions they had for improving the app (EX7). The majority (4, 66.7%) suggested compatibility with screen readers, stating that work should be done on *'incorporating the onscreen instructions into the audio?'* (P8) and that *'it should make sense when using only audio'* (P9). One suggested that this may need to work by *'somehow explain[ing] or describ[ing] the changing scenes a bit like active Alt text'* (P4). Several participants talked about visual alterations, such as *'if you use dynamic text the font will enlarge automatically'* (P1), *'the contrast on the text could've been less harsh on the eyes'* (P7) and *'more VI/Dyslexia friendly font, make text more spaced'* (P6). P6 also mentioned moving *'instructions to a less obstructive place on the screen'*. Finally, the tutorial was mentioned by 2 participants (33.3%) who asked for *'a clearer tutorial'* (P8) and that it *'should include how to manipulate objects and have a clear end'* (P6).

(a) EX6

(b) EX7

Figure 4.14: Word cloud showing the difficulties VI participants faced when using the testing AR app (EX6) and their suggestions for making it easier to use (EX7). Questions are shown in Table 4.4. Created in Qualtrics.

Content creators. Figure 4.12b shows content creators' ratings of how difficult they found the app to use. They found it easier to use than the VI participants, with a mean rating of 2.8 (median 2.0). They were mostly agreed about this with all ratings in the range of 1-2 except one outlier of 5 (standard deviation 1.3). However, they rated XR based on how difficult to use they thought it would be for VI individuals but their rating of the app was on their personal experience using it, so these two data sets are not comparable. The highest rating of 5 was given by the VI content creator.

Figure 4.15a shows a word cloud of content creator responses when they were asked what

(a) EX6 (b) EX7

Figure 4.15: Word cloud showing the difficulties content creators faced when using the testing AR app (EX6) and their suggestions for making it easier to use (EX7). Questions are shown in Table 4.4. Created in Qualtrics.

difficulties they faced when using the app (EX6). Their responses were vastly different to those given by VI participants (see Figure 4.14a). Only the VI content creator suggested screen-reader compatibility, stating *'no access at all for TalkBack'* (CC1). There was little in common across the rest of the responses. Some had difficulties with placing the AR content, such as *'struggling to detect my surface'* (CC3) and *'the animal appeared very, very close to a wall, so it was a bit difficult to interact with it'* (CC4) which was solved by re-starting and placing the playspace elsewhere. Others managed to place objects, but then *'found it quite hard to select AR elements (trees, plants, rocks, etc.) and manipulate them'* (CC5). As with the VI participants, the tutorial was mentioned; *'it didn't seem to explain what to do'* (CC2). Finally, the UI around the AR was disliked by one of them *'the menu for items to place in the forest was partially obscured on my phone, and I didn't like the obstructing pop-ups telling me to move on'* (CC3).

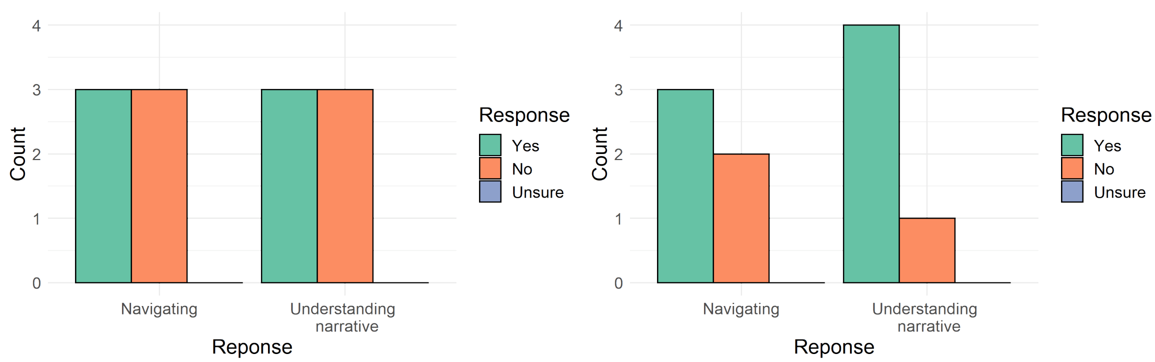
Figure 4.15a shows a word cloud of content creator responses when they were asked what suggestions they had for improving the app (EX7). The blind content creator simply suggested screen-reader compatibility 'to read all text on [the] screen including instructions and names of buttons'. Most (3, 60%) mentioned improving how the playspace was created with suggestions such as *'allow me to calibrate my play area better'* (CC3), *'it'd probably be useful to warn the user about not being too close to walls or other large surfaces'* (CC4) and *'the scene wasn't centred around where I selected, so half of the woods were in my fireplace'* (CC2). There were several suggestions for Scene 1, which allows the user to gradually create their own forest by adding elements (see Section 4.2.1). This included allowing them to *'cluster the objects. It'd also be nice to be able to choose the size of the object, and possibly also to copy entire cluster of objects, in order to rapidly create something nice'* (CC4), *(* CC3) and *'enable the user to unselect an object from the top menu and hence be able to tap on the screen without placing an object'* (CC5). There were also 2 (40%) suggestions about improving the UI, such as *'moving the menu for placing items in the forest'* (CC3) and *'implementing a nicer and more standard menu for selecting and placing elements in the forest'* (CC4) which they clarified as *'using buttons with a representation of the graphical element that then will appear'* (CC4). Finally, several talked about improvements to understanding the tutorial and instructions, suggesting *'more instructions for how to do things'* (CC3) and a *'practice with manipulating AR objects in the tutorial section'* (CC5).

4.5.3 AR App Sounds

Several questions were asked to determine relationships between sounds and the AR app, shown in Table 3.8. Participants were asked if sounds were helpful overall and which categories of sounds were helpful (NV1-3/ST1-3). They were then asked if additional sounds would help and which categories of additional sounds would help (NV4-6/ST4-6). The results are divided into these two areas. These questions were repeated for both navigational sounds (NVx) and sounds for understanding the narrative of the app (STx).

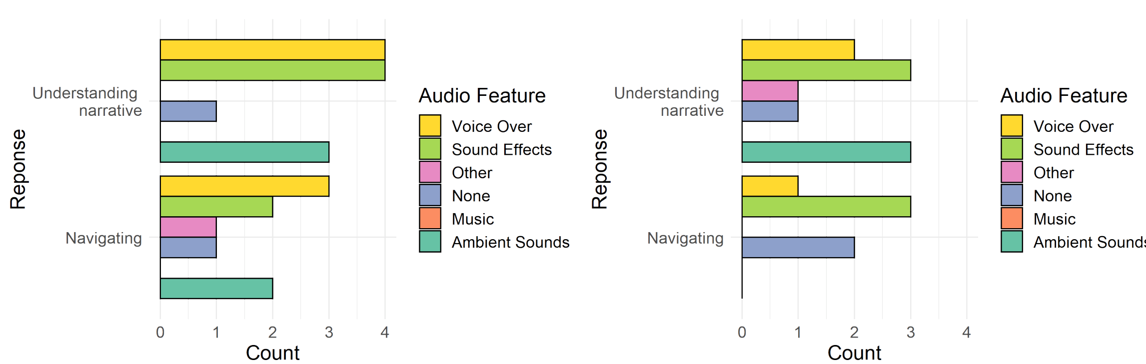
Did sounds help

The below presents participant responses to whether sounds in the app were helpful (NV1/ST1) and which sounds in the app were helpful (NV2-3/ST2-3).



(a) VI Participants (b) Content Creators

Figure 4.16: Participant ratings of whether sounds were helpful in the initial AR testing app. There are 2 categories of responses, for whether sounds helped them to navigate (NV_1) or to understand the narrative (ST_1). Questions from Table 3.8.



(a) VI Participants (b) Content Creators

Figure 4.17: Participant ratings of which sounds were helpful in the initial AR testing app. There are 2 categories of responses, for whether they thought sounds helped them to navigate (NV2) or to understand the narrative (ST2). Questions from Table 3.8.

VI participants. Figure 4.16a shows that the VI participants were divided over whether sounds helped them use the app, with 3 (50%) responding either yes or no, for both purposes of the sounds. Participants were consistent in their answers across both questions (NV1/ST1), showing that sounds either helped them with everything or not at all. Those who said no sounds helped them in the app all self-reported as blind or blind with residual vision, whereas those who said sounds did help self-described as partially-sighted or other. As established, the two (33.3%) who self-described as other were all towards the more sighted end (see Section 4.11). Therefore, it would seem that sounds only helped the more sighted participants. This is perhaps because less sighted participants rely upon screen-readers, as established by their responses to question EX6 in Section 4.5.2. They had no helpful sounds from the screen-readers themselves, and thus were mostly unable to get very far into the app to experience any other sounds.

Figure 4.17a shows what sounds VI participants thought helped them use the app. Voice-over was rated consistently high across both purposes of the sounds, but with fairly high support for both sound effects and ambient sounds. One participant responded that no sounds were helpful across both questions, stating *'I couldn't understand anything about it'* (P9). Another of the participants who didn't find any sounds helpful chose voice-over, ambient sounds and sound effects *'based on the assumption that the app had accessibility support, I would want the features selected to be able to navigate and enjoy the experience'* (P4). The final participant, P1, who selected that they didn't find the sounds helpful selected other across both questions stated *'I only really heard the ambient background/nature sound which was lovely but not particularly helpful'*, but did also choose ambient sounds for ST2, because *'I did feel like I was in a forest though as I could hear leaves and trees rustling and the sound of forest animals'*. Only one of those who found sounds helpful overall said that sound effects helped them to navigate that app (NV2) because they gave them an *'idea where to look/what to look for'* (P8). The others who found sound helpful both chose voice-over as helping navigate, stating *'I took cues from what was said to figure out what to do next which was helped by the written instructions'* (P6) and that it *'was very helpful as sometimes it is hard to read text on a screen and the ambient sounds add to the atmosphere'* (P7).

For understanding the story (ST2), all of those who found sounds helpful overall chose both sound effects and voice-over. Their answers to ST3 seem to present the sound effects as playing a supporting role to the voice-over. Voice-over *'explains well what is happening within the story'* (P6), whilst sound effects *'add to it to help understand things that are mentioned'* (P6) and *'relate the scenes to the narrative'* (P7). One additionally chose ambient sounds, explaining that all three categories of sounds together fill in *'gaps of thing perhaps I just couldn't/didn't see, as well as setting the scene'* (P8).

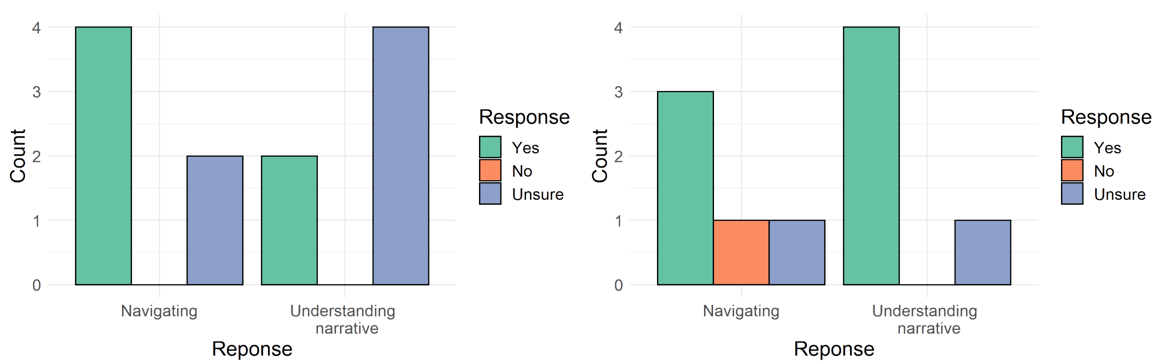
Content creators. Figure 4.16b shows that overall, content creators found sounds to be more helpful than VI participants, especially for understanding the story (ST1). A sighted content creator responded no to both questions. However, the only blind content creator responded no to whether sounds helped them to navigate. They also talked about needing screen-reader functionality in Section 4.5.2.

Figure 4.17b shows what sounds content creators thought helped them use the app. Sound effects were rated highest for both questions, with ambient sounds the next highest for navigating (NV2) and both voice-over sound effects for understanding the story (ST2). The lack of reliance on voice-over compared to VI participants (Figure 4.17b) is not surprising. For navigating (NV2), the two content creators who said no sounds helped them overall stated that none of the categories were helpful. The sighted one clarified by stating *'there* *didn't appear to be a need to navigate. The adding and removing of trees could be performed while stationary'* (CC2), whilst the blind content creator gave no further information to NV3. Those who chose only sound effects for helping navigate found the sounds the creature made in Scene 2 (see Section 4.2.1) the most helpful, saying they *'prompted me to turn my phone, explore the AR scene and look for the monster'* (CC5) and *'it helped me localising the creature and interacting with it'* (CC4). The other chose sound effects and voiceover as the both *'let me know when things were happening that I should pay attention to'* (CC3).

For understanding the story (ST2), there was again a focus on other sounds playing a supporting role to the narrative voice-over in the app, summed the participants who chose sound effects, ambient sounds and voice-over as *'the voiceover told the story, the ambient sounds set the scene, and the sound effects added specific narrative detail'* (CC3) and *'ambient sounds nicely paint the forest scene even without the on-screen visual AR elements. The narration was also very helpful in understanding what's going on. The sound effect of the roaring monster also helped me to understand how large and 'scary' the monster is even in the absence of the visual elements'* (CC5). One participant picked sound effects and ambient sounds, but not voice-over because they were '*were good for being "transported" in the virtual environment'* (CC4). The VI content creator chose the other category, stating that *'the spoken story-line about the forest'* (CC1) was helpful for them understanding the narrative. This could possibly be counted towards voice-over, as they may have not chosen that option thinking that referred to screen-reader technology.

Additional Sounds

The below presents participant responses to whether they thought additional sounds would be helpful in the app (NV4/ST4) and which categories of additional sounds would be helpful (NV5-6, ST5-6).

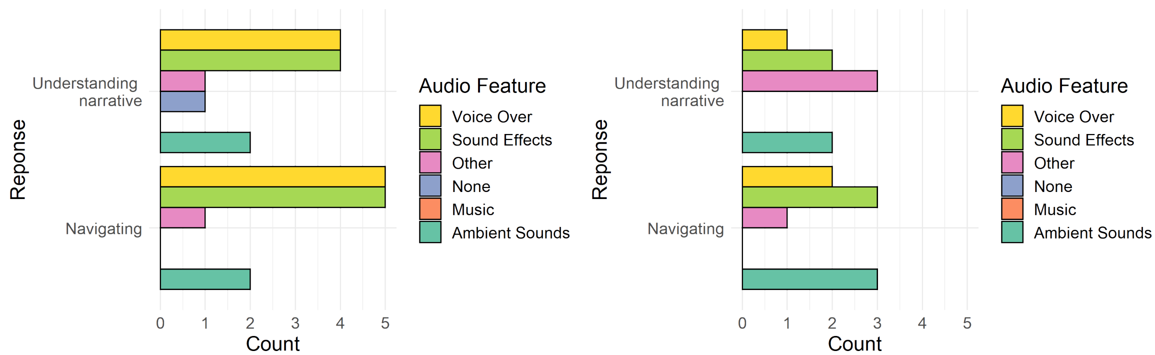


(a) VI Participants (b) Content Creators

Figure 4.18: Participant ratings of whether additional sounds would be helpful in the initial AR testing app. There are 2 categories of responses, for whether they thought sounds would help them to navigate (NV_4) or to understand the narrative (ST_4). Questions from Table 3.8.

VI participants. Figure 4.18a shows that VI participants were fairly uncertain whether extra sounds would help, but were overall more positive towards the idea. They were more sure about additional sounds helping them to navigate (NV4) with 4 (66.7%) responding positively, whilst the majority (4, 66.7%) responded that they were unsure whether additional sounds would help them to understand the story (ST4). Interestingly, the same 2

(33.3%) thought that sounds would help across both questions. Fisher's exact test of independence was used to determine there were not statistically significant associations between sight level (B3), enjoyment (EX4) or whether participants already thought sounds were helpful (NV1/ST1) and whether they thought additional sounds would be helpful (NV4/ST4). Whether participants thought additional sounds would help was independent of their level of sight, whether they enjoyed the app and whether they thought existing sounds helped them in the app already.



(a) VI Participants (b) Content Creators

Figure 4.19: Participant ratings of which additional sounds would be helpful in the initial AR testing app. Split into 2 categories of responses, for whether they thought additional sounds would help them to navigate (NV5) or to understand the narrative (ST5). Questions from Table 3.8.

Figure 4.19a shows what additional sounds VI participants thought would help them to use the app. Across both questions, sound effects and voice-over were chosen by the majority of participants, with some support for ambient sounds. The same participant who didn't find any sounds helpful but chose voice-over, ambient sounds and sound effects in response to NV2/ST2 chose the same again here, answering that this was again on the assumption of what they would want if they were first able to access the app using a screen-reader.

For navigation (NV5), even those who were unsure of whether additional sounds would help (NV4) gave categories of sounds they thought would be helpful. Those who chose voice-over (5, 83.3%) focused on it being the way that they would get *'practical instructions on where/what to look for as well as reading the on screen instructions'* (P8), with one suggesting that voice-over screen-reader functionality should be a default, so that sighted people have to turn it off, rather than *'requiring partially sighted users needed to find the audio description button or menu'* (P1). The participant who didn't choose voice-over only selected sound effects. Interestingly, they self-described as partially sighted so had a greater level of sight than most other participants (see Section 4.5.1). They thought that sound effects could *'indicate when you can do something else/move on'* (P6) such as telling you that a new button had popped up, especially if *'it's not in their immediate field of vision'* (P6). This idea of using sound effects to indicate state changes for navigating was backed up by others, with one suggesting *'include an audio key that can be accessed throughout the game to remind people of what each sound means'* (P1). The participant who thought ambient sounds would help also chose sound effects, voice-over and other and particularly thought that *'a constant sound from a particular source helps to establish orientation'* (P9).

For understanding the story (ST5), one participant chose none, stating that additional sounds were not needed but that it *'could be more descriptive of the monster as it appears so those who can't see it as well are able to understand what it looks like'* (P6). This idea of voice-over filling in the information they might otherwise miss in the story was backed up by another participant who stated *'if I had audio description of any new characters or animals etc I'd know what I'd need to look for'* (P1). Although there was as much support for ambient sounds as voice-over (4, 66.7%), no further ideas were given in response to ST6.

Content creators. The same question was asked of content creators, whether they thought additional sounds would help rather than if additional sounds would specifically help VI users. Overall, content creators were more certain about using additional sounds to help in the app. Especially for helping users to understand the story (ST4), with positive answers from all but one of them (4, 80%). However, one content creator said that no additional sounds were needed to help them navigate, perhaps because they already thought that sounds in the app helped them navigate (NV1) and found the app fairly easy to use, rating it a 2 out of 5 (see Section 4.5.2). The same content creator responded that they were unsure about additional sounds to both questions. They thought that none of the sounds in the app had helped them navigate or understand the story (NV1/ST1), but rather than adding more sounds they wanted improvements to *'prioritise the experience with the existing sounds first'* (CC2).

Figure 4.19b shows what categories of additional sounds content creators thought would be helpful. There was much less focus on adding voice-over in comparison to VI participants, which is similar to the responses both groups gave to which existing sounds they thought were helpful.

For navigating (NV5), content creators thought both sound effects and ambient sounds would be most helpful (3, 60% each). When asked to elaborate on their answers for what sounds would be helpful to navigate (NV6), 4 (80%) content creators suggested spatial audio. One assumed this would be a feature, stating *'I really didn't get very good spatial information from the sound. I struggled to tell where sounds were coming from'* (CC2). The others thought it would be a helpful feature for helping them identify objects, stating that *'localisable sounds, possibly with a direct link with visual objects, would be very, very useful'* (CC4), *'might help me to direct my attention'* (CC3) and would *'enable greater separation between the audio objects and enhance localisation, which could aid navigation around the AR scene'* (CC5). Those who chose sound effects (3, 60%) didn't provide many further ideas on how they could help navigation, except to add more *'when I interact with objects'* (CC3). Interestingly, those who chose voice-over (2, 40%) gave similar ideas despite one being sighted and the other blind. The sighted content creator suggested *'navigation within the app would be easier if pop-up hints and instructions were audible'* (CC5), whilst the blind content creator simply suggested *'spoken instructions from character'* (CC1).

For understanding the story (ST5), other was the most chosen category by content creators (3, 60%). One of these simply thought changes should be made to existing sounds only, as mentioned above. Another suggested being able to *'balance between the ambient sound, sound effects and voice over could be adjusted to make the important elements stand out'* (CC5). The final one was the blind content creator, who also selected voice-over and ambient sounds, asking for *'more spoken content'* (CC1). The remaining 2 (40%) both selected sound effects (with one selecting ambient sounds additionally), talking about them in a supporting role to voice-over in a similar way to their responses on existing sounds. They stated that they should be *'sounds that are linked to visual objects and that are mentioned in the narration'* (CC4) and if they are *'directly related to the story [sound effects] would help* *convey the narrative'* (CC3).

4.5.4 Ranking Task

Participants were then asked to perform ranking tasks on sounds in the app, to inform the developments of personalisable audio controls (see Section 4.4.4). As with the previous section on their thoughts on sounds in the app, the task was repeated twice, once for how important the sounds were to navigating (NIx_NV) and once for how important they were to understanding the narrative (NIx_ST). They were also asked questions about their approach to the ranking task and their opinion of being allowed to alter the relative levels of different sounds (Rx_NV/Rx_ST).

Rankings

The below presents participant responses to the ranking task. It is worth noting that although participants were able to listen to each sound to clarify what it was, they were not provided information on what the categories they were putting sounds into meant, or a preview of how the mix would sound once these categories were applied. The results of Ward's [32] ranking task are included in the Appendices, including both the results from the original sound designer of 'The Turning Forest' in Figure E.1 and the results from content creators surveyed in Figure E.2. Of course, these asked participants to rank sounds on narrative importance, not on navigational cues.

The results of the ranking tasks, Figures 4.20 and 4.21 have been displayed in a similar way to Ward's results (see above). Hierarchical visualisations are used to represent the 4 categories that participants could put sounds into. Although for this work the categories were labelled 1-4 rather than 0-3, they were given the same names (essential/high/medium/low importance). The modal choice was chosen when assigning sounds to a category. However, the small sample size paired with low agreement among participants (see below) means that several sounds had multiple modes. For these, the mean is then used to decide the category and displayed accordingly. To show agreement between participants, the percentage who chose to put the sound in the given category is also displayed. For comparison between navigational and narrative rankings and across VI and content creator results, italicised and coloured font is used.

VI participants. It was important to obtain an indication of how much participants agreed with each other on their rankings. Similarly to Ward's work [32], Fleiss's Kappa was computed to assess the agreement between the participants when performing the ranking tasks. This gives a Kappa value of 1 if participants are in total agreement. In the navigational ranking task, there was poor agreement, Kappa = 0.0348, p = 0.39. In the narrative ranking task, there was agreement worse than chance, Kappa = -0.0273, p = 0.496. This is further displayed by the number of sounds that have percentages below 50 in Figure 4.20. One explanation could be that participants didn't fully understand the task.

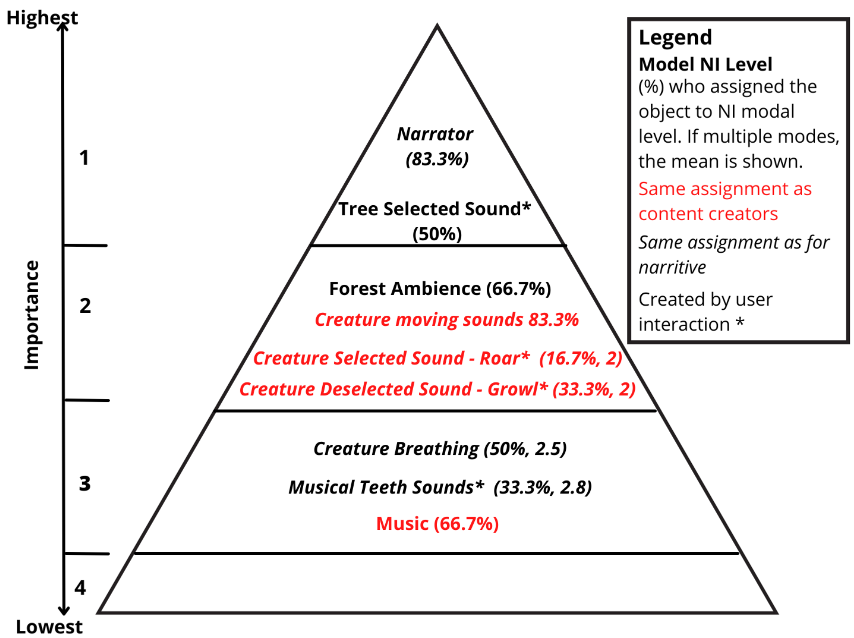
Across both ranking tasks, VI participants put no sounds in the low (4) and few sounds in the essential (1) categories. Ward [32] found a similar trend in her results, that participants mostly put sounds in the middle categories. They were reluctant to rate sounds at the extremes. However for navigation, all but one participant (5, 83.3%) utilised the essential category and all but 2 (4, 66.7%) utilised the low category and for narrative, all participants utilised the essential category and all but one (5, 83.3%) utilised the low category.

Figure 4.20a shows how VI participants rated a sound's importance for helping them to navigate. The sounds they most agreed on were the narrator, rated essential, and the creature moving sounds, rated high (both 5, 83.3%). The high-rated narrator agrees with VI participants responses to questions on which sounds helped them in the app (NV2/ST2) and which additional sounds would help them to use the app (NV5/ST5). Interestingly, the user interaction sounds (denoted with an *) were placed in a variety of categories. This could perhaps be explained by participants who weren't able to get past Scene 1 not encountering further interaction sounds. The same is true of those that can be classed as sound effects (admittedly most of the sounds), which are rated across several categories. Perhaps the constant forest ambience and sounds of the creature moving (both rated high) were used in conjunction with the essential narrator to direct their attention. However, this argument is weakened by the creature's breathing being rated medium, as this is also constant.

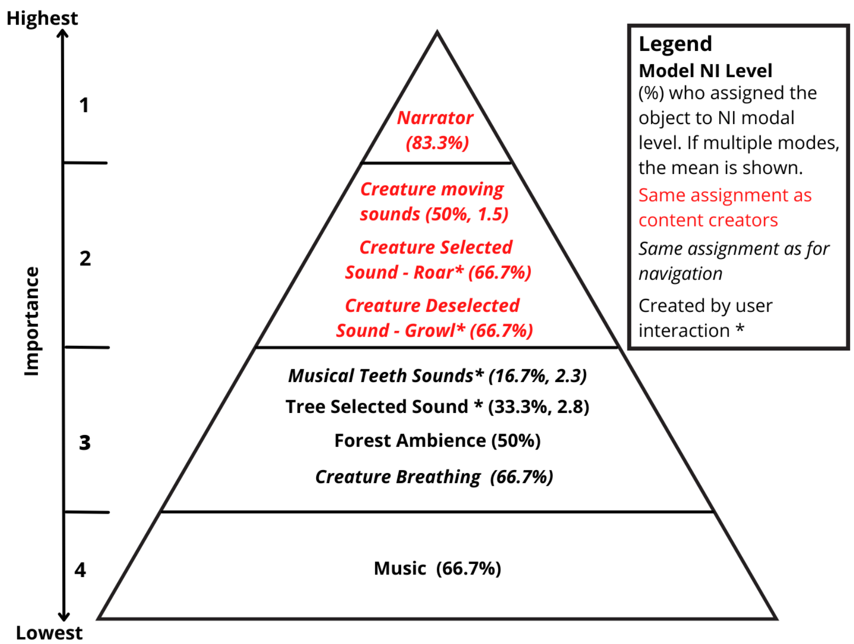
Figure 4.20b shows how VI participants rated a sound's importance for helping them to understand the narrative. A number of the sounds stayed in the same category as for navigating, excepting the tree selected sound and the forest ambience. The narrator was once again rated essential, with the highest level of agreement (5, 83.3%). Interestingly, all sounds rated high were related to the creature, which is a central element of the narrative. Although, other creature sounds were rated in the medium category. It is also interesting that music is placed in the low category with a fair level of agreement (66.7%), as it could be argued that music helps to understand the mood of a narrative. The same could possibly be said of the atmospheric forest ambience, placed in the medium category. Perhaps this gives an insight into VI priorities in understanding AR narratives; that the need to comprehend the flow of the story through the narrator and the main players in it (i.e. the creature) through sound effects are more important than getting a feel for the atmosphere through music or ambience. This could also reflect on their experiences of having content described through Audio Description (AD) and screen readers, meaning spoken descriptions are most important because they are expected.

Content creators. Fleiss's kappa was again computed to assess the agreement between content creators when performing the ranking tasks. In the navigational ranking task, there was poor agreement, Kappa = 0.0944, p = 0.0581. In the narrative ranking task there was also poor agreement, Kappa = 0.0603 , p = 0.261. Content creators also made little use of the essential (1) and low (4) categories, putting only one sound in them across both tasks. For navigation, 3 participants (60%) did not utilise the essential category and 4 (80%) did not utilise the low category and for narrative, all but one participant (4, 80%) utilised the essential category, but 4 (80%) did not utilise the low category. Again, it would seem that participants were reluctant to rate sounds at the extreme ends of the scale.

Figure 4.21a shows how content creators rated a sound's importance for helping them to navigate and Figure 4.20b shows how content creators rated sounds importance's for helping them to understand the narrative. All sounds stayed in the same category as the narrative task, apart from the narrator. All other sounds were rated as important, except for the forest ambience and the music. Perhaps these were viewed as more atmospheric-scene setting sounds than essential to understanding what was happening. Interestingly, the only sound with more than 60% agreement across either task was the creature's breathing being placed in the important category for navigation.

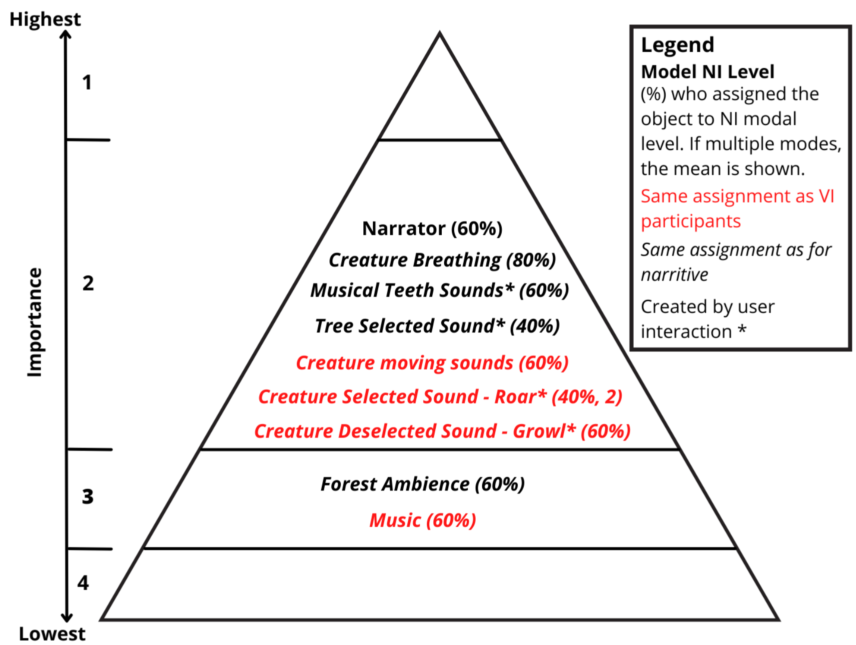


(a) NI_NV

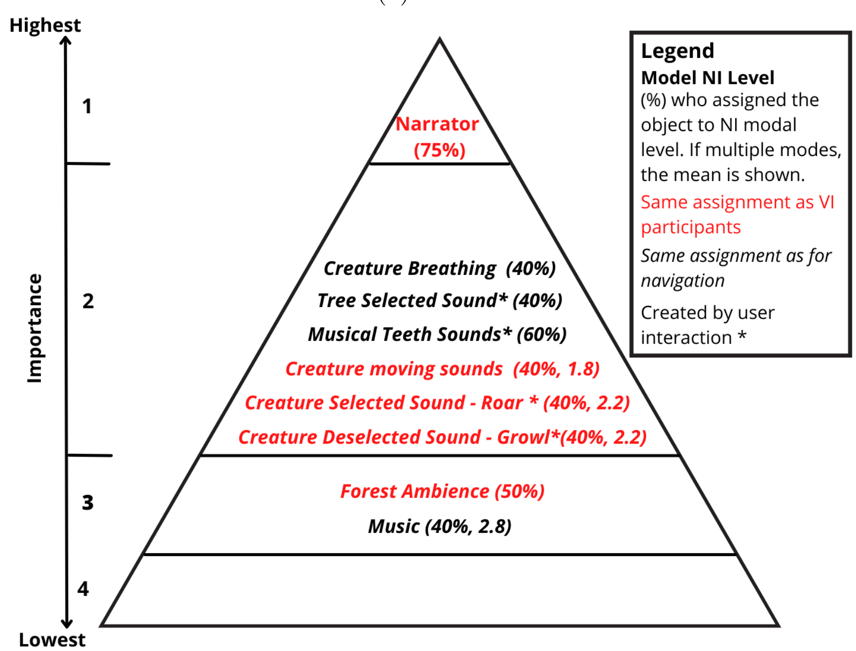


(b) NI_ST

Figure 4.20: Hierarchical visualisation of the modal narrative importance level assigned to each audio object in the AR testing app by VI participants. The percentage of respondents assigning each object to the modal level is noted in parentheses. Where there were multiple modes, the mean is noted and used for ssigning an importance level. Split into 2 categories of responses, for a sounds navigational importance (NI_NV) and a sounds narrative importance (NI_ST). Italicised text denotes objects assigned to the same importance level by VI participants for both navigational importance and the narrative importance. Red text denotes sounds that were also assigned that importance level for the same category (NI_NV/NI_ST) by content creators. Results from ranking task as described in Section 4.4.4



(a) NI_NV



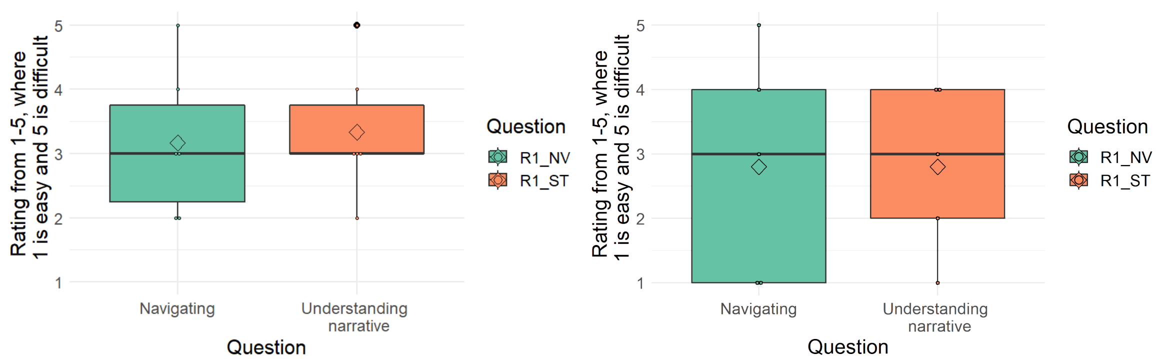
(b) NI_ST

Figure 4.21: Hierarchical visualisation of the modal narrative importance level assigned to each audio object in the AR testing app by content creators. The percentage of respondents assigning each object to the modal level is noted in parentheses. Where there were multiple modes, the mean is noted and used for assigning an importance level. Split into 2 categories of responses, for a sounds navigational importance (NI_NV) and a sounds narrative importance (NI_ST). Italicised text denotes objects assigned to the same importance level by content creators for both navigational importance and the narrative importance. Red text denotes sounds that were also assigned that importance level for the same category (NI_NV/NI_ST) by VI participants. Results from ranking task as described in Section 4.4.4

VI participants and content creators were evenly split on agreeing which categories to put sounds in (shown in red on Figures 4.20 and 4.21). They agreed on 4 out of 9 categories (44.4%) for navigational importance and 5 out of 9 categories (55.5%) for narrative importance. The only sounds they agreed on across both tasks were creature sounds - moving, selected and deselected.

Approach to ranking task

Participants were asked questions about how difficult they found the ranking tasks (R1_NV/R1_ST) and what their approach was to the task. This is interesting given that there was little agreement either within each group of participants or across the two groups.



(a) VI Participants (b) Content Creators

Figure 4.22: Participant ratings of how easy the NI sound ranking task was. Split into 2 categories of responses, for how difficult it was to ranking the sounds navigational importance (R1_NV) and the sounds narrative importance (R1_ST). Questions from Table 4.7. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants. Figure 4.22a shows that overall VI participants found it relatively hard to assign an importance category to each sound. They were more certain about this for the narrative task (R1_ST) than the navigational one (R1_NV). For assigning categories a navigational importance, they rated the task a mean of 3.17 (median 3, standard deviation 1.17) and for assigning a narrative importance, a mean of 3.33 (median 3, standard deviation 1.03). Kruskal-Wallis tests determined that there were no significant relationships between level of sight (B3) and how difficult participants found either ranking task (R1_NV/R1_ST).

From VI participant responses on their approach to the navigation ranking task (R2_NV), two themes seem to emerge. First, whether they tell users where they are or where they need to go, such as letting participants *'understand where you were in the story to carry out the next action'* (P6), that *'I want to keep away from the creature, so the creature sounds are important'* (P9) or that *'I would think of the sounds in relation to movement and my interactions and warnings of threats'* (P4). This leads to the second theme, whether sounds game information on interactions with objects, such as letting users *'know if you've selected/deselected something'* (P6).

When asked to explain their approach to the narrative ranking task (R2_ST), a theme that emerged was whether sounds helped participants make sense of the story, such as *'I thought about what would make sense if it was included or not'* (P1), *'if I actually notice/used it in the app'* (P8) and *'whether you'd know where/who/what was happening in the story'*

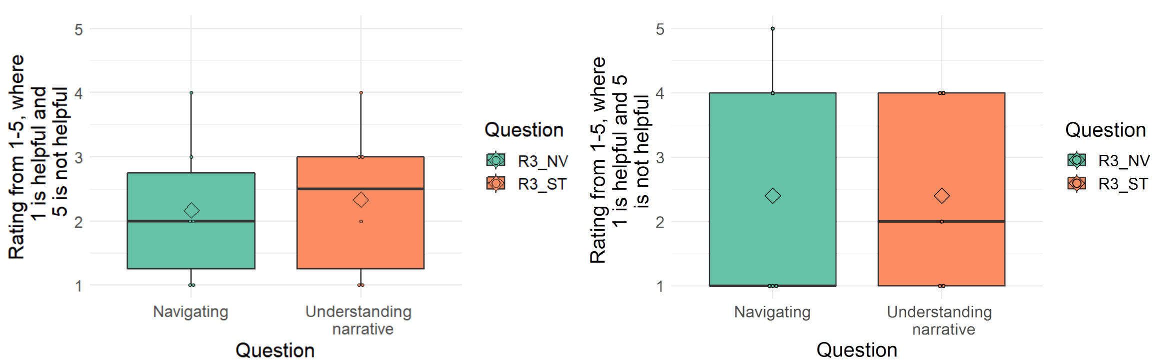
(P6), with one participant stating *'the narrator gave me the best idea of what was going on'* (P9). Others focused on how sounds affected the impact of the story, describing their approach as steered by how *'impactful they are in relation to the scene'* (P7), with two focusing on the creature as it *'gave the best idea of the amount of danger'* (P9) and its *'growl is a threat so high importance'* (P4).

Content creators. Figure 4.22b shows that content creators also found it relatively hard to assign an importance category to each sound. However, there was much more uncertainty in this than among the VI participants. They were especially uncertain about how difficult it was to assign a navigational importance category (R1_NV), with ratings across the full range giving a mean of 2.8 (median 3), with a standard deviation of 1.3. There was slightly less uncertainty about how difficult it was to assign a narrative importance category (R1_ST), with no ratings of 5 and a smaller standard deviation of 1.3 also giving a mean of 2.8 (median 3). The blind content creator found it the easiest, rating both tasks 1 out of 5.

Content creators seemed to have a wide variety of approaches to the navigational importance task (R2_NV). A theme that did emerge was again linked to sounds giving information on interactions with objects, with one stating *'I generally prioritise sounds to do with changes of state and user interaction'* (CC2) and another thinking *'does the sound provide feedback on user's interaction'* (CC5). CC5 also mentioned choosing sounds that *'make you want to move and explore the AR scene'*. Another simply focused on sounds that *'were identifiable through a visual object, and explicitly mentioned in the narration'* (CC4).

When considering their approach to the narrative ranking tasks (R2_ST), content creators seemed to again have a variety of approaches. They mentioned that they chose based on whether the sound *'was linked to a physical object, and explicitly mentioned in the narrative'* (CC4). Others did focus on the *'relevance to the storyline'* (CC1) and whether sounds *'explain what is going on...[or] paint a picture of what the scene looks like even in the absence of visual elements'* (CC5).

Perception of having personalisation control



(a) VI Participants

(b) Content Creators

Figure 4.23: Participant ratings of whether having a personaliable audio control would be helpful. Split into 2 categories of responses, for whether they thought the control would be helpful for navigating (R3_NV) or for understanding the narrative of the app (R3_ST). Questions from Table 4.7. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

Participants were asked whether they thought it would be helpful to be able to change the levels of sounds that helped with navigating (R3_NV) or understanding the narrative (R3_ST) of the app.

VI participants. Although there was some uncertainty, Figure 4.23a shows that VI participants thought that being able to change the levels of sounds would be helpful. They were slightly more certain that being able to change the levels of sounds for navigating the app would be helpful, rating it a mean of 2.17 (median 2, standard deviation 1.17). For being able to change sounds helpful to narrative, they gave a similar mean of 2.33 (median 2.5, standard deviation 1.21). Kruskal-Wallis tests determined there were no significant relationships between level of sight (B3), whether participants though sounds helped in the app (NV1/ST1), whether they thought additional sounds would be helpful in the app (NV4/ST4) and whether participants thought a a personalisation control would be helpful for navigation (R3_NV) or for understanding the narrative (R3_ST).

Content creators. Figure 4.23b suggests opinion was varied about whether being able to change the levels of sounds would be helpful. For navigation (R3_NV), 3 (60%) rated it as 1 (helpful), but the other 2 participants have ratings right down at the other end of the scale. This gave a mean rating of 2.4, but a very different median of 1 and a standard deviation of 1.95, which makes it hard to give an overall judgement. For narrative (R3_ST), there was slightly more spread of results, with the mean of 2.4 closer to the median of 2. However, the wide range of ratings and standard deviation of 1.52 still make it difficult to make a decisive judgement.

4.5.5 Workflow Questions

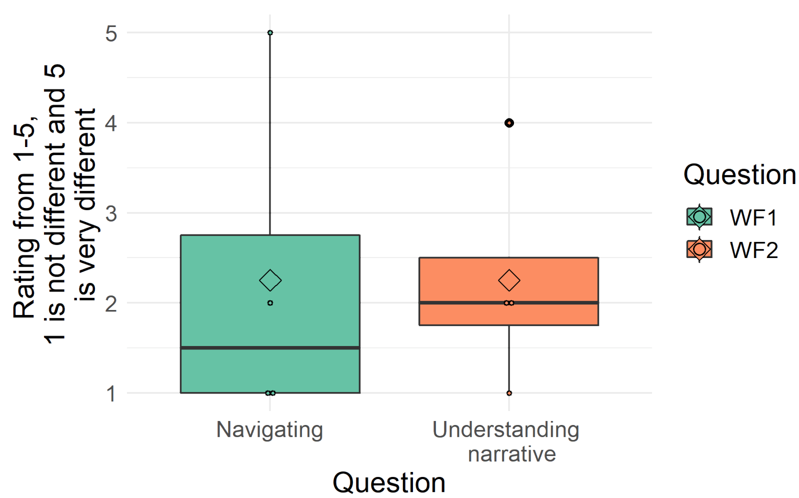


Figure 4.24: Content creator ratings of how different ranking sounds is to their own workflow. Split into 2 categories of responses, for how different ranking sounds due to their navigational importance is from their workflow (WF1) and how different ranking sounds due to their importance for understanding the story is to their workflow (WF2). Questions from Table 4.8. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

Content creators were asked further questions about how they would integrate personalisation into their own workflows, shown in Table 4.8.

Figure 4.24 shows content creator ratings for different ranking sounds due to narrative or navigational importance was to their workflows. It is worth noting that one content creator did not respond to these questions. For those left, it seems that ranking sounds in this way was fairly familiar, especially when considering the narrative (WF2), which gave a mean of 2.25 (median 2.0), with only one ranking (25%) of 3 or above. There was also a mean of 2.25 (median 1.5) for whether ranking sounds due to navigational importance was different to their workflows, but more uncertainty around this rating.

Unfortunately, the content creators did not provide much insight into what tools they would need to allow users to personalise their audio experience (WF3). One suggested *'Subtitles, separate volume controls for ambience, sfx, and music'* (CC2). CC4 suggested that users should be given a chance to explore what difference personalisation would make before they start an experience through *'an initial level calibration of the different sounds'* which would allow them to *'select some options in advance, or even just to explore what the sounds will be "offline"'*.

4.6 Interview Results

After completing the survey, all participants were also invited to take part in an interview. 5 visually impaired participants (P1,4,6,7&8) and 2 content creators (CC1 & 2) were willing to take part. The format for the interviews is described in Section 4.4.6.

The interviews were transcribed using built-in functionality on Microsoft Word for the Web [176]. Corrections were then made by the author. Once the transcripts were created, the author then summarised each interview, creating a list of the interviewee's thoughts on several themes. The themes looked for were their thoughts on their overall app experience and on each enhanced audio feature in turn, as these are the areas of greatest interest. The questions asked, shown in Table 4.10 reflect this. However, as repeated discussion of a wider concept of audio as an accessibility feature occurred, this was later added and participants' thoughts on it summarised. All interview summaries were then compared, with recurring thoughts or ideas noted, along with which participants subscribed to a particular view or approach regarding that theme. To capture the variability of VI experience, ideas are still reported even if only one participant mentioned them. Details of how to access these summarising documents as well as full interview transcripts can be found in Appendix E.

The following summary of the interviewee's thoughts and ideas that came up in the interviews is split up into categories based on these themes. Of course, there is some overlap between the different categories as some ideas branch over multiple. VI participants and content creator responses are aggregated, as there were many overlaps in their ideas. This may have been due to one of the content creators also self-reporting as blind and the small number of content creator respondents. Linking phrases, such as 'you know' or 'like' were removed from quotes so that they are easier to read in this discussion.

Perception and Experience of AR

Participants were asked about their previous experiences of AR and their perceptions of it. This clarified that VI individuals assume they can't use AR because they assume *'it's very visual. A lot of people will just say oh well, you can't do it because you have to be able to see'* (P1). Even those who were able to utilise some sight thought it would be difficult for reasons such as *'it is usually on a small screen in a small environment, so having to try and focus on that [is] difficult and there aren't usually accessibility options in terms of making things bigger'* (P6). In this section of the interview, CC2, who is a game audio programmer said more broadly about the gaming industry *'I don't think games do enough on the accessibility front, I think we could definitely do a better job there overall'* and that when they have programmed/mixed content *'we've not really considered accessibility too much when doing that. We're just trying to get as much information to the player as possible'*.

Participants were also given chance to talk generally about the AR app before more specific questions were asked. Much of the information they gave repeated feedback explored in Section 4.5.2. Those who relied on screen readers found it impossible because they *'couldn't get past the initial screen'* (CC1). CC1 went on to describe how they then *'got a sighted person. I think they got me past, but then things wouldn't work after that'*. If they were able to get in, many described being able to hear the audio but it didn't give them enough information, so *'it didn't make any sense [...] I was able to sort of tap some things on the screen and make them work, but it didn't tell me what they were'* (P4).

Audio helping accessibility

Before discussing the specific strategies that this work aims to explore, participants were first given a chance to talk about what they thought of using audio as a way to enhance the accessibility of the app for everyone. This was met with great positivity. A number of them thought that doing so would improve the experience for everyone, not just VI users. One stated *'audio can make something so much more immersive, I think if you sort of focus on the audio for everybody, not just for a visually impaired person maybe it would benefit everybody'* (P1) and another said *'I think it's good when it's there for everyone'* (P8). A more sighted VI participant (P7) stated that even though they can get a lot of information through their sight, *'sometimes just visual isn't enough'*, so audio is necessary as they *'prefer audio when I'm struggling visually'*. The only sighted interviewee, the gaming content creator (CC2), backed this up by saying *'I really love the idea of audio as something that conveys meaningful gameplay information and helps. Helps you in a way that we generally don't associate with audio'*, sharing that in his past work he had seen this work in a shooting game; *'this guy told us like he couldn't tell which one he was locking onto. So we put a different sound and I think little things like that little details really help you even if you are not visually impaired at all'*.

When asked specifically how audio could actually help, several suggestions were made.

- Much emphasis was placed on the need for spoken descriptions, which is explored in the 'Guide Character' category below. Several participants (P1, P7, CC1) suggested this themselves before being prompted.
- That audio can provide information on the location and type of cues. This is explored in the 'Spatial Audio' and 'Sound effects' categories.
- That audio can highlight new things or state changes, discussed further in the 'Sound effects' and 'NI control' categories.
- That audio can be used to provide information on tasks or interactions, discussed further in the 'Sound effects' and 'NI control' categories.

Spatial audio

All the participants interviewed thought spatial audio would improve the app. CC2 had suggested it previously in the survey and one VI participant (P8) suggested it even before the idea was introduced, *'I don't know if maybe like having being able to hear where it was coming from sort of as well adds again, as someone who can't see it as well. A lot more sort of, I suppose textures'*. For the rest, the idea was introduced to them, with a simple explanation, such as 'it is 3D or 360 degree audio that lets you hear sounds all around you in the location they're coming from'. As an initial response to this, many thought it would add to the experience stating *'it could be could be amazing if it works'* (P8) and *'it'd be more impressive'* (P7). It was also mentioned that it would make it *'so much more interesting and immersive. And I think if that's what the aim of AR is to be as immersive as possible'* (P1).

An interesting theme that came up with the spatial audio was that it would be just be *'tuning into a skill that a lot of people already have'*, as *'most visually impaired people I know have been taught to listen to where things are coming from. Like crossing a street I can't see a car coming. But I can tell you where it's coming from or if its Electric, and when it's going to stop'* (P8) which was an idea backed up by another describing *'I know if there's a car coming. I know which direction it's coming from and roughly how fast it's going. So that sort of thing would be the ideal'* (P4). They thought that *'using audio as a directional thing is the key thing'* (P7) which could make the app more accessible, or even *'not just more accessible. I think it would just be more cool as well'* (P4). This further backs up the idea of audio being an accessibility tool, but one that can be enjoyable regardless of whether a user needs it as an access feature or not.

In terms of how spatial audio could help them, VI participants had few ideas beyond *'you need it to tell you the direction and what it is'* (P4). This is probably because they may not have much experience with it. However, P6 did suggest that for it not to be overwhelming *'I guess maybe if you had the trees nearer to the player make sound. Have a radius maybe'*. CC2 had experience of implementing spatial audio in games and discussed that having fixed points to help the user orient themselves as *'they stood out quite a bit in the soundscape. And yeah, more distinct positions, so it's a sound changing as you move around it and you're getting a lot of positional information'*. They also talked about creating *'distinct areas that have their own signature sounds so you could roughly tell which way you were facing, but without being able to see the screen'*. The idea of the spatial audio being focused on the things nearest to the listener also resonated with them, but *'primarily for CPU performance, we kind of put a lot of limiting into have maybe the nearest three instances of a certain type of sound'*, which also gives a *'clearer mix'*. This idea of not overdoing it, so therefore ensuring the right cues are given at the right time is also discussed in the 'Sound Effects' and 'NI control' categories.

Sound Effects

Participants were positive about the use of sound effects to help make the app more accessible, with P4 and CC2 in particular emphasising that they would help understand what was going on. However, they were divided as to what the main purpose of sound effects should be. P6, P7 and P8 focused on the sound effects being used to help with interactions, such as *'when you click on something, like a distinctive noise for like each thing'* (P7) and *'if they could hear that interactive noise they'll be like oh okay I've done it'* (P6). They also determined that these interaction sounds should be *'distinct from the sort of general forest* *sounds you've got going on [...] audibly, quite different from what you've already got'* (P8). Other participants focused more on sound effects being used to highlight new things or state changes. P1 stated *'if it is something in the story that is going to tell you something new [...] that would definitely be essential'* and CC2 shared that in their work *'we often try to tie sounds to state changes'*. Some of their thoughts on how sounds could be highlighted were linked to the NI control (see category below), with support for the idea that the sound of an object that is new or needs highlighting in a particular moment could have a higher volume level for that moment, then be moved to a lower volume level when it is no longer new or so essential.

Of course, using sound effects for these purposes is not mutually exclusive, they can be used for both. However, another area raised by some participants was not overdoing it, as this could be more confusing than helpful. P4 stated *'obviously don't have audio for every single item. The important stuff'*. This was also linked to the previous discussion with CC2 over spatial audio, with them trying to *'reduce the amount of noise'* by *'limiting it to have maybe the nearest three instances of a certain type of sound'* to both help CPU load and ensure a clearer mix.

Another idea was suggested by P1; *'a key for the audio sounds. So that if I'd forgotten what one of the sounds was I'd be able to, click on the key and have them playback so that I could remind myself'*. This idea was shared with other subsequent participants, with a positive response. P8 was particularly positive, but suggested an extension involving *'having a key like that or even the ability to turn like parts of it off if you didn't feel like you needed all of it. So like if you know when you've touched the tree, you're not going to want to hear a click every time you touch a tree'*.

Guide character

When discussing developments to spoken instructions in the app, the term 'guide character' was adopted. For any developments to be in-line with the ideas of the accessibility research this project is based on (see Section 2.4), it would be preferable for standard screen-reader technology with a computerised voice to not be used. Participants were asked about their opinions of a guide character who could handle both screen reader functionality and provide hints on how to use and complete tasks within the AR world. They were told that this character would be created using a recorded human voice, and designed to fit in with the narrator already in the app. This can be thought of as the nearest equivalent to 'i-voice' used by the EAD project [19, 27–30]. Participants were also given the chance to suggest their own ideas.

Before the idea of the guide character was introduced, several participants had already highlighted the need for spoken cues to help them in the app. P4 suggested '*if I slide my finger in particular directions. I'd like it to say I'm sliding my finger from one area to a nearer tree'* and similarly, P1 suggested *'having the instructions or the tutorial in an audio sort of format*'. P4 also started to allude to these spoken instructions achieving more than standard screen-readers do by saying *'I think a voice-over would help with both knowing what to do and also adding to the narrative'*. CC1 also talked about going beyond just what they expect from TalkBack *'that sort of descriptive detail, so it brings the game alive. You're not just, playing with building blocks, you're knowing what's that, what that block makes up, and that's part of the visual'*. CC2, who is sighted, even suggested having screen-reader functionality enabled by default when the app opens, as *'for most games although they've got these options, you've got to find them hidden away in a deep menu somewhere to turn* *them on'*. Some participants also talked about disliking standard screen readers because *'it's not natural voice, and I think my brain doesn't process it quite the same way as kind of a normal voice'* (P1) and *'it's always the same sort of robot monotone voice. And I feel like if you're gonna try and cater to that audience, at least like put a bit of effort in rather than. I don't know how it works behind the scenes, but just typing it in and getting a robot to say it'* (P8).

When the guide character was suggested, participants shared that they liked the idea. The fact that it wouldn't be a computerised voice was received well. For example, P8 said *'if you could make it more personable, then I think it would definitely sit better'*. P7 suggested going even further in making it interesting by having *'characters that are diverse and so it's distinct, so it's one character is giving you this information and another characters giving you another set of information'*. They also liked the thought that *'if it's written as part of it coz then it doesn't feel like it is just lumped on coz some things like that can just feel like it's an afterthought'* (P6), and that doing it in this way would prevent spoken instructions from feeling *'like some sort of foreign object in this place that they've clearly spent a lot of time and money creating'* (P8). Finally, the thought that it could be enjoyable for everyone was also appealing. P1 said this would change their whole approach to screen readers '*I think if it doesn't feel like an accessible add on if everybody is experiencing the same, then yeah, I think in my mindset probably would be in a completely different place*'. P8 thought that it *'will enhance someone else's experience. But then make it accessible for people who need it to be'*.

The most uncertainty about the guide character came around how hints should be accessed. P8 particularly thought that it should be automatic as *'finding a help button could be quite difficult'* because they rely on screen magnification. They thought it should give *'a good little while to try and figure it out. And if there's some way of it figuring out that you're really actually quite stuck then butting in'*. Most other participants thought that it would be more important to have control over asking the guide characters for help or more details themselves and potentially even having the option to turn some or all of it off depending on their needs. The main reason for this was that they still wanted to be challenged to explore by themselves, stating *'it's not really a game if you're just talked through the whole thing [...] it's gotta be a challenge'* (P4) and *'it's quite thrilling in quite a, you know, a good emotion to feel. Well, I'm on my own now. Now my guides gone. What's this going to lead me into?'*. However, a balance was suggested by P7 who suggested *'maybe a timer for the option'*, which CC1 also agreed with *'after a certain time period or something, the guide could just pop in, even if you haven't asked anything just to sort of give some reassurance or maybe even announce that it's taking you a long time'*. Another suggestion by P4 was having different modes of description for different difficulties of play, or when they discussed the NI control they also suggested an extra control, *'one for the level of the sound. How many sounds and you know how descriptive the sounds are and one for the level of the voice over descriptions'*.

CC1 had further ideas for the guide character. They wanted *'an option somewhere to get detailed description'*. They thought this could be always present with *'some options relevant to around about the time of that game, either something you've just passed or something that might be just coming up around the corner'*. They also suggested the guide character providing details on the AR context, through an *'overview map perhaps in the menu, and then that's there all the time, so I'd like to hear the overall forest described again'* or by having some way to *'find out the distance between you and the monster'*.

Personalisation Control

In a similar manner to the guide character, participants were introduced to the idea of having a single personalisation control in the app, which could be used to reduce the level of sounds that weren't important in a particular moment and increase the level of sounds that were. This is based upon Ward's NI work in [32].

Most participants responded favourably to the idea, although they had many suggestions for how this could work. P1 thought it would be *'really helpful and I think not just for visually impaired people, but maybe people with autism and things like that who can't necessarily cope with lots of sounds'*, whilst P7 responded that it would make the app *'a lot easier because sometimes you still want the background noise, but if the background noises are higher than the narrator, then it's not a good audio mix'*. CC2 did urge caution with how it was presented, that it needed to be explained well and that users shouldn't be given a lot of options as soon as they opened the app, but it should be presented *'in a more appropriate context, possibly, maybe after you've heard your first voice'*. Participants were also happy with it being a simple slider control, as this is something they are familiar with and *'you have to be a little bit careful with gestures unless they're explained as part of the story line'* (P1) especially because various gestures are often used by screen readers.

Multiple participants suggested the idea of having more than one control. P1 suggested *'it could be useful to have maybe sliders for separate genres, like music, sound effects, that sort of thing'* and P8 also said *'so for your general forest [...] do you want that high, medium or low? And then the same for interacting with things, and then a third group for like the actual narration'*. Part of the reason for this was a fear that by reducing control to one slider, you could be taking control away from the user because *'every person is different and telling them what it's like. Either tell them what is the most important or give them the option'* (P6) and *'you definitely want to give the user a bit more control. I mean not completely, because then it can just ruin like how you want the experience to be'* (P7). However, further discussion allowed further opinions to develop, especially once it was reiterated that sound objects wouldn't be in a fixed importance level, but could move to different importance levels depending on what was happening at a particular moment. P7 thought this could be more powerful than having lots of control *'because I could change the settings and then miss something that is essential at one point'*. P8 summed it up in this way: *'I feel like someone's got to make those decisions [...] as long as it's informed by the people you're making the decisions for'* and *'it's just creating options so that people can work as best as they can to suit their own needs or how they want to interact with it'*.

When asked what sounds the control should highlight the most, there was a similar divide to that previously explored for the sound effects. Some (P1, P7, CC2) thought that the most essential sounds should be those tied to state changes or new things. Whereas others (P4, P6, CC1) thought the most essential sounds were those tied to tasks needing completion, with the creature being used an as an example; *'if you're looking for a monster, you need the monster sounds to navigate where it is'* (P6) and *'if the monsters to be avoided you, you just need the sound to fade away as you move away and there may be some other sound that is going to be some area of safety needs to be louder than the monster because that's your cue to where you can go'* (CC1).

An area that was discussed in relation to both spatial audio and the NI control was the idea of highlighting cues by distance. As mentioned, CC2 suggested only having closer instances of sounds active as a way of both reducing CPU load and increasing the clarity of the overall mix. P6 suggested a similar idea of having a 'radius' of sounds nearer to the player. Whether coding this in as a feature on top of the increases and decreases in level that are created by proximity in spatial audio is an interesting area for discussion. However, one potential problem pointed out by CC1 is that *'because you use [the] volume of sound to kind of get a feel for how far something is away from you. So if you start messing about without an increase in the volume of the monster sound or something, he's gonna sound like he's like just right by you'*. This could be a problem with implementing a NI personalisation control where users can create their own level changes and focus on sounds by the way they explore within the spatialised auditory world. But it also may not be, as most users are unlikely to adjust sliders whilst also moving around. This, the two are unlikely to be altered at the same time causing confusion.

The content creators were further asked about the idea of the control in relation to their work. CC2 in particular liked the simplicity of the idea, saying *'hundreds of sliders is in itself slightly less accessible'*. Given that CC2 codes (unlike CC1), they had more thoughts on implementing it in their own work. They started to think about how they would do this in Wwise [177], saying it would be simple as '*this sort of doesn't even need to be a plug in because it is built in functionality of a parameter driving the volume of a thing is something that's used in lots of different ways already'* and *'the game would send a single parameter to the middleware library, and then in there you'd say for each event you've got a basically a slider to control that number to drive parameters that you could say okay, this parameter drives volume and you know there's a curve, so when it's one it's 0dB and when it's at zero it's minus 100dB'*. They also thought it would be a *'trivial amount of processing'*.

4.6.1 Interview Recommendations

From the summary of interview responses and themes discussed, it's possible to pull out several recommendations:

- Audio can be an accessibility tool, but one that can be enjoyed regardless of whether a user needs it as an accessibility feature or not. Examples of this could include spatial audio, spoken cues or sound effects which create a sonic world in which all information is provided, meaning it's more immersive for everyone but VI users get all information they need to understand what is happening.

- Audio should not be overwhelming and should ensure that the right information is presented at the right time. This may be managed by user control, such as the personalisation slider idea. However, the intuitive interaction afforded by spatial audio may also provide a natural way for users to highlight what they need by turning to give it their focus and/or moving to greater proximity.

- There are several things that audio needs to provide information on, including state changes, interactions, the appearance of new things and information on tasks that need completing.

- Allowing users the opportunity to request more information if they want it can be a way of allowing them to tailor the experience to their access needs and also enjoy the challenge of using it. For example, having a key so they can find out what individual sound effects mean if they are relying on them, or allowing them to ask the guide character for more information on the experience around them or the task that they are doing.

- Having level changes due to proximity happening in conjunction with NI control level changes could potentially be confusing.

4.7 Discussion of Study Two

Study Two presents real insights into VI experience of the testing AR application. It establishes that VI participants find it a similar difficulty to the currently existing AR apps assessed in Chapter 3. As with those apps, there does not seem to be a relationship between a VI individual's level of sight and how difficult they find the AR testing app to use. This contrasts with content creators, who were mostly sighted and found the app both easier to use and more enjoyable, which could because of their jobs and higher level of previous experience with XR.

Whilst content creators mostly suggested minor improvements to how the app worked, VI participants highlighted the vital importance of spoken cues and screen readers in particular. This thread ran through their overall feedback of the app, the sounds they thought were helpful, the additional sounds they thought would be helpful and the audio ranking task. This is consistent with the results of Study One. The suggestions content creators made are perhaps representative of the challenges VI users would face if the app had a base level of accessibility for VI users.

Both VI participants and content creators were positive about the accessibility strategies suggested. There was much support for the idea that well-designed audio could create a more enjoyable experience for everyone, but also aid accessibility for VI users. It was particularly noted that audio can be used to highlight state changes, interactions, the appearance of new things and information on tasks that need completing. Content creators were particularly enthusiastic about the use of sound effects and spatial audio to achieve these means. VI participants agreed with this, but with a priority of first having well-designed spoken cues. They were particularly supportive of the idea of having a guide character that was designed specifically for the AR app, able to provide both screen reader functionality and hints within the app in a real human voice.

Both VI participants and content creators were supportive of the idea of having a control to allow them to highlight important sounds and filter out unimportant ones. There were some concerns that the simplicity of it could take away control from users, but *'as long as it's informed by the people you're making the decisions for'*, the simplicity of it aids accessibility over many controls being given. Content creators did not think the idea was too different from their own workflows and thought it would be simple for them to implement.

Finally, it was interesting to note that across the sections dealing with helpful sounds in the app, additional sounds that could be added and the audio ranking task, there was little distinction between how sounds could help with navigation or understanding the story. In Study One, this was explained by the supreme need for screen readers to access AR apps at all meaning that other sounds were irrelevant by comparison. However, the fact that this also seems to apply to sighted content creators means that further thought is needed. Perhaps the two are simply so interwoven in the AR medium that the impact a sound, or lack of a sound has is as important to one as it is to another. Without comprehending the narrative, understanding the meaning and purpose of navigational tasks is impossible. Similarly, without understanding the how, where and what of navigation tasks, exploration and progression through a narrative is impossible. This makes sense, as they are attempting to capture both the interestingness and interactivity elements of an immersive system, as defined by Hyunkook Lee [112]. In this model, all the attributes of the system are intrinsically linked, impacting the immersive experience of the user. However, even if they are deeply linked, responses to the ranking task and interviews seem to indicate that considering both is important, especially when considering adding, altering or allowing users control to specifically create more accessible experiences.

4.8 Chapter Summary

This chapter has addresssed the third objective of this work, developing a mobile-based AR testing app. In doing so, it has begun to address the fourth objective of this work, which is to assess the efficacy of the enhanced and personalisable audio accessibility strategies proposed. It has done so by meeting the 2nd and 3rd sub-objectives outlined in Section 1.2.2.

The third main objective was to create a mobile-based, narrative AR app as a testing platform with accessibility comparable to currently available AR apps. This app is outlined in Section 4.2:

- The AR testing app is an adaptation of a narrative VR experience called 'The Turning Forest' [164].
- Unity's AR Foundation [178] was used to develop the AR testing app for both iOS and Android.
- As with most currently available mobile AR apps (see Chapter 3), the AR testing app supports OS built-in magnification and colour changing settings, but does not support OS built-in audio description (screen readers) or subtitles. Some spoken cues are provided by the narrator, but this is in no way equivalent to screen reader support.

The second of the sub-objectives was to confirm the level of accessibility that the testing app has, through an initial user experience study. Study Two evaluated this, particularly in Section 4.5.2, with the following results:

- VI participants found the AR testing app to be as difficult to use as existing AR apps.
- Screen reader functionality was once again the most important feature needed to make the AR app accessible.

The third of the sub-objectives was to obtain initial feedback on VI and content creator perceptions of the accessibility strategies proposed, as well as allowing discussion of other improvements and ideas. Study Two evaluated this, particularly through the survey in Sections 4.5.3 to 4.5.5 and through the interviews in Section 4.6. These gave the following results:

- Both content creators and VI participants were positive about well-designed audio being used as an accessibility tool.
- Although other sounds and spatial audio may strongly aid it, spoken cues are of the highest priority for the AR testing app to be made accessible to VI users.
- Sound design needs to not overwhelm, but highlight state changes, interactions, the appearance of new things and information on tasks that need completing. This can be aided by user control, or by exploration within spatial audio.

Chapter 5

Creating an Accessible AR Application

5.1 Introduction

This chapter presents the changes made to 'The Turning Forest' Augmented Reality (AR) testing app, based on the feedback given by participants in Study Two (see Chapter 4). Creating an updated app with enhanced audio features allowed for the final study to take place (Chapter 6), so that the fourth objective of this work could be completed, assessing the efficacy of the accessibility strategies proposed (Section 1.2.2). To support this chapter, Appendix G contains comprehensive documentation of the changes made. It includes scripts, documents with further details of changes made, screenshots of both the Unity project and the final updated AR testing app and walkthrough videos of the updated app.

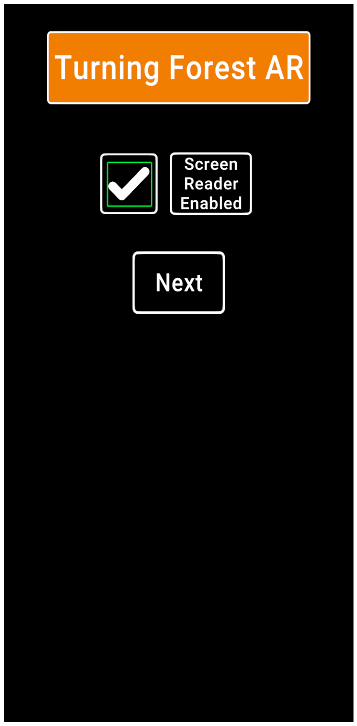
5.2 Enhanced Audio Features

Participants were very positive about the potential of audio to create AR that was both more accessible, but also more enjoyable for all users regardless of their level of sight. The enhanced audio features added to the app are examined in turn in the following sections. It's worth noting however that they work best in tandem.

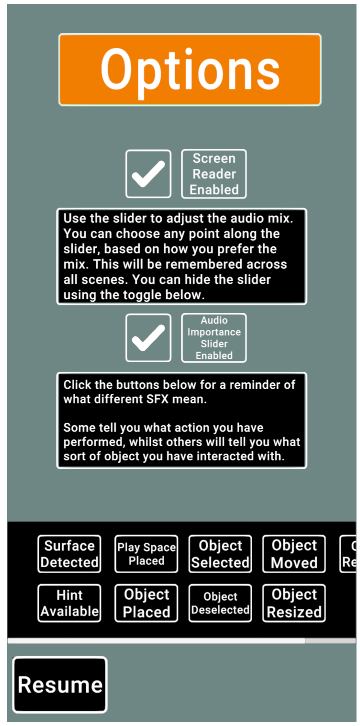
5.2.1 Guide Character

Participants were particularly enthusiastic about the addition of more spoken cues, both to add screen reader functionality but also to provide further hints within the AR scenes. They were also positive about this being designed as a character within the app, which tried to fit in with the narrative and also used a real human voice that complimented the feel of the app. These ideas are derived from the use of 'i-voice' by the Enhancing Audio Description (EAD) project [19, 27–30] to create spoken cues that felt part of the narrative and encouraged listeners to identify with the describer. However, given the non-linear and interactive nature of AR, alterations have been made to the original format. See Section 2.4.2 for further information on EAD methods.

To this end, a 'guide character' was added to the app. Even if users didn't use the screen reader functionality, it was still available, guiding them through the tutorial (see Section 5.4) and providing hints in the AR scenes, as shown in Figure 5.1c.



(a) Start screen



(b) Options Menu

(c) Hint button and personalisation control within an AR scene

Figure 5.1: Screenshots of new features in 'The Turning Forest' updated AR testing app.

To create the spoken content of the guide character, a list was drawn up of all the user interface elements (for screen reading) and all the tasks within the AR experience. This determined what needed to be communicated. Several further considerations were used to develop the guide characters script:

- Ensuring that the character fitted with the narrative of the app. This was achieved by: First, providing the guide character with a back story as the sister of the existing narrator, which was explained in the tutorial and alluded to in other scenes. Second, using friendly and informal language for hints. For example, many hints are structured as questions or suggestions, such as 'Can you hear something? Let's search the forest.' or 'I'll leave you to keep on exploring. When you're ready to find out what happens next, just press the next button'.

- Making sure that users felt comfortable using the screen reader functionality for user interface elements. For these, the structure of spoken content followed a more standard screen reader format, providing detail first on what the name of the element, then following this with a hint on what it is and how to interact with it. These hints followed the suggested format of beginning with plural verbs so that they do not sound like an imperative [179]. For example, 'Next scene *pause for 3 seconds* Button, double-tap to activate.

- The guide character needed to particularly help with locating objects, as many of the tasks to be completed within the scenes involved the user locating and then selecting the relevant AR object. Both the structure of the hints and the use of spatial audio to place the guide character in the intended location were used to achieve this.

- Numerous participants suggested that they would like to be able to try exploring themselves, but be able to ask for more help if they got more stuck. Therefore, for each task within the scenes, the guide character needed to provide multiple hints. These began by providing a simple prompt of what the expected task was, then if the user tapped the hint button again, it would add more details. For example:

  1. On first press - 'What happens if you touch the creature?'
  2. On second press - 'Can you hear the creature breathing? What happens if you go up and touch it?
  3. 'Come over here! The creature is still, just staring at us. Follow the sound of its breath and of my voice. When you get close, touch the creature to see what happens.'

These considerations and the list of needed content were used to create a script for the guide character. This was then recorded by a voice-over artist, Katherine M Tucker, who also helped create the script, particularly by suggesting methods to make the character fit in with the narrative, feel more friendly and also get information across more succinctly.[1] Details of how to access this script can be found in Appendix G.

Once the audio was recorded, it was then added to the app. For screen reader content, see below. For the rest of the content, every scene had a 'hint manager' script which ensured that the correct audio clip was either being played automatically, or attached to the hint button to be triggered by users when needed. When relevant, these scripts set the spatial location of the guide character. It also ensured that hints would not play at the same time as the narrator was speaking to avoid confusion.

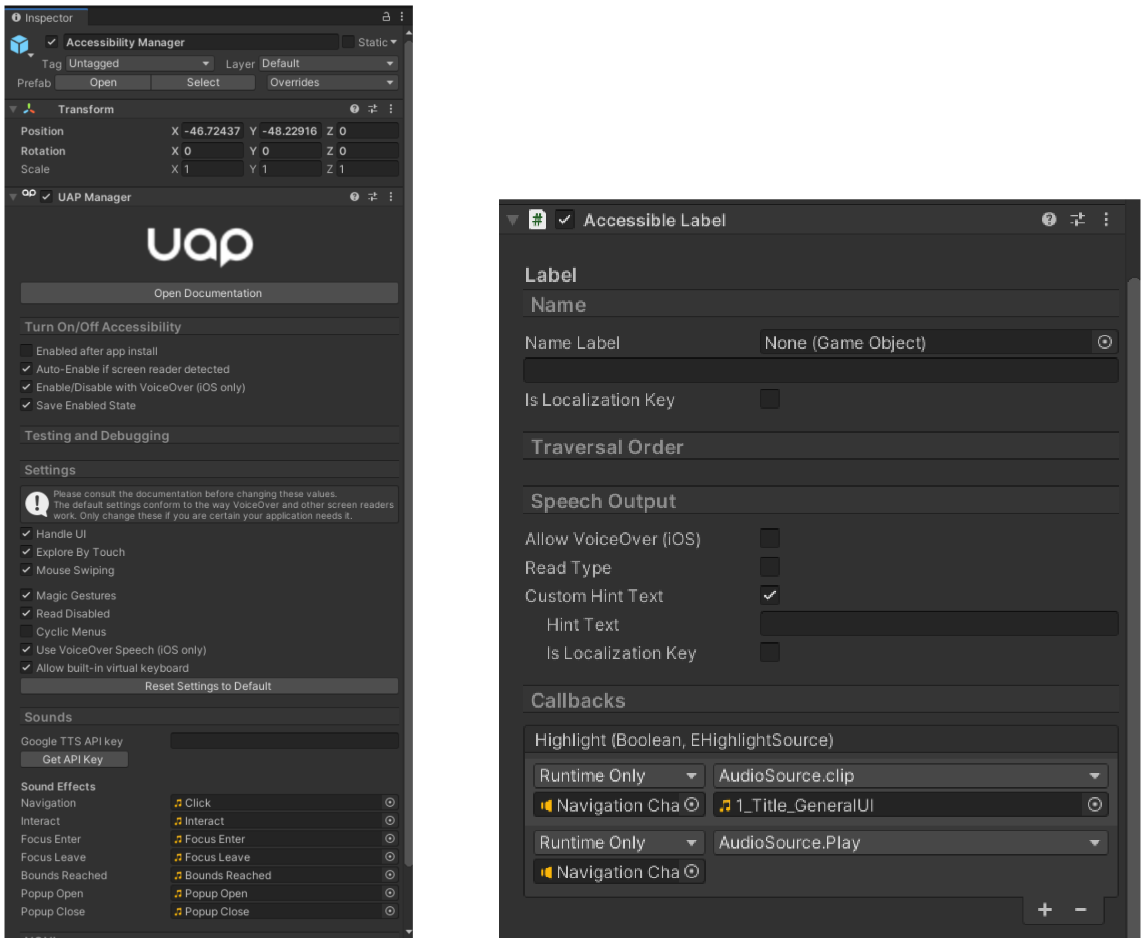
Screen reader functionality

As Unity does not provide built-in support for enabling screen reader functionality on mobile devices, a third-party solution had to be found. Metalpop Games have developed the User Interface Accessibility Plugin (UAP) [180]. This Unity package allows the developer to implement a custom screen reader for both Android and iOS devices. As far as the author is aware, no other tools exist to provide developers with the means to implement screen reader functionality in Unity.

How UAP works By adding the relevant accessibility component (i.e Labels, Buttons, Toggles, Drop Down Lists, Sliders, Images and Input Edit Fields) to all elements within a scene, as shown in Figure 5.2b, an 'Accessibility Manager', shown in Figure 5.2a can then access and read them out. The accessibility manager uses the information provided in the components to determine the hierarchy of elements on the screen and therefore in what order they should be read. Elements are grouped by an 'Accessible UI Group Root' put at the root containing all the user interface elements. By default, UAP uses text-to-speech functionality built into the operating system it is used on, so will sound just like the OS built-in screen reader. The developer can customise the exact method the screen reader works as shown in Figure 5.2. By default, users can explore in the same manner as default screen readers, either

[1] The voice-over artist can be found at `https://www.katherinemtucker.com/`

by directly tapping on the element wanted or by swiping through the determined hierarchy in order. For the AR testing app default functionality like this was kept. However, UAP does not always do this perfectly. At times, the option to directly tap does not work and the user is forced to navigate using only gestures.



(a) UAP [180] Accessibility Manager

(b) Custom voice-over added to a user interface element

Figure 5.2: Unity screenshots showing implementation of the Guide Character for the updated AR testing app.

How users activate UAP. Given that UAP is not a built-in screen reader, users are not able to use both at the same time. How this therefore works is determined by the operating system used:

- On Android the built-in screen reader (Talkback) works as an overlay over the whole screen, blocking all touch input and deciding what to let through, which prevents apps from getting around it [180]. This is because non-physical home buttons are commonly used, so if an app disabled TalkBack the user would be unable to leave the app. UAP automatically detects that the user is using Talkback when they open the app. To use the app, they, therefore, have to turn Talkback off temporarily.
- On iOS, the built-in screen reader (VoiceOver) does not work as an overlay, in part because many iPhones still have physical home buttons. Previously, it was possible

for iOS to automatically detect UAP and hand over management of screen reader functionality to it. However, iOS 14.0 changed the default behaviour for non-native user interface views. Therefore, UAP does not automatically work on iOS any longer. Users who open the app will therefore be confronted with the phrase ‘Direct touch area. Use the rotor to enable direct touch for this app’. Rotor is a function within VoieOver that allows the user to interact with apps without VoiceOver interfering, meaning that customer screen readers can pick up inputs.[2] UAP still picks up on the fact that VoiceOver is enabled and automatically enables itself.

Customising UAP. So that custom voice-over can be used, UAP had to be altered to not call upon text-to-speech functionality to read out the labels of elements. Instead, when the element is highlighted, the relevant recorded content is passed to the guide character and played. This is shown in Figure 5.2b.

Problems with UAP. The fact that the screen reader requires some setup for both iOS and Android is not ideal. This has the potential to cause problems for users. When testing, it also seemed that UAP did not always detect that screen readers are in use, on either Android or iOS. To help mitigate this, the screen reader in the app is set to always be turned on by default. The first screen the user is presented with shows this and allows them to choose whether to turn it off, as shown in Figure 5.1a. Users having to turn screen reader functionality off, rather than struggling to turn it on was suggested by a content creator in Study Two.

A further problem found was that although UAP is supposed to detect the user directly selecting an element, it sometimes does not and only lets the user navigate with swiping gestures. Although this is unfortunate, these gestures are part of built-in screen readers, so users who rely on them are used to them.

Although screen reader functionality can be used as a method for the user to select AR objects, this was not used. Doing so would essentially treat them as another user interface element, meaning the user can simply swipe between them. This does not seem to allow for intuitive exploration of the virtual world created. This choice had the potential to make the app much more difficult for those who fully rely on screen readers. However, it was hoped that the enhanced audio features outlined in this chapter would allow a more intuitive, natural exploration of the virtual AR world, which is more similar to the experience afforded to sighted users. However, in a method similar to how screen readers highlight the currently selected element, all AR objects are visually highlighted when selected. The use of sound effects to inform the user what they are interacting with and how they are interacting further seeks to provide an enhanced replacement to the use of screen reader functionality with AR objects.

5.2.2 Spatial Audio

Participants were enthusiastic about the idea that spatial audio could provide them with a way to explore the virtual world of the app intuitively. In some ways, it allows users a method of prioritising what sounds are important to them by the way that they turn towards and focus on interesting elements.

Google’s Resonance Audio Software Development Kit [181] was used to add this functionality to the app. It enables the creation of highly customisable high-quality binaural

[2] More can be found on this here: `https://developer.apple.com/forums/thread/663529`

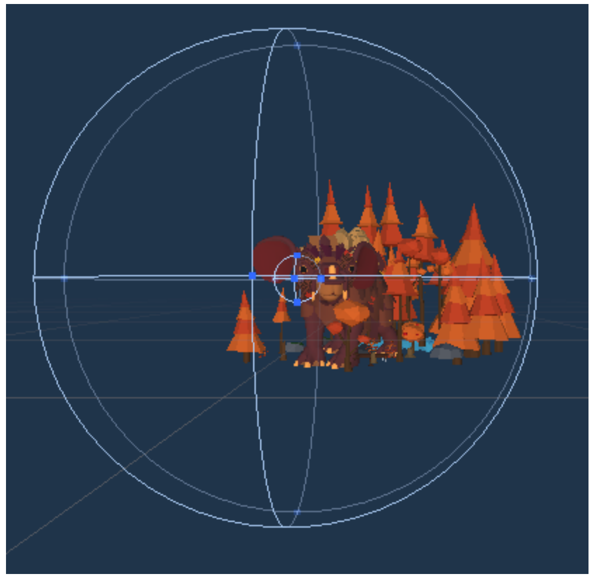
soundscapes but is also optimised for the limited computational resources of mobile devices. It is compatible with a number of development platforms and audio tools, including those which may be used for AR creation including Unity, Unreal, FMOD, Wwise and Android or iOS development suites [182]. The one downside of using it in Unity without any of the more advanced audio integrations (such as FMOD or Wwise) is that it requires routing of the outputs of all audio sources to the master bus of the 'ResonanceAudioMixer'. This means that buses can not be used to sub-mix audio sources, which affects the method used to implement the personalisation control described in Section 5.3.

With the exception of the music and the forest ambience (which was already head-locked binaural), all audio sources were spatialised. This includes the guide character, who was placed next to the objects that she was guiding the user towards when relevant. Figure 5.3b shows the process for spatialising an individual game object. Much of the workflow for implementing this feature focused on ensuring that each object was spatialised accurately by altering the spread, volume roll-off and min/max distances at which an object could be heard. Figure 5.3a shows an example of the final implementation of this for the creature's breathing. The way spatial audio was implemented was almost hyper-realistic so that it could aid the accessibility of the app. Particular attention was paid to the maximum distance that the user could hear different objects, attempting to prevent the soundscape from becoming overcrowded while still giving the user information about other elements being around them. It needed to be obvious that an object was becoming more distinct as the user got closer to it to help guide them towards it and focus on any interactions. But at the same time, users needed to remain aware of other objects around them.

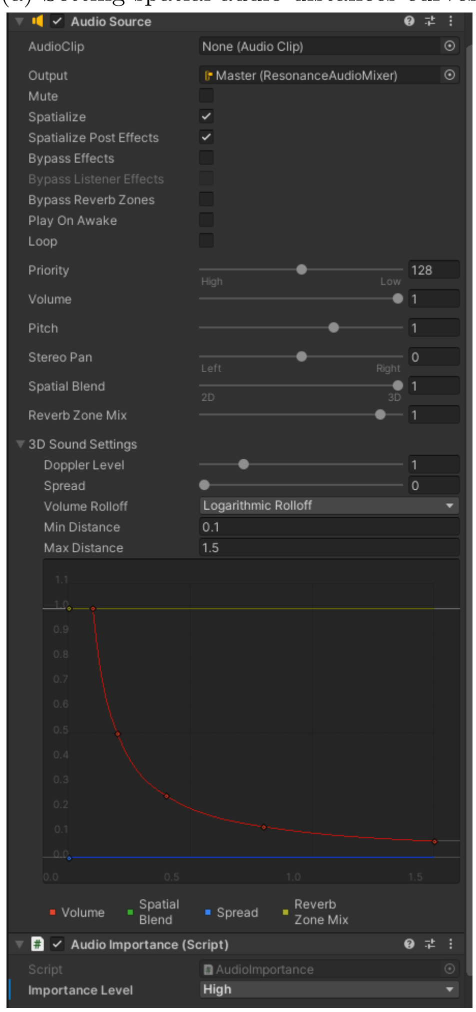
5.2.3 Sound Effects

After spoken cues, participants rated the current and possible additional use of sound effects the most helpful. Given the large number of sounds already in the app, ensuring that sound effects enhanced the accessibility and weren't confusing was important. The spatial separation afforded by the use of binaural audio helps with this by placing sounds in distinct locations (see Section 5.2.2). Also, several features were added to aid users' understanding of the sound effects and simplify the soundscape when needed, as described below. However, the main way of preventing confusion was by using deliberate sound design, utilising the sounds already in the app and only adding sounds that were necessary and helpful. This means that many of the sounds already in the app were simply boosted (by around 9dB) so that participants could pick them out more easily. One key example of this is the creature breathing, which some participants had already mentioned helping them to locate the creature. Other examples include the creatures growling sounds or the musical notes and tweeting made by birds, both of which are triggered when users select them.

Visual cues already existed in the app to tell users if they had collided with an AR object, shown in Figure 4.5b. Audio cues were added to make this more accessible. First, during the collision all sounds are low-pass filtered, to give the user the feeling of being in a constricted space that they should leave. Second, the guide character gives the hint 'Oops! Looks like you've walked into something *(pause for 3 seconds)*. Maybe try moving back a bit'. This hint is not low-pass filtered so the user can easily understand what is being said. If the guide character is currently speaking out screen-reader content, this is stopped, but if the guide character is giving the user a hint about how to complete an AR task, this is paused and then continues when the user moves back from the AR object. Details of how to access a Unity Package to replicate this functionality can be found in Appendix G.



(a) Setting spatial audio distances curves



(b) Resonance audio settings.

Figure 5.3: Unity screenshots showing changes made to the spatial audio for the updated AR testing app.

Table 5.1: Sound effects added to the AR testing app. All extra sounds were sourced from Zapsplat [183]. Details of how to access a document with links to all sound effects added can be found in Appendix G.

Item	Sound effects added
Interactions	
Surface detected	Chime like sound, continues until first object or playspace is placed
Object placed	Impact sound of a tree being set down
Playspace placed	Louder impact sound, a sandbag being dropped
Object selected	Low pitched tone
Object deselected	High pitched tone
Object removed	An axe chopping wood
Object moved	A tree being dragged across the forest floor. For rocks this was changed to the sound of rocks scraping against each other
Object resized	A tree branch being twister to create a creaking sound
Object specific	
Trees and plants constant	Various creaking, rustling sounds
Trees and plants selected	Various rustling sounds
Rocks constant	Pebbles falling on a hard surface
Rocks selected	A stone being dropped into a rock pool

Table 5.1 shows the sound effects that were added to the app. These fit into two categories. First, interaction sounds were added to all objects regardless of what sounds they already had. These can be thought of as being ‘earcons’, a sonification technique discussed in Section 2.3.1. They were consistent across the whole app so that the user could learn to identify if and what type of interactions they were undertaking. Where possible, sounds were chosen that intuitively link to the action being performed. Second, object-specific sounds were added. These can be thought of as ‘auditory icons’, another sonification technique discussed in Section 2.3.1. All object types had an identifying sound that played constantly, meaning that when paired with spatial audio, the user could locate and identify types of objects around them. Now, a constant tree creaking and rustling sound could guide users towards a specific tree as well. As an additional layer to the selected interaction sounds, different types of objects had different sounds that played when they were selected so that the user could identify what they were. The creature growling was an existing example of this, but now an element like a rock would also make an identifiable sound. There are of course difficulties in choosing sounds that are distinct and recognisable for very similar elements within a forest scene, such as using different rustling sounds to identify one type of tree from another. However, the differences between different groups of objects (i.e. the creature, trees, birds, rocks) afford enough variety in themselves for an interesting virtual sonic world.

Several features were added to assist users’ understanding of the sound effects:

- Several participants suggested having a key to explain what different sound effects meant. This was added to the options menu, which was easily available in all AR scenes (see Figure 5.1b). It included interaction sounds, as shown in Table 5.1. Object specific sounds were not added, both to prevent the number of sounds in the key from becoming overwhelming and because they were intentionally designed to be intuitive.

- Explanation of how sound effects are used is given by the guide character in the tutorial (see Section 5.4), including allowing users to explore the sound effects key as part of the tutorial (see Figure 5.4b).

- A script was added that mutes all other sound effects when a particular object is selected so that the user can focus on how they are interacting with that object. This uses Unity's game object 'Tags' functionality to search for all other objects in the scene tagged as containing sound effects and stops their audio sources from playing any sounds until the selected object is deselected again. Details of how to access a Unity Package for this can be found in Appendix G. The general ambience, music and spoken cues from the narrator and guide character continued whilst interacting with the object.

5.3 Personalisation Control

Participants were positive about the idea of having a single control with which they could personalise the audio mix of the app. This was derived from Ward's Narrative Importance (NI) work, described in Section 2.3.1.

Creating this was a two staged process. First, finding an easily implementable method for coding the functionality into the app. The design considerations and final implementation of it are described in Section 5.3.1. Second, the considerations of participants needed to be taken to account in setting how the control affected the mix. How this was achieved is shown in Figure 5.3.2. Figure 5.1c shows the control in place in the app during an AR scene and Figure 5.4c shows the control being explained in the updated app tutorial.

5.3.1 Creating the Personalisation Control

Before creating the control, there were several considerations to take into account:

- Ease of implementation. It needs to be easy to apply to a complex AR app with many audio objects, both for use within the 'Turning Forest AR' application but also if this coding framework were to be adopted by future content creators. To aid this it needs to have:
  - Minimal initial set-up and additional coding.
  - An easy method for choosing how important the different sounds within the app are. Although there is an in-built 'Priority' level, which could be referenced, this may already be used by developers to determine which audio sources are virtualised first if the number of audio sources is greater than the available audio channels. This is also an 0-255 integer, so asking users to choose a number along this scale isn't as easy as letting them choose from a simple drop-down menu with the 4 levels of importance as previously defined (see Section 2.4.4).

  - Must have the ability to work across multiple Game Scenes, so that the user can set and forget their desired mix.

- Compatibility with existing workflows
  - The main consideration for this was at what point in the audio path changes made by the control should be implemented. It was decided to be implemented on the individual audio sources, rather than on groups used the in the Unity audio mixer. This also means that is compatible with Unity without the need for any extra audio tools such as Wwise or FMOD. Another deciding factor for this was that Unity only allows users to send audio to one mixer group, without further sub-mixing level changes applied by any other buses. Applying the personalisation level to each audio source means that developers do not have to change their current mixing workflows but can continue using existing mixing groups. Further, Resonance Audio, which is used for the spatial audio in this and many other apps requires that all audio is only sent to the Resonance Audio master, without the use of any mixer groups. This means that using mixer groups for each importance level would not be compatible with Resonance Audio.
  - Flexibility. It needed to be easy to refer to when scripting so that the importance levels of different audio sources can be altered depending on their importance in a given moment. It also needed to be easy for developers to potentially change the look of the actual control to fit in with their projects.

- Processing power
  - Scripts created for the personalisation control should not use costly functions like ‘update’, which run every frame.
  - If the then importance is going to be chosen for each audio source, extra consideration needs to be put into ensuring the script that does this is not costly.

- Effectiveness
  - It needs to be easy for the end-user to use the control.
  - The effect of the control needs to be both noticeable and sound good. This was of particular concern as neither default sliders nor the levels applied by audio sources to audio in Unity are logarithmic. This was solved by careful calculation of the functions which create multipliers to the level of the audio source.

After considering all these points, the personalisation control was implemented using the following C# scripts:

1. **Slider Controller**. This script works when attached to a slider component in Unity. For the second script to work, the slider must be tagged in Unity as an ‘ImportanceSlider’. When the scene is first opened, the slider finds if there is an ‘Access Level’ assigned in Unity’s ‘Player Prefs’ and sets the slider to that value. Using the ‘Player Prefs’ class allows the set value to be stored across all scenes and between different game sessions. Therefore, the level set by the user will stay consistent throughout the app regardless of which scene the user is in. If the player alters the value of the slider, this will change the value of ‘Access Level’, in the range of 0-1.

2. **Audio Importance**. This script can be applied to any Unity game object which also contains an audio source. It essentially applied Ward's narrative importance equation (Equation 2.1) to the level of the audio source. It first searches to find any game object tagged an 'ImportanceSlider'. This means that the update function can be avoided, as functions to the level of the audio source can be called when the slider is changed, using Unity's built-in 'ValueChangeCheck'. function. Multipliers for each importance level are defined in Section 2.4.4. As a linear range of 0-1 is used on Unity sliders, these were adjusted from logarithmic values. The script sets up a drop-down menu to be added to the Unity game object so that the user can select the importance level there, as shown at the bottom of Figure 5.3b. The chosen level is also public, so can be altered by other scripts. When a user changes the slider, the value of the slider (N) is multiplied by the importance level of the object (M), setting a value between 0 and -1, which is then added to the original level of the audio source, attenuating it by the needed amount.

An option was also added to the options menu so that users could hide the control to prevent it from being distracting.

A positive of the personalisation control being coded in this way is that because the developer has to add it to every audio source and set the level, it forces them to think about the importance of every audio object as they add them. The downside of this is that it may prove time-consuming to manage in complex scenes. However, the fact that each 'Audio Importance' script automatically searches for the slider means that the only set-up needed is to choose the importance level of the object. Another downside of this implementation is that use of Unity's built-in audio source and slider components limits the control to linear, not logarithmic changes. This means that changes made will not sound as natural as the user slider across the slider.

Appendix G provides details on how to access this personalisation control through two Unity packages. The first contains a game object with a slider and the 'Slider Controller' script set up. The second contains a game object with an audio source and the 'Audio Importance' script set up.

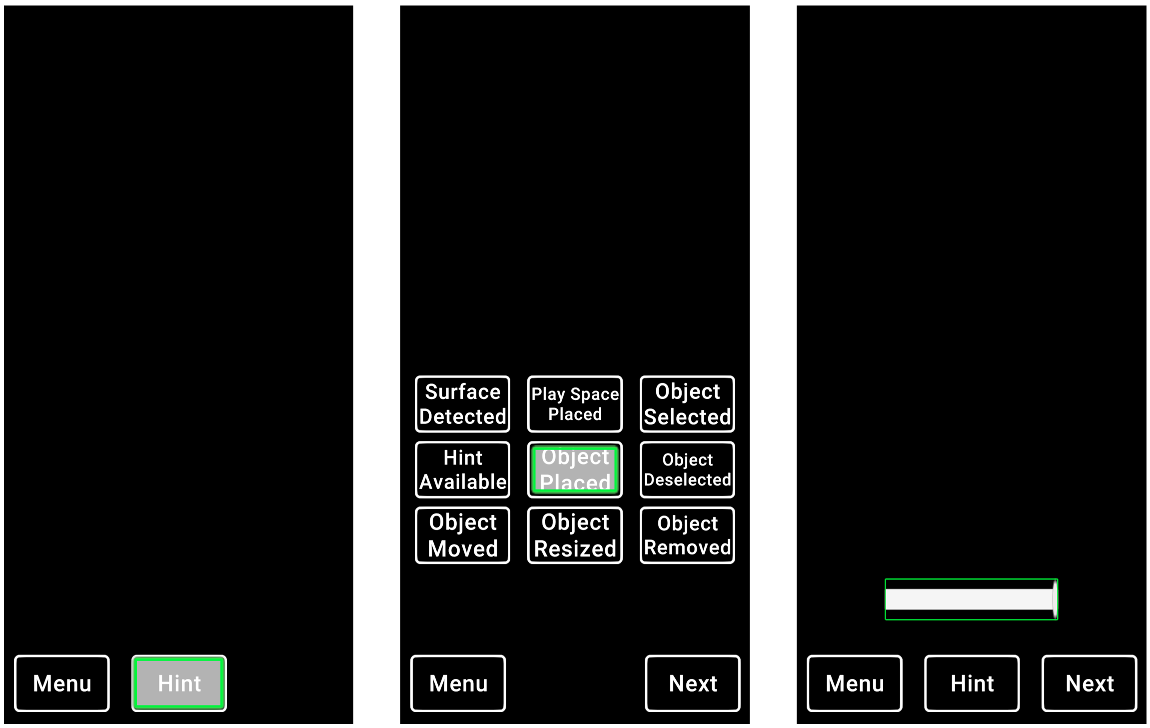
5.3.2 Setting Personalisation Control Levels

Once the control was created, the importance level of each audio source needed to be set. To do so, the feedback from Visually Impaired (VI) participants (Figure 4.20) and content creators (Figure 4.21) in Study Two was combined with the feedback Ward obtained from the original sound designer (Figure E.1) and a study with content creators (Figure E.1) in [32]. Not all sounds were rated by all of these groups, so the following weightings were given to their choices:

- For sounds that were rated in Study Two, the narrative importance level and navigational importance levels of VI participants and content creators were each given 20% weighting, meaning 80% of the weighting was given to those who had used the sounds in an AR context. The remaining 20% was split evenly between the original sound designed and content creator ratings obtained by Ward [32].
- For sounds not rated in Study Two, the original sound designer and content creator ratings obtained by Ward [32] were weighted 50% each. However, the final choice was compared to similar sounds rated in Study Two to ensure they matched up with choices

made by those who had used them in an AR context. However, as the ratings obtained by Ward were often similar to those in Study Two no further changes were needed.

These choices were then implemented to all sounds in the app. A full spreadsheet with these weightings and the final level applied to every sound in the app can be found via Appendix G.



(a) Explaining how to access hints

(b) Explaining sound effects

(c) Explaining the personalisation control

Figure 5.4: Screenshots of the tutorial in 'The Turning Forest' updated AR testing app.

5.4 Tutorial

To help users become familiar with the newly added features, the tutorial previously in the app was expanded (see Section 4.2.1). The tutorial now follows the following course:

1. The guide character introduces themselves. They explain to the user how the hints work, getting them to access a hint to move on, as shown in Figure 5.4a.

2. The concept of spatial audio is introduced to the user, with the guide character moving around to demonstrate this whilst speaking the phrase 'as this app uses 3D audio, I'll move so that the location of my voice will help you find your way'.

3. The sound effects are introduced, including the idea of their being sounds that give information on interactions and sounds to provide information on what each object is saying 'you will hear these sound effects where the object is, and additional sounds might tell you what it is. For example, this combination tells you that you have

selected a tree'. Users are then allowed to explore the key giving details on what different interactions mean, shown in Figure 5.1b.

4. The personalisation control is then explained and a soundscape is played so that users can have a go at using it to set a mix for themselves.

5. Users are then guided through placing their first AR object. This follows the same structure as the original tutorial detailed in Section 4.2.1, but with the guide character and all other enhanced audio features added.

The script of the guide character during the tutorial and a walkthrough video of it can be found via Appendix G.

5.5 Chapter Summary

This chapter has presented the changes made to 'The Turning Forest' AR testing app, based upon the access strategies this work has proposed in Chapter 2 and the feedback of VI individuals and content creators in Study Two (Chapter 4). These changes included

- A guide character who provides both screen reader functionality and additional hints in a real human voice, using a script that is designed to fit in with the narrative of the app. The hints provided develop as the user asks for more help.
- Spatial audio using Google's Resonance Audio [181].
- Additional sound effects to both aid interactions and the users' recognition of objects within the app.
- Further advanced sound effect features which include low-pass filtering and hints on collision with AR objects and the turning off of all other sound effects once an object is selected so that the user can focus on the object.
- A simple personalisation control, which allows participants to alter the mix of the app so that they can focus on the most important sounds easily. This remembers their mixing choice across all app scenes and different app usage sessions. It is designed using two lightweight, simple scripts that could be easily implemented in any Unity project without the need for additional software tools.

Chapter 6

Study Three - Evaluating the Accessible AR App

6.1 Introduction

This chapter presents the culmination of the work done to meet the fourth objective of this work, assessing the efficacy of the enhanced audio personalisable audio accessibility strategies proposed (Section 1.2.2). The Augmented Reality (AR) testing app first presented and assessed via Study Two in Chapter 4 was updated to include dedicated accessibility features, as described in Chapter 5. These include a guide character with screen reader functionality, spatial audio, extra sound effects, a personalisation control and an expanded tutorial.

In this Chapter, Study Three asks participants to assess how successful these updates have been in making the app more accessible to them. As with Study Two, both Visually Impaired (VI) users and content creators were asked to participate, with a similar format of both surveys and interviews used. Many of the questions used are similar to those in Study Two, so that changes can be easily compared. Further additional questions are also used to assess the features added in detail.

Study Three thus provides an initial evaluation of how effective the accessibility strategies proposed are, answering the research questions that this work has posed (Section 2.7.3).

6.2 Study Three Methodology

This section outlines the design of Study Three, including what information was given to and collected from participants. Full ethical approval was obtained from the University of York Physical Sciences Ethics Committee. This can be found in Appendix A. Details of how to access the full survey instrument can be found in Appendix F. Participants for this study were recruited through the same networks utilised in Studies One and Two, which are detailed in Section 3.5.1.

6.2.1 Study Design

Study Three was intended to be as similar as possible to Study Two so that they could easily be compared. Much of the study design is the same as previously laid out in Section 4.3. Throughout the survey and interviews the 'sounds that provide you with cues for navigating' and the 'sounds that provide you with cues to help you with understanding the narrative, or story' are considered separately. This helps delineate how effective accessibility solutions have been in the previously established dual purposes of soundscapes in AR (see Chapter 2).

Distributing the app

Participants were first provided with a Qualtrics link to an information sheet on the purpose of the study and what participation involved. The main instructions given to them are shown in Table 4.2. In the same way as Study Two, this initial link led to questions collecting what OS their mobile smartphone uses, the email address they use on their app store and whether or not they are a content creator. This allowed for the correct distribution of the app via either Google Play Console [172] or iOS Testflight [173].

For those who wanted to use a screen reader, further instructions were provided on how to make the screen reader in the app work. On Android, this involved turning TalkBack off temporarily. On iOS users had to activate 'Direct Touch'. The app screen-reader should auto-detect that TalkBack/VoiceOver was used when it was first opened and thus turn itself on as soon as set up was complete.

For any user who has played games, or used apps with their own built-in screen reader these processes would likely be familiar. As not all participants may have had this kind of experience, detailed instructions were provided. As turning off or altering their screen reader in such a way may be unnerving for users who rely upon it, reassurance was provided to them and further technical support over email if needed.

Survey

Separate Qualtrics surveys were distributed to VI users and content creators. The questions asked are detailed in Sections 6.2.2 to 6.2.6. Content creators were asked all the same questions as VI participants, with additional questions on their professional experience, shown in Section 6.2.2 and about integrating the audio accessibility features like those used in the app into their own workflows, shown in Section 6.2.6.

Interviews

After filling out the survey, participants were invited to an interview over Zoom. A final Qualtrics link was used to obtain consent for the interviews and find what times participants were available. Details on the format of the interviews can be found in Section 6.2.7.

6.2.2 Demographic and Past Experience with AR

All participants

As in Studies One and Two, B1-B4.1, shown in Table 3.4, were asked to establish some basic information on the demographic of the participants. As the study was advertised to participants who may not have already taken part in the previous studies, Questions B6-7 were also asked to confirm which previous studies participants had previously taken part in. As the studies took place over the course of a year, it was possible that participants unique identifiers (last 3 digits of phone number/postcode) could have changed, so this meant that linking of information across studies could be confirmed. B5 assessed whether participants were already familiar with the 'Turning Forest AR'. This question was most important for new participants or those who had not taken part in Study Two.

As in Studies One and Two, questions XR1 & XR2, shown in Table 3.5 were used to determine participant's previous experience with Extended Reality (XR). Questions XR3 was also used in this study alone, to establish whether their perception of XR had changed now that participants had used the updated app.

Additional questions for Content Creators

As in Studies One and Two, additional questions were asked of the content creators. These are shown in Table 4.3. Questions PRO6-PRO8 were not used in this study. They were less relevant, as similar questions on their workflows go more in-depth into how compatible personalisation is with their workflows (Section 6.2.6).

6.2.3 AR App Experience

Table 6.1 shows the questions participants were asked about their general experience using the app. Table 4.4 shows those also used in Study Two and the similarities to Study One. Questions UX2-3 would have been helpful to have in Study Two, so were added. They allow a clear picture of what scenes the participants used and whether they managed to get all the way through the scenes. Question UX10 was added as it was realised that questions UX6-9 focused on 'using' the app, which may lead to answers given reflecting how easy it was to navigate and interact with the app and not necessarily how much participants had understood the narrative of the app. Although question UX11 is similar to XR3 (see Table 3.5), it was added to focus in more on whether the accessibility solutions themselves had made an impact on the participants' perception of XR rather than any other factors.

6.2.4 AR App Sounds

Table 6.2 shows the questions participants were asked about the sounds in the app. To comprehensively assess the enhanced audio-based accessibility features added, a more in-depth breakdown by each feature was needed. The same questions were asked for each of the features added; sound effects, spatial audio, and the guide character. The guide character had additional questions (GC1-2) to determine whether participants used it as a screen reader and whether using it as a screen reader was effective.

The questions begin with a simple evaluation of whether participants enjoyed the feature or not (SFX1/BN1/GC3). The remaining questions then focus on how helpful the sound feature was across several areas. SFX2/BN2/GC4 asks whether the sound features helped them feel more immersed in the sound. In particular, this is attempting to understand the interestingness element of the immersive system and thus their narrative engagement with the system, as defined by Hyunkook's work [112] (see Section 2.5.5). SFX3-5/BN3-5/GC4-7 focus on the interactivity of the immersive system and thus participants' task/motor engagement. The different focuses of SFX3-5/BN3-5/GC4-7 cover all the building block tasks Herskovitz [18] identified for interacting with AR (see Section 2.6.3). SFX3/BN3/GC5 link with the tasks 'Observing AR content' and 'Establishing Physical/Visual Correspondence' by asking how helpful the sound was for identifying objects. SFX4/BN4/GC6 link with the tasks 'Observing AR content' and 'Creating Virtual Content' by asking how helpful the sound feature was for locating the objects. SFX5/BN5/GC7 link with the tasks 'Activating Virtual Content' and 'Transforming Virtual Content' by asking if the sound feature helped with interacting with the objects.

Table 6.1: Survey items in Study Three addressing key concepts about the participants experience of the AR app including enjoyment and ease of use. Questions with a * were also used in Study 2, so provide easy comparison points.

ID	Questions about app experience	Type of response
UX1*	How long did you spend using the AR app?	Free text
UX2	Which parts of the app did you use? You can select multiple.	Tutorial/Scene 1 The Autumn forest/Scene 2 An Unusual Meeting/Scene 3 Water Rising/Scene 4 The Winter Forest
UX3	Which parts of the app did you complete? Completion means that you pressed the 'next' button to go to the next scene. Except in the Tutorial and Scene 4 The Winter Forest, where you are left to explore until you want to move on via the menu. You can select multiple.	Tutorial/Scene 1 The Autumn forest/Scene 2 An Unusual Meeting/Scene 3 Water Rising/Scene 4 The Winter Forest
UX4*	What sort of headphones did you use for listening to the app? Please provide make and model.	Free text
UX5*	What sort of environment did you use the AR app in? For example, outside or sat down at a table.	Free text
UX6*	Did you enjoy using the AR app?	Yes/No/Unsure
UX7*	Was the AR App difficult for you to use? Please provide a ranking between 1 and 5, where 1 is Easy and 5 is difficult.	1 →5
UX8*	What difficulties did you encounter while using the app?	Free text
UX9*	Do you have any suggestions about how the app could be made easier for you to use?	Free text
UX10	How easy was it for you to follow the story of the AR app? For a reminder of the narrative, see here. Please provide a ranking between 1 and 5, where 1 is Easy and 5 is Difficult.	1 →5
UX11	If the accessibility techniques used in our app were used in other Extended Reality, would it make you more or less likely to use Extended Reality? Please provide a ranking between 1 and 5, where 1 is much more likely to use Extended Reality and 5 is much less likely to use Extended Reality.	1 →5

Table 6.2: Survey items in Study Three addressing how helpful sounds where for navigation and narrative comprehension, and how helpful each enhanced audio feature was. Those marked with a * were also used in Study Two.

ID	Questions about Navigational Sounds	Response
The following questions will ask you about sounds within the application that provide you with cues for navigating through the AR app. Navigation may include finding your way around the app, exploring AR experiences and interacting with AR objects.		
NV1*	Did any sounds within the app help you with navigating?	Yes/No
NV2*	Which sounds were particularly helpful for navigating? You can select multiple.	Sound Effects/Narrator/Guide Character/Music/Ambient Sounds/Other
ID	Questions about Narrative Sounds	Response
The following questions will ask you about sounds within the application that provide you with cues to help you understand the narrative, or story, that the application follows. For a short summary of the story in the scenes you experienced, click here (link to summary of story).		
ST1*	Did any sounds within the app help you with understanding its narrative?	Yes/No
ST2*	Which sounds were particularly helpful for understanding its narrative? You can select multiple.	Sound Effects/Narrator/Guide Character/Music/Ambient Sounds/Other
ID	Questions about Sound Effects	Response
The following questions will ask you about the sound effects within the application.		
SFX1	Did you enjoy the sound effects?	Yes/No/Unsure
SFX2	Did the sound effects help you to feel immersed in the virtual world of the app? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 $\rightarrow$5
SFX3	Did the sound effects help you to identify what type of objects were around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 $\rightarrow$5
SFX4	Did the sound effects help you to identify the location of objects around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 $\rightarrow$5
SFX5	Did the sound effects help you to identify how you were interacting with objects? For example selecting, moving or resizing them. Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 $\rightarrow$5

ID	Questions about Spatial Audio	Response
The following questions will ask you about the spatial audio within the application. For more information on the type of spatial audio used in the app, see here (link to [61]).		
BN1	Did you enjoy the spatial audio?	Yes/No/Unsure
BN2	Did the spatial audio help you to feel immersed in the virtual world of the app? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
BN3	Did the spatial audio help you to identify what type of objects were around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
BN4	Did the spatial audio help you to identify the location of objects around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
BN5	Did the spatial audio help you to identify how you were interacting with objects? For example selecting, moving or resizing them.Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
ID	Questions about Guide Character	Response
The following questions will ask you about the guide character within the application.		
GC1	Did you use the screen reader functionality of the guide character to help you with using the menus and buttons in the app?	Yes/No
GC2	*Only asked if answer to GC1 was 'Yes'* Did the guide character help you with navigating through the user interface of the app? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
GC3	Did you enjoy the guide character?	Yes/No/Unsure
GC4	Did the guide character help you to feel immersed in the virtual world of the app? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
GC5	Did the guide character help you to identify what type of objects were around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
GC6	Did the guide character help you to identify the location of objects around you? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
GC7	Did the guide character help you to identify how you were interacting with objects? For example selecting, moving or resizing them. Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5

6.2.5 Personalisation Control

Table 6.3 shows questions about the personalisation control used in the app. C1 first established whether participants used the control. If they didn't then Qualtrics skipped to the next set of questions. Questions C2-4 were adapted from work done by Lauren Ward, establishing whether the similar personalisation control she used was effective [32]. They attempt to understand how much participants used the control and the extent to which they utilised the accessible mix provided by the control. Figure 6.1 shows the Qualtrics interface participants used to demonstrate what level they mostly left the slider on.

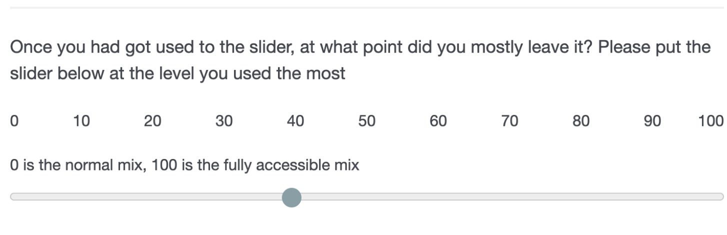


Figure 6.1: Personalisation Control Slider for survey instrument in Study Three.

Questions C5-6 aim to understand what purpose participants used the control for, to understand the narrative or aid navigation and whether it was successful in doing so. C8-9 establish whether the way the control highlighted sounds was effective, which highlighted sounds were helpful and whether sounds could have been highlighted that weren't.

6.2.6 Workflow

Content Creators were asked some additional questions to determine whether the enhanced audio accessibility features used in the app would be compatible with their workflows. These are shown in Table 6.4. As with previous questions, these are split into each of the audio features; sound effects, spatial audio, the guide character and the audio personalisation control. The questions first ask whether they already use such features in their own work (WS1/WB1/WG1/WC1) to establish whether the features are compatible with their existing design processes, tools and workflows. Second, they are asked whether they would consider using the audio features to help with the accessibility of their own work. This gives another measure for the success of the features and an idea of whether content creators would adopt them.

6.2.7 Interview Format

The interviews took the same format as in Study Two. They were undertaken one on one over Zoom by the author. Participants were given the same information, shown in Table 4.9, before recording was started using built-in Zoom functionality. A set of questions were used to guide the interviewer, ensuring consistency between interviews and that all areas of interest were covered. These are shown in Table 6.5. However, the interviews were very informal and participant-led, to allow them the freedom to explore their own experiences and ideas. As with the survey, there was a focus on whether the enhanced audio features helped, with participants able to provide more in-depth reflections on particular points where they did or didn't help. Participants were encouraged to suggest and explain ideas for making AR accessible using the features suggested or their own novel ideas.

Table 6.3: Survey items in Study Three addressing participants use of the personalisation control and how helpful it was. If they answered 'No' to C1, the survey skipped to the end of this question block.

ID	Questions about Personalisation Control	Type of response
The following questions will ask you about the audio control, which allowed you to change the mix of the sounds within the app.		
C1	Did you use the personalisation slider?	Yes/No
C2	What difference did the control make? You can select multiple.	Easier to understand/More enjoyable/Did not change/Less Enjoyable/Harder to understand/Not Sure
C3	Once you had got used to the audio control, how much did you keep on adjusting it throughout the experience?	I left it in one place throughout/I adjusted it a couple of times/I adjusted it continuously/Not sure
C4	Once you had got used to the slider, at what point did you mostly leave it? Please put the slider below at the level you used the most	Slider with range 0 →100 *Shown in Figure 6.1.*
C5	Was using the audio control to change the levels of the sounds helpful for being able to navigate the app? Navigation may include finding your way around the app, exploring AR experiences and interacting with AR objects. Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
C6	Was using the audio control to change the levels of the sounds helpful for being able to understand the story? For a reminder of the narrative, see here (link to summary of story). Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
C7	Were the sounds that the audio control highlighted helpful? Please provide a ranking between 1 and 5, where 1 is Helpful and 5 is Not Helpful.	1 →5
C8	Which sounds were particularly helpful?	Free text
C9	Were there any sounds that were not highlighted that would have been helpful if they had been highlighted?	Free text

Table 6.4: Survey items in Study Three addressing workflow tools for content creators.

ID	Questions about workflow	Response
The following questions will ask you about the sound effects within the application.		
WS1	How different was the use of sound effects to your own work? Please provide a ranking between 1 and 5, where 1 is Not Different At All and 5 is Very Different.	1 →5
WS2	How likely would you be to use sound effects to help with the accessibility of your own work in the future? Please provide a ranking between 1 and 5, where 1 is Likely and 5 is Unlikely.	1 →5
The following questions will ask you about the spatial audio within the application. For more information on the type of spatial audio used in the app, see here (link to [61]).		
WB1	How different was the use of spatial audio to your own work? Please provide a ranking between 1 and 5, where 1 is Not Different At All and 5 is Very Different.	1 →5
WB2	How likely would you be to use spatial to help with the accessibility of your own work in the future? Please provide a ranking between 1 and 5, where 1 is Likely and 5 is Unlikely.	1 →5
The following questions will ask you about the guide character within the application.		
WG1	How different was the use of a guide character to your own work? Please provide a ranking between 1 and 5, where 1 is Not Different At All and 5 is Very Different.	1 →5
WG2	How likely would you be to use a guide character to help with the accessibility of your own work in the future? Please provide a ranking between 1 and 5, where 1 is Likely and 5 is Unlikely.	1 →5
The following questions will ask you about the audio control within the application.		
WC1	How different was the use of an audio control to your own work? Please provide a ranking between 1 and 5, where 1 is Not Different At All and 5 is Very Different.	1 →5
WC2	Having used our audio control, how likely would you be to use a similar control in your own work? Please provide a ranking between 1 and 5, where 1 is Likely and 5 is Unlikely.	1 →5

Table 6.5: Questions used in Study Three to guide the interviews. Questions 1-3 were not asked if a participant took part in interviews previously in Study Two. Where there are multiple questions under the same number, those after the first are follow up questions.

Interview Questions

1. Tell me a bit about yourself, your visual impairment and how you use technology and make technology accessible for you in your everyday life?
2. Have you used AR before?
3. Do you think extended reality is something that is hard for someone with visual impairment to access?
4. What did you think of the AR application?
5. Talk through sounds in the app - repeat these questions for spatial audio, sound effects and the guide character
   (a) Was *category of sounds* enjoyable?
   (b) Did *category of sounds* make it easier to use the app?
   (c) Did *category of sounds* provide you with helpful information?
   (d) How would you change *category of sounds* to make them more helpful to you?
   (e) *Specific to the guide character.* Was the way that you could access information from the guide character intuitive? Would you structure it differently?
   (f) *Specific to the guide character.* Was the amount of information that the guide character provided helpful? Did you still enjoy the challenge of exploring by yourself?
6. Work through Scene by Scene
   (a) How far did you get in this Scene?
   (b) What caused problems?
   (c) Can you think of any particular solutions to the problems you encountered?
7. Did you use the audio control within the app?
   (a) Was the audio control helpful? How did you use it?
   (b) What would you change about the control to make it better?
   (c) Is anything else in the app that you wish you had control over?
8. Do you have any other suggestions for how the app could be made easier for you to access, that you haven't been able to discuss already?

6.3 Survey Results

6.3.1 Demographic and Past Experience with AR

The full aggregated, anonymised demographic of all participants who were recruited across the studies is shown in Tables B.1 and B.2.

Table 6.6: VI participant demographic information for Study Three.

Question (ID in brackets)	Responses	Count	%
Visual Impairment (B3)	Partially sighted	2	16.7
	Blind with residual vision	5	41.7
	Blind	4	33.3
	Other	1	8.3
Familiarity with XR (B4)	I'm familiar with it and have used it a couple of times	5	41.7
	I'm familiar with it, but have never experienced it myself	5	41.7
	I've never heard of it before	2	16.7
Heard The Turning Forest before (XR2)	Yes, more than once	2	16.7
	Yes, once	3	25
	No	8	58.3

VI participants. A total of 12 VI participants were recruited for this study; P1, 3-5, 7 & 10-16. Table 6.6 shows their key demographic information. Out of these, 5 went on to take part in interviews (Section 6.4). None reported having hearing loss (B2). Their ages ranged from 27-65, with a mean of 40 (median 37.5). This study was more weighted towards those with the least sight than the other studies, with 4 (33%) participants self-reporting as fully blind and 5 (41.7%) as blind with residual vision. The participant who reported as 'other' was once again P7, who has strabismus and has been previously established as being towards the more sighted end of participants.

Several participants had taken part in the previous studies of this work. 5 of them (41.5%, P1, P3-5 and P7) took part in Study One and 3 (25%, P1, P4 and P7) took part in Study Two. Two additional participants therefore stated they had heard 'The Turning Forest' before, one only once and one multiple times. Therefore, most participants (8, 58.3%) were unfamiliar with 'The Turning Forest' material, and most (9, 75%) had not used the previous version of the app. This means that most were coming into it without a comparison to the original material, but were approaching this as something new.

Unlike Study Two, some participants (2, 16.7%) had never heard of XR before (B4). Interestingly, 2 participants who had taken part in Study One reported they were familiar with XR but had never used it themselves. This may be because if they rely on screen readers, they may have been unable to use the actual AR functionality of the apps. However, all those who took part in Study Two reported having used XR a couple of times, meaning that 2 additional participants were recruited who had used XR before. No participants gave further information on what sort of XR they had used (B4.1). Overall, the VI participants can be thought of as both inexperienced with XR, with 58.4% reportedly having not used it before. Fisher's exact test of independence was used to determine that there was not a statistically significant association between participants' level of sight (B3) and their familiarity with XR (B4) ($p = 0.1727$).

Table 6.7: Content Creator Demographic information for Study Three.

Question (ID in brackets)	Responses	Count
Familiarity with XR (B4)	I'm familiar with it and use it regularly	1
	I'm familiar with it and have used it a couple of times	1
	I'm familiar with it, but have never experienced it myself	1
Medium (PRO1)	XR (including VR, AR and MR)	1
	Gaming	1
	Film	1
Genre(s) (PRO2)	Action/Adventure	2
	Documentary/Simulation	1
	Drama/Entertainment/News/Sports/Comedy/Music	1
Organisation (PRO3)	Small Independent Production house	1
	Other	2
Majority of work (PRO4)	Audio Programmer	1
	Audio Designer	1
	Producer	1
Years in content creation (PRO 5)	1-9	1
	10+	2
Worked on an XR production (PRO9)	Yes	1
	No	2
Heard The Turning Forest before (XR2)	Yes, once	3

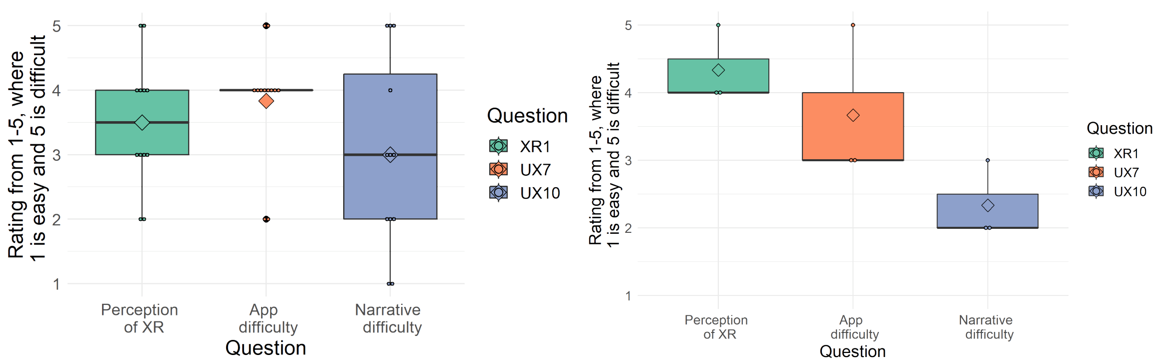
Content Creators. Only 3 content creators (CC1,2 & 6) were able to take part in this study. Table 6.7, shows an overview of their demographic information. They had ages ranging from 34-60. None reported having any hearing loss. 2 of them also took part in Study Two, including CC1 who is blind. However, the 1 who didn't take part in Study Two still stated that they had used 'The Turning Forest' once before. Therefore, content creators were fairly familiar both with XR and with the content of the testing AR app.

The content creators had quite a wide range in their years of experience (PRO5), between 3 - 14 years. Unlike Study Two, CC6 took part, who works in XR production as their main medium (PRO1, PRO9). They answered documentation/simulation for what genre(s) they work with, providing more information that they work with *'training in VR'*. They were therefore the only content creator who uses XR regularly (B4). The gaming content creator (CC2) had used it a couple of times for *'console games'*, whilst the visually impaired film content creator stated that they had never used it themselves. This is because they shared that they fully rely on screen readers in Study 2, so therefore were unable to explore the actual AR content of the app.

Whilst a smaller cohort, they represented a variety of mediums and experiences. Unfortunately, only CC1 was available for the interviews reported in Section 6.4.

6.3.2 Perception of XR

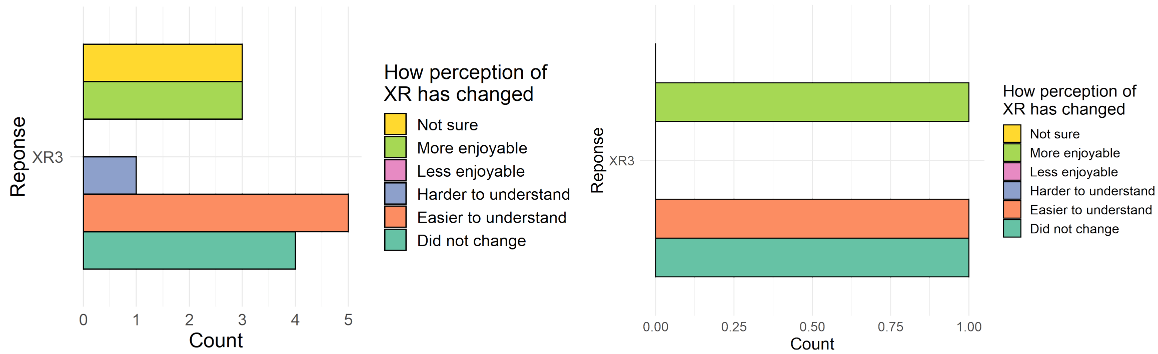
Participants were asked about their perception of XR as a whole (XR1), whether the app used had changed their perception of XR (XR3) and how likely they would be to use XR with the same accessibility techniques in future (UX11).



(a) VI Participants (b) Content Creators

Figure 6.2: Participant ratings of how easy to use the updated AR testing app was (UX7), how easy it was to follow the narrative (UX10) and how easy XR is as a whole (XR1). Questions XR1 from Table 3.5 and UX7 from Table 6.1. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

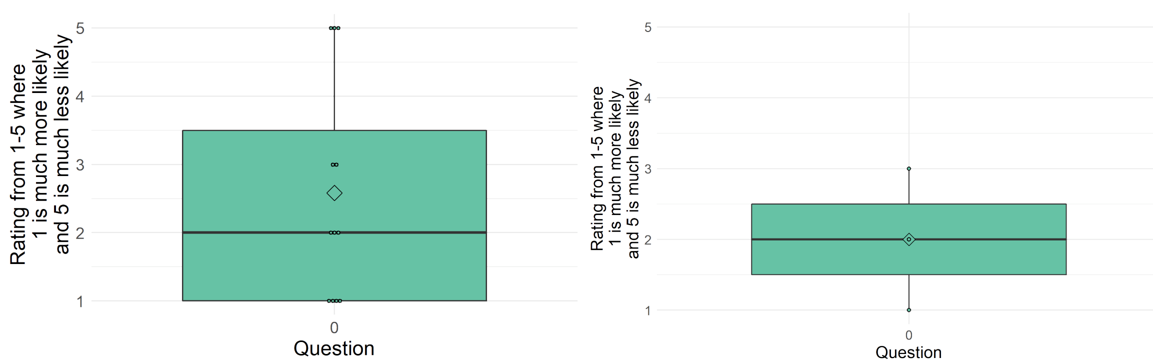
VI participants. Figure 6.2a shows how participants rated the difficulty of XR (XR1) alongside their ratings for how difficult the app was. It was rated a mean of 3.5 (median 3.5) with fair certainty compared to many other ratings in this study (standard deviation 1). This is the same as the mean rating of 3.5 given by VI participants in Study Two. This means that overall they perceived XR as fairly difficult to use. All participants who self-reported as fully blind did rate it a 4 or higher, whilst those in other categories varied more in their ratings. Whether participants had taken part in previous studies also seemed to have no impact on their rating of the difficulty of XR, with choices ranging from 2 to 4. Kruskal-Wallis tests determined that there were no significant relationships between ratings of the difficulty of XR (XR1) and sight level (B3) or familiarity with XR (B4).
How VI participants' perception of XR had changed (XR3) after using the testing AR app is shown in Figure 6.3a. Interestingly, one chose both 'did not change' and 'harder'. By a slim margin, 'easier to understand' was the most chosen option (5, 41.7%) and all those who



(a) VI Participants (b) Content Creators

Figure 6.3: Participant responses for how their perception of XR has changed after using the testing AR app. Question XR3 from Table 3.5

picked 'more enjoyable' had also picked 'easier to understand'. However, nearly as many (4, 33.3%) said that their perception had not changed, which given responses to XR1 means that they still perceive it as difficult for them to access. There does not seem to be any relationship between how perceptions of XR had changed and their level of sight (B3) or how they rated the difficulty of XR overall (XR1). Interestingly, those who rated it easier to understand were also split as to how difficult they found the AR app (more on this in Section 6.3.3), with 3 out of 5 (60%) rating it a 4 out of 5 difficulty and the rest rating it a 2. Perhaps despite finding it difficult to use, it was still less difficult than they had previously imagined.



(a) VI Participants (b) Content Creators

Figure 6.4: Participant ratings of how likely they would be to use XR if it used the same accessibility techniques as the updated AR testing app. Questions UX11 from Table 6.1. Content creators answered the same question as VI participants, rating it for themselves. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

Figure 6.4a shows that VI participants were uncertain about whether they would use XR with the same accessibility techniques as the AR testing app (UX11) (standard deviation 1.62). However, they were somewhat likely to overall, giving a mean rating of 2.58 (median 2).

Spearman's rank correlation was used to assess the relationship between participants' difficulty using the app (UX7) and how likely they were to use XR with the same accessibility techniques (UX11). There was a positive correlation between the two variables (r = 0.5842) with some statistical significance (p = 0.04608). This means that overall, those who found the app difficult would not be likely to use future apps with the same accessibility solutions and visa versa. However, there were some outliers, with 3 participants notably rating the app a 4 out of 5 difficulty, but giving a 1 ('much more likely') UX11. All those who responded 5 ('much less likely') to UX11 did not enjoy the app. Interestingly, some of those who responded with 1 ('much more likely') also did not enjoy the app. Therefore it is possible that even if participants found the app difficult and not necessarily enjoyable to use, they still saw the potential of the accessibility solutions used.

Kruskal-Wallis determined there were no significant relationships between How likely they were to use XR with the same accessibility techniques (UX11) and level of sight (B3) or familiarity with XR (B4)

Content creators. Figure 6.3b shows how content creators rated the difficulty of XR (XR1) alongside their ratings for how difficult the app was. Question XR1 was slightly altered for content creators, to ask how difficult they perceived XR to be for VI individuals, rather than themselves. They seemed to perceive it as more difficult than VI participants did, giving it a mean rating of 4.3 which they were fairly agreed on. This is also higher than the rating of 3.5 given by the content creators in Study One. The fact that they all rated it a 4 or higher somewhat negates the fact that the VI content creator was the only one to rate it a 5, all found it difficult regardless of sight level.

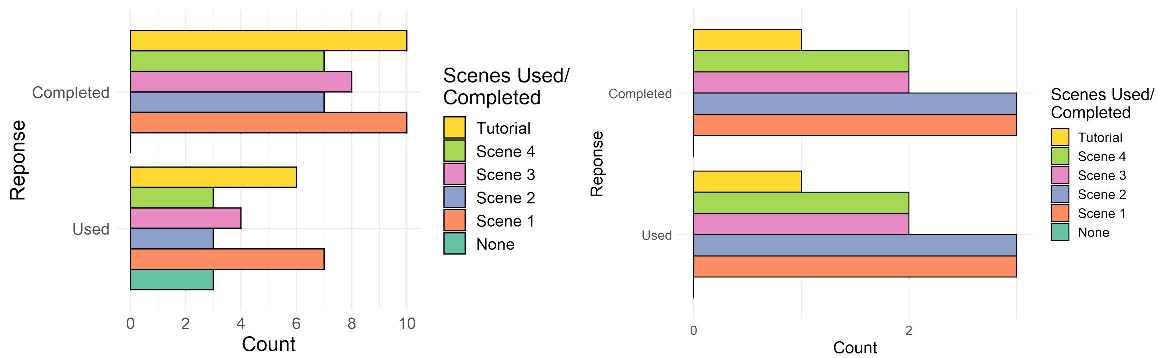
Figure 6.3b shows how the content creators' perception of XR had changed after using the AR testing app (XR3). Except for one whose perception did not change, they were positive. Interestingly, even though the VI content creator rated both XR as a whole (XR1) and the AR app (UX7) a 5 difficulty, they selected that the app had made them perceive XR as 'more enjoyable'. Encouragingly, the content creator who works on XR productions responded that they thought it was 'easier to understand'.

Figure 6.4 shows that content creators would be fairly likely to use XR with the same accessibility techniques again (UX11), giving a mean rating of 2. The VI content creator who rated the app most difficult (5 out of 5 difficulty) responded 1 ('much more likely') to UX11. Whilst the other 2 content creators both found the app easier to use (both a 3 out of 5 difficulty), they were less sure that they would use XR with the same accessibility techniques in future.

6.3.3 AR App Experience

Several questions were asked to determine participants experience using the app, shown in Table 4.4. Here, the most significant results are examined in turn; whether participants enjoyed the app (UX6) and how difficult the app was to use (UX7).

VI participants spent a wide-ranging amount of time using the app (UX1), between 25 and 210 minutes, giving an average time of over an hour (mean 69 minutes, median 50). All wore headphones (UX4), with the majority using a version of Apple AirPods (8. 66.7%). There were some more unusual choices used by others, including 2 (16.7%) using Aftershokz bone-conduction headphones and 1 participant (8.3%) using audio glasses. All but



(a) VI Participants

(b) Content Creators

Figure 6.5: Participant responses of which scenes they used (UX2) and which scenes they completed (UX3) in the updated AR testing app. Questions from Table 6.1

one participant (11, 91.7%) mentioned that they were sat down whilst using the app, and the remaining participant wasn't specific, saying only *'in my living room'*. Most participants (6, 50%) also mentioned being at a table whilst using the app. Only 1 participant seemed to use the app more actively by walking around to explore it, mentioning that they *'tried walking indoors'* (P14).

Participants were asked which scenes they tried to use (UX2) and which scenes they completed (UX3). The results for this are shown in Figure 6.5a. Overall, those who self-reported as fully blind attempted fewer scenes than all other groups. The majority (6, 50%) of participants tried all scenes, with the tutorial and Scene 1 being the most used (10, 83.3% each). It seems that most participants worked through them in order, as the later scenes are the least used. However, one participant did skip straight to Scenes 3 and 4, without even using the tutorial. This may be because they did take part in Study 2, so even if they hadn't seen those scenes before, they were familiar with the app.

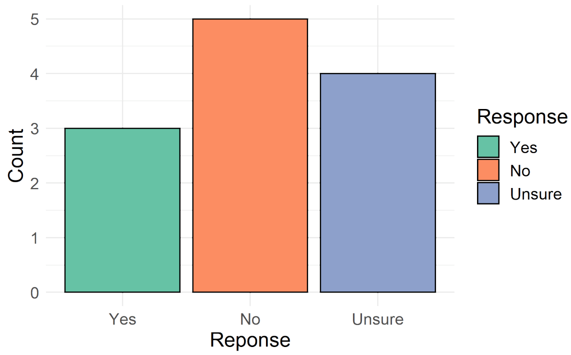
What is very noticeable in Figure 6.5a is that not all participants were able to complete the scenes they used, with 3 (25%) unable to complete any scenes. However, 5 participants (41.7%) were able to complete all the scenes they attempted. There was no clear correlation between participants' level of sight and whether they could complete the scenes they attempted. Interestingly, 4 out of the 5 who managed to complete all scenes they attempted rated the app a 4 out of 5 difficulty, so must have persevered despite finding it hard to use.

Content creators. Only 2 of the content creators reported how long they spent using the AR testing app (UX1). The VI content creator spent 2 hours, whilst the XR content creator spent 20 minutes. All content creators used headphones (UX4), all using in-ear models including Apple Air Pods and SoundMagic E10. All of them mentioned using it sat down (UX5), with 2 (66.7%) using the floor and the remaining 1 using a table.

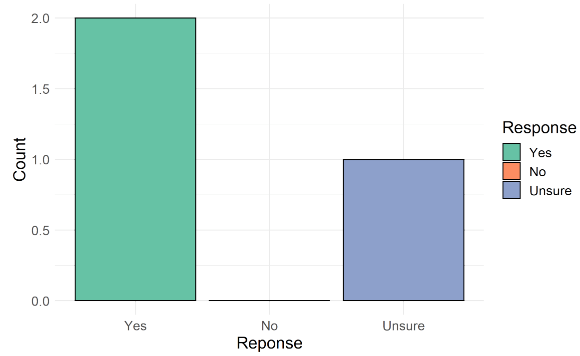
Figure 6.5b shows which scenes content creators used (UX2) and completed (UX3). It can be clearly seen that all the scenes content creators attempted they also completed. What is interesting is that both the sighted content creators skipped the tutorial, but used all other scenes, whilst the VI content creator only used the tutorial and Scene 1. This makes sense, given that they rated the app much more difficult to use than the others. The fact that the sighted content creators didn't use the tutorial means that they may have been less aware of some of the enhanced features, which are detailed in Section 5.4.

App Enjoyment

The below presents participant responses to whether they enjoyed the app (UX6).



(a) VI Participants



(b) Content Creators

Figure 6.6: Participant ratings of whether they enjoyed using the updated AR testing app. Question UX6 from Table 6.1.

VI participants. Figure 6.6a shows VI participants' responses for whether they enjoyed the AR testing app. Only 3 (25%) stated that they did enjoy the app, with the rest split between stating they didn't enjoy it (5, 41.7&) and being unsure (4, 33.3%). However, this does mean that a larger proportion of participants did not enjoy the app than in Study Two (Section 4.5.2). Fisher's exact test of independence determined there were no significant associations between app enjoyment (UX6) and level of sight (B3) or familiarity with XR (B4).

Content creators. Figure 6.6b shows content creator ratings for whether they enjoyed the app. All of them enjoyed the app, except 1 who was unsure about it. Notably, the VI content creator did enjoy the app.

App Difficulty

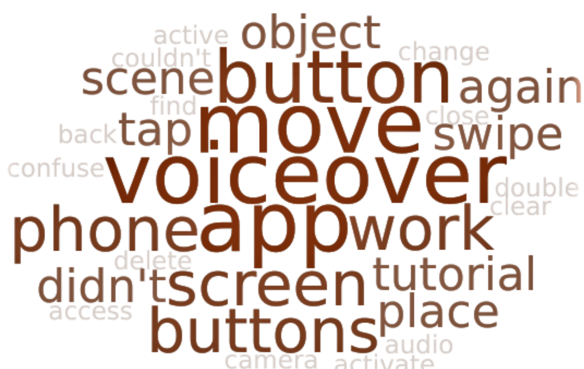
The below presents participant responses to how difficult to use the app was (UX7) and how difficult it was to follow the narrative (UX10). Free text responses gave them a chance to say what difficulties they faced when using it (UX8) and their suggestions for how it could be made easier (UX9).

VI participants Figure 6.2a shows that VI participants found the app difficult to use (UX7) and more difficult than their prior perception of XR (XR1), rating it a mean of 3.83 (median 4). There was a high amount of agreement on this (standard deviation 0.94), with most (8, 66.7%) participants rating it a 4, with outliers at both 2 and 5 (2, 16.7% each). This is very similar to the mean rating of 4 for the initial AR testing app in Study Two (Section 4.5.2) and the ratings ranging 4-4.5 of the apps investigated in Study One (Section 3.6.2). None of those who self-reported as fully blind rated the app easier than a 4 out of 5 difficulties, whilst those in other categories varied more in their ratings.

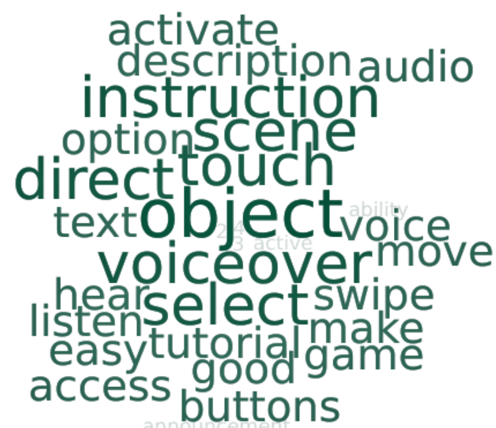
VI participants found it slightly easier to understand the narrative of the AR testing app (UX10), giving a mean rating of 3 (median 3). However, there was much more uncertainty in this, with responses spanning across the whole scale (standard deviation 1.48). Those who self-reported as fully blind varied as much in their ratings for how easy it was to understand the narrative as all other categories, with one participant even giving it a 1 (easy). Spearman's rank correlation was used to assess the relationship between participants' difficulty using the app (UX7) and how difficult they found understanding the narrative (UX10). There was a positive correlation between the two variables r = 0.2996. However, this was not statistically significant p = 0.3442. The app being rated this difficult to use and understand was not the expected outcome of this study. Further results in Section 6.3.4 particularly and of course, the final discussion of this study in Section 6.5 seek to explain the reasons for this and what it could mean for the goals of this work.

Statistical testing found no significant relationships participants difficulty rating of the app (UX7) or of the narrative (UX10) and whether they enjoyed the app (UX6), their level of sight (B3) or their familiarity with XR (B4).

What is clear across both these results is that it cannot be assumed that just because an individual has a particular level of visual impairment, they will find something the same difficulty as someone within the same 'category'. This once again strengthens the need for personalised accessibility solutions.



(a) UX8



(b) UX9

Figure 6.7: Word cloud showing the difficulties VI participants faced when using the updated AR testing app (UX8) and their suggestions for making it easier to use (UX9). Questions are shown in Table 6.1. Created in Qualtrics.

Figure 6.7a shows a word cloud of VI participant responses to what difficulties they faced when using the AR testing app (UX8). At a first glance, this would seem similar to the equivalent word clouds in Study Two, Figure 4.14. VoiceOver (screen reader functionality) was once again mentioned as the most mentioned feature, by 7 participants (58.3%). Many of the problems they highlighted were with the functionality of the Unity Accessibility Plugin [180] used, as described in Section 5.2.1. Although extra instructions were provided, a number of them were unable to make their built-in screen-reader work with the one in the app (by turning on direct touch on iOS, turning off Talkback on Android). The response to this ranged from *'getting it set up [...] was difficult'* (P5) to not being able to get it to work *'no matter what I tried'* (P10), with one participant having to *'delete the app and reinstall'* (P1) to get it to work. Those who did use screen readers found other issues, such as not being able to directly select User Interface (UI) elements, but being forced to use gesture-navigation with one mentioning that it was *'not clear how to navigate screen'* (P16) and other issues such as *'it took several seconds for the button to become active'* (P3) and *'buttons didn't always activate when I double tapped on them'* (P11). As well as gestures not working the way they expected, 2 participants mentioned the fact that they couldn't customise the screen reader in the same way as they usually would with VoiceOver, saying *'because I'm used to VoiceOver it felt very slow to me'* (P4) and that *'the tutorial voice was so slow and there was no skip option, so eventually I got tired of restarting'* (P3). P12 did not want screen reader functionality, so struggled with the fact that the screen reader was enabled by default, as it did not allow them to simply select options, but forced gesture based navigation. However, once they disabled this they were *'able to navigate freely'*. Other than screen-reader functionality, the other most mentioned area was that participants thought *'instructions [were] not always clear'* (p13), meaning that one described being *'unsure of what I was to do to complete each task'* (p14), with another similarly saying *'I found it a bit confusing'* (p15). There were some issues with the screen reader speaking at *'the same time as the narrator'* (P4), which led to further confusion. Some participants did not fundamentally understand how AR worked based on the instructions given, saying things such as *'I thought I had to move the camera on my phone around a table but I couldn't understand how the AR was supposed to work'* (P11), *'when trying out the scenes I could place objects but then had no idea what to do with them'* (P3), *'can't hear objects I've placed and hence can't interact with them'* and *'when I move closer I never get a button on the screen that allows me to interact with them'* (P16). For 2 participants, this led to them *'getting stuck in a message that said "you have walked into something"'* (P3), meaning they just kept colliding with AR objects and not knowing what to do next. One participant did not mention many issues, only that *'it is quite small [so] it can strain my eyes'* (P7).

Figure 6.7b shows a world cloud of VI participant responses to what their suggestions were for making the AR testing app easier to use (UX9). Several participants (4, 33.3%) mentioned changes to the screen reader. This ranged from simply stating *'don't depend on direct touch'* (P10) because it is *'my last option when accessing a device'* (P14) to wanting better *'instructions on how to activate direct touch'* (P3). Several also suggested that more spoken cues would help them to understand the AR, from understanding the *'task & outcome'* (P14) and having *'a voiced option when selecting objects'* (P1) to *'setting scenes with good description'* (P15) and having a clearer idea that they were in an AR scene by having *'some narration so you know you've immediately landed into the scene'* (P4). One even suggested needing clearer instructions on *'how to hold my phone - horizontally or vertically'* (P16). 2 participants (16.7%) suggested ways that the guide character could be better structured. They had suggestions from not having *'overlap in audio dialogue'* (P11) with the narrator, with a suggestion that screen reader functionality should be delayed by the onset of narration, as *'it would be best to have VO not land on that text so you can clearly hear the narration and then listen to VO for navigation'* (P4). There were 3 participants (25%) who did not suggest improvements to the spoken cues, but simply improvements to the user interface such as *'different colours for the text/instructions'* (P7) and having *'a trial before playing the game so [that I was] confident how to move things'* (P13).

Content creators Figure 6.2b shows how difficult content creators found the app. They found it easier to use than VI participants, rating it a mean of 3.67. Difficulty was clearly related to sight as the two sighted content creators rated it a 3 and the VI content creator rated it a 5. However, it is interesting that the sighted content creators thus did not find it easy to use and that therefore the overall rating is higher than the 2.8 given to the initial testing app in Study Two. All the content creators found it easier to understand the narrative of the app (UX10), giving it a mean rating of 2.33. In this case, the VI content creators rating was more similar to the 2's given by both others, rating it a 3.

When asked what difficulties they faced in the app (UX8), there were clear differences between the sighted and VI content creators. The sighted content creators suggested tweaks to the AR such as making it easier *'visualising the scale and understanding of what to anticipate'* (CC6) and CC2 finding it hard that they *'couldn't easily undo mistakes'* as they had placed objects *'blocking the view'*. The VI content creator (CC1) had more difficulty understanding what was happening. They weren't sure if they *'should be tapping once or twice to place objects'* and didn't fully understand audio cues, both for finding AR objects, *'only found out that when I found a rock I had successfully place[d], that it would make a pebble cascading sound'* and for asking for help *'did not realise that the hints would change/develop'*. Similarly to the VI participants they mentioned *'conflict between the AI and the speech system'*.

CC2 had various suggestions to simplify the look of the app, by reducing the *'amount of text'* and information on the screen, including moving the personalisation slider to a settings menu. They also suggested improvements to the AR, such as previewing where objects will be placed such as having a mechanism to *'drag them into the world'*. In a similar vein to VI participants, the VI content creator (CC1) suggested *'much more comprehensive documentation'* including *'sample audio of someone playing the game'*. Their main other suggestion was that it should build up much more slowly with *'far fewer types of objects to be placed in the forest at the start'* so they could have *'a chance to get used to the game, before overloading the choices'*. They also suggested removing *'the option to keep tinkering with the option'* as they thought *'how useful is it, really'*.

6.3.4 AR App Sounds

Several questions were asked to determine relationships between sounds and the AR app, shown in Table 6.2. Participants were asked if sounds were helpful (NV1/ST1) overall and which categories of sounds were most helpful (NV2/ST2). These questions were repeated for both navigational sounds (NVx) and sounds for understanding the narrative of the app (STx).

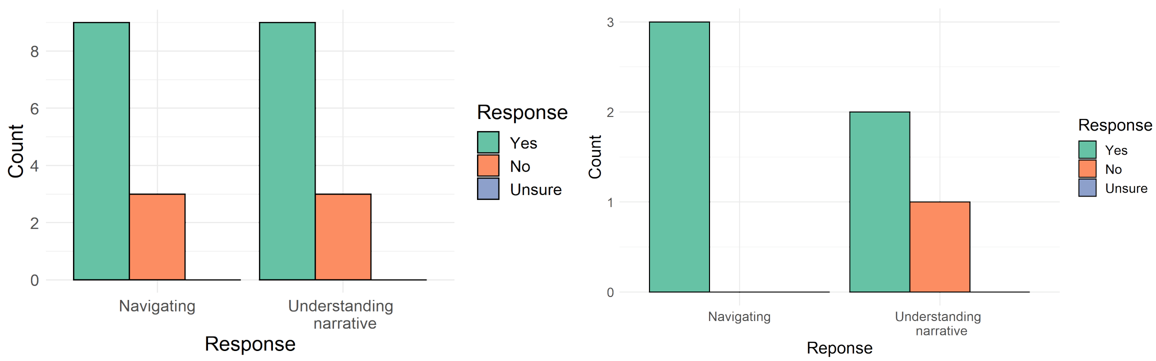
Did sounds help?

The below presents participant responses to whether sounds in the app were helpful (NV1/ST1) and which sounds in the app were helpful (NV2/ST2).

VI participants Figure 6.8a shows that the majority of VI participants (8, 66.7%) thought that sounds helped them to both navigate (NV1) and understand the story (ST1). Participant ratings stayed consistent across the two categories, showing that sounds either helped them with everything or not at all. Statistical testing using Fisher's exact test of independence and Kruskal-Wallis tests determined there were no significant relationships associations between whether they thought sounds helped (NV1/ST1) and level of sight (B3), whether they enjoyed the app (UX6) or how difficult they rated the app or its narrative (UX7/UX10).

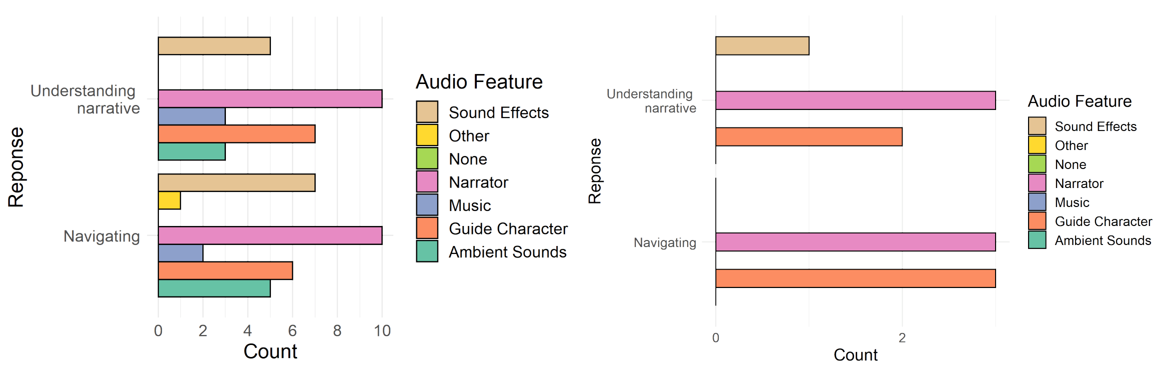
VI participants did give different ratings for which sounds were helpful across the two purposes of the sounds, shown in Figure 6.9a. However, across both categories, the narrator came out as the most important (10, 83.3%). All of those (3, 25%) who said that no sounds helped in the app (NV1/ST1) only selected the narrator for both categories, even though one self-identified to each of the 3 sight categories (blind/blind with residual vision/partially sighted). There are some similarities with Study Two (Figure 4.17a), with spoken cues being the most important, followed by sound effects. However, unlike Study Two, some participants found music helpful across both categories.

Interestingly, for navigating the app (NV2), sound effects were rated the second most highly (7, 58.3%), above the guide character (6, 50%). A significant number of participants (5, 41.7%) found the ambient sounds useful for navigating. What is curious is that the role of sound effects (5, 41.7%) and the guide character (7, 58.3%) seem to reverse for helping participants to understand the story (ST2). Ambient sounds were also less helpful for this, rating the same as music (3, 25%).



(a) VI Participants (b) Content Creators

Figure 6.8: Participant ratings of whether sounds were helpful in the updated AR testing app. There are 2 categories of responses, for whether sounds helped them to navigate (NV_1) or to understand the narrative (ST_1). Questions from Table 6.2



(a) VI Participants (b) Content Creators

Figure 6.9: Participant ratings of which sounds were helpful in the updated AR testing app. There are 2 categories of responses, for whether they thought sounds helped them to navigate (NV2) or to understand the narrative (ST2). Questions from Table 6.2.

Content creators. Figure 6.8b shows that overall, content creators also thought sounds were helpful in the app, although the VI content creator did not think any sounds helped them to understand the story (ST1). Given that they all rated the same, there were no differences between whether content creators thought sounds helped navigate (ST1) and how easy they found the app to use (UX7). However, the content creator who said no sounds helped them to understand the narrative (ST1) did find it harder to understand the narrative of the app (UX10). There does not seem to be a relationship between enjoyment and whether sounds helped, given that the content creator who thought sounds didn't help to navigate (NV1) enjoyed the app (UX6) whilst the content creator who was unsure whether they enjoyed the app stated sounds did help across both categories.

Similarly to VI participants, the narrator was the most important sound, with unanimous consent across both categories shown in Figure 6.9b. However, the guide character came out as just important for understanding the story (ST2). The VI content creator kept their selection the same across both categories (guide character and narrator). However, what is interesting is that one of the others went from choosing both spoken cues for NV2 to choosing only the narrator for ST2, whilst the other added sound effects. Whilst there are some similarities to content creator ratings in Study Two (Figure 4.17b), with spoken cues most important and sound effects helping understand the story more, the lack of rating for any other sound categories and sound effects not helping navigating is notable.

6.3.5 AR App Enhanced Audio

In this study only, participants were asked further questions to determine the impact of the enhanced audio features added to the updated AR testing app (Table 6.2). They were first asked whether they enjoyed the features (SFX1, SP1, GC3) and then whether the features helped them with several key AR tasks (SFX2-5, SP2-5, GC4-7). Statistical testing was done to determine relationships between whether the features helped (SFX2-5, SP2-5, GC4-7) and level of sight, app enjoyment (UX6) or how difficult they found the app as a whole (UX7). Due to limited sampler size most relationships are not significant. Where significant relationships are founded they are reported in the text.

For these ratings across how enhanced audio features helped, Fleiss's Kappa was computed to assess the agreement between participants. VI participants had poor agreement (Kappa = 0.206) which was statistically significant ($p < .0005$). Interestingly, this is higher than the level of agreement Ward found between content creators when they were completing the Narrative Importance (NI) categorisation task (Kappa = 0.11).

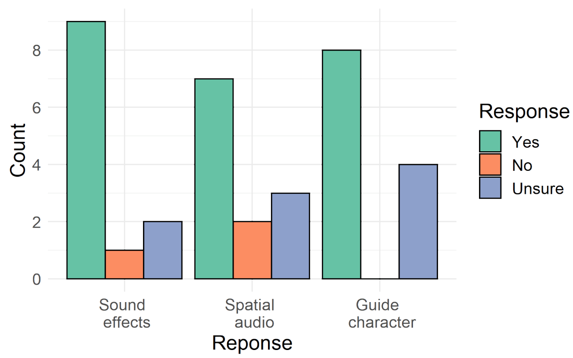
The content creators also had poor agreement (Kappa = 0.151) which was statistically significant ($p < .0005$), although this should be viewed with caution given the small size of the data set. Additional questions were asked specifically about the screen reader functionality of the guide character (GC1-2) for those that used it.

Did participants enjoy the enhanced audio?

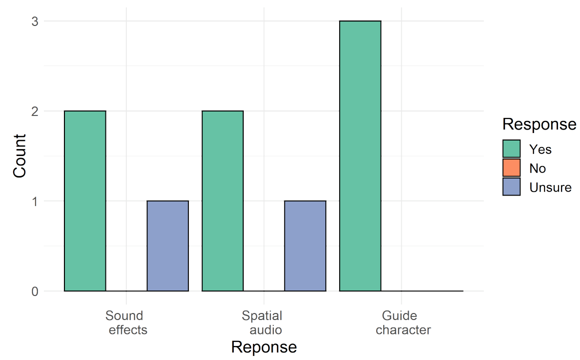
The below presents participant responses for whether they enjoyed the enhanced audio featured added to the updated AR testing app.

VI participants Figure 6.10a shows that VI participants enjoyed all of the enhanced audio features in the updated AR testing app, with the most enjoying sound effects (SFX1, 9, 75%) and the least enjoying the spatial audio (SP1, 7, 58.3%). Interestingly, most of the remaining participants were unsure, rather than stating that they didn't enjoy the features. Several participants who were unsure about whether they enjoyed the app as a whole did enjoy the various enhanced audio features.

Content creators Figure 6.10b shows that overall the content creators also enjoyed the enhanced audio features in the updated AR testing app. Interestingly, the same sighted content created (CC2) responded that they were unsure about both the sound effects (SFX1) and spatial audio (SP1). They were also the only content creator who didn't say they enjoyed the app, responding 'unsure' to UX6. This means that there doesn't seem to be an association between content creators' sight level and whether they enjoyed the enhanced audio features, but there could be an association between whether content creators enjoyed the sound effects (SFX1) and spatial audio (SP1) and whether they enjoyed the app (UX6)



(a) VI Participants



(b) Content Creators

Figure 6.10: Participant ratings of whether they enjoyed enhanced audio features in the updated AR testing app. The figure shows responses for whether they enjoyed the sound effects (SFX_1), the spatial audio (SP_1) and the guide character (GC_3). Questions from Table 6.2

Did the guide character aid navigation?

The below presents participant responses to whether they used the guide character (GC1). Those who used it were then asked if the guide character helped navigate through the user interface of the updated AR testing app.

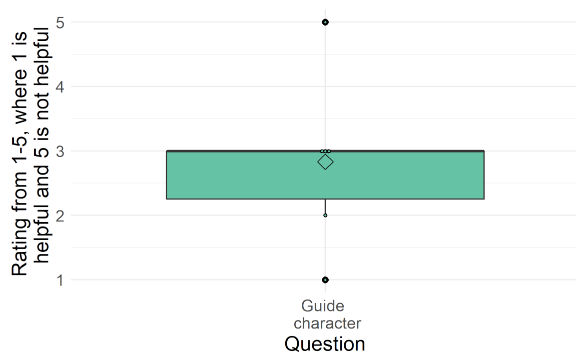


Figure 6.11: VI participant ratings of whether the guide character helped them to navigate the updated AR testing app. Question GC2 from Table 6.4. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

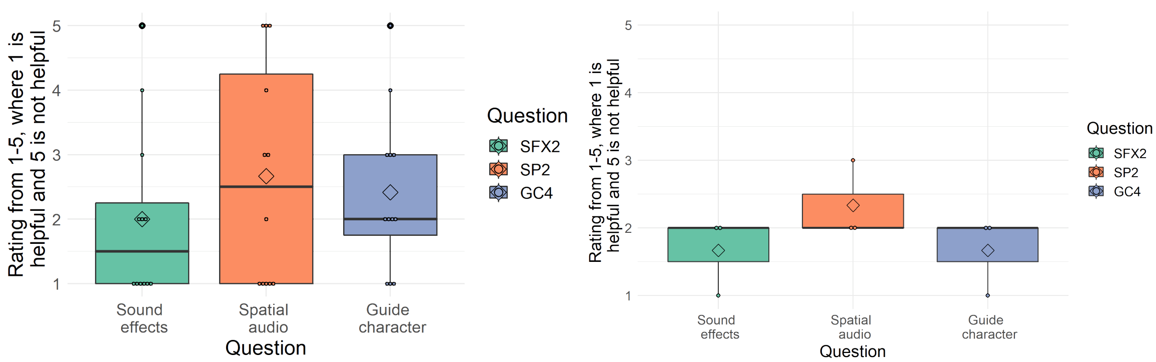
VI participants. Only half (6, 50%) of the VI participants actually used the screen reader functionality of the guide character (GC1). All those who self-identified all fully blind used it and none of those who self-identified as partially sighted or 'other'. Those who self-identified as blind with residual vision were mixed, with 2 out of the 5 (40%) using the screen reader functionality.

Those who used the screen reader functionality rated it as fairly helpful in Figure 6.11, with a mean of 2.83 (median 3). Although there was a low interquartile range, there were outliers at both extremes of the scale. Many of the difficulties that were faced in using the guide character as a screen reader were extensively described by participants in their responses to UX8-9, detailed earlier in Section 6.3.3. What is interesting is that 2 of the 6 participants (33.3%) rated the guide character as being helpful (2 or less) but rated the app as difficult to use overall (4 or above). Those who rated the app easiest to use (2 or below) didn't use the guide character at all.

Content creators. Only the VI content creator used the screen reader functionality of the guide character (GC1). It's a shame that the sighted content creators didn't try it out and provide feedback. The content creator who used it rated it as helpful (1), despite rating the app as a whole as difficult to use (5).

Did participants feel immersed by the enhanced audio?

The below presents participant responses to whether the enhanced audio features helped them to feel immersed in the virtual world of the updated AR testing app (SFX2, SP2, GC4).



(a) VI Participants (b) Content Creators

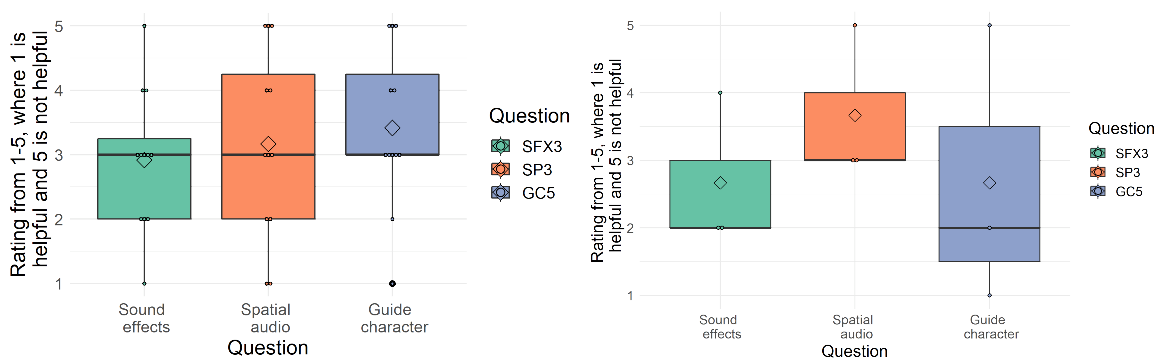
Figure 6.12: Participant ratings of whether enhanced audio features helped them to feel immersed in the AR testing app. The figure shows responses for the sound effects (SFX_2), the spatial audio (SP_2) and the guide character (GC_4). Questions from Table 6.2. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants Figure 6.12a shows that VI participants were unsure whether the enhanced audio features helped them to feel immersed, with ratings across the whole scale for all the features. They were most agreed on the sound effects (SFX2, standard deviation 1.348) and found that it aided immersion the most, with 6 (50%) rating it helpful (1). This gave a mean of 2 (median 1.5). They were the most uncertain about the spatial audio (SP2), with an interquartile range stretching across most of the scale and a standard deviation of 1.723. Looking at the averages, this also aided immersion the least with a mean of 2.67 (median 2.5). The guide character (GC4) was marginally more helpful for participants feeling immersed, with a mean rating of 2.42 (median 2, standard deviation 1.24). Therefore, overall the enhanced audio features did aid immersion for VI participants, but spatial audio was divisive in this regard.

Content creators. Figure 6.12b shows that the content creators agreed more on whether the enhanced audio features helped them to feel immersed. Overall, they thought they were fairly helpful at aiding immersion and thought they aided immersion more than VI participants did. The sound effects (SFX2) and guide character (GC4) were the most helpful, with both rated a mean of 1.67. In concordance with VI participants, the spatial audio (SP2) was the least helpful, with a mean of 2.33. The VI content creator rated all the features as fairly helpful (2). However, the other two rated sound effects and the guide character the same, but shifted their rating up by 1 towards less helpful for the spatial audio. The content creator who was unsure whether they enjoyed the app (UX6) did always rate the enhanced audio features the least helpful, giving the highest rating of 3 to the spatial audio (SP3). This may have therefore affected their enjoyment of the app.

Did the enhanced audio help participants identify what objects were?

The below presents participant responses to whether the enhanced audio features helped them to identify the type of objects around them in the updated AR testing app (SFX3, SP3, GC5).



(a) VI Participants

(b) Content Creators

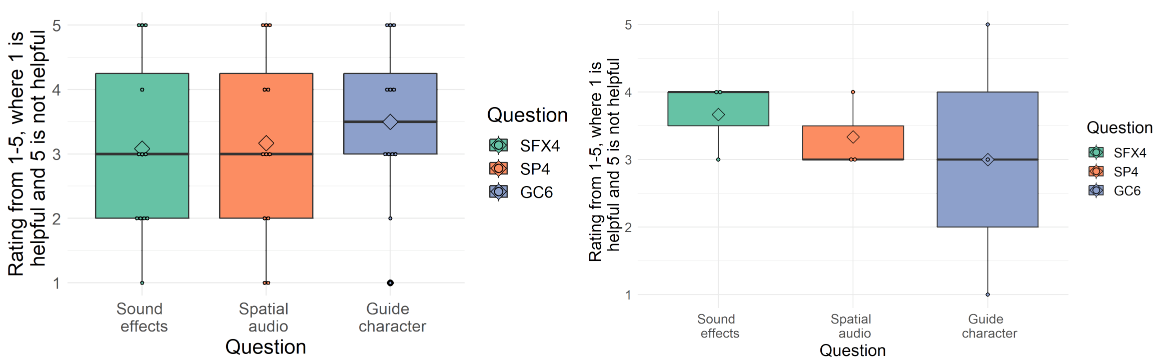
Figure 6.13: Participant ratings of whether enhanced audio features helped them to identify what type of objects were around them in the AR testing app. The figure shows responses for the sound effects (SFX_3), the spatial audio (SP_3) and the guide character (GC_5). Questions from Table 6.2. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants Figure 6.13a shows that overall, VI participants didn't find the enhanced audio features helpful for identifying the type of objects around them, but they weren't necessarily unhelpful either. The sound effects were rated the most helpful with a mean of 2.9 (median 3, standard deviation 1.084), whilst the guide character was the least helpful, with a mean of 3.42 (median 3, standard deviation 1.24). VI participants were far from unanimous in their ratings of all the features, with ratings across the whole scale and particular disagreement about the spatial audio, which had a mean rating of 3.17 (median 3, standard deviation 1.467). How helpful the enhanced audio features were for identifying the type of objects around VI participants in the app did have a limited impact on how difficult they found the app was as a whole.

Content creators Figure 6.13b shows that content creators found the enhanced audio features less helpful for identifying the type of objects than VI participants did. They also rated sound effects the most helpful, but it was rated the same as the guide character with a mean of 2.67. There was more uncertainty about the guide character. Content creators rated the spatial audio the least helpful, giving it a mean of 3.67. The VI content creator consistently rated all features as less helpful than the sighted content creators, giving a 4 for sound effects and 5 for both spatial audio and the guide character. How helpful the enhanced audio features were for identifying the type of objects around content creators in the app did have a limited impact on how difficult they found the app was as a whole.

Did the enhanced audio help participants identify the location of objects?

The below presents participant responses to whether the enhanced audio features helped them to identify the location of objects around them in the updated AR testing app (SFX4, SP4, GC6).



(a) VI Participants (b) Content Creators

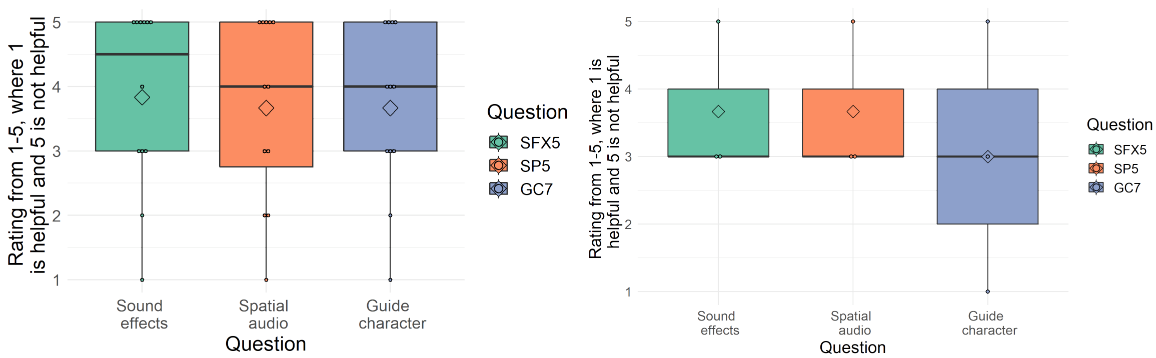
Figure 6.14: Participant ratings of whether enhanced audio features helped them to identify the location of objects around them in the AR testing app. The figure shows responses for the sound effects (SFX_4), the spatial audio (SP_4) and the guide character (GC_6). Questions from Table 6.2. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants Figure 6.14a shows that overall, VI participants didn't find the enhanced audio features helpful for identifying the location of objects around them, but they weren't necessarily unhelpful either. However, there was little agreement between participants with ratings across the whole scale for all features. The differences between average ratings were marginal. Sound effects were rated the most helpful with a mean of 3.08 (median 3, standard deviation 1.38). The spatial audio was rated very similarly to this, with a mean of 3.17 (median 3, standard deviation 1.47). Participants were most agreed about the guide character, rating it the least helpful with a mean of 3.5 (median 3.5, standard deviation 1.24). How helpful the enhanced audio features were for identifying the location of objects around VI participants in the app did have a limited impact on how difficult the app was as a whole.

Content creators. Figure 6.14b shows that content creators thought the enhanced audio features were less helpful for identifying the location of objects than VI participants did, rating them unhelpful overall. The differences between average ratings were marginal. They were in the most disagreement over the guide character, with 1 at each end of the scale, giving a mean of 3. The spatial audio was rated the next most helpful, with a mean of 3.33. Finally, the sound effects were rated a mean of 3.67. The VI content creator consistently rated the enhanced sounds as less helpful than the sighted content creators, always selecting a 4 or higher. How helpful the enhanced audio features were for identifying the location of objects around content creators participants in the app did have a limited impact on how difficult the app was as a whole.

Did the enhanced audio help participants identify interactions with objects?

The below presents participant responses to whether the enhanced audio features helped them to identify how they were interacting with the objects around them in the updated AR testing app (SFX5, SP5, GC7).



(a) VI Participants

(b) Content Creators

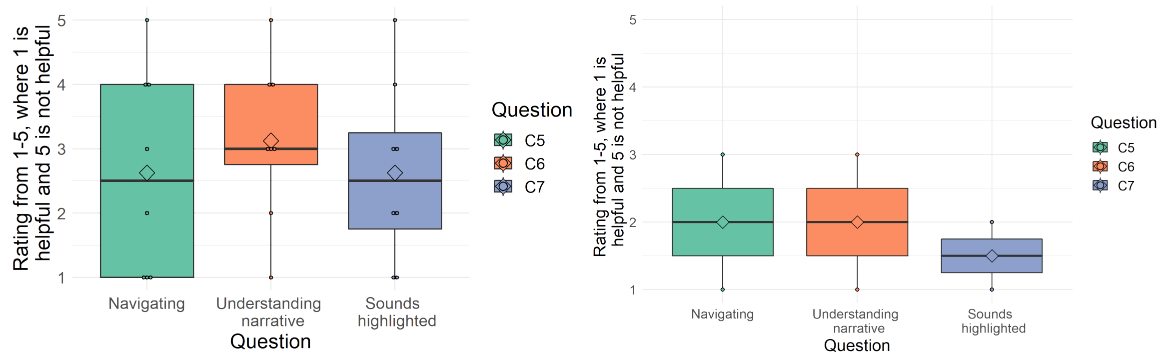
Figure 6.15: Participant ratings of whether enhanced audio features helped them to identify how they were interacting with objects in the AR testing app. The figure shows responses for the sound effects (SFX_5), the spatial audio (SP_5) and the guide character (GC_7). Questions from Table 6.2. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants. Figure 6.15a shows that of all the tasks of the enhanced audio features, VI participants found them the least helpful for identifying how they were interacting with objects. However, there was still much disagreement about this with ratings across the scale. The differences between average ratings were marginal. Unlike the other task, sound effects were the least helpful, rated with a mean of 3.83 (median 4.5, standard deviation 1.4). Both the spatial audio and guide character were rated a mean of 3.67 (median 4). Spearman's rank correlation determined that there was a statistically significant ($p = 0.02815$) positive correlation ($r = 0.63$) between whether sound effects (SFX5) helped VI participants identify how they were interacting with objects around them and how difficult they found the app as a whole (UX7). How helpful the enhanced audio features were for identifying how VI participants were interacting with objects around them in the app did have a limited impact on how difficult the app was as a whole.

Content creators. Figure 6.15b shows that content creators rated enhanced audio features fairly similarly for this tasks to their ratings for other tasks. Overall, they were rated unhelpful, but with marginal differences between ratings. The sound effects and spatial audio were both given a mean of 3.67 (median 3, standard deviation 1.15), whilst the guide character had a mean of 3 (median 4, standard deviation 2). The VI content creator consistently rated the enhanced audio features as less helpful than the sighted content creators, choosing a 5 for all features. How helpful the enhanced audio features were for identifying how they were interacting with objects around them in the app did have a limited impact on how difficult the app was as a whole.

6.3.6 Personalisation Control

Participants were asked questions about the personalisation control within the updated AR testing app (Table 6.3). After establishing whether they used the control (C1), the questions determined how they used the control (C3-4), what difference it made (C2) and whether it was helpful (C5-9).



(a) VI Participants (b) Content Creators

Figure 6.16: Participant ratings of whether the personalisation control in the AR testing app was helpful. The figure shows responses for whether the control helped them to navigate (C5), understand the narrative (C6) and whether the sounds the control highlighted were helpful (C7). Questions from Table 6.3. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

VI participants Most VI participants (8, 66.7%) did use the personalisation control (C1). All those who did not use it self-identified as being blind with residual vision or fully blind. Most participants (5 out of 8, 62.5%) adjusted it a couple of times throughout the app (C3), with all but one of the remaining participants leaving it in one place. The other wasn't sure how they used it. Not all of those who used the control provided a response for what level they left it at (C4). Despite the slider in Qualtrics being WCAG 2.0 compliant and therefore working with screen readers, it may still have been difficult to use. Those who did respond used a wide range of levels, between 0 and 90 out of 100. Their mean level was 39.6 (median 30), so more towards the standard mix. But it would seem that they did try to use the control to help improve their experience, given the range of levels and that most adjusted it a couple of times throughout. VI participants were split as to whether the control made the app 'more enjoyable' (C2) or whether it 'did not change' anything (C2), with each chosen by 4 (50% of those who used the control). Of the remaining 2 participants, 1 responded they were 'not sure' what difference it made and the other selected 'easier to understand'.

Figure 6.16a shows how helpful VI participants thought the personalisation control was. There was much disagreement about it, with ratings across the scale. They were particularly unsure about whether it helped them navigate the app (C5), rating it neither helpful nor unhelpful with a mean of 2.63 (median 2.5). For helping understand the narrative (C6), the control was rated as even less helpful with a mean of 3.13 (median 3). Spearman's rank correlation determined that there were no statistically significant correlations between how helpful participants found the personalisation control (C5/C6) and how difficult they found the app (UX7) or its narrative (UX10). It would seem that the control did not help participants use the app or understand its narrative. For those who found the app easy, the control was helpful and for those who found it difficult, the control was not helpful.

VI participants were neutral on whether the sounds the control highlighted were helpful, rating it a mean of 2.63 (median 2.5). When asked which highlighted sounds were helpful (C6), the most common theme was interaction sounds with participants mentioning *'sound effects, selected sound'* (P1), *'when pressing an object and getting feedback'* (P13) and those most *'helpful in a way to navigate and differentiate the action[s]'* (P12). 2 out of the 8 (25%) mentioned spoken hints, with one clarifying that it was helpful to make *'the music a bit quieter to hear the narrator'* (P7). A further 2 also motioned the more ambient sounds, *'such as trees or rocks so you know where they are'* (P4). When asked if there were any sounds that were not highlighted that would have been helpful (C9), most participants seemed unsure. 3 (37.5%) did not respond and a further 4 (50%) gave responses to indicate they were not sure. The remaining 1 (12.5%) instead gave feedback based on what sounds could be added to the whole app, suggesting *'it would be helpful to have the characters speak the item I've selected when adding or removing items'* (P1).

Content creators. Both the sighted content creators used the personalisation control (C1), but the VI content creator did not. They used it very differently, with one leaving it in one place throughout at 29 out of 100 and one adjusting it a couple of times, but mostly leaving it at 0 (C2&3). The one who did not adjust it was unsure about the impact the control had, but the one who played with it more thought the control made the app more enjoyable.

Figure 6.16b shows that the 2 content creators who used the control thought that it was thought the control more helpful than the VI participants did. They both kept the same ratings for whether it helped with navigation (C5) or for understanding the story (C6), with one rating it a 1 and one a 3, giving an average of 2. Given that they both gave the same rating of 3 for how hard the app was to use (UX7) and 2 for how hard it was to understand the narrative (UX10), it is not possible to say if the control helped them to use the app, as they found it fairly easy anyway.

Both content creators rated the sounds that the control highlighted as helpful (C7), with an average rating of 1.5. They gave no answers to whether additional sounds should have been highlighted (C9). When asked what highlighted sounds were helpful (C8), one mentioned that the *'guide was the strongest assistance'* (CC6), whilst the other said that it made it *'a calmer and more pleasant experience reducing the noise levels, but I cranked it back up for I didn't miss anything'* (CC2).

6.3.7 Workflow Questions

The content creators were asked further questions to determine how different the features added to the updated AR testing app were from their own workflows (WS1, WB1, WG1, WC1) and whether they would be likely to use similar features in their own future work (WS2, WB2, WG2, WC2).

How different from current workflows

Figure 6.17 shows the content creators' responses to how similar the additional features in the updated testing AR app were to their own work. Perhaps unsurprisingly, the sound effects (WS1, mean 2.3) and spatial audio (WB1, mean 3.33) were most similar. The content creator who rated spatial audio as very different (5) was the VI film producer. However, the use of the guide character (WG1, mean 4.33) and personalisation control (WC1) were very different to the content creators own work. The gaming audio programmer found all the features the most similar to their own work, not rating any higher than a 3. However, it is interesting that apart from one outlier, there was no rating of the features as not different at all (1). This perhaps shows that these features are used, but not necessarily as accessibility features.

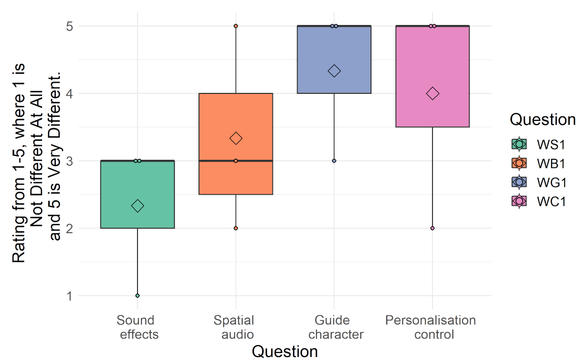


Figure 6.17: Content creator ratings of how different the enhanced audio features in the updated AR testing app were to their own workflows. The figure shows responses for the sound effects (WS1), the spatial audio (WB1), the guide character (WG1) and the personalisation control (WC1). Questions from Table 6.4. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

How likely to adopt

Encouragingly, Figure 6.18 shows that some content creators would be willing to adopt most of the added features in the updated AR testing app into their own workflows to help with accessibility, except possibly the personalisation control. They were unanimous in saying that they would be likely to use similar sound effects (WS2) and spatial audio (WB2), all responding with a 1. Whilst the others kept the same rating for the guide character (WG2), the XR content creator changed their rating to a 3, meaning this had a mean of 1.67. Interestingly, both of those who worked with interactive mediums were willing to adopt a similar personalisation control (WC2), whilst the VI film producer said they were unlikely to adopt this, meaning it had a mean rating of 2.33. Encouragingly, despite rating them both as very different to their own work (responding 5 to WG1 and WC1), the XR content creator was very willing to adopt them in their own work (responding 1 to WG2 and WC2). This is a positive results overall, but limited by the very small cohort size.

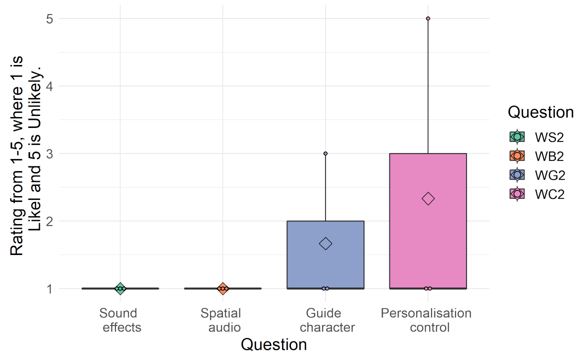


Figure 6.18: Content creator ratings of how likely they would be to adopt the enhanced audio features used in the updated AR testing app. The figure shows responses for the sound effects (WS1), the spatial audio (WB1), the guide character (WG1) and the personalisation control (WC1). Questions from Table 6.4. As an addition to standard box plot format, means are shown as diamonds and data points are shown as circles.

6.4 Interview Results

After completing the survey, all participants were invited to take part in an interview. 5 VI participants were willing to take part (P1,5,13,14&16). Unfortunately, only one content creator (CC1) was available, who is also fully blind.

The interviews were transcribed using built-in functionality on Microsoft Word for the Web [176]. Corrections were then made by the author. Once the transcripts were created, the author then summarised each interview, creating a list of the interviewee's thoughts on several themes. The themes looked for were their thoughts on their overall app experience and on each enhanced audio feature in turn, as these are the areas of greatest interest. The questions asked, shown in Table 6.5 reflect this. All interview summaries were then compared, with recurring thoughts or ideas noted, along with which participants subscribed to a particular view or approach regarding that theme. To capture the variability of VI experience, ideas are still reported even if only one participant mentioned them. Details of how to access these summarising documents as well as full interview transcripts can be found in Appendix F.

The following summary of the interviewee's thoughts and ideas that came up in the interviews is split up into categories based on these themes. Of course, there is some overlap between the different categories as some ideas branch over multiple. VI participants and content creator responses are aggregated, as there were many overlaps in their ideas, and only one VI content creator took part. Linking phrases, such as 'you know' or 'like' were removed from quotes so that they are easier to read in this discussion.

Overall App experience

Survey feedback has already shown that most participants found the app difficult to use. In particular, those who relied on screen readers further mentioned that *'direct touch is a problem'* (P14). For most this seemed to be a hurdle they were able to get over when setting up *'once I got into it, it was all right'* (P5), but it did mean they were starting off the experience negatively before even getting to AR functionality. Once they got into the app, many found it very difficult to use AR functionality. Some of this was a fundamental problem

with understanding the concepts and terminology of AR, which are explored in greater depth in the 'Guide Character' category of responses below. This meant that several participants (CC1, P5, P14, P16) had difficulty finding AR objects, either after they had placed them *'the hardest bit was finding them again'* (P5), or knowing how to approach them so they could then interact *'I couldn't approach the trees and things and select them'* (P16). Those who did manage to find AR objects and select them then struggled with more advanced interactions, such as moving or resizing them (P5, P13), often managing these interactions *'by accident the first time'*.
However, some participants did have a more positive experience with the app. For example, P1, who used the initial version in Study 2 found it much easier to use, ending up spending *'quite a bit of time using it, but not because I had to but because I was enjoying it'*. They even intended to use the screen reader but found that they didn't need to because *'the additional features allowed me to be able to navigate without it'*.

Although participants had difficulties with much of the app, this didn't stop them from being positive about the overall concept and the potential of the features that were used (P14, P5, P1, P13, P16). P14, who was fully blind mentioned that even though they found interaction with AR almost impossible, *'you could see that a lot of thought and a lot of you know activity had gone into the app, the app is quite thorough'*, that they *'think it's great potential'* and that they thought *'you've got a lot of it right... more right than wrong, just the fine tuning'*. Several also mentioned that *'it was something I felt I could participate in'* and that their feedback was making a real difference.

Participants had lots of suggestions about how the app could be improved to be more accessible for them. Some of these don't fit within the following categories and some draw from multiple categories, so they are included here:

- Making the visuals in the app more accessible (P1, P5, P13). Participants talked about the fact that *'although it's very audio, we quite often sometimes you just have to rely on the little bit of vision you've got'* (P5). They even mentioned enjoying using their residual vision, but only if this was possible with the content provided. The audio then supplements this and provides a fallback to provide missing information.

- Given how difficult it was, P13 wanted *'something that would indicate a bar to how much you've done with that'*. This progress bar would help motivate them by showing them *'how far along you are'*. No other participants explicitly mentioned this idea, but several did mention the idea of having positive feedback to help them know when they were progressing through tasks.

- Having specific sections of the screen that are used only for user interface or for AR elements. CC1 and P16, who were both fully blind liked this idea as it would make it a *'completely separate area for different things, so it's broken down'* meaning they would find it easier to explore their screen with their fingers (as screen reader users often do) as they have clear starting points for finding what they needed. However, P5, who is blind with residual vision but does rely on screen readers thought *'that just over complicates it'*.

- Having some method of building up a mental map of the AR playspace. Although some participants did find sound effects and spatial audio helpful to remind them *'that's what I did, I put that there'* (P1), this was not enough for many. Especially those with less sight. Whilst the extent to which different audio categories helped participants find AR objects is discussed later, several participants (CC1, P1, P5, P14) suggested the idea that they *'like to see things mentally in my head'* (P14) and that *'What you're needing to build in your mind is a map, a screen map'* (P5). Although many of the solutions suggested using the guide character more and are discussed there, they also involved higher-level changes to the app, for example:

- *'An inventory of what's in your forest and you just go down and it's just the numbers of things that you put in'* (CC1)
- Providing richer contextual descriptions such as *'you might say what, like scan my surroundings and that could be whatever parameter that you set as the game designer and that's going to tell you what's within 3 metres of you, all around you'* (CC1), which could *'use a clock as a compass, so if there was a tree at 3 o'clock, it would be good to hear tree, 3 o'clock you know rather than beep beep'* (P14). P5 and P16 also agreed with this idea.

Spatial Audio

Whilst several participants talked about the spatial audio helping to immerse them in the app, only one found it useful as *'a reminder for me of where I'd put things'* (P1). There were two main reasons why the rest didn't find it helpful. The first was that there was *'too much spatial noise, sometimes there [were] children chattering in the background and there [were] other things going on'* (P14), which made it hard to pick out the object they were trying to locate. Similar ideas are discussed in greater depth later under the sound effects category. The second reason was that it was not accurate enough, that *'the indication of distance wasn't that good'* (P16), and *'I don't know if they actually change as in getting loud or not to me I didn't get that impression at all'* (P13) and *'I can't distinguish that. I can work out that it is all coming from somewhere. But I can't place it'.* There are several possible explanations for this, which may be interlinked. The first two relate to the accuracy of the technology itself and the second two relate to users' ability to accurately localise sounds:

- The binaural technology used may simply not be accurate enough for a small, tabletop sized AR playspace, with objects all within an approximate distance of 50cm. The fact that participants struggled even in Scene 1 where they had control over how many audio objects were in the playspace weakens this argument as they could place fewer objects at greater distances from each other. However, such small differences in location can be hard to pinpoint for even expert listeners.

- As generalised Head-Related Transfer Functions (HRTFs, see Section 2.3.2) are used, it could also be argued that the quality of the spatial audio is sub-optimal for accurate localisation, although this is an ongoing debate [183].

- Spatial accuracy can be difficult to establish if you are unfamiliar with spatial audio, especially if lower-quality listening methods are used, such as the Apple Airpods used by most participants.

- Auditory spatial accuracy is also strongly affected by visuals, to the point where visual cues dominate the perceived position of a linked sound. This is known as 'the ventriloquist effect' [184–186]. As visuals cues are either non-existent or limited for VI users, they are less able to build up a mental map with visuals and then correspond sounds to that. Further work is needed on this, as it may be key to understanding how VI individuals can accurately understand immersive media.

Sound Effects

Discussion of sound effects was further split into two types; interaction sound effects. Most participants (P1, P5, P13, P16) mentioned that they enjoyed the sound effects. However, across both types, many felt *'there were so many other sounds I was watching for it just added another layer on top of another layer'* (P14). The number of sound effects of both types became overwhelming and confusing at points. P14 summarised it in this way: *'I think the idea of having the noise, the sound identifiers is a plus one, but once you have too many objects. Then you have a learning curve'* and likened it to having to *'to learn Morse [code] rather than a simple audio clip'*. This links to there not being an established language for explaining AR to VI users, as discussed later.

Across both types of sound effects, the suggested solution (P5, P14, CC1) was to use provide more spoken information. Specific thoughts on what this could look like are discussed in the 'Guide Character' section, but participants did think that *'you can do a lot of the interactive stuff without using too many words'* (P14). However, P16 did argue that *'if you just have everything voice, it's not as real'* especially as *'this is trying to be in an immersive experience and so using real sounds makes it feel more immersive'*.

Interaction Sound Effects. The main debate about this use of sound effects was over whether they immediately made sense, without necessarily having to use the sound effects key in the options menu. Several participants (P5, P13, P14) thought that they *'made perfect sense. I thought they were quite good 'cause I didn't need an explanation for what they were. So they must have been logical'* (P5) and that once you had used them once *'you know to look out for that or listen out for that for the next time'* (P13). However, P1 thought that not all of them were easy to understand, only those that *'had a bit of context to them as sound'*, using *'the resizing and the remove one'* which were a stretching and a wood chopping sound respectively. Both P1, P5 and P14 thought that *'some of the issue was having too much'*, that even if sounds made sense there were simply too many different interaction sounds to remember, especially when combined with all other sounds in the app. P5 suggested that the solution was *'an announcement again'*.

Object-specific Sound Effects. Participants were in greater agreement over the fact that it was harder to understand what was happening with this use of sound effects. They were split as to whether they could distinguish easily between different objects, for example telling different types of elements added to their forest in Scene 1. P5 and P14 thought that this was easy, whilst P1 and CC1 did not. The main problem was simply remembering what everything was (P1, P5, P14, CC1), CC1 stated that *'it would need many more hours I think to get my head around you know what the noises meant'* and P14 thought that it was because there was too much *'sound clutter'* meaning it ended up just being *'like interference'*. Given this feedback, perhaps use of less overlapping sounds if needed.

However, it seems to have value as many still mentioned that they were enjoyable and helped to immerse them in the virtual world. CC1 suggested the solution for this was to use spoken cues at least while *'you're getting used to the game and announcing what you've got in your forest, naming it would be great'*. P1 suggested that sound effects for each object should be played before adding that object in Scene 1, then they'd *'be able to kind of then know as I'm going through and navigating what that sound represents'*. The most novel suggestion was made by P16, who thought that there should be a way of *'a beacon on a particular thing that you were trying to get to so that it gives a very high pitched sound so* *you can really focus in on that particular thing'*, they also thought that this idea would be strengthened by the use of the spatial audio.

Guide Character

All participants except P14 enjoyed the guide character. In particular, they mentioned *'the fact that it's a proper voice, it's not a computerised thing'* (P1) which was *'very warm, friendly voice and not patronising. It sounded professional, as if you could trust what she's going to be telling you'* (CC1). They also mentioned that it *'makes a lot of sense actually, and it was good'* (P13) to have the same voice providing both screen reader functionality and hints within the app. However, as already established through the survey, the guide character didn't necessarily provide all the help it needed to (Section 6.3.4) and many participants initially struggled with its screen reader functionality (Section 6.3.3). Participants had a few main suggestions that would improve it for them:

- Clearer instructions and terminology: many of the completely blind participants (CC1, P13, P14, P16) didn't fully understand how AR worked or how to interact with it. They emphasised the need for everything to be explained, down to the *'angle of the phone being held'* (CC1). In particular, being clear on what different actions meant. For example, there was confusion around how to place AR objects; *'I thought what you were supposed to do is just select it just double tap it and that placed it'* (P16) and how to select AR objects. This was because the app only ever described selecting them, which was done by a single tap interaction. However, *'often when things say tap for screen readers, you need to do a double-tap so there was just a little bit of some sort of ambiguity there'* (CC1). This would be solved by having much clearer definitions of how different elements of the app could be interacted with, using a consistent terminology that was either *'used elsewhere [...] so people already know what you're talking about'* and if not was clearly stated, such as *'when you say tap you're always going to say double-tap or single-tap or whatever to sort of prompt us'* (CC1). In summary, the guide character needed to *'anticipate where somebody might misunderstand'* (CC1).

- Given their difficulties using the AR, multiple participants (CC1, P5, P13, P14) suggested that they needed to be eased into it much more slowly, having *'a little practice in what it meant'*. This would mean *'stripping it right, right, right, right back and then just sort of building it up'* (CC1) by *'adding three and finding three. I'd have had a better experience and then the next time 5 and then the next time 6 and then in five days time I could do 10'* (P5). This would allow them to *'confirm that your technology is working and also mean that your instructions could be simpler'* (CC1), without just using instructions because *'you can't retain the information [...] if you have written instructions, people who are sighted can flick backwards and forwards to them, you can't'* (P5). This would allow a better understanding of other sounds in the app as *'your brain or your ears can then build it up if that makes sense'* (P5). This makes sense regardless of a user's level of sight, as AR can be difficult to interact with for the first time for anybody.

- Providing feedback on what objects are within a certain vicinity (CC1, P14, P16). This links back to and is further explored in the 'Overall App Experience' discussion category.

- Once AR objects have been located, provide spoken feedback on what they are (P1, P5, P13, P14, CC2, P14, P16) *'you can put a label on some of the trees, says that's a beech tree or this is a red oak'*. This could also provide information on how to interact with the found object, and feedback as interactions occurred such as *'if something said fern selected that that would help'* (P14).

- Having the option to speed it up, whilst keeping the custom voice. This was particularly suggested for the screen reader functionality of the app as they are used to *'having an option. Having a choice, of a man or a woman or fast speed, slow speed snail pace'* (P5). However, P16 stated that *'I'd put this change in the 'nice to have but not essential' category'*.

- P1, who does not rely on screen reader functionality suggested that it would be helpful to have *'a compromise in between that if we clicked on something it would give you audio feedback'* so that they could ask the guide character for help with interactions and user interface elements, but without having to use it fully as a screen reader.

The debate over whether the guide character should automatically provide help or not was continued from Study Two. P1 asked *'I don't know if that's an option to have in the options as in, it's enabled or something automatic hints enabled'*. CC1 also stated that they were *'definitely in the camp of the people that would like that guide voice talking lots and really looking over my shoulder. Now that is partly because it's a brand new game to me'*, but suggested having a '*minimal help mode or maximum help mode'* to allow some customisation of this. However, most (P1, P13, P14, P16) thought that the method used of being able to ask for hints was right *'as long as the hint button is always there, you can always tap it. Leave it to the individual'* (P14) because *'is quite a good principle to kind of have, let someone try to figure it out themselves'* (P1) and *'sometimes you just want to do things on your own'*. They thought it was *'nice to know that it's there'* (P5) and that *'as you got further into it, they did change, so I kept going back to them'* (P16) meaning that *'when they had more information, we obviously used it, but if it didn't, if it's just the same repeat message you said all right, I'll move on a bit here and see anything changes that way'* (P14).

Personalisation Control

The participants didn't discuss the personalisation control much. This is perhaps an indication that it was less crucial to making the testing AR app accessible to them than other features discussed. Further to this, some participants (P4, P13) thought it was too subtle *'I thought that's a really good idea and it may be, but it didn't. It didn't really make any [difference]'* (P13), whilst one *'didn't see a need for it'* (P5). They instead suggested that *'if I was going to use the slider for anything I would quite like to be able to use it to speed the voice up'*. However, P1 did find it useful when they got stuck in the app, as they *'had most of the sounds on, like pretty much all of the time'* but *'at certain points I turned it down a little bit'*, particularly when they were in unfamiliar scenes or struggling with tasks. They did want to be able to choose sound categories for the control to leave unaffected, which is an interesting idea, although it does undermine some of the simplicity central to the original concept. P13 did think it was a good idea because it made the app *'like it's my game, it's what I'm controlling'*.

6.4.1 Interview Recommendations

From the summary of interview responses and themes discussed, it's possible to formulate several recommendations:

- Participants struggled with all areas of the app, from screen reader functionality to finding AR objects and then interacting with AR objects once they had been found.
- Despite the difficulties faced, participants were positive about the strategies used, or rather the potential of the strategies to be effective once fundamental issues were resolved.
- Spatial audio did not help users to precisely locate objects they were looking for. This could be an issue with spatial accuracy, or with difficulties caused by an overly busy soundscape.
- Whilst sound effects were enjoyable, there were too many, which made them confusing rather than useful. A smaller number of sound effects that overlap less and make sense without having to be explained should be used. Sound effects can also not be relied upon to identify different objects, especially if the objects are fairly similar, so spoken cues may be more understandable.
- The guide character did not provide enough information to actually understand how to interact with AR. Clearer instructions and explanations of what terminology means is needed.
- Spoken instructions should form the basis of explaining the virtual world; such as describing objects within a certain vicinity, providing descriptions of what selected objects are and how they can be interacted with, and giving feedback as interactions occur.
- Although participants enjoyed the guide character being a real human voice, with a script custom-made to fit in with the narrative of the app, they still wanted to be able to personalise it in ways they are used to, such as speeding it up in a manner similar to standard screen readers or being able to choose how much help it provided.
- The personalisation control did not particularly help with making the app more accessible, although some participants found it enjoyable and appreciated being able to customise the app in a novel way.
- Although it would be easy to assume that solutions should just present all information to VI users immediately, the idea of there still being some challenges so that experiences are engaging was important. All information on AR fundamentals must be provided clearly, but VI users should be allowed agency to calibrate the difficulty of experiences when then completing specific tasks.

6.5 Discussion of Study Three

In Study Three quantitative assessments of difficulty demonstrated that the updated AR testing app with enhanced audio accessibility features added was no easier to use than the initial AR testing app examined and even faced many of the same difficulties. One thing that would be easy to focus on is the failings of the screen reader functionality of the app. In particular that it was difficult for participants who needed it to set it up and it didn't always allow users to explore their screens in the natural ways expected. Although the conclusion from this could be that it highlights a problem with Unity not having comprehensive screen reader support, the fact that the majority of AR content is made in Unity [14] means that it is therefore a problem for creating accessible mobile AR as well. However, the fact that how difficult participants found the app was not linked to sight (and therefore wasn't linked to screen reader usage) demonstrates that the problems they faced with using AR run deeper. Therefore this study once again strengthens the argument for allowing personalisation in accessibility methods. Further investigation of whether the enhanced audio features helped with various AR tasks showed that they simply did not provide the information needed. They were most helpful to both VI participants and content creators for identifying the type of AR objects around them and least helpful for identifying how they were interacting with objects. Interestingly, sound effects seemed to be most helpful for many of these tasks (still within the range of not being helpful overall), but the guide character was rated the most helpful in general and for navigating the app. Enhanced audio features that are becoming increasingly common in immersive media did not go far enough to provide meaningful accessibility. This highlights a problem for the development of immersive media, that more fundamental considerations are needed than assuming that creative use of audio alone can provide meaningful accessibility for VI users.

There was a marked difference between the responses of sighted content creators and VI participants to what difficulties they faced and their suggestions for improving the app. This highlights a strong disconnect between how content creators and those with access needs view experiences.

However, there were a number of encouraging signs. Participants did think that sounds within the app helped and enjoyed the added enhanced audio features. It did change the perception of the majority to think that it was easier to understand. The interviews further revealed that the VI participants thought that although the accessibility features used did not provide the information they needed this time, they did hold the potential to, with many suggestions on methods for this. This also demonstrates the value of qualitative interviews. The survey shows that the app was difficult to use, but the interviews allowed participants to clarify the strategies are positive, but they don't solve the more fundamental problems.

Perhaps the most important suggestion is that although sound effects and spatial audio are enjoyable, spoken cues are by far the most important for actually providing the information needed. Mascetti et al [54] had a similar conclusion when investigating sonification of guidance data to help VI individuals cross roads, that although people like sonification it needs frequent explanation. Although both previous studies did also conclude spoken cues were the most important, this study demonstrates this is the case even with other enhanced audio features. It does show that VI individuals will find it more enjoyable and immersive if these spoken cues are designed as part of the virtual world being explored. The suggestions provided by participants about what information it should provide, how terminology can be made clearer and that richer contextual information may hold keys for a future guide character that can provide all the information needed. It was also encouraging that content creators also saw the potential of the strategies, and were very willing to adopt them into their own work.

In terms of allowing VI users to personalise their experience, many enjoyed the personalisation control and some did use it to highlight sounds and help with tasks. However, it did not make the experience accessible in meaningful ways. Other personalisation options were suggested, such as choosing the level of help provided by the guide character or being able to alter the speed of their voice.

6.6 Chapter Summary

This chapter addressed the final objective of this work, assessing the efficacy of the accessibility strategies proposed. It has done so by completing the 5th and 6th sub-objectives outlined in Section 1.2.2, evaluating the updated AR testing app to obtain VI and content creator feedback that establishes the feasibility of the strategies proposed with the following results:

- The updated AR testing app was no easier to use with enhanced audio features added, meaning that the manner in which strategies were used here does not create more accessible AR for VI users.
- The strategies have potential to be feasible, but more fundamental considerations are first needed. Positivity around their potential and the wealth of suggestions given show that with further work, they could create more accessible AR for VI users in enjoyable, immersive ways.
- Although spatial audio and sound effects are enjoyable and immersive, using them to provide the information VI users need in AR is currently difficult.
- Spoken cues are by far the most important for VI users and how they are most used to receiving information. However, to work in an immersive, interactive AR medium they need to have extremely clear instructions and terminology. Further, creating cues that respond to the interactive, explorative nature of the AR medium with rich contextual information may be paramount to making AR content accessible.

Chapter 7

Conclusions

7.1 Introduction

This work set out to establish whether enhanced and personalisable audio strategies could be used to create more accessible Augmented Reality (AR) for Visually Impaired (VI) users. What was discovered in this process was that AR is not only difficult for VI individuals but that the accessibility challenges that exist are much more fundamental. The needs and voices of VI individuals are not being considered even at the fundamental level of the game engines and tools used to create AR. Therefore, they are being excluded. This work has highlighted that much more needs to be done in AR, with VI individuals central to developments before a true consideration of the strategies proposed can be completed. The following sections demonstrate how these conclusions have developed, expound upon them and provide thoughts on how future investigations may begin to solve the problems discovered.

7.2 Summary of Objectives Addressed

This section contains a self-contained summary of the objectives addressed in work, including the methodologies and results used in doing so.

The first objective of this work was to establish the current state of accessibility research in Augmented Reality (AR) for Visually Impaired (VI) users. Chapter 2 addressed this objective, with the following contributions:

- A consideration of the needs of VI users themselves and an overview of existing accessibility solutions that aim to meet these needs. This highlighted that their needs are often not considered, with much of society remaining inaccessible. Technology can either provide solutions to mitigate this by making the world more accessible, or it can become yet another barrier. This depends upon whether accessibility is considered from the beginning of design processes or not. It also highlighted just how diverse the needs of VI individuals are. A fundamental principle of this work is derived from this, that accessibility attempts should take this diversity into account by providing personalisable services.
- An examination of existing research, technologies and accessibility standards concluded that creative use of enhanced audio is a powerful, but as yet under-utilised tool for the inclusion of VI individuals. In particular, spoken cues are the primary method

for communicating information. Consideration of the Enhancing Audio Description (EAD) Project [19, 27–30], Ward's Narrative Importance (NI) work [21, 31, 32], Audio Games and Audio Augmented Reality (AAR) demonstrated this potential. However, accessibility guidance for AR, mobile AR accessibility research, assistive AR research and assistive mobile AR apps show that although excellent progress is being made, it is mostly limited to using AR as an assistive tool and not considering ways that all AR content can be made accessible to VI users. They also show spatial audio and personalisation techniques remain under-explored. This is surprising, given the adaptive, multi-sensory and immersive nature of Extended Reality (XR) technologies such as AR.

- Based on these points, it was proposed that 'it is possible to use creative and personalisable audio practices to develop mobile AR content which is accessible to VI users and includes them in enjoyable, immersive ways', which is the central hypothesis of this work.

The second objective of this work was to establish the accessibility of currently available AR apps. Chapter 3 addressed this with the following contributions:

- An assessment of 37 current mobile AR applications (apps) against both standard accessibility features and enhanced audio features found many shortcomings. Crucially, compatibility with built-in screen reader features is low and alternate solutions do not go far enough to provide meaningful support. Further, almost non-existent use of spatial audio and minimal use of sound effects demonstrates that creative use of audio, both in general and as an accessibility tool is simply not a consideration in most mobile AR app development.

- VI users validated these results by using Angry Birds AR, Pokémon Go and WWF free rivers in Study One. This highlighted that VI individuals perceive XR to be difficult, visual and not designed with their needs in mind. Participants found the apps extremely difficult to use and strongly emphasised the need for spoken cues. They were also optimistic about other sounds helping them, especially sound effects. Finally, Study One strengthened the argument for personalisable accessibility solutions, as participants' level of sight did seem to impact how difficult or enjoyable they found the apps.

The fourth objective of this work was to assess the efficacy of the accessibility strategies proposed, answering the central hypothesis of this work. Chapters 4 - 6 addressed this with the following contributions:

- A bespoke mobile AR testing app was developed, based on 'The Turning Forest'. Study Two confirmed that this was equivalent in accessibility to the currently available apps assessed in Study One, with spoken cues once again the most requested feature. Both content creators and VI participants supported the idea that well-designed audio could both aid accessibility for those who need it and also create a more enjoyable experience for everyone. There was much support for having a guide character that was designed specifically for the AR app, with sound effects, spatial audio and a personalisation control supporting it by highlighting state changes, interactions, the appearance of new things and information on tasks that need completing. Their suggestions were used to update the AR testing app with enhanced audio features.

- Study Three demonstrated that the AR testing app was quantitatively no easier to use than its previous iteration in Study Two, or the currently available apps assessed in Study One. It determined that the enhanced audio features and personalisation control did not go far enough in providing the information needed for VI users of AR. Interview feedback in particular determined that this does not mean that these strategies are not feasible, but that they are not feasible in the way that they are implemented here. Participants were positive about their potential, contributing a wealth of suggestions for possible future implementations. In short, enhanced audio features must support clear spoken instructions which have well designed and consistent terminology. Returning to the central hypothesis, these strategies are not enough to make AR accessible by themselves at this point. Further fundamental understanding of the information needed by VI users to make AR accessible to them is necessary. Once this is available these strategies may be viable to aid the accessibility of AR content. Especially as there was support for their ability to aid immersion and enjoyment.

Based on these contributions, it was possible to answer the two research questions posed:

1. *What challenges do the dual purposes of the soundscapes in XR content, narrative and navigation, present for this approach to accessibility*

   This work showed that these dual purposes of soundscapes in AR are so interlinked that the effects of accessibility solutions on them must be considered on both concurrently and with equal consequence. The results in these studies also highlight a further and equally interlinked dimension: interaction. The solutions used did not give VI participants the information needed on how they were interacting with AR. Without being able to interact with it, they were unable to navigate or progress through narrative content. In previous applications of EAD strategies [19, 27–30] to television, VI users both have a higher level of pre-existing familiarity with the medium and do not need to participate in the medium in the same way as for interactive AR. At the moment AR is a difficult medium for all users, but particularly VI users, and for soundscapes to provide accessibility, precise consideration of how they can give clear, responsive and contextual information is needed.

2. *How can the concept of Narrative Importance (NI)[21, 31, 32] for personalisable accessible audio be extended into the audio description domain for blind and partially sighted viewers?*

   The central premise of this method, providing users with personalisable control of their audio world is applicable to VI users and AR. VI participants enjoyed using an NI-based control and some did use it to help their progress. However, for this control to be meaningful, further options such as choosing the level of help provided by a guide character or being able to alter the speed of spoken cues were suggested for further exploration.

7.3 Contributions

By addressing the objectives set out, this work presents three main contributions to the field:

1. **Evidence of the inaccessibility of AR**. This work has highlighted just how far there is still to come to create mobile AR content that is truly accessible to VI in-

dividuals, in interesting and immersive ways. The review of existing apps and Study one particularly highlighted this, showing that the lack of support for even standard accessibility features means that AR is difficult and unenjoyable for VI users. The fact that tools such as Unity, which is used to create the majority of AR content [14], do not support screen readers shows that at present, the needs of VI individuals are not being considered, meaning that they are being excluded. There is some hope in the acknowledgement of access needs by emerging organisations advocating and providing standards for XR. The development of assistive AR research and assistive mobile AR apps provides a glimpse into just how powerful AR could be for AR, but ultimately highlight just how far there is for the majority of AR content to come.

2. **A reinforcement of the importance of personalisation**. Across all studies, the difficulties VI participants faced and their suggested solutions were demonstrably not linked to their level of sight. Accessibility is therefore personal and any attempted solutions must provide for this individuality by being personalisable and VI individuals themselves should be central to design processes. Further, personalisation may afford greater engagement with experiences, as giving users control can give them greater ownership and allow them to adjust experiences to be more enjoyable to them. For example, multiple interviewees in Study Three mentioned that they would like to alter the difficulty of the experience by calibrating the hints provided to their needs so that there could still be a challenge and therefore enjoyment in the experience for them.

3. **The need for a fundamental understanding of the needs of VI users in AR**. It is perhaps unsurprising that this work's attempts to use enhanced and personalisable audio practices to create more accessible mobile AR were not successful. What was highlighted is that a much more fundamental understanding is needed of the ways that VI users interact with technology, how these ways could match up with the features of the AR medium and how accessibility solutions could be provided that make the most of the immersive, adaptable and interactive nature of XR media. Herskovitz's [18] taxonomy of AR building block tasks provides a strong starting point for this and VI individuals themselves need to be involved in future developments. The enthusiasm that the creative access solutions received suggests that if this fundamental problem is addressed, rapid improvements could be made to make AR accessible and immersive for VI users, such as using the strategies explored in this work. However, the results across all studies show that strategies cannot be taken from linear media such as film and TV and applied to immersive, interactive media such as AR. In particular, the results highlight just how important spoken cues are. They cannot be replaced by other audio features, but well-designed sound effects and spatial audio can support them. The importance of clear spoken cues, particularly for technology users new to the medium was the missing link in the final study. If a fundamental understanding of the needs of VI users in AR is reached, then it will be possible to create clear and consistent instructions and terminology for spoken cues and to have a solid appreciation for how other enhanced audio features could meaningfully support accessibility and augment how immersive and enjoyable experiences are. A participant interviewed in Study Three put it this way: *'it was an excellent piece of work. With lots of endeavour and lots of well thought out solutions that need just a little bit of tweaking'*.[1]

[1] P14 in Study Three Interviews.

7.4 Further work

Further research directions are suggested below.

To build on this work and exploit the potential of the strategies proposed, the following suggested avenues of investigation are made. However, these rely on the wider and more fundamental areas proposed later:

- Herskovitz's [18] AR building block tasks should be matched up with enhanced and personalisable audio features so that their potential can be maximised by communicating more effectively the information needed in AR.
- As spoken cues are the most necessary and helpful to VI users, they need to have clear instructions and terminology. This needs to be based on the common language mentioned above. Further, creating cues that give rich, responsive contextual information will make the most of the interactive, explorative nature of the AR medium. For example, describing objects within a certain vicinity, providing descriptions of what selected objects are and how they can be interacted with, and giving feedback as interactions occur.
- Investigation is needed into how sound effects can effectively communicate information to the user. Both in terms of the initial learning curve, and preventing numerous layers from becoming confusing. Directions for this may include research into whether specific frequency separations are needed, or whether for VI users spearcons [49] (sped up spoken cues) are more effective.
- Consideration of how spatial audio can provide useful information in AR situations. This may mean further research is needed into how whether spatial audio can provide accurate localisation cues in all sized AR playspaces. It may also be linked to simplifying soundscapes so the cues given are distinct.
- Further work is needed to understand exactly what personalisable audio options would be most helpful. Altering relative levels based on how important sounds are may be more useful if the other suggestions above prove effective. Or it may be even less necessary if the ability of spatial audio to allow users to focus on sounds of interest is improved. However, other options should be explored, such as providing control over the level of enhanced audio features, the number and types of informational hints provided or the speed of spoken cues (if custom screen readers which do not provide this option are used).
- Application of the workflows created beyond using basic Unity workflows, as many audio programmers for gaming and AR use tools such as Wwise [177] or FMOD [187].

At a more fundamental level than applying enhanced and personalisable audio features to create more accessible AR, several areas must be investigated and paradigms changed:

- VI individuals must be involved in any and all research and creative processes aiming to make AR more accessible, and even just those aiming to make AR. Their individual needs and experience mean that accessibility solutions cannot be considered after the fact, and cannot be designed with a 'one size fits all' mindset. Iterative design processes may be especially needed as the challenges faced are complex. As shown by this work, it is only once initial problems are solved that further challenges will present themselves.

- Our understanding of how the features of the AR medium match up with how VI users interact with technology is far too limited. Herskovitz's [18] work creating a taxonomy of AR building block tasks has provided a starting point. This is needed before the creation of any accessibility strategies, whether they used enhanced audio or not.
- Based on the above, a common language needs to be built up to aid VI users' understanding of AR. This needs to communicate everything from fundamental usage considerations upwards.[2] Anticipating the issues every user may face is impossible, but ensuring that all the information is available to them in clear, understandable ways is vital for inclusion.
- The mobile AR testing app in this work only attempted solutions in a tabletop sized AR playspace [110]. Play-space size affects the types of interaction and levels of exploration, such as moving around physical space. Therefore consideration is needed of whether unique accessibility (and possibly safety) solutions are needed for each size of playspace.
- Given that AR aims to be an immersive medium, it's not enough to just make it 'work' for VI individuals. Thought is needed on what it means for a VI individual to be immersed in a virtual world. The conceptual model of immersive experiences put forward by Hyunkook Lee [112] provides a starting point for this. Further work should assess all elements of VI individuals' immersion in an immersive system (physical presence, social/self-presence, involvement).

7.5 Concluding Thoughts

After using the final AR app in Study Three, a fully blind participant stated that *'the small steps we take now, maybe in 10 years might make a difference to somebody else. I might even be able to use it better than somebody else'*.[3] The small step taken by this work is to highlight just how inaccessible AR is for VI users and attempting to explore ways that it could be made accessible using enhanced and personalisable audio features. It is hoped that many more small steps will follow.

[2] P16 in Study Three highlighted that they need to know 'how to hold my phone - horizontally or vertically'.

[3] P14 in Study Three Interviews.

Appendices

Appendices C - G contain information on the contents and (where relevant) the usage of the digital appendices accompanying this work. For ease of reference, Appendix B contains aggregated details of participants, linked across all 3 studies.

Appendix A

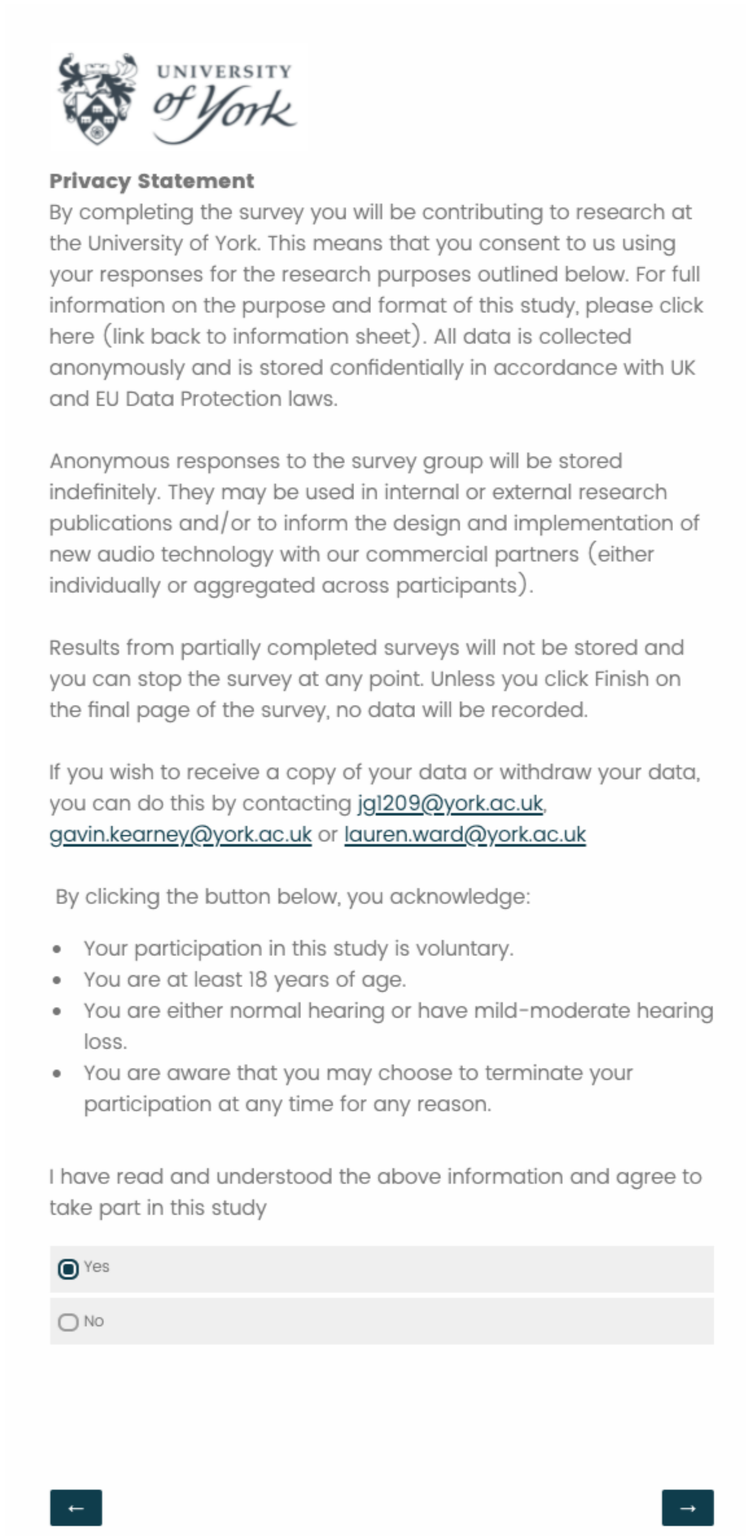
Ethical Approval

Ethical approval for all the participatory research carried out in this MSc was obtained from the University of York Physical Sciences Ethics Committee with the following approval codes:

- Study One - Gregg300321
- Study Two - Gregg241220
- Study Three - Gregg240621

In all studies, users were given information on their privacy, anonymity and data protection. This is shown in Figure A.1. Phrasing will have changed slightly depending on what information was being collected (i.e. survey interview). All versions can be seen with the relevant forms/surveys for each study, details of which can be found in Appendices D, E and F.

To allow for anonymous storage and identification of their data, unique identifiers were collected, these are shown inn Figure A.2.



UNIVERSITY of York

Privacy Statement

By completing the survey you will be contributing to research at the University of York. This means that you consent to us using your responses for the research purposes outlined below. For full information on the purpose and format of this study, please click here (link back to information sheet). All data is collected anonymously and is stored confidentially in accordance with UK and EU Data Protection laws.

Anonymous responses to the survey group will be stored indefinitely. They may be used in internal or external research publications and/or to inform the design and implementation of new audio technology with our commercial partners (either individually or aggregated across participants).

Results from partially completed surveys will not be stored and you can stop the survey at any point. Unless you click Finish on the final page of the survey, no data will be recorded.

If you wish to receive a copy of your data or withdraw your data, you can do this by contacting jg1209@york.ac.uk, gavin.kearney@york.ac.uk or lauren.ward@york.ac.uk

By clicking the button below, you acknowledge:

- Your participation in this study is voluntary.
- You are at least 18 years of age.
- You are either normal hearing or have mild-moderate hearing loss.
- You are aware that you may choose to terminate your participation at any time for any reason.

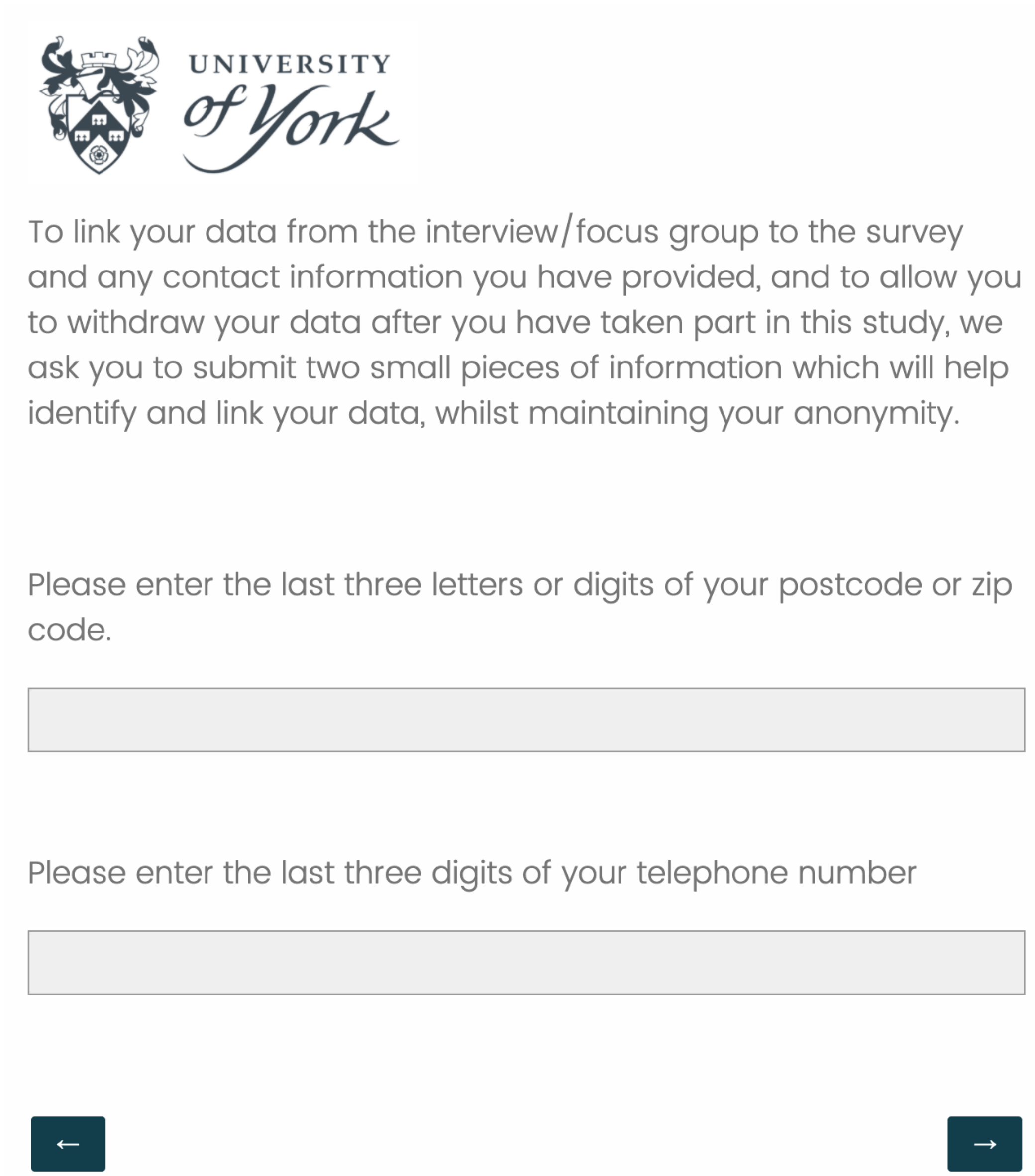
I have read and understood the above information and agree to take part in this study

Yes

No

← →

Figure A.1: An example of the privacy statement given in each study at every point data was collected. Shown here is the privacy statement from Study One. Small changes were made depending on the kinds of data being collected.



UNIVERSITY of York

To link your data from the interview/focus group to the survey and any contact information you have provided, and to allow you to withdraw your data after you have taken part in this study, we ask you to submit two small pieces of information which will help identify and link your data, whilst maintaining your anonymity.

Please enter the last three letters or digits of your postcode or zip code.

Please enter the last three digits of your telephone number

← →

Figure A.2: Screen showing the unique identifiers participants were asked for in each study.

Appendix B

Participants

This appendix contains anonymised details of the participants who took part in all three studies in this work. Information on their demographic information is aggregated across the studies, to allow for ease of comparison across studies.

Table B.1 details the VI people who took part in Study One (see Chapter 3), Study Two (see Chapter 4) and Study Three (see Chapter 6).

Table B.2 details the demographic of the content creators and Table B.3. These content creators were recruited to take part in Study Two (see Chapter 4) and Study Three (see Chapter 6).

Table B.1: Demographic of participants recruited for all studies. Text in italics shows where participants gave extra information. A * in the studies column indicated that this participant took part in an interview as well as the survey. The responses shown are taken from the survey they filled out first, as ages (B1) and perceived familiarity with XR (B4-4.1) may have changed throughout the project period

ID	B1	B2	B3, *B3.1*	B4, *B4.1*	
	Age	Hearing Loss	Visual Impairment	Familiarity with XR	Studies
P1	32	No	Blind with residual vision	I'm familiar with it and have used it a couple of times, *Virtual reality headsets for some games*	1,2*,3*
P2	23	No	Blind	I'm familiar with it, but have never experienced it myself	1
P3	63	No	Blind with residual vision	I'm familiar with it, but have never experienced it myself	1,3
P4	39	No	Blind with residual vision, *no central vision, some peripheral so can tell when a screen changes*	I'm familiar with it, but have never experienced it myself	1,2*,3
P5	39	No	Blind with residual vision	I've never heard of it before	1,3*
P6	21	No	Partially sighted	I'm familiar with it and have used it a couple of times, *VR headsets and some mobile AR games*	1,2*
P7	19	No	Other, *Strabismus*	I'm familiar with it, but have never experienced it myself	1,2*,3
P8	20	No	Other, *short sighted in one eye and long sighted in the other, registered sight impaired in one eye and severely sight impaired in the other. Eyes also shake making text and depth difficult*	I'm familiar with it and have used it a couple of times, *AR, unsure what exactly*	2*

Table B.1 - continued from previous page

P9	61	No	Blind	I'm familiar with it, but have never experienced it myself	2
P10	29	No	Blind	I'm familiar with it, but have never experienced it myself	3
P11	35	No	Blind	I've never heard of it before	3
P12	20	No	Partially sighted	I'm familiar with it and have used it a couple of times	3
P13	53	No	Partially sighted	I've never heard of it before	3*
P14	65	No	Blind with residual vision	I'm familiar with it and have used it a couple of times	3*
P15	21	No	Blind	I'm familiar with it, but have never experienced it myself	3
P16	63	No	Blind	I'm familiar with it, but have never experienced it myself	3*

Table B.2: Demographic of Content Creators recruited for all studies. Text in italics shows where participants gave extra information. A * in the studies column indicated that this participant took part in an interview as well as the survey. The responses shown are taken from the survey they filled out first, as ages (B1) and perceived familiarity with XR (B4-4.1) may have changed throughout the project period

ID	B1	B2	B3, *B3.1*	B4, *B4.1*	
	Age	Hearing Loss	Visual Impairment	Familiarity with XR	Studies
CC1	45	No	Blind	I'm familiar with it, but have never experienced it myself	1,2*,3*
CC2	37	No	No	I'm familiar with it and have used it a couple of times, *Played a few VR games. Tried a friend's AR simulation software aimed at offshore oil workers.*	2*,3
CC3	23	No	No	I'm familiar with it and have used it a couple of times *Mostly VR games - Beat Saber*	2
CC4	39	No	No	I'm familiar with it and use it regularly, *I have experienced and designed audio-based virtual and augmented reality "experiences", mainly perceptual/sensory training and assessment routines.*	2
CC5	24	No	No	I'm familiar with it and have used it a couple of times *Audio-visual AR and VR, audio AR*	2
CC6	34	No	None	I'm familiar with it and have used it a couple of times, *console games*	3

Table B.3: Professional Experience of Content Creators recruited for all studies. Text in italics shows where participants gave extra information. Questions PRO6-8 are not shown here, as they were only asked in Study 2, so are analysed in that chapter.

ID	PRO1, *PRO1.1*	PRO2, *PRO2.1*	PRO3	PRO4, *PRO4.1*	PRO5	PRO9
	Medium	Genre(s)	Organisation	Majority of work	Years as content creator	Worked on XR production
CC1	Television	Drama, Entertainment, News, Sports, Comedy, Music, Action, Adventure	Other, *not for profit*	Sound Recordist/-Producer	3	No
CC2	Gaming	Action, Adventure	Other, *Game development studio*	Audio Programmer	15	No
CC3	XR (including VR, AR and MR)	Entertainment	National or Large Broadcaster	Researcher	0.5	No
CC4	XR (including VR, AR and MR)	Music, Simulation, Other, *Health-related (mainly audiology), i.e. implementation in AR/VR of clinical audiological assessments*	University	Researcher	5	No answer
CC5	XR (including VR, AR and MR)	Other *Don't work with content, more so experiences*	National or Large Broadcaster	Researcher	0.5	No
CC6	XR (including VR, AR and MR)	Documentary, Simulation, Other, *VR training*	Small Independent Production house	Audio Designer	10	Yes

Appendix C

Context

This supplementary resource contains full feature comparison tables of the apps and research projects investigated whilst establishing the current state of AR accessibility, which are detailed in Section 2.6 and Section 3.2.

Within this directory is:

- ***/Mobile AR Apps:*** *this folder contains a spreadsheet with full details of the 37 mobile AR apps investigated in Section 3.2. This provides further information on how they match up with the assessment criteria. It also contains the counts of how many apps used each feature and image files of the bar charts used for ease of comparing the usage of the assessed features. Also included in the spreadsheet are details of the redeveloped original (without accessibility features) version of the Turning Forest AR testing app, as detailed in Chapter 4 to allow for assessment of how accessible it is compared to currently available Mobile AR apps. The details of the Turning Forest are not included in the counts investigated in Chapter chap:StudyOne.*

- ***/AR Research Projects:*** *this folder contains the original spreadsheet with details of the assistive AR research projects investigated in Section 2.6.4, with the counts of how many used each feature and image files of the bar charts used for ease of comparing the usage of the assessed features.*

- ***/Assistive Mobile AR Apps:*** *this folder contains a spreadsheet with full details of the assistive mobile AR apps investigated in Section 2.6.5. Also included in the spreadsheet are details of the redeveloped, accessible version of the Turning Forest AR testing app, as detailed in Chapter 5 to allow for assessment of how accessible it is compared to currently available assistive Mobile AR apps. The details of the Turning Forest are not included in the counts investigated in Chapter chap:LitReview.*

Appendix D

Study One

This appendix includes supplementary information for Study One, as detailed in Chapter 3. Within this directory is:

- ***/Apps Evaluated_Study One:*** *this folder contains a spreadsheet with details of the 3 apps used in Study One and their accessibility and enhanced audio features. Also included in the spreadsheet are details of the redeveloped original (without accessibility features) version of the Turning Forest AR testing app, as detailed in Chapter 4 to allow for assessment of how accessible it is compared to currently available Mobile AR apps.*
- ***/Survey Instrument_Study One:*** *this folder contains the full survey used.*
- ***/Results and Analysis_Study One:*** *this folder contains spreadsheets with the results of Study One. There is a spreadsheet with the results for all questions in it, and for ease, the results are split up into several spreadsheets for demographics and each of the apps that participants used. A sub-folder contains .CSV with responses coded as numerical values and the R-code used to analyse results, along with image for for graphs made.*

Appendix E

Study Two

This appendix includes supplementary information for Study Two, as detailed in Chapter 4.

The following figures contain the original Narrative Importance model importance assignations from the studies done by Ward in [32].

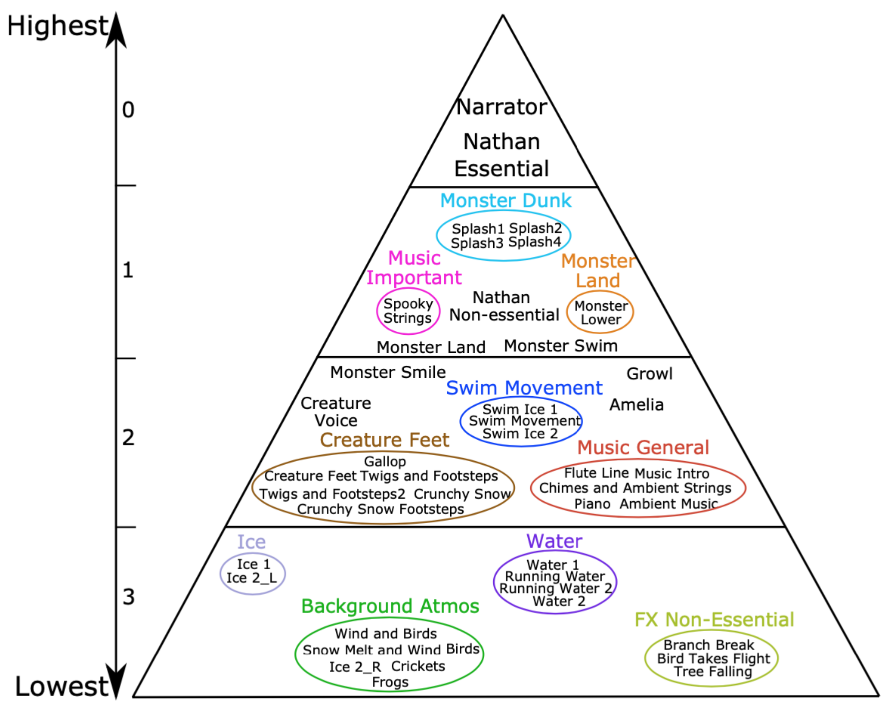


Figure E.1: Hierarchical visualisation of the NI metadata assigned to each audio object in 'The Turning Forest' with the groups set by the sound designer denoted by coloured circles and the group titles noted in corresponding coloured text. Reproduced from Lauren Ward [32].

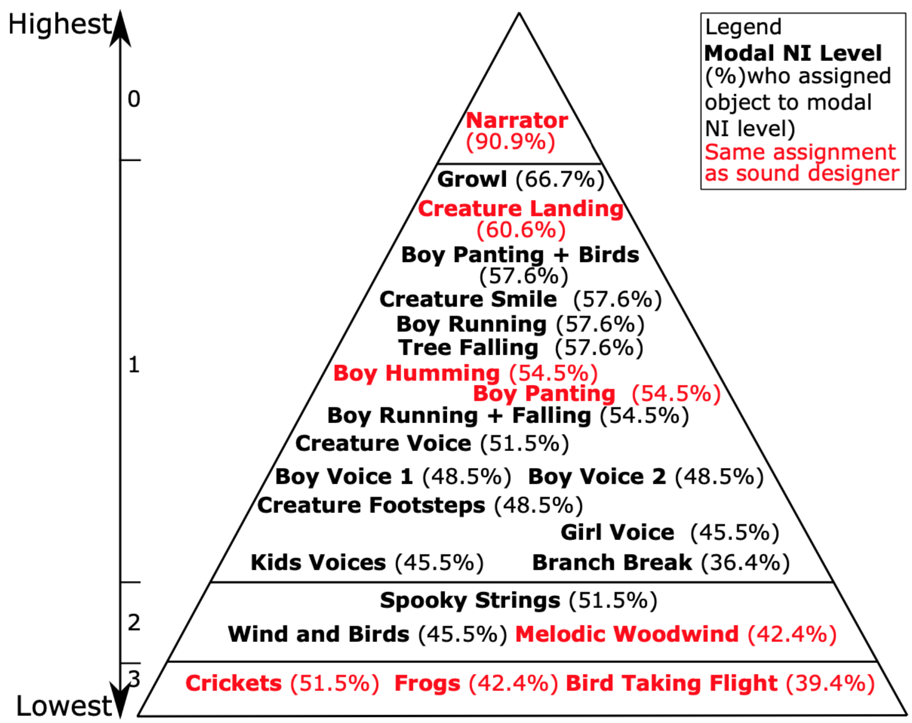


Figure E.2: Hierarchical visualisation of the modal narrative importance level assigned to each audio object in 'The Turning Forest'. The percentage of respondents assigning each object to the modal level noted in parentheses and red text denotes objects which were assigned most commonly by participants to the same narrative importance level as assigned by the original sound designer. Reproduced from Lauren Ward [32].

The remainder of this appendix described the digital appendix directory for Study Two. Within this directory is:

- ***/Information Sheet_Study Two:*** *this folder contains the information given to participants before they took part in the study. This includes details of how they can take part, what the study is investigating and what risks are involved. A privacy statement details what data will be collected and how it will be used. Once participants read this and provided details, they were given instructions for downloading and using the app. It also contains the anonymised consent given by participants for taking part.*

- ***/Interviews_Study Two:*** this folder contains all interview transcripts for Study Two, as detailed below. It also contains a summary of the feedback from the interviews.

  - ***/Content Creators Interviews_Study Two:*** *this folder contains transcripts of the interviews conducted with content creator participants for Study Two. Each interview also has a summary document of the most interesting points discussed.*
  - ***/Interview Consent_Study Two:*** *this folder contains the privacy statement given to participants before they took part in an interview. It also contains the anonymised consent given by participants for taking part.*
  - ***/VI Interviews_Study Two:*** *this folder contains transcripts of the interviews conducted with VI participants for Study Two. Each interview also has a summary document of the most interesting points discussed.*

- ***/Survey_Study Two:***.

  - ***/Content Creators Survey_Study Two:*** *this folder contains spreadsheets with the content creator survey results for Study Two. There is a spreadsheet with the results for all questions in it, and for ease, the results are split up into several spreadsheets for each of the sections detailed in 4.3. A sub-folder contains .CSV with responses coded as numerical values and the R-code used to analyse results, along with image for for graphs made.*
  - ***/Survey Instrument_Study Two:*** *this folder contains the full survey used.*
  - ***/VI Survey_Study Two:*** *this folder contains spreadsheets with the VI survey results for Study Two. There is a spreadsheet with the results for all questions in it, and for ease, the results are split up into several spreadsheets for each of the sections detailed in 4.3. A sub-folder contains .CSV with responses coded as numerical values and the R-code used to analyse results, along with image for for graphs made.*

Appendix F

Study Three

This appendix includes supplementary information for Study Three, as detailed in Chapter 6. Within this directory is:

- ***/Information Sheet_Study Three:*** *this folder contains the information given to participants before they took part in the study. This includes details of how they can take part, what the study is investigating and what risks are involved. A privacy statement details what data will be collected and how it will be used. Once participants read this and provided details, they were given instructions for downloading and using the app. It also contains the anonymised consent given by participants for taking part.*
- ***/Interviews_Study Three:*** *this folder contains all interview transcripts for Study Three, as detailed below. It also contains a summary of the feedback from the interviews.*
  - ***/Content Creators Interviews_Study Three:*** *this folder contains transcripts of the interviews conducted with content creator participants for Study Three. Each interview also has a summary document of the most interesting points discussed.*
  - ***/Interview Consent_Study Three:*** *this folder contains the privacy statement given to participants before they took part in an interview. It also contains the anonymised consent given by participants for taking part.*
  - ***/VI Interviews_Study Three:*** *this folder contains transcripts of the interviews conducted with VI participants for Study Three. Each interview also has a summary document of the most interesting points discussed.*
- ***/Survey_Study Three:***.
  - ***/Content Creators Survey_Study Three:*** *this folder contains spreadsheets with the content creator survey results for Study Three. There is a spreadsheet with the results for all questions in it, and for ease, the results are split up into several spreadsheets for each of the sections detailed in 6.2. A sub-folder contains .CSV with responses coded as numerical values and the R-code used to analyse results, along with image for for graphs made.*
  - ***/Survey Instrument_Study Three:*** *this folder contains the full survey used.*
  - ***/VI Survey_Study Three:*** *this folder contains spreadsheets with the VI survey results for Study Three. There is a spreadsheet with the results for all questions in it, and for ease, the results are split up into several spreadsheets for each of*

the sections detailed in 6.2. A sub-folder contains .CSV with responses coded as numerical values and the R-code used to analyse results, along with image for for graphs made.

Appendix G

AR Testing Application

Within this directory is

- ***/Final Accessible AR App*** this folder contains details of the final updated Turning Forest AR application used in Study Three (see Chapters 5 and 6). Contains several sub-folders:
  - */Guide character script* - contains the script used by the voice over artist for recording the Guide Character.
  - */Personalisation control* - contains an excel sheet combining the importance ratings from both this work and Ward's work [32] so that choices could be made about what importance levels were assigned to all sounds in the app. Also contains two Unity Packages. First 'Personalisation slider' creates a slider that can be added to scenes to set how accessible or not the mix is globally. This contains the 'Slider Controller' script described in Section 5.3.1. Second, 'Personalised Audio Source' combines this information with a set importance level and the predetermined level of the audio source to alter the level of the audio source as the user changes the slider. This contains the 'Audio Importance' script described in Section 5.3.1.
  - */Screenshots* - comprehensive screenshots of all of the user interface and key tasks within the AR scenes.
  - */Sound Effects* - contains a document with details of all sound effects added. Contains 2 Unity Packages. First, 'Collision SFX and Hint' low pass filters all audio and plays a hint when the user collides with a game object. Second, 'Highlight SFX' mutes all other game objects tagged as containing sound effects when it is called.
  - */Unity Screenshots* - comprehensive screenshots of key parts of the development of the app, including spatial audio enhancements and screen reader functionality.
  - *Walkthrough video* - contains videos running through all user interface and scenes in the updated AR testing app. Some walk through videos are recorded with, and some without the guide character as a screen reader to show play with either option.
- ***/Preview AR App*** this folder contains details of the the 'preview' Turning Forest AR application (without accessibility features) used in Study Two (see Chapter 4). A walk through video is included, and several sub-folders:

  - */Screenshots* - comprehensive screenshots of all of the user interface and key tasks within the AR scenes.
  - */Unity Screenshots* - comprehensive screenshots of key parts of the development of the app, including scenes, optimisations that were made to materials, textures, animations and the overall quality.

Unfortunately, Android screen recording functionality records in mono only, so the spatial audio is not demonstrated by the walk through video. Using screen recording functionality at the same time as the AR app unfortunately causes it to run much slower and lag at points that it wouldn't in normal usage.

Usage:

- The Android builds for both apps can be directly copied and installed onto any Android device compatible with AR Core. This means the device must have access to the Google Play store and be compatible with Google Play Services for AR. To see the full criteria and list compatible devices, follow this link (in general, you must be running Android 7.0 or newer): `https://developers.google.com/ar/devices`. Please note that these are for personal use in demonstrating the work presented here. They cannot be distributed or used for further research without further permission. All intellectual property right for Turning Forest content belongs with the BBC.
- The personalisation control assets allow for a personalisation slider, as used in the final app, to be added to any Unity project.

Abbreviations and Acronyms

Table 1 contains abbreviations and acronyms used throughout this thesis. The table on which each one first used or defined is also given.

Table 1: Abbreviations and acronyms used throughout this thesis

Abbreviation	Meaning	Page first used	Defined
AD	Audio Description	p25	p25
app	application	p2	p2
AAR	Audio Augmented Reality	p29	p33
AR	Augmented Reality	p2	p30
EAD	Enhancing Audio Description	p15	p27
HMD	Head Mounted Display	p14	p31
ITIF	Information Technology and Innovation Foundation	p35	p35
MR	Mixed Reality	p30	p30
NI	Narrative Importance	p15	p29
OS	Operating System	p14	p14
RNIB	Royal National Institute of Blind People	p18	p18
TTS	Text To Speech	p21	p21
UI	User Interface	p45	p45
VI	Visually Impaired	p2	p18
VR	Virtual Reality	p14	p30
WHO	World Health Organisation	p18	p18
XR	Extended Reality	p2	p30

References

[1] L. Pezzullo, J. Streatfeild, P. Simkiss and D. Shickle, 'The economic impact of sight loss and blindness in the UK adult population,' *BMC Health Services Research*, vol. 18, no. 1, pp. 1–13, 2018, ISSN: 14726963. DOI: 10.1186/s12913-018-2836-0.

[2] WHO, 'World report on vision,' World Health Organisation, Geneva, Tech. Rep. 14, 2019, pp. 1–160. [Online]. Available: https://www.who.int/publications/i/item/world-report-on-vision.

[3] J. Slade and R. Edwards, 'My Voice 2015,' RNIB, Tech. Rep., 2015. [Online]. Available: https://www.rnib.org.uk/sites/default/files/My%20Voice%20UK%20Report-FINAL%7B%5C_%7D0.PDF.

[4] B. Y. E. Dick, *Risks and Challenges for Inclusive and Equitable Immersive Experiences — ITIF*, 2021. [Online]. Available: https://itif.org/publications/2021/06/01/risks-and-challenges-inclusive-and-equitable-immersive-experiences (visited on 23/06/2021).

[5] B. Agulló, M. Montagud and I. Fraile, 'Making interaction with virtual reality accessible: Rendering and guiding methods for subtitles,' *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, vol. 33, no. 4, pp. 416–428, 2019, ISSN: 14691760. DOI: 10.1017/S0890060419000362.

[6] E. Dick, *Current and Potential Uses of AR / VR for Equity and Inclusion*, 2021. [Online]. Available: https://itif.org/publications/2021/06/01/current-and-potential-uses-arvr-equity-and-inclusion (visited on 23/06/2021).

[7] J. M. Coughlan and J. Miele, 'AR4VI: AR as an Accessibility Tool for People with Visual Impairments,' in *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*, IEEE, 2017, pp. 288–292, ISBN: 9780769563275. DOI: 10.1109/ISMAR-Adjunct.2017.89.

[8] Perkins Coie LLP, '2020 Augmented and Virtual Reality survey report: Industry insights into the future of immersive technology,' Perkins Coie LLP, Tech. Rep., 2020. [Online]. Available: https://www.perkinscoie.com/images/content/2/3/231654/2020-AR-VR-Survey-v3.pdf.

[9] T. Merel, *After Mixed Year, Mobile AR to Drive \$108 Billion VR/AR Market by 2021 (Guest Blog)*, 2017. [Online]. Available: https://www.thewrap.com/mixed-year-mobile-ar-drive-108-billion-vrar-market-2021-guest-column/ (visited on 02/06/2021).

[10] Cyber-XR Coalition, 'Immersive Technology Standards for accessibility, inclusion, ethics, and safety,' Cyber-XR Coalition, Tech. Rep., 2020. DOI: 10.4018/978-1-7998-3250-8.ch006.

[11] Ericsson ConsumerLab, 'Ready, steady, game! Is augmented reality (AR) the next level of gaming? An Ericsson ConsumerLab insight report,' Stockholm, Sweden, Tech. Rep. March, 2019.

[12] V. Geroimenko, Ed., *Augmented Reality Games. I , Understanding the Pokémon GO Phenomenon*. Cham : Springer, 2019, ISBN: 9783030156169.

[13] V. Geroimenko, Ed., *Augmented reality games. II, The gamification of education, medicine and art*. Cham, Switzerland : Springer, 2019, ISBN: 9783030156206.

[14] Unity Technologies, 'Top 2020 Trends: Enterprise AR \& VR,' *Unity Technologies*, 2020.

[15] Android, *Made for the way you want to use it*, 2022. [Online]. Available: `https://www.android.com/accessibility/%7B%5C#%7Dvisibility-view-the-world` (visited on 27/01/2022).

[16] Apple, *Vision - For every point of view*, 2022. [Online]. Available: `https://www.apple.com/uk/accessibility/vision/` (visited on 27/01/2022).

[17] A. S. Ross, X. Zhang, J. Fogarty and J. O. Wobbrock, 'Epidemiology as a framework for large-scale mobile application accessibility assessment,' *ASSETS 2017 - Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 2–11, 2017. DOI: `10.1145/3132525.3132547`.

[18] J. Herskovitz *et al.*, 'Making Mobile Augmented Reality Applications Accessible,' *ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, 2020. DOI: `10.1145/3373625.3417006`.

[19] M. Lopez, G. Kearney, K. Hofstädter and G. Balla, 'Enhancing audio description: accessible filmmaking, sound design and the importance of educating filmmakers,' *Media Practice and Education*, vol. 21, no. 4, pp. 289–304, 2020, ISSN: 2574-1136. DOI: `10.1080/25741136.2020.1832830`. [Online]. Available: `https://doi.org/10.1080/25741136.2020.1832830`.

[20] E. Rovithis, N. Moustakas, A. Floros and K. Vogklis, 'Audio legends: Investigating sonic interaction in an augmented reality audio game,' *Multimodal Technologies and Interaction*, vol. 3, no. 4, pp. 1–18, 2019, ISSN: 24144088. DOI: `10.3390/mti3040073`.

[21] L. A. Ward and B. G. Shirley, 'Personalization in object-based audio for accessibility: A review of advancements for hearing impaired listeners,' *AES: Journal of the Audio Engineering Society*, vol. 67, no. 7-8, pp. 584–597, 2019, ISSN: 15494950. DOI: `10.17743/jaes.2019.0021`.

[22] C. Simon, M. Torcoli and J. Paulus, 'MPEG-H Audio for Improving Accessibility in Broadcasting and Streaming [White Paper],' *Fraunhofer Institute for Integrated Circuits IIS Management*, 2019. [Online]. Available: `https://www.iis.fraunhofer.de/content/dam/iis/de/doc/ame/wp/20191001%7B%5C_%7DMPEG-H%7B%5C_%7DAccessibility%7B%5C_%7DWhite%7B%5C_%7DPaper.pdf`.

[23] G. Ren, S. Wei, E. O'Neill and F. Chen, 'Towards the Design of Effective Haptic and Audio Displays for Augmented Reality and Mixed Reality Applications,' *Advances in Multimedia*, vol. 2018, pp. 1–11, 2018, ISSN: 16875699. DOI: `10.1155/2018/4517150`. [Online]. Available: `https://www.hindawi.com/journals/am/2018/4517150/`.

[24] A. N. Nagele *et al.*, 'Interactive Audio Augmented Reality in Participatory Performance,' *Frontiers in Virtual Reality*, vol. 1, no. February, pp. 1–14, 2021. DOI: 10.3389/frvir.2020.610320. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2020.610320/full.

[25] M. Montagud, P. Orero and A. Matamala, 'Culture 4 all: accessibility-enabled cultural experiences through immersive VR360 content,' *Personal and Ubiquitous Computing*, vol. 24, no. 6, pp. 887–905, 2020, ISSN: 16174917. DOI: 10.1007/s00779-019-01357-3.

[26] P. Romero-Fresco, 'In support of a wide notion of media accessibility: Access to content and access to creation,' *Journal of Audiovisual Translation*, vol. 1, no. 1, pp. 187–204, 2018, ISSN: 2617-9148. DOI: 10.47476/jat.v1i1.53.

[27] M. Lopez, G. Kearney and K. Hofstädter, 'Enhancing audio description: Sound design, spatialisation and accessibility in film and television,' *Proceedings of the Institute of Acoustics*, vol. 38, pp. 174–182, 2016, ISSN: 14786095. [Online]. Available: http://enhancingaudiodescription.com/publications/.

[28] M. Lopez, G. Kearney and K. Hofstädter, 'Audio Description in the UK: What works, what doesn't, and understanding the need for personalising access,' *British Journal of Visual Impairment*, vol. 2050, 2018, ISSN: 17445809. DOI: 10.1177/0264619618794750.

[29] ——, 'Seeing films through sound: Sound design, spatial audio, and accessibility for visually impaired audiences,' *British Journal of Visual Impairment*, 2020, ISSN: 17445809. DOI: 10.1177/0264619620935935.

[30] M. Lopez, G. Kearney and K. Hofstadter, 'Enhancing Audio Description: Inclusive Cinematic Experiences Through Sound Design,' *Journal of Audiovisual Translation*, vol. 4, no. 1, pp. 157–182, 2021. DOI: 10.47476/jat.v4i1.2021.154.

[31] B. Shirley, L. Ward and E. T. Chourdakis, 'Personalization of object-based audio for accessibility using narrative importance,' *CEUR Workshop Proceedings*, vol. 2423, no. June, 2019, ISSN: 16130073.

[32] L. A. Ward, 'Improving Broadcast Accessibility for Hard of Hearing Individuals : using Improving broadcast accessibility for hard of hearing individuals using object-based audio personalisation and narrative,' PhD, Salford, 2020. DOI: 10.13140/RG.2.2.31454.46405.

[33] N. Bosanquet and P. Mehta, 'Evidence base to support the UK Vision Strategy,' 2008.

[34] Z. Cattaneo and V. Tomaso, *Blind Vision: The Neuroscience of Visual Impairment*. MIT Press, 2011, ISBN: 9780262295819.

[35] Royal National Institute of Blind People, *The criteria for certification*, 2021. [Online]. Available: https://www.rnib.org.uk/eye-health/registering-your-sight-loss/criteria-certification (visited on 16/11/2021).

[36] Specsavers, *What does 20/20 vision mean?* 2022. [Online]. Available: https://www.specsavers.co.uk/help-and-faqs/what-does-2020-vision-mean (visited on 13/04/2022).

[37] M. A. Hersh and M. A. Johnson, *Assistive Technology for Visually Impaired and Blind People*. Springer, 2008. DOI: 10.1007/978-1-84628-867-8.

[38] I. Tor-carroggio and P. Orero, 'User profiling in audio description reception studies : questionnaires for all,' no. Edler 2009, pp. 1–8, 2020.

[39] M. Oliver, B. Sapey and P. Thomas, *Social work with disabled people*. Palgrave Macmillan, 2012.

[40] M. Crabb, M. Heron, R. Jones, M. Armstrong, H. Reid and A. Wilson, 'Developing accessible services understanding current knowledge and areas for future support,' *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1–12, 2019. DOI: 10.1145/3290605.3300446.

[41] S. Rai, 'Audio description app user trial: Report, in partnership with movie reading,' RNIB, London, England, Tech. Rep. September, 2015. [Online]. Available: http://www.rnib.org.uk/sites/default/files/Audio%7B%5C_%7Ddescription%7B%5C_%7Dapp%7B%5C_%7Dtrial%7B%5C_%7Dsummary.pdf.

[42] J. T. Richards, K. Montague and V. L. Hanson, 'Web Accessibility as a Side Effect,' in *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12)*, New York, NY, USA: Association for Computing Machinery, 2012, pp. 79–86, ISBN: 9781450313216. DOI: 10.1145/2384916.2384931. [Online]. Available: https://doi.org/10.1145/2384916.2384931.

[43] Royal National Institute of Blind People, *Assistive technology*, 2022. [Online]. Available: https://www.rnib.org.uk/sight-loss-advice/equality-rights-and-employment/staying-work/assistive-technology (visited on 05/01/2022).

[44] A. Rodrigues, H. Nicolau, K. Montague, J. Guerreiro and T. Guerreiro, 'Open Challenges of Blind People Using Smartphones,' *International Journal of Human-Computer Interaction*, vol. 36, no. 17, pp. 1605–1622, 2020, ISSN: 15327590. DOI: 10.1080/10447318.2020.1768672. arXiv: 1909.09078.

[45] T. Hermann and A. Hunt, 'An introduction to interactive sonification,' *IEEE Multimedia*, vol. 12, no. 2, pp. 20–24, 2005, ISSN: 1070986X. DOI: 10.1109/MMUL.2005.26.

[46] T. Hermann, A. Hunt and J. G. Neuhoff, *The sonification handbook*. Berlin:Logos, 2011, ISBN: 9783832528195 (pbk.)

[47] W. W. Gaver, 'Auditory Icons: Using Sound in Computer Interfaces,' *Human-computer interaction*, vol. 2, pp. 167–177, 1986. DOI: 10.1207/s15327051hci0202_3.

[48] M. M. Blattner, D. A. Sumikawa and M. Robert, 'Earcons and Icons : Their Structure and Common Design Principles,' *Human–Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989. DOI: 10.1207/s15327051hci0401. [Online]. Available: https://doi.org/10.1207/s15327051hci0401%7B%5C_%7D1.

[49] B. N. Walker *et al.*, 'Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus,' *Human Factors*, vol. 55, no. 1, pp. 157–182, 2013, ISSN: 00187208. DOI: 10.1177/0018720812450587.

[50] B. F. G. Katz *et al.*, 'NAVIG: augmented reality guidance system for the visually impaired: Combining object localization, GNSS, and spatial audio,' *Virtual reality : the journal of the Virtual Reality Society*, vol. 16, no. 4, pp. 253–269, 2012, ISSN: 1359-4338.

[51] S. Mascetti, A. Gerino, C. Bernareggi and L. Picinali, 'On the evaluation of novel sonification techniques for non-visual shape exploration,' *ACM Transactions on Accessible Computing*, vol. 9, no. 4, 2017, ISSN: 19367228. DOI: 10.1145/3046789.

[52] K. Drossos, N. Zormpas, A. Floros and G. Giannakopoulos, 'Accessible games for blind children, empowered by binaural sound,' in *PETRA '15: Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2015, pp. 1–8, ISBN: 9781450334525. DOI: 10.1145/2769493.2769546. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2769493.2769546.

[53] L. B. Merabet, E. C. Connors, M. A. Halko and J. T. Sánchez, 'Teaching the Blind to Find Their Way by Playing Video Games,' *PloS one*, vol. 7, no. 4, pp. 1–10, 2012. DOI: 10.1371/Citation.

[54] S. Mascetti, L. Picinali, A. Gerino, D. Ahmetovic and C. Bernareggi, 'Sonification of guidance data during road crossing for people with visual impairments or blindness,' *International Journal of Human Computer Studies*, vol. 85, pp. 16–26, 2016, ISSN: 10959300. DOI: 10.1016/j.ijhcs.2015.08.003. arXiv: 1506.07272. [Online]. Available: http://dx.doi.org/10.1016/j.ijhcs.2015.08.003.

[55] F. Rumsey, *Spatial Audio*. Oxford:Focal Press, 2001, ISBN: 0240516230.

[56] Dolby Laboratories Inc, 'Dolby Atmos^{\tiny\textregistered} for the Home Theater,' *White Paper*, 2016.

[57] Apple, *Apple Music announces Spatial Audio with Dolby Atmos; will bring Lossless Audio to entire catalog*, 2021. [Online]. Available: https://www.apple.com/uk/newsroom/2021/05/apple-music-announces-spatial-audio-and-lossless-audio/ (visited on 25/05/2021).

[58] J. Friberg and D. Gärdenfors, 'Audio games: New perspectives on game audio,' *ACE '04: Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, vol. 74, pp. 148–154, 2004. [Online]. Available: https://doi.org/10.1145/1067343.1067361.

[59] M. Lombard and T. Ditton, 'At the heart of it all: The concept of presence,' *Journal of Computer-Mediated Communication*, vol. 3, no. 2, pp. 1–67, 1997, ISSN: 10836101. DOI: 10.1111/j.1083-6101.1997.tb00072.x. [Online]. Available: https://doi.org/10.1111/j.1083-6101.1997.tb00072.x.

[60] P. Skalski and R. Whitbred, 'Image versus sound: A comparison of formal feature effects on presence and video game enjoyment,' *PsychNology Journal*, vol. 8, no. 1, pp. 67–84, 2010, ISSN: 17207525.

[61] A. Smith, *Binaural audio: What is it? How can you get it?* 2017. [Online]. Available: https://www.whathifi.com/advice/binaural-audio-what-it-how-can-you-get-it (visited on 29/07/2021).

[62] T. Potisk, 'Head-Related Transfer Function,' *University of Ljubljana Faculty of Mathematics and Physics*, vol. Seminar Ia, 2015.

[63] J. Sodnik and S. Tomažič, *Spatial Auditory Human-Computer Interfaces*, 1st ed. 20. Springer International Publishing AG, 2015, ISBN: 9783319221106.

[64] F. Rumsey, 'Binaural Audio and Virtual Acoustics,' *J. Audio Eng. Soc*, vol. 65, no. 6, pp. 524–528, 2017. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18784.

[65] B. FG and L. Picinali, 'Spatial Audio Applied to Research with the Blind,' *Advances in Sound Localization*, no. 1991, 2011. DOI: 10.5772/15206.

[66] C. Fink, *How Spatial Audio Can Change Everything For The Blind*, 2020. [Online]. Available: https://www.forbes.com/sites/charliefink/2020/06/30/how-spatial-audio-can-change-everything-for-the-blind/?sh=755500515080.

[67] J. Popp, M. Neuendorf, H. Fuchs, C. Forster and A. Heuberger, *Recent advances in broadcast audio coding*, 2013. DOI: 10.1109/BMSB.2013.6621702.

[68] P. Coleman *et al.*, 'An Audio-Visual System for Object-Based Audio : From Recording to Listening,' *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 1919–1931, 2018. DOI: 10.1109/TMM.2018.2794780.

[69] C. Pike, R. Taylor, T. Parnell and F. Melchior, 'Object-based spatial audio production for virtual reality using the Audio Definition Model,' *Proceedings of the AES International Conference*, vol. 2016-Septe, pp. 1–7, 2016.

[70] R. Oldfield, B. Shirley and J. Spille, 'Object-based audio for interactive football broadcast,' *Multimedia Tools and Applications*, no. May 2013, pp. 2717–2741, 2013. DOI: 10.1007/s11042-013-1472-2.

[71] C. Pike, P. Taylour and F. Melchior, 'Delivering Object-Based 3D Audio Using The Web Audio API And The Audio Definition Model,' in *1st Web Audio Conference (WAC-2015)*, 2015, pp. 2–6, ISBN: 0199537992. [Online]. Available: https://wac.ircam.fr/pdf/wac15%7B%5C_%7Dsubmission%7B%5C_%7D24.pdf.

[72] S. Rai, 'RNIB international AD exchange study: Observations from a focus group study,' RNIB, London, England: Media and Culture Department, Tech. Rep. May, 2011.

[73] J. Snyder, 'Audio description: The visual made verbal,' *International Congress Series*, vol. 1282, pp. 935–939, 2005, ISSN: 05315131. DOI: 10.1016/j.ics.2005.05.215.

[74] J. Whitehead, 'What is audio description,' *International Congress Series*, vol. 1282, pp. 960–963, 2005, ISSN: 05315131. DOI: 10.1016/j.ics.2005.05.194.

[75] Ofcom, 'Ofcom' s Guidelines on the Provision of Television Access Services Guidelines on the provision of television access services,' Office of Communications [Ofcom], Tech. Rep. February, 2021. [Online]. Available: https://www.ofcoma.org.uk/tv-radio-and-on-demand/broadcast-codes/tv-access-services.

[76] S. Rai, J. Greening and L. Petré, 'A Comparative Study of Audio Description Guidelines Prevalent in Different Countries,' RNIB, London, England, Tech. Rep., 2010, p. 112. [Online]. Available: http://audiodescription.co.uk/uploads/general/RNIB.%7B%5C_%7DAD%7B%5C_%7Dstandards1.pdf.

[77] Ofcom, 'Ofcom's Code on Television Access Services,' Office of Communications [Ofcom], Tech. Rep. February, 2019. [Online]. Available: https://www.ofcoma.org.uk/tv-radio-and-on-demand/broadcast-codes/tv-access-services.

[78] UK Government, *Digital Economy Act 2017*, 2017. [Online]. Available: https://www.gov.uk/government/collections/digital-economy-bill-2016 (visited on 13/06/2021).

[79] G. F. Williams, *Why Accessibility Policies Are Important, And What Can Be Learned From The BBC*, 2021. [Online]. Available: https://levelup.gitconnected.com/why-accessibility-policies-are-important-and-what-can-be-learned-from-the-bbc-1561df1378d0 (visited on 08/10/2021).

[80] A. Remael, 'For the use of sound. Film sound analysis for audio-description: some key issues,' *MonTI: Monografías de Traducción e Interpretación*, vol. 4, no. 4, pp. 255–276, 2012, ISSN: 18894178. DOI: 10.6035/monti.2012.4.11.

[81] J. P. Udo and D. I. Fels, 'The rogue poster-children of universal design: Closed captioning and audio description,' *Journal of Engineering Design*, vol. 21, no. 2-3, pp. 207–221, 2010, ISSN: 09544828. DOI: 10.1080/09544820903310691.

[82] L. Fryer and J. Freeman, 'Presence in those with and without sight: Audio description and its potential for virtual reality applications,' *Journal of CyberTherapy \& Rehabilitation*, vol. 5, no. 1, pp. 15–23, 2012.

[83] L. Fryer, 'Putting it into Words : The Impact of Visual Impairment on Perception , Experience and Presence By Louise Fryer Thesis submitted for the degree of Doctor of Philosophy Department of Psychology Goldsmiths College University of London,' Ph.D. dissertation, Goldsmiths College (University of London), 2013, pp. 1–372. [Online]. Available: https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.601742.

[84] L. Fryer and J. Freeman, 'Visual impairment and presence: Measuring the effect of audio description,' in *ACM International Conference Proceeding Series. Inputs-Outputs '13*, Brighton, England, 2013, pp. 1–5, ISBN: 9781450325813. DOI: 10.1145/2557595.2557599. [Online]. Available: https://doi.org/10.1145/2557595.2557599.

[85] ——, 'Can you feel what I'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment,' *Challeging Presence. Proceedings of the 15th International Conference on Presence*, no. March 2014, pp. 99–107, 2014.

[86] L. Fryer, 'Staging the Audio Describer: An Exploration of Integrated Audio Description,' *Disability Studies Quarterly*, vol. 38, no. 3, pp. 1–13, 2018, ISSN: 1041-5718. DOI: 10.18061/dsq.v38i3.6490.

[87] ——, 'The independent audio describer is dead: Long live audio description!' *Journal of Audiovisual Translation*, vol. 1, no. 1, pp. 170–186, 2018, ISSN: 2617-9148. DOI: 10.47476/jat.v1i1.52.

[88] ——, 'Quality Assessment in Audio Description: Lessons Learned from Interpreting,' *Quality Assurance and Assessment Practices in Translation and Interpreting*, pp. 155–177, 2019.

[89] P. Romero-Fresco, 'Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking,' *Journal of Specialised Translation*, no. 20, pp. 201–223, 2013, ISSN: 1740357X.

[90] ——, *Accessible filmmaking.* 2018, pp. 498–515, ISBN: 9780429053771. DOI: 10.4324/9781315717166-31.

[91] A. Cavallo and L. Fryer, 'Integrated access inquiry 2017-18 report,' *Retrieved March*, vol. 10, p. 2020, 2018.

[92] E. I. Fernández, S. M. Martínez and A. Chica, 'Cross-fertilization between Reception Studies in Audio Description and Interpreting Quality Assessment: The Role of the Describer's Voice.,' in *Audiovisual Translation in a Global Context*, September, R. B. Piñero and J. Díaz-Cintas, Eds., Palgrave Macmillan, London, 2015, ch. 5, pp. 72–95, ISBN: 9781137552891. DOI: 10.1057/9781137552891_5. [Online]. Available: https://doi.org/10.1057/9781137552891%7B%5C_%7D5.

[93] A. Cavallo, 'Seeing the word, hearing the image: the artistic possibilities of audio description in theatrical performance,' *Research in Drama Education: The Journal of Applied Theatre and Performance*, vol. 20, no. 1, pp. 125–134, 2015, ISSN: 1470112X. DOI: 10.1080/13569783.2014.983892. [Online]. Available: http://dx.doi.org/10.1080/13569783.2014.983892.

[94] L. N. Viswanathan, T. McDaniel, S. Krishna and S. Panchanathan, 'Haptics in audio described movies,' *HAVE 2010 - 2010 IEEE International Symposium on Haptic Audio-Visual Environments and Games, Proceedings*, pp. 200–201, 2010. DOI: 10.1109/HAVE.2010.5623958.

[95] M. Chion, *The Voice In Cinema*. New York : Columbia University Press, 1999, ISBN: 0231108222.

[96] E. Rovithis, 'A classification of audio-based games in terms of sonic gameplay and the introduction of the audio-role-playing-game: Kronos,' in *Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound*, 2012, pp. 160–164, ISBN: 9781450315692. DOI: 10.1145/2371456.2371483.

[97] Creative Heroes, *AudioGames.net*, 2022. [Online]. Available: https://audiogames.net/.

[98] B. Shirley, M. Meadows, F. Malak, J. Woodcock and A. Tidball, 'Personalized object-based audio for hearing impaired TV viewers,' *AES: Journal of the Audio Engineering Society*, vol. 65, no. 4, pp. 293–303, 2017, ISSN: 15494950. DOI: 10.17743/jaes.2017.0005.

[99] XR Association, 'XRA's Developers Guide, Chapter One: Fundamental Design Principles for Immersive Experiences,' Tech. Rep., 2018. [Online]. Available: https://xra.org/research/xr-primer-1-0-a-starter-guide-for-developers/.

[100] T. Leland, *Extended reality: A convergence is coming — Qualcomm*, 2017. [Online]. Available: https://www.qualcomm.com/news/onq/2017/05/31/extended-reality-convergence (visited on 19/01/2022).

[101] XR Association, *XR at a Glance*, 2022. [Online]. Available: https://xra.org/xr-at-a-glance/ (visited on 19/01/2022).

[102] J. Dalton and J. Gillham, 'Seeing is believing. How virtual reality and augmented reality are transforming business and the economy,' Tech. Rep., 2019. [Online]. Available: https://www.pwc.com/gx/en/technology/publications/assets/how-virtual-reality-and-augmented-reality.pdf.

[103] PwC, *Why we believe VR/AR will boost global GDP by \$1.5 trillion*, 2019. [Online]. Available: https://www.pwc.co.uk/services/economics/insights/vr-ar-to-boost-global-gdp.html (visited on 05/01/2022).

[104] Microsoft, *Microsoft HoloLens 2*, 2022. [Online]. Available: https://www.microsoft.com/en-gb/hololens?SilentAuth=1%7B%5C&%7Dwa=wsignin1.0 (visited on 20/01/2022).

[105] Magic Leap, *Magic Leap 1 is a wearable computer for enterprise productivity*, 2022. [Online]. Available: https://www.magicleap.com/en-us/magic-leap-1 (visited on 20/01/2022).

[106] Google, *Discover Glass Enterprise Edition*, 2022. [Online]. Available: https://www.google.com/glass/start/ (visited on 20/01/2022).

[107] K. Hommaru and J. Tanaka, 'Walking support for visually impaired using ar/mr and virtual braille block,' in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12188 LNCS, Springer Nature Switzerland AG 2020, 2020, pp. 336–354, ISBN: 9783030492816. DOI: 10.1007/978-3-030-49282-3_24.

[108] S. Colquhoun, *Are you ready to augment your reality?* 2020. [Online]. Available: https://www.mercedes-benz.com.au/passengercars/experience/mercedes-me-magazine/innovation/articles/augmented-reality/story-content.module.html (visited on 20/01/2022).

[109] M. Rohrer and B. Sagges, *Panasonic Automotive Unveils Driveable AR HUD 2.0 with Patented Eye-Tracking- System at CES 2022*, 2022. [Online]. Available: https://na.panasonic.com/explore/press-release/panasonic-automotive-unveils-driveable-ar-hud-20-patented-eye-tracking-system-ces-2022 (visited on 20/01/2022).

[110] Google Play Developers Guide, *Experience size*, 2021. [Online]. Available: https://developers.google.com/ar/design/environment/experience-size (visited on 21/04/2021).

[111] J. C. Kuang, B. Delaney, N. Yue and C. Mercer, '2019 Augmented Reality Industry Report,' Greenlight Insights, Tech. Rep., 2019. [Online]. Available: https://greenlightinsights.com/reports/2019-augmented-reality-industry-report/.

[112] H. Lee, 'A Conceptual Model of Immersive Experience in Extended Reality,' no. September, pp. 1–20, 2020. DOI: 10.31234/osf.io/sefkh.

[113] Facebook, *Announcing Project Aria: a research project on the future of wearable AR*, 2020. [Online]. Available: https://tech.fb.com/announcing-project-aria-a-research-project-on-the-future-of-wearable-ar/ (visited on 24/05/2021).

[114] Bose Corporation, *Bose Frames Audio Sunglasses*, 2022. [Online]. Available: https://www.bose.co.uk/en%7B%5C_%7Dgb/products/frames.html (visited on 21/01/2022).

[115] J. Isaac Engel and L. Picinali, 'Long-term user adaptation to an audio augmented reality system,' *24th International Congress on Sound and Vibration, ICSV 2017*, no. July, 2017.

[116] H. Cooke, *Looking for Nigel - an audio augmented reality experience*, 2021. [Online]. Available: https://www.bbc.co.uk/rd/blog/2021-01-audio-augmented-reality-spatial-voice (visited on 24/05/2021).

[117] R. Mehra, O. Brimijoin, P. Robinson and T. Lunner, 'Potential of Augmented Reality Platforms to Improve Individual Hearing Aids and to Support More Ecologically Valid Research,' *Ear and hearing*, vol. 41, 140S–146S, 2020, ISSN: 15384667. DOI: 10.1097/AUD.0000000000000961.

[118] Bose Corportation, *Bose SoundControl ™ Hearing Aids*, 2021. [Online]. Available: https://www.bose.com/en%7B%5C_%7Dus/products/headphones/earbuds/soundcontrol-hearing-aids.html%7B%5C#%7Dv=soundcontrol%7B%5C_%7Dhearing%7B%5C_%7Daids%7B%5C_%7Dgray (visited on 20/10/2021).

[119] L. B. Jaloza, *Inside Facebook Reality Labs: Research updates and the future of social connection*, 2020. [Online]. Available: https://tech.fb.com/inside-facebook-reality-labs-research-updates-and-the-future-of-social-connection/ (visited on 24/05/2021).

[120] P. P. Rusu, M. D. Schipor and R. D. Vatavu, 'A lead-in study on well-being, visual functioning, and desires for augmented reality assisted vision for people with visual impairments,' *2019 7th E-Health and Bioengineering Conference, EHB 2019*, pp. 19–22, 2019. DOI: 10.1109/EHB47216.2019.8970074.

[121] E. Dick, *Principles and Policies to Unlock the Potential of AR / VR for Equity and Inclusion*, 2021. [Online]. Available: https://itif.org/publications/2021/06/01/principles-and-policies-unlock-potential-arvr-equity-and-inclusion (visited on 23/06/2021).

[122] B. Froissard, H. Konik, A. Trémeau and É. Dinet, 'Contribution of augmented reality solutions to assist visually impaired people in their mobility,' in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8516 LNCS, Springer International Publishing Swizerland 2014, 2014, pp. 182–191, ISBN: 9783319075082. DOI: 10.1007/978-3-319-07509-9_18.

[123] Accessible Platform Architectures Working Group, *XR Accessibility User Requirements*, 2020. [Online]. Available: https://www.w3.org/TR/xaur/ (visited on 05/03/2021).

[124] XR Association, 'XRA's Developers Guide, Chapter Two: Creating Safe, Inclusive, and Respectful Immersive Experiences,' Tech. Rep., 2018. [Online]. Available: https://xra.org/research/xr-primer-2-0-a-starter-guide-for-developers/.

[125] ——, 'XRA's Developers Guide, Chapter Three: Accessibility \& Inclusive Design in Immersive Experiences,' Tech. Rep., 2018. [Online]. Available: https://xra.org/research/xra-developers-guide-accessibility-and-inclusive-design/.

[126] Oculus, *Designing Accessible VR*. [Online]. Available: https://developer.oculus.com/resources/design-accessible-vr/ (visited on 25/01/2022).

[127] Magic Leap, *Accessibility*, 2018. [Online]. Available: https://developer.magicleap.com/en-us/learn/guides/bp-for-accessibility (visited on 25/01/2022).

[128] IEEE, *The IEEE Global Initiative on Ethics of Extended Reality: Industry Connections Activity Initiation Document (ICAID)*, 2021. [Online]. Available: https://standards.ieee.org/content/dam/ieee-standards/standards/web/governance/iccom/IC20-016-Global%7B%5C_%7DInitiative%7B%5C_%7DEthics%7B%5C_%7DExtended%7B%5C_%7DReality.pdf (visited on 10/08/2021).

[129] Partnership on Employment \& Accessible Technology (PEAT), *The Accessibility Playbook for Emerging Technology Initiatives*. [Online]. Available: https://www.peatworks.org/futureofwork/playbook/ (visited on 10/08/2021).

[130] Xracess.org, *The XR Access Initiative*, 2021. [Online]. Available: https://xraccess.org/ (visited on 20/08/2020).

[131] T. Pilgrim, *Major push to make virtual worlds more inclusive launched Reported by :* 2019. [Online]. Available: https://www.brunel.ac.uk/news-and-events/news/articles/Major-push-to-make-virtual-worlds-more-inclusive-launched (visited on 10/08/2021).

[132] University of Cambridge, *Inclusive Design Toolkit Home*, 2013. [Online]. Available: http://www.inclusivedesigntoolkit.com/ (visited on 10/08/2021).

[133] Y. Zhao, M. Hu, S. Hashash and S. Azenkot, 'Understanding low vision people's visual perception on commercial augmented reality glasses,' *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2017-May, pp. 4170–4181, 2017. DOI: 10.1145/3025453.3025949.

[134] A. Fidyka and A. Matamala, 'Audio description in 360º videos. Results from focus groups in Barcelona and Kraków,' vol. 44, no. 8, pp. 1–8, 2011. DOI: https://doi.org/10.1075/ts.18018.fid. [Online]. Available: https://rightsstatements.org/vocab/InC/1.0.

[135] ——, 'Production of access services in immersive content: Understanding the needs of audio describers,' *Hikma*, vol. 18, no. 2, pp. 277–300, 2019, ISSN: 24454559. DOI: 10.21071/hikma.v18i2.11683.

[136] L. Picinali, A. Etherington, C. Feakes and T. Lloyd, 'VR interactive environments for the blind: Preliminary comparative studies,' in *VTT Symposium (Valtion Teknillinen Tutkimuskeskus)*, Technical Research Centre Finland, 2011, pp. 113–115, ISBN: 9789513876029.

[137] L. Picinali, A. Afonso, M. Denis and B. F. Katz, 'Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge,' *International Journal of Human Computer Studies*, vol. 72, no. 4, pp. 393–407, 2014, ISSN: 10715819. DOI: 10.1016/j.ijhcs.2013.12.008. [Online]. Available: http://dx.doi.org/10.1016/j.ijhcs.2013.12.008.

[138] A. F. Siu, M. Sinclair, R. Kovacs, E. Ofek, C. Holz and E. Cutrell, 'Virtual Reality Without Vision: A Haptic and Auditory White Cane to Navigate Complex Virtual Worlds,' *Conference on Human Factors in Computing Systems - Proceedings*, 2020. DOI: 10.1145/3313831.3376353.

[139] C. Morrison *et al.*, 'Social sensemaking with ai: Designing an open-ended ai experience with a blind child,' *Conference on Human Factors in Computing Systems - Proceedings*, 2021. DOI: 10.1145/3411764.3445290.

[140] A. Brock and M. Hachet, 'Teaching spatial thinking to visually impaired students using augmented reality: Introducing the VISTE project,' in *GET'17 - Journées Multimédia, Géomatique, Enseignement \& Apprentissage*, 2017.

[141] Y. Zhao, S. Szpiro, J. Knighten and S. Azenkot, 'CueSee: Exploring Visual cues for people with low vision to facilitate a visual search task,' *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 73–84, 2016. DOI: 10.1145/2971648.2971730.

[142] A. Guo *et al.*, 'Vizlens: A robust and interactive screen reader for interfaces in the real world,' *UIST 2016 - Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 651–664, 2016. DOI: 10.1145/2984511.2984518.

[143] D. Sato, U. Oh, K. Naito, H. Takagi, K. Kitani and C. Asakawa, 'NavCog3: An evaluation of a smartphone-based blindindoor navigation assistant with semantic features in a large-scale environment,' *ASSETS 2017 - Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 270–279, 2017. DOI: 10.1145/3132525.3132535.

[144] J. Huang, M. Kinateder, M. J. Dunn, W. Jarosz, X. D. Yang and E. A. Cooper, 'An augmented reality sign-reading assistant for users with reduced vision,' *PLoS ONE*, vol. 14, no. 1, pp. 1–9, 2019, ISSN: 19326203. DOI: 10.1371/journal.pone.0210630.

[145] Y. Zhao, E. Kupferstein, B. V. Castro, S. Feiner and S. Azenkot, 'Designing AR visualizations to facilitate stair navigation for people with low vision,' *UIST 2019 - Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, no. August, pp. 387–402, 2019. DOI: 10.1145/3332165.3347906.

[146] J. M. Coughlan, B. Biggs, M.-A. Rivière and H. Shen, 'An Audio-Based 3D Spatial Guidance AR System for Blind Users,' *Computers Helping People with Special Needs*, vol. 12376, no. 12, pp. 475–484, 2020, ISSN: 1873507X. DOI: 10.1007/978-3-030-58796-3_55.

[147] L. Gupta, N. Varma, S. Agrawal, V. Verma, N. Kalra and S. Sharma, 'Approaches in assistive technology: A survey on existing assistive wearable technology for the visually impaired,' in *Computer Networks, Big Data and IoT*, June, vol. 66, 2021, pp. 541–556. DOI: 10.1007/978-981-16-0965-7_42.

[148] Microsoft, *Microsoft Soundscape*. [Online]. Available: https://www.microsoft.com/en-us/research/product/soundscape/ (visited on 06/10/2021).

[149] Envision, *Envision App*, 2021. [Online]. Available: https://www.letsenvision.com/envision-app (visited on 03/11/2021).

[150] Google LLC, *Lookout by Google*, 2021. [Online]. Available: https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal%7B%5C&%7Dhl=en%7B%5C_%7DUS%7B%5C&%7Dgl=US.

[151] Microsoft, *Seeing AI App from Microsoft*. [Online]. Available: https://www.microsoft.com/en-us/ai/seeing-ai.

[152] Mediate, *Supersense*, 2020. [Online]. Available: https://www.supersense.app/ (visited on 06/10/2021).

[153] Lazarillo, *Lazarillo*, 2021. [Online]. Available: https://lazarillo.app/ (visited on 06/10/2021).

[154] Tuat Corp, *Sullivan Plus*, 2021. [Online]. Available: https://www.mysullivan.org/ (visited on 06/10/2021).

[155] Cloudsight Inc, *TapTap See*, 2021. [Online]. Available: https://taptapseeapp.com/ (visited on 03/11/2021).

[156] Be My Eyes, *Be My Eyes*, 2021. [Online]. Available: https://www.bemyeyes.com/ (visited on 12/10/2021).

[157] Google, *Accessibility Scanner*. [Online]. Available: https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.auditor%7B%5C&%7Dhl=en (visited on 01/03/2021).

[158] Rovio Entertainment Corporation, *Angry Birds AR:Isle of Pigs – Apps on Google Play*. [Online]. Available: https://play.google.com/store/apps/details?id=com.rovio.abar.

[159] ——, *Angry Birds AR:Isle of Pigs - App Store Preview*. [Online]. Available: https://apps.apple.com/gb/app/angry-birds-ar-isle-of-pigs/id1456120259.

[160] Niantic Inc, *Pokémon GO - Apps on Google Play*. [Online]. Available: https://play.google.com/store/apps/details?id=com.nianticlabs.pokemongo.

[161] ——, *Pokémon GO - App Store Preview*. [Online]. Available: https://apps.apple.com/gb/app/pok%7B%5C'%7Be%7D%7Dmon-go/id1094591345.

[162] World Wildlife Fund, *WWF Free Rivers - Apps on Google Play*. [Online]. Available: https://play.google.com/store/apps/details?id=org.worldwildlife.FreeRiver%7B%5C&%7Dhl=en%7B%5C_%7DUS%7B%5C&%7Dgl=US.

[163] ——, *WWF Free Rivers - App Store Preview*. [Online]. Available: https://apps.apple.com/us/app/wwf-free-rivers/id1349935575.

[164] J. Woodcock, C. Pike, F. Melchior, P. Coleman, A. Franck and A. Hilton, 'Presenting the S3A object-based audio drama dataset,' *Audio Engineering Society 140th Convetion*, 2016. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18159.

[165] EBU, 'Audio Definition Model (ADM),' *Tech 3364*, vol. 2, no. June, 2018. [Online]. Available: https://tech.ebu.ch/docs/tech/tech3364.pdf.

[166] VRTOV, *The Turning Forest*, 2016. [Online]. Available: https://vrtov.com/projects/turning-forest/ (visited on 02/11/2021).

[167] Unity Technologies, *Augmented Reality Development Software*, 2021. [Online]. Available: https://unity.com/unity/features/ar (visited on 02/06/2021).

[168] ——, *Wondering what Unity is?* [Online]. Available: https://unity.com/our-company.

[169] D. Stenfors, *Low Poly Forest Pack*, 2019. [Online]. Available: https://assetstore.unity.com/packages/3d/environments/low-poly-forest-pack-60137%7B%5C#%7Ddescription.

[170] Acorn Bringer, *Simplistic Low Poly Nature*, 2018. [Online]. Available: https://assetstore.unity.com/packages/3d/environments/simplistic-low-poly-nature-93894.

[171] E. Dogan, *LowPoly Water*, 2018. [Online]. Available: https://assetstore.unity.com/packages/tools/particles-effects/lowpoly-water-107563.

[172] Google, *Google Play Console*. [Online]. Available: https://play.google.com/console (visited on 25/01/2021).

[173] Apple, *Testing Apps with TestFlight*. [Online]. Available: https://testflight.apple.com/ (visited on 25/01/2021).

[174] Zoom Video Communications Inc, *Zoom*. [Online]. Available: https://zoom.us/ (visited on 17/02/2021).

[175] Aira Tech Corp, *Visual Interpreting – Get Live, On-demand Access to Visual Information*, 2022. [Online]. Available: https://aira.io/ (visited on 25/01/2022).

[176] Microsoft, *Transcribe your recordings*, 2021. [Online]. Available: https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57 (visited on 15/05/2021).

[177] Audiokinetic, *Wwise*. [Online]. Available: https://www.audiokinetic.com/products/wwise/ (visited on 09/06/2021).

[178] Unity Technologies, *Unity's AR Foundation Framework*. [Online]. Available: https://unity.com/unity/features/arfoundation (visited on 02/06/2021).

[179] M. Gemmell, *Accessibility for iPhone and iPad apps*, 2010. [Online]. Available: https://mattgemmell.com/accessibility-for-iphone-and-ipad-apps/ (visited on 24/05/2021).

[180] MetalPop Games, *UI Accessibility Plugin Version*. [Online]. Available: http://www.metalpopgames.com/assetstore/accessibility/doc/index.html.

[181] M. Gorzel *et al.*, 'Efficient encoding and decoding of binaural sound with resonance audio,' *Proceedings of the AES International Conference*, vol. 2019-March, 2019. [Online]. Available: https://www.aes.org/e-lib/browse.cfm?elib=20446.

[182] Google, *Resonance Audio*, 2018. [Online]. Available: https://developers.google.com/resonance-audio/ (visited on 07/04/2022).

[183] C. Armstrong, L. Thresh, D. Murphy and G. Kearney, 'A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database,' *Applied Sciences (Switzerland)*, vol. 8, no. 11, 2018, ISSN: 20763417. DOI: 10.3390/app8112029.

[184] W. D. Hairston, M. T. Wallace, J. W. Vaughan, B. E. Stein, J. L. Norris and J. A. Schirillo, 'Visual localization ability influences cross-modal bias,' *Journal of Cognitive Neuroscience*, vol. 15, no. 1, pp. 20–29, 2003, ISSN: 0898929X. DOI: 10.1162/089892903321107792.

[185] C. E. Jack and W. R. Thurlow, 'Effects of degree of visual association and angle of displacement on the 'ventriloquism' effect,' *Perceptual and Motor Skills*, vol. 37, no. 3, pp. 967–979, 1973, ISSN: 00315125. DOI: 10.2466/pms.1973.37.3.967.

[186] H. A. Witkin, S. Wapner and T. Leventhal, 'Sound localization with conflicting visual and auditory cues,' *Journal of Experimental Psychology*, vol. 43, no. 1, pp. 58–67, 1952, ISSN: 00221015. DOI: 10.1037/h0055889.

[187] Firelight Technologies Pty Ltd, *Fmod*. [Online]. Available: https://fmod.com/ (visited on 09/06/2021).