The use of Interactive Spatial Audio as a Means of Addressing Auditory Hypersensitivity in Autistic Young People

Daniel Johnston

Doctor of Philosophy

University of York
Electronic Engineering

October 2020

Abstract

Individuals diagnosed with autism spectrum disorder (ASD) are characterised as having impairments in social-emotion interaction and communication, alongside displaying repetitive behaviours and a limited range of interests. Additionally, they can frequently experience difficulties in processing sensory information with prevalence in the auditory domain. Often triggered by everyday environmental sounds, auditory hypersensitivity can lead to self-regulatory fear responses such as crying and isolation from sounds. Rather than a physiological pain reaction, literature suggests that these hypersensitivities are established in an irrational fear of the sounds. This has subsequently led to several successful interventions that increase habituation to problematic auditory stimuli through controlled, safe and graduated exposure. In recent years this therapy has been embedded into serious computer game mechanics, delivering exposure-based training within an engaging and motivating virtual environment. The research presented in this thesis investigates the use of three-dimensional spatial audio as an approach to rendering realistic simulations of adverse stimuli within a virtual reality game environment developed to reduce auditory hypersensitivity in autistic children and adolescents ($8 - 19$ years old). This was achieved through empirical studies assessing the effectiveness of reducing the negative emotional associations towards auditory stimuli presented using binaural based spatial audio. With additional emphasis evaluating the VR intervention environment and game mechanics. Findings indicate that playing the game produced as part of this doctoral project over the course of 4-weeks can be effective in reducing negative emotions linked to problematic sounds in autistic young people. Furthermore, comparative results suggest that these reductions can be significantly improved upon when using binaural based spatial audio instead of stereo audio rendering. This therefore supports the use of three-dimensional sound as an audio rendering technique to target auditory hypersensitivity in autistic individuals, ultimately improving their quality of life.

Contents

List of Figures

List of Tables

Acknowledgements

I would first like to extend my deepest appreciation to my supervisors Dr Gavin Kearney and Dr Hauke Egermann, for their guidance and continued support throughout this research journey. Their expertise, enthusiasm for the project and kindness have been a critical contribution to the completion of this thesis and my personal development as a researcher. I will be always grateful to them for providing me this opportunity.

Furthermore, I would like to thank Dr Mariana Lopez and Professor Damian Murphy for their advice over the three years that helped provide extra direction for the research. I would also like to thank everyone in the Audio Lab for helping to create a fun, creative and motivating environment to be in. Everybody there is truly wonderful and it is an amazing place to work and grow.

The research would also not have been possible if it was not for the support of the schools and organisations that took part in the investigations. You were always accommodating and enthusiastic to contribute to the projects, providing invaluable assistance and advice to help the work come to fruition.

Finally, I would like to thank my friends and family who were always interested in my work and always keen to hear about my progress. Especially my partner and best friend Sally, who's endless love, support, and above all patience has been the reason I made it to the finish line.

Thank you.

Declaration of Authorship

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. I also declare that parts of this research have been presented in previous conference and journal publications, which are listed as follows:

- Johnston, D., Egermann, H., Kearney, G. (2018). Innovative computer technology in music-based interventions for individuals with autism moving beyond traditional interactive music therapy techniques. *Cogent Psychology,* 5(1), 1554773.
- Johnston, D., Egermann, H., Kearney, G. (2018, October). SoundFields: A Mixed Reality Spatial Audio Game for Children with Autism Spectrum Disorder. In *Audio Engineering Society Convention 145.* Audio Engineering Society.
- Johnston, D., Egermann, H., Kearney, G. (2019, March). An Interactive Spatial Audio Experience for Children with Autism Spectrum Disorder. In *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio.* Audio Engineering Society.
- Johnston, D., Egermann, H., Kearney, G. (2019). Measuring the Behavioral Response to Spatial Audio within a Multi-Modal Virtual Reality Environment in Children with Autism Spectrum Disorder. *Applied Sciences,* 9(15), 3152.
- Johnston, D., Egermann, H., Kearney, G. (2020). SoundFields: A Virtual Reality Game Designed to Address Auditory Hypersensitivity in Individuals with Autism Spectrum Disorder. *Applied Sciences*, 10(9), 2996.

Chapter 1

Introduction

1.1 Introduction

ASD is a complex pervasive neurological condition that influences perception, cognition and emotion. Characterised through impaired development in social interaction, communication and repetitive behaviours and interests, standard diagnosis manuals such as the The Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-IV-TR) [1] have partitioned ASD into a number of sub-types which accommodates multiple levels of cognitive and social abilities [2].

Commonly, autistic children experience complications in tactile, visual and auditory processing, with over 96% of children with ASD experiencing hypersensitivities in multiple domains. Whilst difficulties in processing sensory information is not specifically limited to individuals on the autistic spectrum, it is more prevalent within this population when compared to other developmental disabilities [3], [4]. Sound sensitivity is a specifically poignant and common issue for those with ASD. Diagnostic questionnaires collected from 17,000 familes by the Autism Research Institute revealed that approximately 40% of children displayed symptoms associated with auditory hypersensitivity [5]. These difficulties in sensory processing can provoke atypical self-regulatory behaviours which may be observed as aggressive or autonomic fear responses such as covering ears, crying and self injury from blows to the ears [6]. Unfortunately these profound aversions are reported to be provoked by common environmental sounds [7]. By avoiding challenging acoustic environments, autistic children can experience increased isolation and further impairment in social communication.

Atypical and extreme behavioural responses to auditory sensory simulation within the autistic population can be attributed to an irrational fear of the sound in place of a physiological pain reaction associated with the stimulus [7]. In addition, there is extensive literature that observes specific phobias as one of the most frequent causes of anxiety with autism. The prevalence rates of phobias recorded in ASD is at 30-64% [8], [9] compared to the 5-18% recorded within the typically developed population [10]. With this in mind, Cognitive Behavioural Therapy (CBT) interventions such as exposure based training have gained the most empirical support for addressing auditory hypersensitivity in autism. Such interventions aim to develop new and positive emotional associations with aversive sounds through safe graduated exposure during play based activities.

It is highly recommended that early intervention is essential for autistic children, with evidence supporting very early intervention for children as young as 3 years old [11], [12]. Early brain development not only includes the emergence of associated abnormal neural connections and an exponential rate of brain growth, but also a period of increased brain plasticity [13]. Early therapy can therefore influence the development of brain systems implicated in social, emotional and communication behaviours [13], [14]. As well as pharmaceutical [15] and dietary [16] treatments, there are a wide range of interventions that include speech/language therapy [17], occupational therapy [18] and sensory integration [19].

Computer based interventions for ASD have been the subject of extensive research for the last decade [20]. With noted success in addressing social [21], communication [22], behavioural [23] and adaptive skills [24]. Autistic people respond well to computerised therapies, as they provide a predictable, repeatable and immediate learning environment that can deliver clearly defined tasks away from unnecessary stimuli [25]. Furthermore, computers are free from the social constraints that those with ASD can often find challenging a confusing [26].

Novel approaches to computer based interventions have moved passed the traditional desktop computer, with examples incorporating the use of serious games [27] and Virtual Reality (VR) [28]. These innovative uses of technology not only address the core symptoms of autism, but they emphasise the importance of integrating realistic multi-sensory stimuli in enhancing the therapy experience for anxiety and phobias. Recently, serious computer games have demonstrated success in reducing auditory hypersensitivity symptoms in children with autism by integrating CBT into the core game mechanics [29], [30]. During game-play, players are rewarded for gradually being exposed to reproductions of problematic environmental sounds, resulting in increased habituation and decreased negative emotional responses [29], [30].

The use of virtual reality based interventions have been successful in reducing anxiety based disorders in those with and without ASD [31]–[34]. The technology is capable of rendering realistic 3D environments that accurately simulate feared stimuli and environments [35]. Further to this, an audio rendering technique known as Ambisonics has the ability to reproduce immersive 360° soundfields that include dynamically moving virtual sound sources which respond to the users head rotation and position within a virtual environment [36].

The aim of the research presented in this thesis is to examine how binaural based Ambisonics can be used to help reduce auditory hypersensitivity in autistic children and adolescents. The prominent audio rendering technique used within contemporary VR systems, binaural Ambisonics, is capable of simulating realistic auditory environments which rotate and adjust based upon the listener's head rotation and position within a virtual environment [37]. Improved realism could therefore be applied to the reproduction of averse auditory stimuli within exposure based therapy games deployed to VR. Enhancing established empirically supported intervention techniques and increasing accessibility to therapy.

1.2 Statement of Hypothesis

The overall hypothesis that forms the primary motivation of the research presented in this thesis is as follows:

> *The use of binaural ambisonic rendering to deliver auditory stimuli within an exposure based serious game will have a positive impact on reducing associated negative emotions of autistic children and adolescents experiencing auditory hypersensitivity.*

The primary terms of this hypothesis and how they relate to the thesis are explained as follows:

- **Binaural Ambisonic Rendering**: The reproduction of immersive 3D virtual auditory environments over headphones. Commonly used in VR systems, this audio rendering technique can rotate a virtual soundfield based on the listener's movements to allow for localisation of sound sources.
- **Autism Spectrum Disorder**: A complex and lifelong developmental disability that typically appears during early childhood. Individuals diagnosed with autism are characterised as having impairments in social emotional interaction and communication, alongside displaying repetitive behaviours and interests. Additionally, they can frequently experience difficulties in processing sensory information.
- **Auditory Hypersensitivity**: Often experienced by autistic people, this can manifest as an intense aversion to common environmental sounds. Reactions can be displayed as covering ears, verbalisation and sound avoidance. Observation in research suggest that auditory hypersensitivity is caused by a psycho-emotional response to sound rather than malformations within the auditory canal.

1.3 Structure of Thesis

The remainder of this thesis is divided into seven chapters. The first two will provide detailed reviews of existing literature that will inform the original research presented in the subsequent chapters. A summary description of the chapters are as follows:

Chapter 2 explores some of the unique ways in which individuals with autism process sensory information within the auditory domain, with a focus on auditory hypersensitivities. This is followed by a review of the current intervention approaches used to target the issue. Finally, the chapter will discuss the benefits of computer driven therapy, in particular serious games and VR, and how these technologies can be used to enhance existing intervention frameworks.

Chapter 3 describes the fundamentals of spatial audio. First binaural listening, including interaural time and level difference and spectral cues caused by the pinnae are explored. Following this, the chapter will discuss spatial audio rendering techniques with a focus on

Ambisonics. Finally, the chapter will explore psychological and emotional responses to spatialised sound and how it can be applied within exposure based therapy frameworks.

Chapter 4 presents a novel VR based experiment that aims to investigate how individuals with autism respond to spatial audio events delivered with a multi-modal virtual environment. Based upon the atypical auditory processing outlined in Chapter 2, a two phase experiment is presented that compares behavioural reactions to spatialised and non-spatialsed audio and the effect of background noise on the participant's interactions with the environment.

Chapter 5 describes a novel VR game named SoundFields. A game developed to reduce the negative emotions associated with specific auditory stimuli experience by autistic children and adolescents who display auditory hypersensitivity. A detailed explanation will be given about the underlying therapeutic schemes that have been embedded into the games core mechanics in order to motivate the player to interact and engage with the framework.

Chapter 6 is the first of two original studies which evaluates the use of 3^{rd} order Ambisonics as an audio rendering tool to target auditory hypersensitivity in autism. The 4-week investigation exposed participants to problematic sounds embedded into the serious game mechanics and VR environment discussed in Chapter 5. The primary outcome of the experiment compared pre-and-post measurements of participants self-reported emotional associations towards problematic stimuli. This was supported by a secondary outcome which tracked voluntary interactions with exposure stimuli across the experimental sessions. Alongside the spatial audio rendering technique, how participants engaged with the VR game is also evaluated.

Chapter 7 expands upon the study presented in Chapter 6 with a longitudinal 8-week study to compare 3^{rd} order Ambisonics to stereo rendering of problematic stimuli. Replicating the outcome measurements of the previous chapter, comparisons were made between the two audio conditions. In addition to a 4-week experimental period, a control period was implemented to increase the validity of the results gathered from the experimental sessions. Finally, a follow-up measurement was added to asses if any increase tolerance was maintained.

Chapter 8 summarises the experimental studies presented throughout the thesis followed by conclusions derived from the core hypothesis. Finally, suggestions are made for how future studies can expand on the outcomes of this research.

1.4 The description of autism within the thesis

The way in which autism is described has been the subject of much discussion for many years [38]. Disagreement occurs within the autism community and is frequently attributed to the wide variations in which autism touches peoples lives [39]. Furthermore, the increasing influence of the neurodiversity movements has generated discussion within the research community concerning the most appropriate approach to describing autism within literature. This is a crucial topic to consider as the terminology used within research can influence how people perceive autism, clinical practice, public policy and future research directions [38]. Furthermore, language can have a direct impact on how individuals with an autism diagnosis perceive autism. This self-perception has links to a person's self-esteem, self-identity and mental health [40]. Finally, due to the the interdisciplinary nature of this thesis (see Section 1.5), it is crucial to highlight the importance of this issue to those unfamiliar with autism research. This section aims to briefly outline the key issues in the language used for the description of autism and their links to the different models of disability, Following this, it will identify the terminology used throughout this thesis.

Within literature there are two different and contested approaches, Person First Language (PFL) (i.e person with autism) and Identity First Language (IFL) (i.e autistic person). PFL is taught in health profession programs and often finds usages within literature that places autism within the medical model of disability [41], [42]. The medical model of disability examines the pathology of a condition and attempts to prevent, cure or care for a disabled person through medical intervention [43]. In terms of autism, medical research has aimed to explore the neurological causes of autism such as excess neuron overgrowth [44] or a dysfunctional mirror neuron system [45].

Person first language was first introduced in the 1970s as an alternative to terminology such as 'handicapped' and 'disabled' [38]. By doing so it aimed to humanise disabled people by putting the person before the disability, highlighting that a disability is not their only defining feature [46]. In addition, it recognised that a disabled person is a 'person first' and therefore should be provided the same rights and protections as any other person [38].

However, in recent years there has been opposition to the use of PFL. Gernsbacher [47] suggests that person-first language can increase stigma as it is often used with more stigmatised disabilities. For example 'children with autism' compared to 'blind children'. In addition, autistic scholars such as Sinclair explain that PFL suggests that autism can be separated from the individual, when in fact this is not the case [48]. This also brings into focus the potential limitations of placing autism into the medical model of disability. Medical intervention has been crucial in the treatment of many physical co-moralities associated with autism such as epilepsy [49], [50]. However, critics argue that it hyper-focuses on the deficits of autism rather than what they can accomplish through characteristic traits [41] such as strengths in pattern recognition [51] and enhanced perception of sound [52]. The deficit-focused discourse of the medical model has also been observed as having a significant impact upon how autistic individuals build an identity outside of a diagnosis [53].

Gernsbacher [47] and other autistic scholars and advocates champion the use of IFL [48], [54], [55]. This support has been documented within the autistic community with research conducted by Bury *et al* [56] observing 49% of autistic participants electing for IFL. One of the central arguments for its use is that it does not separate the individual from the diagnosis, minimising the societal view that autism is 'bad' [38]. Instead, IFL assists in expanding upon the view that autism should be celebrated as central part of a person's identity [57]. This perspective of oneself has been echoed in a recent study which investigated the attributes autistic people perceived as having a positive or negative impact upon their identity and wellbeing [40]. The survey identified a wide range of attributes of autism that participants were proud of (e.g "gifted", "uniqueness"), creating a positive affiliation with their autism identity. However, despite the fact that IFL is preferred, saying that a person is "on the spectrum" is still considered less offensive than "with autism" [56]. This is despite it essentially being a person first expression.

Identity first language has been linked to the social model of disability. The social model describes disability not as an individual's impairment but as a social construction and a result of an oppressive and excluding environment [43]. Therefore, any improvement made to a disabled person's life will be from a successful adaption of the social environment which originally failed to account for their differences [43]. Under the social model, common behaviours associated with autism can be seen not as medical symptoms but as unconventional forms of behaviour that are not yet accepted by society [41]. One example is the syntactliy faithful repetitions of overheard words, known as echolalia [58], which has previously been regarded as meaningless vocalisations [59]. However, research has now shown that these repetition of phrases may act as an important communicative function between the individual and their caregivers [59].

Throughout this thesis, identity first language will be used when referencing individuals with an autism diagnosis. Firstly, this decision was based upon the support in literature calling for the use of IFL. Secondly, and possibly more importantly, is the supporting empirical evidence that this terminology is the preferred approach within the autistic community [39], [56], with PFL being considered as offensive [60]. Additionally it is important to discuss how the heterogeneous nature of autism will be described in the proceeding research. Past research has used the spectrum metaphor to describe an autistic individual as being either 'high functioning' or 'low functioning'. It is believed that the categorisation based on what an individual cannot do is an oversimplification of the individual autism experience [61]. With this in mind, 'high' and 'low' high functioning will not be used. Instead, as recommended by Botha *et al* [62], language will be used that highlights the particular constellation of characteristics within the context of the research being described.

Although the terminology used reflects that of the social disability and neuro-diversity movements [38], [62], how the research in this thesis is categorised is somewhat different. Anderson [41] suggests that autism may not fall under either the medical or social model of disability, but rather a third model: the predicament model. This paradigm recognises that disability is both biologically based and socially constructed. In which a person with atypical functionality will confront restrictions within a society which has not taken into consideration atypical function [41], [63]. The predicament model encompasses the unique and subjective experiences of each individual and how they identify with their disability [41], [64]. Anderson argues that autism is highly individualised and both an ability and disability [41]. Therefore, the predicament model encompasses positive autism experiences such as visual-spatial aptitude, and the negative experiences such as auditory hypersensitivity.

As explored in Chapter 2, auditory hypersensitivity can mean autistic people will experience specific auditory stimuli in unique ways. Unfortunately, due to the commonality of the sounds it is impossible for society to adapt in order to mitigate these negative experiences. This therefore necessitates the need for interventions to help people on the autistic spectrum to become habituated to sounds which may normally cause distress [7], [29], [65]. The objective of the research in Chapters 6 and 7 is to reach this goal by delivering an intervention through engaging interactive game (see Chapter 5). The game takes advantage of the general positive experiences provided by computer games [66]–[69], whilst also addressing the unique and nuanced sensory processing of the individual.

This section aimed to briefly outline the key considerations when describing autism, particularly within research. The issue is a complicated one and can directly impact the lives of autistic people. The author recognises the arguments surrounding correct terminology necessitates exploration beyond this section of the thesis. The author therefore recommends the following literature for more a comprehensive description [38], [39], [43], [61], [62].

1.5 Positioning of Research

The design, development and empirical evaluation of an interactive technology based intervention for autistic people requires the input of a number of key fields of study. Furthermore, these interventions will benefit from the diverse strengths and viewpoints that each discipline brings to the singular research project [70]. The research presented is this thesis adopts this approach, and so therefore at this stage it is important to examine its inter-disciplinary nature and its positioning as a contribution to the wider research community.

Firstly, the application of spatial audio in reducing auditory hypersensitivity in autistic individuals relies upon established psychological theories and practices which address this issue using CBT [7], [65], [71]. The application of CBT is grounded in extensive psychological and neurological investigations conducted by the autism research community spanning over five decades (see Chapter 2). In addition, the thesis examines the psychological concepts of immersion and presence [72] when neurotypical (NT) and autistic people experience virtual environments (see Chapters 2 and 3). These models form the foundation for the support and utilisation of spatial sound within the research described later.

Secondly, this research relies on areas of empirical investigations which explore the use of computer games as tools for delivering cognitive behavioural therapy [29], [73]–[76] (see Chapters 2 and 5). Research which focuses on using computer game mechanics to address the difficulties experienced by autistic people in areas such as social interaction, communication and auditory hypersensitivity, has shown that this is an engaging and motivating approach to therapy. Furthermore, the development of consumer virtual reality technology means that computer based interventions can now benefit from the immersive and realistic visual and auditory stimuli VR provides.

Finally, the proceeding research employs technology and conventions from the research field of audio engineering. The reproduction of realistic immersive 3D virtual environments through binaural Ambisonic rendering is one of the core components of the hypothesis. How Ambisonic sound is recorded, produced and delivered to an individual is founded upon ongoing research originating in the 1970s [77] (see Chapter 3). A culmination of nearly half a century of research and advancements in technology now allows users to experience spatial audio environments on consumer platforms such as virtual reality [78], [79]. This increased accessibility therefore creates new opportunities for Ambisonic technology to be utilised in novel applications outside of the realm of entertainment. Applications such as the proposed intervention presented in this thesis.

Further to outlining the research areas this thesis is rooted in, Figure 1.1 also illustrates its contribution to the same research fields. Virtual reality has been a focus of research within the autism research community for over two decades [80]. Despite this, the use of spatial audio within interactive interventions has not yet been explored. The primary use of spatial audio within the thesis is a tool to address auditory hypersensitivity through CBT techniques. However, the investigation in Chapter 4 explores the behavioural response of autistic children to auditory stimuli rendered through binaural Ambisonics within VR. Therefore, this work may provide a valuable contribution to an active research area that investigates and develops VR based interventions for autistic individuals. Furthermore, it provides further empirical support for the use of virtual interactive technology as viable tools for delivering psychology established therapy for autistic children and adolescents.

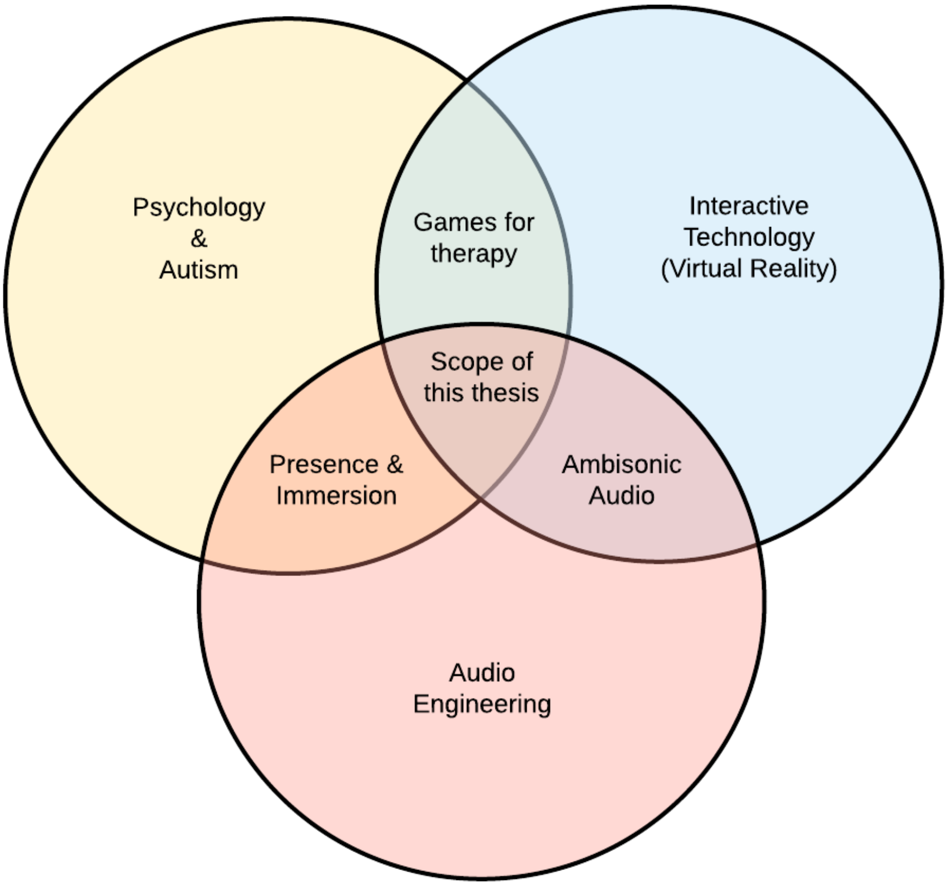


Figure 1.1: The Multi-disciplinary background and contributions of this thesis

Finally, by examining an use of Ambisonic audio technology the scope of this research also extends into the domain of audio engineering. Although the sound production and rendering techniques used throughout the thesis are not novel (see Chapter 5), the application in which they are experienced is. As previously mentioned, digital interventions benefit from multi-disciplinary engagement [70]. With that in mind, it is hoped that this work will inspire future collaboration between the audio engineering and autism research communities.

1.6 Research Contributions

The main aim of this research is to investigate how binaural based ambisonic audio can be used as an tool to target auditory hypersensitivity in autistic children and adolescents. Below is a list of novel research contributions presented in this thesis:

- **A Review of literature investigating computer based technology to deliver therapy for auditory hypersensitivity experienced by autistic people**: This provides an in depth analysis of the use of computer driven technology such as serious games and VR as tools to deliver CBT interventions for autistic people. The review discusses how this technology can be used to bypass some of the core symptoms of autism to increase engagement, therapeutic outcomes and accessibility to treatment.
- **Measurement of the behavioural response to spatial audio within a multi-modal VR environment in autistic children**: Literature has observed atypical auditory processing displayed in autistic individuals such as deficits in auditory scene analysis and sound source localisation (see Section 2.2). This novel experiment investigates if these difficulties in processing auditory information would directly impact how autistic school children interact with a presented virtual spatial audio environment presented within VR. Two experiments were conducted with participants diagnosed with ASD ($n = 29$) that compared: (1) behavioural reaction between spatialised and non-spatialised audio; and (2) the effect of background noise on participant interaction. Participants listening to binaural-based spatial audio showed higher spatial attention towards target auditory events. In addition, the amount of competing background audio was reported to influence spatial attention and interaction. These findings indicate that despite suggested associated sensory processing difficulties, those on the autistic spectrum can correctly decode the auditory cues simulated with current spatial audio rendering techniques.
- **Development of a VR serious game that targets auditory hypersensitivity in autistic children and adolescents**: Developed using exposure based mechanics found in serious games designed to target auditory hypersensitivity in autistic individuals (see Section 2.5.1), this novel game uses head-tracked binaural based spatial audio to deliver realistic auditory stimuli.
- **The application and evaluation of binaural based Ambisonics to help reduce auditory hypersensitivity in autistic children**: Finally, the use of 3^{rd} order ambisonic audio is evaluated as a tool for the reproduction of problematic auditory stimuli in order to reduce auditory hypersensitivity in ASD. Direct comparisons were made between ambisonic and stereo audio rendering techniques. Research outcomes were achieved by measuring emotional responses to problematic stimuli before and after the experimental period. In addition, voluntary interactions with stimuli were recorded within the virtual environment.

1.7 Statement of Ethics

The experimental procedures and data management presented throughout this thesis were approved by the University of York Physical Sciences Ethics Committee. Reference numbers for each study are as follows: Chapter 4 Johnston030518, Chapters 6 & 7 Johnston220219.

Chapter 2

Auditory Hypersensitivity in Autism Spectrum Disorder: Current approaches to therapy

2.1 Introduction

This literature review begins with an examination of atypical auditory processing in autism, with particular focus on hypersensitivity to sound. Before evaluating current therapy frameworks, research will be presented that investigates its potential physiological, neurological and psychological causes. Following this, there is an in-depth investigation into how new technologies such as VR and serious games are able to enhance traditional cognitive behaviour therapy interventions for autism, improving engagement with therapeutic sessions and increasing generalisation of new skills following an intervention. Finally, the implications of current research are explored with suggestions for both future investigations and important design considerations for interventions addressing auditory hypersensitivity for individuals on the autism spectrum.

2.2 Auditory Processing in Autism

Research into sensory perception in autism has revealed an interesting profile of strengths and weaknesses in processing auditory information [81]. Observations are diverse, ranging from enhanced musical abilities [82] to an impaired ability to separate speech sounds from environmental background noise [83]. For the context of this thesis it is important to consider the complex ways autistic individuals process auditory sensory information. Primarily this will provide a foundation of empirical evidence for using spatial audio to target auditory hypersensitivity. Moreover, understanding these unique auditory processing patterns will aid in making informed sound design choices, tailoring the audio environments to maximise the potential therapy outcomes of the intervention presented later in this thesis (see Chapter 5). This section will therefore focus on studies which use both behavioural and neurological analysis techniques to investigate the various mechanisms behind the processing of auditory stimuli in ASD.

2.2.1 Enhanced perception of sound

How an autistic person perceives sound has been the focus of a large amount of research, beginning with investigations into autistic savants. Savant, which is derived from the french work savoir, which means ‘to know’ [84], refers to an individual who in spite of low general cognitive abilities, shows remarkable skills in isolated domains such as mathematical calculations [85] , drawing [86], [87] and musical memory and interpretation [84], [88], [89]. Although savants are rare within populations of individuals with cognitive impairments [90], incidences of these special abilities appears to be higher within the autistic population [52].

However, increased sound perception and pitch discrimination in ASD is not limited to savants and most recent research into this area has concentrated on musically naive individuals with no formal training in music theory or performance. Research conducted by Heaton *et al.* [91] compared ten autistic children to neurotypical controls in their capacity to identify and recall musical tones. Results showed that autistic children with no specialised musical training or talent had superior abilities in discriminating and labelling signal music notes. A further study by Heaton *et al.* [92] examined the capacity to discriminate pitch in speech stimuli in autistic and NT children. Results showed that when compared to controls, the autistic participants displayed exceptional sensitivity to the non-semantic pitch information of speech. Subsequent research in which children were asked to match sentence stimuli and graphical representations to pitch shape found that autistic children were able to provide significantly more perceptual interpretations of auditory stimuli than NT controls[93].

These investigations identify that enhanced pitch perception may contribute to language impairment in the early stages of development. This is supported in a study that reported a significant proportion of ASD participants who displayed both accurate auditory discrimination and early language delays [94]. An intriguing argument for this is the reduced engagement with linguistic stimuli and a biased attention towards the musical non-speech aspects, resulting in an over-specialisation of pitch-processing mechanisms [95]–[98]. It has also been suggested that those on the autistic spectrum have a brain that focuses on right hemispheric cognitive functions such as spectral processing, at the disadvantage of left hemispheric cognitive mechanisms like temporal, speech and language processing [99]. However, a study by Boets *et al* [100] that compared temporal amplitude modulation detection and pitch perception in both autistic and TD controls, found no convincing evidence to support enhanced right and decreased left hemisphere processing is ASD. Instead they concluded, that the atypical pitch processing observed in ASD is not due to an enhanced perceptual ability, rather a higher focus and attention to simple and local auditory information.

2.2.2 Processing sound in background noise

Commonly within a natural listening environment speech will coexist with competing everyday background sounds, creating an amalgamation of acoustic signals arriving at the ear [101]. In order to perceive grouped elements as separate acoustic events, the auditory system must perform Auditory Scene Analysis (ASA) [102]. A process which is heavily reliant upon synchronous and successive organisation of acoustic information, Bregman [102] conceptualised that ASA was a two staged operation that involved primitive and schema-driven mechanisms. During the first stage, simultaneous processing of psycho-acoustic cues such as fundamental frequency, changes in amplitude and temporal differences begin to decode the incoming sounds into discrete auditory objects. The second stage schema-driven mechanism requires the listener to utilise top down processes such as attention, working memory of a sound source and language to correctly identify the auditory object.

It has been reported that in the presence of competing background noise, those with autism often have difficulties in separating relevant auditory information (i.e., speech and target noises) into discrete auditory objects connected to the different sound sources within the environment [103]. Interestingly, this is despite literature that observes a superior ability in processing the frequency information of pure tones and musical stimuli [92]. These complications in processing sound information in noise have links to a deficit in auditory spatial attention, which is often exhibited as a failure to orient to speech and other social stimuli [104] and being distracted or having trouble functioning in the presence of increased noise [105].

These difficulties in understanding speech in noisy environments for those with autism has been attributed to a reduction in the use of spectral and temporal dips. In addition to auditory scene analysis, retrieving speech signals from background noise can also be supported by exploiting the minute spectral reductions in the masking of target speech frequencies (spectral dips), and reductions in the intensity of background noise (temporal dips) [106]. When a Signal-to-noise ratio (SNR) increases substantially, this allows the listener to successfully distinguish segments of the target speech source. Following this, the higher-order cognitive mechanics such as memory and language can replace the missing information and infer what has been said [83], [106]. In a study that compared the perception of speech amongst background noise with spectral and temporal dips, autistic participants displayed a significantly higher speech reception threshold when compared to TD controls [83]. Suggesting an impaired ability to utilise spectral and temporal dips to decode speech in noise. Similarly, Groen *et al* [107] observed that autistic children performed significantly worse at identifying words in pink noise without temporal dips than in noise containing dips, relative to NT participants. However, recent research by Schelinski and colleagues [108] noted contradictory findings. Despite autistic participants demonstrating lower performance at speech in constant noise recognition, the introduction of temporal gaps into the noise signal did generate a significant improvement. Nonetheless, all three studies indicated that small changes in SNR can significantly change speech recognition in the presence of persistent background noise [83], [107], [108].

It has also been observed that the presence of competing noise can have an impact upon the localisation and spatial attention towards non-speech stimuli. A study conducted by Teder-Salejarvi *et al* [109] presented pink-noise stimuli accompanied by distractor signals from multiple positions across two free-field loudspeaker arrays. Relative to NT controls, autistic adults showed a significantly reduced ability to spatially attend to a singular sound source in the presence of multiple environmental sounds.

2.2.3 Sound Source Localisation

Biological mechanisms within the human hearing system allow healthy listeners to make perceptual inferences about sound sources within a space, which will discussed in detail in Chapter 3. However, there is neurological evidence suggesting that impaired auditory localisation is prevalent within the autistic population. Postmortem analyses of brain tissue of autistic individuals have indicated neurological differences in the size and structure of the medial superior olive, an area of the auditory pathway associated with auditory localisation [110], [111]. Further to this, electroencephalogram recordings of brain activity in autistic people found reduced neural processing of presented sound stimuli within the cortical regions connected to spatial processing and selective attention. This suggests poor sound source localisation and diffuse spatial attention [109].

There has also been research that suggests that autistic people experience significant deficits in processing the binaural cues of Interaural Time Differences (ITD) and Interaural Level Differences (ILD). [103]. These cues are discussed in detail in Chapter 3. This impairment in decoding the binaural cues essential to localisation across the horizontal plane would therefore suggest poor sound localisation. Furthermore, impaired sensitivity to ITD and ILD cues are also consistent with malformations in the medial superior olive [110], [111]. Despite these findings, Visser *et al* [112] found sound localisation was not impaired in autistic participants along the horizontal plane, however they did perform significantly worse than TD controls in vertical localisation.

2.2.4 Orientation towards speech stimuli

During typical development, an infant of 6-weeks can display a significant sensitivity to social stimuli, with particular focus to the features of the human face and speech [113],

[114]. Autistic children on the other hand have been reported to display atypical cortical responses to synthetic speech-like stimuli, culminating in reduced involuntary attention to social auditory events [115]. These early limited reactions to social stimuli represent the core social impairments in the earliest and most basic form and have negative impacts on future social and communication development [116]. Soskey et al. [117] found that compared to TD controls, autistic children showed a significant impairment in attending to sounds, in particular speech sounds.

2.3 Auditory Hypersensitivity

Atypical responses to sensory stimulation have been documented in autistic people since its original descriptions in the 1940s by Kanner [118] and Asperger [119], with particular prevalence being exhibited within the auditory domain [5], [120]–[123]. Often, autistic individuals can display negative emotional or extreme behavioural responses when in the presence of everyday environmental sounds such as vocalisations or self injury through hitting the ears[7]. These reactions to perceptually averse auditory stimuli will have an impact upon the daily lives of the individual, resulting in activity limitations and participation restrictions brought on by sound avoidance behaviours [6]. However, despite the extensive literature that reports these atypical reactions to sound there has been debate over the underlying causes, with both physiological and psycho-emotional investigations.

Auditory hypersensitivity has been the focus of much research aiming to detect some physiological evidence that could demonstrate a difference in auditory processing between autistic people and their NT peers. This has been investigated by measuring auditory thresholds as a means of searching for evidence that supports oversensitive hearing. Audiology examinations conducted by Tharpe *et al* [124] of a group of 22 autistic children demonstrated no differences in auditory brainstem responses, distortion product otoacoustic emissions, or acoustic reflex when compared to NT controls. These findings were also supported by Gravel *et al* [125] who reported no significant differences in auditory thresholds between 37 autistic and control participants. Interestingly, parental reports noted that 6 of the autistic children experienced hypersensitivity to sounds.

Further studies have also tested the tolerance of those on the autistic spectrum towards free-field warble tones. Gomes *et al.* [126], who investigated 46 autistic people, found that only 2 of the 11 participants clinically diagnosed with auditory hypersensitivity displayed negative behaviours towards auditory test stimuli rendered at 90dB and playing between between frequencies of 500Hz to 6000Hz. Suggesting that auditory hypersensitivity behaviours are not the result of the auditory pathway. More recently, an investigation involving 50 autistic children and auditory hypersensitivity measured tolerance of warble tones presented at varying loudness levels [127]. When compared to NT controls, there were no significant differences in the tolerance of sounds up to 100 dBHL, with 68% of the autistic participants being able to tolerate sounds at 110 dBHL. Furthermore, rather than an all-inclusive response to sounds of a particular intensity or frequency, empirical evidence also reports that negative behaviours are activated by specific sounds [6], [7], [65], [128]. With this in mind, researchers have suggested that rather than an auditory based problem, adverse reactions may be a psycho-emotional response caused by impairments in the limbic system [126]–[128]. A mechanism referred to as phonophobia [129].

Phonophobia is described as a pathological fear of a sound [130]. This manifests as an abnormally strong activation within the limbic system without a parallel activation of the auditory system, caused by atypical connections between the two mechanisms. An individual who is displaying phonophobia will exhibit strong sound avoidance behaviours in response to specific auditory stimuli [129]. This can also commonly be experienced by younger NT children, who's overactive auditory system alongside an absence of understanding of the experience can activate the limbic and autonomic nervous systems, resulting in a maladaptive fight or flight response [121], [131], [132]. However, whilst this commonly subsides in NT children by early school age, it can persist in autistic people until adulthood [65].

Reactions to auditory stimuli within the limbic system have been ascribed to the non-classical auditory pathway, otherwise called polysensory, non-specific or extralemniscal pathways [133]. Rather than transmitting sensory information from the cochlea to the auditory cortex via the inferior colliculus, resulting in highly processed information controlled by intrinsic brain activity. The non-classical pathway diverts non-processed information at the level of the lateral lemniscus towards the thalamus, which is linked to the auditory association cortex. These association areas of the brain project to the limbic system, including the amygdala, which is involved with emotional memories and responsiveness.

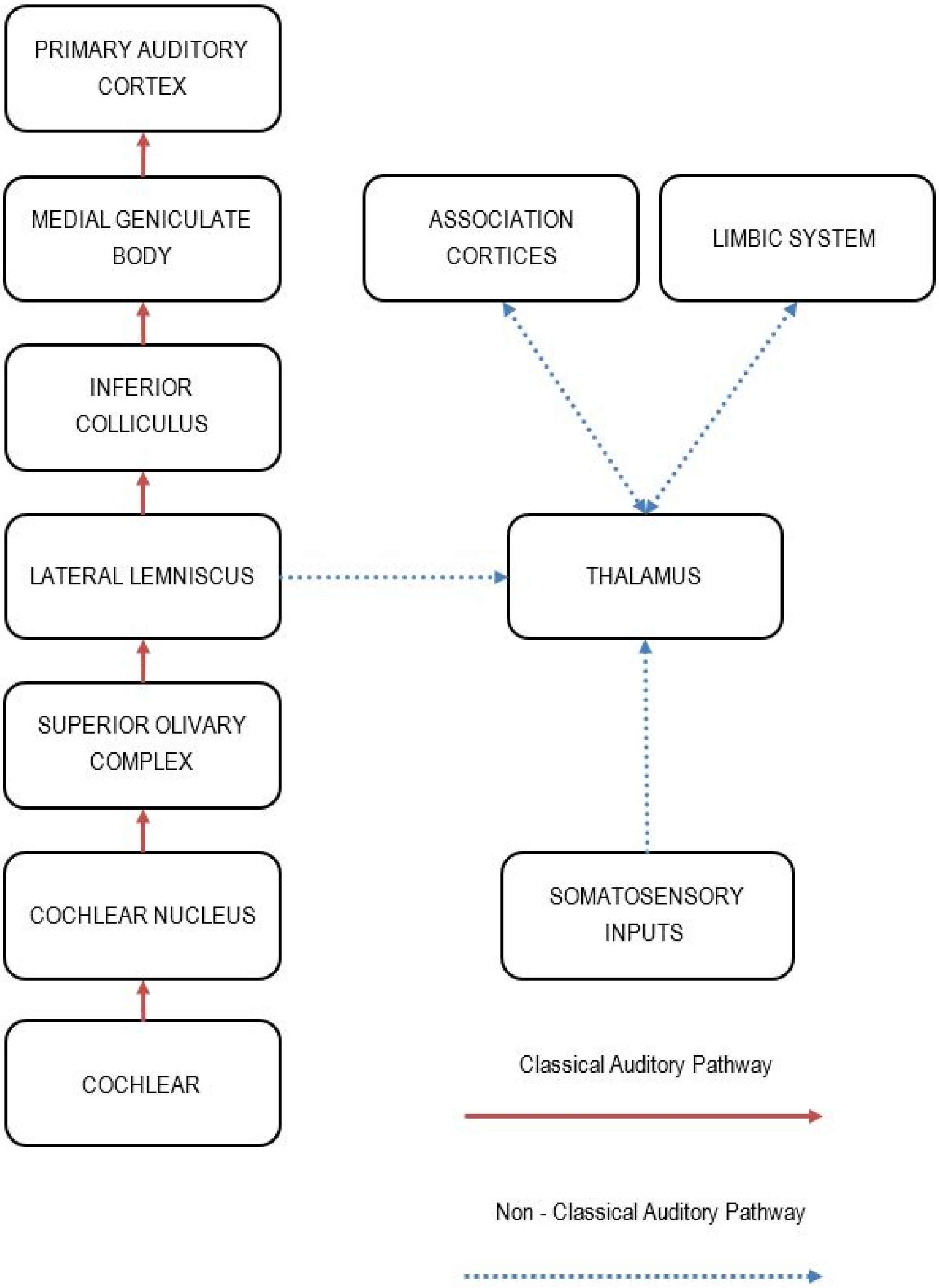


FIGURE 2.1: Diagram of the non-classical auditory pathways, showing communication across the pathways between the auditory system, limbic system and to the auditory cortex. Diagram adapted from [134].

Furthermore, the non-classical pathway also receives information from the somoatosensory systems [134], [135].

A study by Moller and Rollins [136], investigated the link between the somatosensory systems and the non classical pathway. An increased perception of loudness could be induced through electrical stimulation of the median nerve only in NT children up to the age of 8 years old. After this age it was concluded that the classical auditory pathway takes over. In further research, Moller *et al* [133] discovered that the same increased perception of loudness was identified in autistic adults up to 38-years-old. This was recently supported by a magnetic brain imaging study by Shen *et al* [137], which found abnormal amygdala connectivity resulting in sensory hypersensitivity.

An atypical reaction to particular stimuli may also be due to a reduced top-down control of the amygdala from the prefrontal cortex and the hippocampus, which has been linked to anxiety-related disorders such as Post-Traumatic Stress Disorder (PTSD) and social anxiety disorder [138]. Neurological studies have reported evidence of reduced reciprocal connectivity between the pre-frontal cortex and the limbic system in autistic people. This lack of cortical interaction could therefore impair complex information processing from emotional systems, resulting in impairments in emotional regulation [139]–[142]. Further to this, problematic emotional responses to stimuli and situations are well documented in autism research with empirical evidence reporting reduced emotional functioning [143]–[147]. By taking this into consideration, Stiegler *et al* [6] suggests that auditory hypersensitivity could be caused by misophonia. A sound processing impairment in which common sounds can trigger strong negative emotions such as anxiety or anger, not necessarily a phobic reaction [148]. If an autistic person is angered or annoyed by a sound, an intense or extreme emotional reaction could be misinterpreted as dysfunctional auditory pathway.

2.4 Therapy techniques for auditory hypersensitivity

The literature presented in Section 2.3 indicates that rather than physiological differences in the auditory system, auditory hypersensitivity displayed by those on the autistic spectrum is a psycho-emotional response to sounds. If left untreated individuals will continue to display negative emotional reactions and sound avoidance behaviours which can have a significant effect on the quality of life of themselves and family [7]. Therefore, it is suggested that early interventions are crucial in order to reduce any impact [6]. Intervention methods for addressing auditory hypersensitivity are varied in approach and background, mirroring the mixed views of possible causes. However, the most successful approaches aim to habituate an individual to sounds that would normally be considered abhorrent. This section will discuss the varying approaches to therapy for this issue.

2.4.1 Ear protection and noise cancelling headphones

For autistic children who experience auditory hypersensitivity one suggested intervention involves sound isolation using ear protection or noise cancelling headphones [149]. By reducing the level of auditory stimulation experienced by the user, this intervention could be an effective temporary approach to managing problematic auditory stimuli in everyday situations [149]. There is however, a lack of research examining the degree to which ear protection can reduce the associative behaviours of auditory hypersensitivity. A study by Ikuta *et al* [150] included 17 autistic children who exhibited auditory hypersensitivity and associated negative behaviours. The investigation was comprised of two sequenced two week test periods in which subjects would wear PELTOR Optime 1 Earmuffs followed by SONY Digital MDR-NC500D Noise Cancelling Headphones. All participants were permitted to choose to wear ear protection either full time or part time. In order to evaluate any changes in behaviour resulting from either form of ear protection, Goal Attainment Scaling (GAS) was used. GAS is a standardised means of assessing functional goal attainment of a patient over a certain period of time [151], and has been used to evaluate a number of sensory interventions for children with autism [4]. A rating scale of -2 to +2 is used to rate goal achievement. Zero represents no change in performance, with negative numbers given for less-than-expected outcomes, and positive representing greater-than-expected outcomes [152].

In the case of the study conducted by Ikuta *et al* [150], GAS scores were taken to gauge a child's reaction to auditory stimuli during the baseline control period, with comparative measurements recorded throughout both ear protection periods. For noise cancelling headphones, results showed no significant differences in auditory reactive behaviour. However, GAS T-scores taken during the earmuff use showed a reduction in behavioural problems related to sound hypersensitivity. These results are to be expected, the earmuffs used in this experiment are manufactured for industrial purpose and attenuate frequency content across the spectrum (see Table 2.1). Therefore, by considerably reducing the volume of environmental auditory stimuli, an individual will not experience sounds that they find

Table 2.1: Attenuation Values of 3M PELTOR Optime 1 Earmuffs **3m**.

Frequency (Hz)	**63**	**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
Mean Attenuation (dB)	14.1	11.6	18.7	27.5	32.9	33.6	36.1	35.8

intolerable. This will provide an effective and immediate intervention for individuals with serious cases of auditory hypersensitivity.

In more recent research, Pfeiffer *et al* [153] investigated the use of noise cancelling headphones in reducing skin conductance levels and frequency of non-specific conductance response as a result of high levels of hypersensitivities to sound. A small group of autistic participants wore both in-ear and over-ear noise cancelling headphones over an experimental period integrated into their everyday lives. Results from the study indicated that both in-ear and over-ear equipment led to a significant decrease in anxiety related physiological responses when presented with problematic auditory stimuli. However, the use of sound isolators and noise cancelling headphones should be carefully considered when recommending them for individuals with hypersensitivity. By creating a blanket attenuation of environmental sounds, important speech sounds are also affected. As discussed in Section 2.2, autistic individuals have difficulties in processing speech, especially in noisy environments [83]. Ear protection presents an additional obstacle to verbal communication, and could in consequence contribute to social and communicative impairments. Furthermore, it has been warned that ear protection should be considered a temporary intervention, as overuse has been found to increase auditory hypersensitivity by decreasing opportunities for habituation [6], [129].

2.4.2 Auditory integration therapy

Although there is increasing research in both neurological and psychology literature that supports emotional neural mechanisms being involved in oversensitivity to sounds, the topic still remains unclear and greater understanding is needed. However, interventions based upon possible physiological malformations present in the autistic auditory canal exist. Auditory Integration Training (AIT), a therapy first developed by Berard in 1982 [154], is built upon the theory that individuals have a more acute perception of particular frequencies over others, associating the distortions in hearing with cognitive and behavioural problems in autism, dyslexia and depression. The intervention begins with an audiogram which identifies the certain frequencies that elicit atypical reactions. These results are used to psychoacoustically modify classical music, filtering out and removing problematic frequencies. Music is usually listened to over headphones during two half-hour daily sessions over 10 days [155]. During this time, the audio filtering is used to dampen the peak frequencies that the individual is hypersensitive too. Presenting modulated sound that aims to re-educate the hearing system by comprehensively “massaging” the ossicles, eardrum and cochlea [154]. Therefore, correcting any hypothesised distortions in the auditory canal [154]. AIT has gained support in literature [5], [121], [156], however these results are either anecdotal or from small sample groups. In contrast there is evidence from controlled randomised trials [157], [158] and a systematic review [6] that AIT is an ineffective intervention for auditory hypersensitivity within the autistic population. Due to this, it remains unsupported by key organisations such the American Academy of Audiology, the American Speech-Language-Hearing Association, and the American Academy of Paediatrics, who have categorised AIT as a investigational non evidence-based technique [6].

2.4.3 Cognitive behavioural therapy

With evidence suggesting that auditory hypersensitivity within the ASD population is caused by pycho-emotional problems rather than specific physiological manifestations within the auditory system, interventions have been developed which focus on the neural mechanisms involved with emotional processing, emotional memory and emotional reactions. Techniques such as systematic desensitisation can reduce pathological fear and its related emotions by exposing the patient to the feared or conditioned stimulus whilst in the presence of relaxation or play activity. However, this does not alter any existing pathological structures, rather it forms completely new and competing structures that contain no pathological associations [159], [160].

This approach has been demonstrated to reduce fear in children with ASD who have specific phobias such as eating, dogs and visits to the dentist [161]–[163]. Importantly, desensitisation has been used to treat the hypersensitivity of autistic children to specific environmental auditory stimuli. An early example of which is a study by Jackson & King [71] in which a desensitisation process combined with positive reinforcement was used to reduce phobic responses to the sound of a toilet flushing in a 4-year-old autistic boy. An In-Vivo exposure desensitisation technique was used in which the child would be tickled under the arm during using the toilet and flushing. If no negative behaviours were displayed the child would receive verbal praise and an edible. Following a 15 day program the phobic reactions to the toilet flushing were eliminated, these results were maintained at three and six month follow ups.

Another example of desensitisation successfully used to reduce auditory hypersensitivity in autistic children was conducted by Koegal *et al* [7]. The investigation involved a small group of three participants who exhibited extreme aversion to sounds including toilet flushing, household appliances and animal noises. All displayed extreme behaviours such as sound avoidance, covering the ears and screaming. The systematic desensitisation process involved hierarchical exposure towards both real-world stimuli and sounds played through speakers during play activities. At each stage of the intervention the sound source would move closer to the child. After each hierarchical step was completed observers judged to see if the child felt comfortable with the stimulus. When compared to baseline measurements all participants showed a considerable decrease in anxiety levels and adverse reactions to auditory stimuli following the intervention period. This resulted in the children being able to tolerate the sounds and display some positive behaviours, for example laughing and smiling. Another important outcome from this study is that the extinction of negative behaviour was maintained at a 34 week follow-up measurement.

Passive music listening has also been used as a tool for auditory desensitisation for when the conditioned stimulus is broader than a specific sound. In a study by Steigler and Ruhlin [65], a 26 year old male diagnosed with autism presented hypersensitive symptoms towards low pitched voices, describing hearing them as uncomfortable and having a similar effect to fingernails on a chalkboard. As well as the apparent discomfort, the participant also displayed a concomitant vocal disorder which manifested as a high vocal pitch, a self-regulatory behaviour to avoid the lower frequencies of his own voice. For the first six months of the intervention the subject listened to instrumental music he found enjoyable for 30-60 minutes per day. Bass frequencies were initially reduced to a just audible level before being progressively increased across the subsequent weeks until reaching an amplitude perceptually consistent with the middle and higher frequencies. However, for the second six month period the music was substituted for recordings of thunder storms and male singing voices. This was due to the participant's self-reported pessimism in relation to the prognosis and treatment, resulting in him requesting a change of stimuli. Directly following the intervention, the participant was able to lower the pitch of his own voice as he became more comfortable with particular voices and musical notes. Unfortunately these improvements were noted to be impermanent and failed to generalise beyond the experimental conditions. The clinicians surmised that this was due to the participant's reluctance to accept prognosis and goals of the intervention, negatively impacting the outcomes of the investigation.

Similar to the behavioural desensitisation techniques used in vivo exposure, sound-based interventions use modified music and environmental sounds which have a limited relationship to the autistic people, specific audio over-sensitivity. One particular intervention, The Listening Program (TLP) [164], claims to lessen the negative hypersensitivity related behaviours through psychoacoustically altered classical music via audio filtering, whilst also delivering additional therapy to key functional impairments such as social-emotional regulation, language, creativity and stress response [165], [166]. The individual therapy sessions span a total of 20 weeks and require passive listening through specially designed headphones that allow bone conduction. The modulation of low (0 - 750 Hz), mid-range (750-4000 Hz) and high sound wave frequencies (4000 - 8000 Hz), gradually exposes and habituates participants to frequencies that they may find abhorrent [167]. Although there is reported success of the TLP in peer-reviewed articles [168], [169], the study groups remain small. Gee at al. [167] conducted a case control study with three autistic children aged 5-10 years old who exhibited sound hypersensitivity, in order to report any improvements in behaviour and sensory tolerance following the TLP intervention. Each participant listened to 15 minute sessions of the modulated classical music twice a day, five days per week for ten weeks. Measurements taken included the SenSOR scale which rates the over-responsivness of a child to auditory, tactile, vestibular, visual, proprioceptive and gustatory stimuli [170]. As well as improvements on the SenSOR scale for all three case studies, results showed a reduction in sensory processing difficulties and negative self-regulatory behaviours [167].

Innovative multi-modal user interfaces are also offering treatment for sensory processing disorders via Sensory Integration Therapy (SIT). Through play and sensory rich activities, SIT increases the child's ability to integrate sensory information resulting in more adaptive and organised responses to stimuli and environmental challenges [171], [172]. Traditionally

SIT takes part in clinic settings such as a specially designed Snoezelen, multi-sensory environments are physical spaces which house a variety of tools to deliver audio, visual and tactile stimuli such as ball pits and fibre optics string [173]. However, tangible multi-modal interfaces are enhancing musical and sound therapies by augmenting sensory integration therapy through providing added visual and auditory stimuli [174]. *SensoryPaint* [175], an interactive painting tool, administers a multi-sensory experience by allowing users to freely paint digital artwork on a projected surface with complementary sounds which react with the user's movement. The system was evaluated within a laboratory setting with a total of 15 autistic children aged between 10 and 14, with intervention spanning five sessions lasting a maximum of 20 minutes. Results reported improvements in socialisation, motor functions and engagement. Additionally, parents and therapists reported positive impacts in how children reacted to not only auditory inputs, but also visual and somatosensory stimuli.

The previous sections have explored the atypical processing of sound experience by people of the autistic spectrum, with particular focus upon auditory hypersensitivity. Although the cause of negative responses to specific environmental sounds is not fully understood, it is believed that auditory hypersensitivity is rooted in psycho-emotional reactions to sound [127]. Therefore it is believed that the most effective approach is to utilise CBT with the goal of habituation [7]. The following sections will explore how interactive technologies can assist in delivering such interventions, creating engaging therapy environments for autistic young people.

2.5 Computer assisted cognitive behavioural therapy for ASD

Since the early 2000s there has been an exponential growth in the use of Computer Assisted Interventions (CAI) supporting research and the development of children with ASD in social, communicative, language, and educational domains [176], [177]. This has in part been due to technological developments in computer based multimedia technology such as mobile devices and motion capture sensors (e.g. Microsoft Kinect), which has brought advanced and novel approaches to computer interaction into the public domain, moving far beyond the traditional desktop computer.

For an autistic individual that might find the world confusing and unpredictable, computers are free from social constraints and provide an experience that delivers consistent and predictable responses that contribute to maintain interest and increase motivation levels in autistic children [178]. In addition, CAIs can construct a safe environment through graphically displayed cues that promote multi-sensory exploration, which can be modified and personalised to maximise experience towards the individual's specific needs [176], [179]. Finally, computer based environments have been shown to reduce anxiety for autistic people by limiting the multi-sensory distractions that occur in the real world, increasing independence, social interaction and enjoyment [180].

Most existing technology developed for autistic children is a combination of game play and an intervention framework [181]–[183]. For neurotypical children, play activities form an important part in language and social development. But for autistic children, play is heavily affected by cognitive and emotional impairments, thus limiting its associated learning processes [184]. Through technology-driven therapies which offer new forms of interaction via motion capture, children on the autistic spectrum can explore the relationship between physical activities and cognitive mechanisms which stem from sensorimotor interactions with a physical environment [185]. These exploratory interventions have had reported success in helping autistic children learn social cues [186], learn real world scenarios (e.g. crossing the street) [187] and aid in motor skill development [188]. However, there are limited investigations into computer technology's application to CBT for autism.

There is extensive literature that supports the use of CBT as an effective intervention to reduce anxieties in autism [22]. However, there are arguments based on the core symptoms of autism that would suggest that traditional CBT techniques are not appropriate for individuals within this population. Firstly, CBT sessions often involve face-to-face verbal exchanges. These aim to teach the child to not only become knowingly aware of their pathological associations with a conditioned stimulus, but also challenges them to develop new positive associations [189]. This would be challenging for a child with cognitive and social impairments, resulting in diminished motivation and engagement in traditional CBT [190]. Secondly, CBT techniques that use imaginal exposure are not contextualised and do not fully characterise the medium in which the patient experiences their anxiety. This is also problematic for ASD, as they need frequent practice and contextualised exposure to increase chances of real life generalisation [191]. Another important consideration of

Table 2.2: Benefits of computers as tools in psychotherapy [194].

Computers in psychotherapy:
Reduced cost of treatment
Improve access to psychotherapy
Promote engagement in the therapy process
Give systematic feedback to user
Promote self-monitoring
Rehearse coping skills
Encourage self-help
Provide psychoeducation
Store, analyse, and display data
Provide built-in outcome measures
Function reliably without fatigue

current CBT is a lack of accessibility. A recent systematic review by Ince *et al.* [192] has noted that the rates of implementation for cognitive behavioural therapy are below the recommended levels for the United Kingdom. This gap in delivering successful mental health interventions has stemmed from a number of influences including, lack of resources, limited dedicated therapy time, and a lack of specialist training [193].

Computer technology can offer new approaches to therapy and provide solutions to barriers associated with the provision of mental health care. By addressing these possible impairments, computer based therapies could therefore increase the amount of patients gaining access to successful treatment. However, the reasoning for computer assisted therapy extends further than positive impacts on health care resources and accessibility (see Table 2.2).

The present chapter will discuss two approaches to providing computer driven CBT to autistic people, VR and serious games, and provide examples of existing interventions that target auditory hypersensitivity. It will be argued that the current technological state of these computerised interventions could administer a realistic, engaging, safe and repeatable intervention, that is both accessible and cost effective.

2.5.1 Serious games

Unlike conventional computer games which are developed for the purpose of entertainment, serious games are designed with integrated educational goals delivered through in-game mechanics that are proven to support knowledge acquisition and support generalisation of skills [75]. By adopting traditional computer game motivational features such as point rewards and progression systems, they can encourage learning of target skills that the user can either find difficult or unrewarding [195] (eg. tackling medication adherence in adolescents with cancer through behavioural training [76]). In addition to making use of the fundamentals of video game design to produce an immersive and entertaining learning environment, in-game features are also based upon empirical research that recommends that learning outcomes are best achieved when they take place in a context relevant to the user[196], [197].

Table 2.3: Differences between entertainment games and serious games [198]

	Serious Games	**Entertainment Games**
Task vs. rich experience	Problem solving in focus	Rich experiences preferred
Focus	Defined learning or theraputic outcomes	To have fun
Game Mechanics	Challenge may be limited to maintain motivation	Winning and losing is part of the experience
Virtual Environments	Realistic virtual environments	Simplified, realistic or hyper-realistic environments

Within treatment for autism, serious games have been shown to improve emotion and facial recognition capabilities, language and social skills and teach literacy [199]–[201]. This is in part due to a preference to technology based therapy within the ASD population. Furthermore, the individualised virtual environment allows users to acquire new skills and repeatedly practice them within a safe medium, away from the social and emotional contexts they may find difficult [73], [202]. But before it is discussed how serious games can be used in cognitive behaviour interventions for ASD, it is important to understand what serious games are and how their unique design elements and game mechanisms can be targeted to address key impairments within the autistic population.

Further to enhancing education frameworks, serious video games have had noted success in providing an alternative approach to therapy. For example, Arshia *et al.* [203] developed a computer game that uses digital story telling to teach autistic children how to use money whilst simultaneously employing a therapy protocol to address social impairments. Another game by Barajas *et al.* [204] combined both a 3D onscreen Graphical User Interface (GUI), with a Tangible User Interface (TUI) consisting of Lego like blocks augmented with electronic components. This created an interesting integration between the virtual and physical world and was designed to develop the social and cognitive skills of 9 autistic children. The project was evaluated based on social interaction between participants and compared the computer game with a non-computer game comprised of just the TUI. Results showed that the computer based game encouraged more social interactions and collaborative play. Furthermore, the interactivity and visual feedback of the GUI helped the users follow instructions which significantly improved cognitive performance as opposed to the non-computer game.

To date there is limited research investigating CBT interventions delivered through serious games for autistic children. There is an ongoing study by Wijnhoven *et al.* [189] which investigates the effects of an serious game named *MindLight* in decreasing clinical anxiety symptoms in children with ASD, however, the results are not yet published. *MindLight* is grounded in the principles of cognitive-behavioural therapy [205]. During play, users are gradually exposed to 'threatening' cues through structured auditory and visual sensory information, whilst certain game mechanics are based on Alpha and Beta brainwave readings from an EEG-headset. These mechanics aim to teach participants to control anxiety through focused relaxation and concentration.

An important therapy mechanism employed by *MindLight* is Attention Bias Modification (ABM) [189]. This relatively new form of therapy is based on evidence that suggests that individuals with an anxiety disorder have an implicit bias in attention towards a stimulus that they experience as threatening [206], [207]. ABM therapy aims to correct these cognitive biases towards threat related attention through computer based treatment. This is achieved by directing attention away from threatening cues and focusing on the positive features of the virtual environment or by performing tasks [208], [209]. This differs from CBT which exposes patients to feared stimulus so that they can learn and understand the feared stimulus as safe (see Section 2.4.3). As CBT and ABM target different cognitive features of anxiety, it is believed that attention bias modification could augment existing cognitive behavioural therapy frameworks [209].

In the case of *MindLight*, children focus attention on specially designed puzzles, learning to concentrate and respond to positive stimuli. Therefore, distracting themselves from negative stimuli. This is also scaffolded with CBT exposure techniques controlled by Alpha waves. When the child repeatedly encounters threatening stimuli they learn to habituate themselves to them. The recorded Alpha waves are used to measure relaxation which awards points and reduces the virtual threats.

At this time there is only one published study examining this area. *Sinbad and the Magic Cure* [29] is a mobile game, in which during game play children were exposed to sounds that were found to be particularly disturbing to those diagnosed with autism. Audio files were chosen based on an auditory questionnaire completed by ten male children diagnosed with ASD. The game was evaluated using a small cohort of 7 autistic participants aged between 8-11 who experience hypersensitivity to sound. The intervention involved playing the game over multiple sessions over a course of seven days, with their responsiveness to the sounds measured daily through an auditory questionnaire. Although testing spanned a shorter time than most desensitisation methods, the results showed that the children started to develop a tolerance to the sounds. Despite its success there is no mention of the extreme self regulatory behaviours that are associated with auditory hypersensitivity, such as vocalisations or hitting the ears. Instead, a negative response was represented by a series of graphical faces on a scale from angry to happy. This suggests that the participants did not experience the same level of hypersensitivity as the subjects in the study by Koegal *et al* [7]. Furthermore, unlike the sounds used by Koegal *et al.* and Steigler and Ruhlin [65], the sounds are not specific to the individual and a more generalised stimulus would possibly also yield results pointing towards an improved tolerance to the auditory stimulus. Nonetheless, the evaluation of *Sinbad and the Magic Cure* does warrant future examination into how serious computer games can be developed to tackle auditory hypersensitivity.

2.5.2 Virtual Reality

VR is accurately described as a ”computer-mediated simulation that is three-dimensional, multi-sensory, and interactive, so that the user’s experience is “as if” they are inhabiting and acting within an external environment” [35]. In terms of technology the modes in which VR can be delivered are broad, using Collaborative Virtual Environment (CVE), HMDs or desktop computer displays. All of which aim to create an accurate simulation of the physical world in order to promote sensory and behavioural responses comparable to that of real-world experiences [210]. Outside of the realms of entertainment, this has therefore given rise to a wide range of applications for psychological interventions including eating disorders [211], attention-deficit disorders[212], specific phobias [213] and autism [214].

VR has been an active area of research in the development of interventions for the autistic population for over two decades. Empirical studies have recognised VR as an important therapy tool for addressing the core symptoms associated with ASD such as difficulties in social interaction [215] and communication [216]. In addition, the dynamic and controllable environments have been utilised for educational and training practices such as disaster awareness [28], driving [217] and independent functioning [218] training to take place for autistic individuals.

Early work conducted by Strickland [80] in 1996 suggested key aspects of this technology which can be applied to effective and engaging therapeutic solutions for both autistic children and adults, and despite extensive advancements in how VR is delivered these can still be applicative to contemporary research practices. Firstly, many autistic people respond to and enjoy using VR technology [219]. This translates to users actively engaging with interventions, with the enjoyment providing motivation to progress with the therapy program [220]. Secondly, virtual environments provide an experience with consistent and predictable interactions and responses [190]. This can be in the form of limiting social and communication pressures that can often cause anxiety in real world situations [178]. Furthermore, people on the autistic spectrum can frequently experience complications in processing sensory information, with particular prevalence in the auditory and visual domains. This atypical sensory receptiveness can often be a catalyst for a deterioration in behaviour observed as aggressive or autonomic fear responses [6]. VR can control all audio and visual stimuli that are encountered during an intervention session, simplifying complex sensory arrays and thus reducing the elicitation of anxiety [80], [180]. The control of VEs and interactions within them can also construct important opportunities for individualised treatment [80]. Re-adjusting the sensory and learning parameters of the intervention to compensate for the complex needs of the individual [219]. Finally, VR technology can replicate realistic representations of real-world situations in which participants can learn and practice new behaviours and real-life skills. The increased feeling of presence and ecological validity as a result of this realism can have a positive impact on the generalisation of newly acquired skills from the virtual intervention to their real-world applications [219]. However, the visual experience of the user is often the focus of research when investigating

VR interventions for autism. Although what the user sees can increase generalisation via realistic visual similarities with the real world, the sense of realism is reliant on the extent and fidelity of the VR technology delivering the multi-sensory stimuli to the user, including audio.

Similar to serious games, VR has the facility to control audio and visual stimuli and deliver a consistent error free learning environment [221]. However, a significant difference between VR and serious games is its specific ability to deliver immersive three-dimensional virtual surroundings based on either real life or simulated environments. These realistic VEs can be interpreted by young autistic people as real, making decisions based on the similarity between the virtual and real world. This rationalisation is further demonstrated in studies that observe generalisation of skills and information learned with VEs to real world context [222]–[224].

2.5.3 Virtual Reality Exposure Therapy

Outside of autism research, there is extensive literature that supports the use of VR as a tool for delivering exposure based therapy for phobias treatments such as fear of heights [31], agoraphobia [32], and combat-related post traumatic stress disorder [33]. Virtual Reality Exposure Therapy (VRET), sometimes known as In-Virtuo Exposure, makes the most of the technology outlined above to provide precise and safe control over the feared stimulus or situation. One notable example would be in the treatment of a fear of heights, in which a therapist can place the user at different heights whilst removing any chance of them falling. In addition, the same situation can be repeated, offering multiple chances to practise until any avoidance behaviour has become extinct.

VRET relies upon the assumption that virtual environments are able to evoke feelings of anxiety and in turn, through graded exposure, provide the opportunity for the patient to become habituated towards a feared stimulus [225]. This approach to CBT provides a viable alternative to In-Vivo exposure exposure therapy, but with the added benefits of computer driven technology outlined in Table 2.2. In addition, a meta analysis review by Powers *et al* [226]. which included studies on specific phobias, PTSD, and social and panic disorder, concluded In-Vivo exposure exposure interventions were not significantly more effective than VRET.

One crucial advantage to using VR technology such as HMDs of CVEs, is the high levels of presence felt by the user. Presence, defined as a psychological state or subjective perception of a mediated environment as real [72], is regarded as the mechanism responsible for eliciting emotions such as anxiety in response to a virtual stimulus [213], [227], [228]. For the most part, an individual is able to correctly identify when they are using VR technology. Yet presence caused by immersion will cause them to overlook the knowledge that technology is involved with the experience [229]. There is however, contradicting results in studies investigating the relationship between presence and emotions. Some literature has observed a significant positive correlation [228], [230], whilst some have found a negative correlation [231], [232]. Ling *et al.*'s meta analysis [233] which identified 33 studies, explored the relationship between self reported presence and anxiety within immersive virtual environments. They concluded a positive interrelation between presence experienced within VR and anxiety. In addition, it was found that particular phobias were affected more than others, such as fear of animals and flying. However, the relationship was not as significant in social anxiety disorders. It is important to note that these investigations into presence in VR involved neurotypical study groups. Despite the recognition of VR in the treatment and education of ASD, there is surprisingly limited studies that explore presence in virtual environments in autism. Investigations conducted by Wallace *et al* have compared self-reported presence between IQ and age-matched autistic and TD participants, and found no significant differences in reported feelings of spatial presence and ecological validity [234], [235]. These results are important as they show that young autistic people feel spatially present within virtual environments and will engage with what they believe is a realistic experience [235]. In addition, Newbutt *et al* [219] investigated the sense of presence felt in VR in a group of autistic adults whilst wearing a head-mounted display. Observations from this study recorded high levels of self-reported presence, with the authors suggesting this led to the participants having virtual experiences close to those in the real world. Taking these studies into account, it could be considered that VR exposure interventions will have similar levels of success for both autistic people and their neuro-typical peers.

Maskey *et al.* [34] developed a treatment for phobias in autistic children that combined VR environments and CBT. Unlike the research previously mentioned which used HMDs, this project made use of an immersive environment named *The Blue Room*, comprised of

4.1 surround sound and visual images projected over walls and ceilings of a 360 degree seamless screened room. This allowed the participants to freely move and interact with the therapy session. An initial study in 2014 included nine male autistic participants aged 7–13 years old, each with a specific phobia such as shopping, pigeons or busy roads. Each participant received four 20–30 minute sessions within the virtual environment in which a therapist would gradually increase exposure to the virtual representation of the feared stimulus. For example, treatment for the shopping experience phobia revolved around increasing the length of the verbal exchange with the virtual shopping assistant. However, this increase would only occur at a rate in which the child could remain relaxed and comfortable. The children were encouraged to use breathing and stretching exercises as a method of negating any feelings of anxiety associated with their specific phobia. After each session the therapist and parent would then discuss any progress made, and based on this goals would be set that would aim to gradually increase exposure to the specific fear in real world contexts.

FIGURE 2.2: Young person with ASD and therapist in the Blue Room [34].

Throughout the evaluation of the *The Blue Room* participant anxiety was measured using the *Spence Children's Anxiety Scale-parent (SCAS–P) version and child version (SCAS-C)*. Questionnaires that were developed as a means to assess the anxiety symptoms in neurotypical children [236]. Measurements recorded how the autistic children reacted to feared stimulus pre-treatment and at regular intervals post-treatment up to 12–16 months. Seven out of the nine children showed improvements of confidence in approaching the target phobia (both virtually and in real life) by the end of the fourth session. However, the more significant results arose from SCAS scores taken from the 12–16 month post treatment assessments. These results indicated that the majority of participants had the functional ability to cope with previously feared situations or stimuli [34]. The same system was evaluated in a randomised control trial in 2019 [237], which involved 32 autistic young people. Results showed that the treatment group displayed statistically significant improvement on target behaviour ratings when comparing baseline measurements to both the two-week and 6-month follow up sessions. In addition, one-third of the group generalised their confidence in dealing with their specific phobia to real-world contexts.

At this time there is limited literature that explores VRET for reducing phobias using HMD technology within the autistic population. An interesting study conducted by Meindl *et al* [238] aimed to use VR to reduce extreme aversion to blood taking in a 26-year-old autistic male. Prior to the investigation, needle based blood draws required the participant to be physically restrained by five or more adults. The presented intervention was delivered using a low-cost HMD mounted onto a mobile phone, during the experimental sessions the participant was shown a 360° point of view video capture of his father completing a blood draw procedure. Following the experimental period the participant could successfully have blood taken, with this generalised behaviour being maintained at the 1-month follow up session. Although not used to target a specific phobia, the same low-cost technology has been used to expose autistic children to virtual airport environments in order to reduce associated negative and disruptive behaviours [239]. The small group study showed that after three VR sessions and one In-Vivo exposure airport session, participants showed an improvement in behaviours based on pre-post measurements of parental questionnaires and clinical observations.

2.6 Discussion

The complexity of autism means that alongside core impairments, those diagnosed with the disorder can often display extreme and adverse reactions to sensory stimuli, particularly sound [1]. There is still debate over the potential reasons behind auditory hypersensitivity in ASD, with research presenting evidence for either problems within the auditory system [240] or an emotional based response [134]. The auditory system processes the frequency, intensity and temporal properties of sound. Therefore if hypersensitivity is based on a dysfunctional auditory system, autistic people would react to all loud sounds. Research by Lucker [128] has found this not to be the case. In fact, literature provided in this chapter suggests that adverse reactions to particular sounds are based on negative emotional associations. Moreover, atypical neural connections between the auditory system and limbic system can contribute further to negative emotional reactions [134].

The evaluation of therapy techniques is also helpful when considering the underlying roots of hypersensitivity to sound. Auditory Integration Therapy [156] which aims to address potential malformations in the auditory system by massaging the ossicles, eardrum and cochlea has been observed as an ineffective intervention for sound sensitivity. However, successful treatments aim to alter the negative associations with a particular sound in the hope of creating a more neutral or positive response. A reduction in fear based responses can be achieved with cognitive behavioural therapy sessions that desensitise the individual through systematic exposure to problematic sounds during positive experiences such as play [7], [65].

Early intervention is necessary in order to take advantage of brain plasticity [12]. However, it has been suggested that autistic people may not respond positively to traditional CBT techniques. It is believed that impaired social, communication and cognitive abilities may have negative effects on how an individual would knowingly become aware of their pathological associations and then challenge them to form new positive relationships. These difficulties may result in a decrease in engagement and a lack of motivation for therapy [189].

Computer assisted CBT in the form of serious games or VR have had positive results in addressing anxiety and phobias within the autistic population. Applications such as *Mind Light* [189] and *The Blue Room* [34] provide an engaging, safe and consistent environment in which users can practise newly acquired skills at a pace specific to their own complex needs. Furthermore, realistic interactive virtual environments contextualise exposure to stimuli which increase the chances of generalised habituation to post clinic situations [189].

When evaluating systematic computer driven desensitisation for audio hypersensitivity such as *Sinbad and the Magic Cure* [74], it is apparent that there is a lack of contextualised exposure, with negative sounds being triggered during a platformer game. Sound experienced in everyday life is often dynamic, moving between varying locations of the environment and frequently occurring without its visual counterpart [241]. Headphone based 3D audio (see Chapter 3) is a rendering technique capable of recreating a listening experience closest to what is experienced in real life. It simulates the psychoacoustic properties of sound wave propagation into the ear, resulting in the sound source location appearing outside the observer's head across the horizontal and vertical planes [242].

A combination of serious game driven CBT and realistic soundscapes created by spatial audio production techniques could be a viable solution to sound hypersensitivity interventions. Firstly, much like their NT peers, people on the autistic spectrum enjoy playing computer games. Creating a fun and interactive environment would have a positive impact on engagement and motivation [199]. Secondly, by introducing therapy mechanisms such as attention bias modification, an individual can learn to become habituated to an adverse sound by focusing attention on more positive aspects of the environment [189]. Finally, the sense of presence and immersion experienced in VR aid in provoking the feelings of anxiety necessary for successful VRET [213]. Although there is a small amount of literature that links spatial audio to the provocation of fear responses [243], [244], there are investigations supporting its influence on presence [241], [245], [246]. This warrants future examination into the application of spatial audio used to tackle auditory hypersensitivity for autistic people.

2.6.1 Additional research considerations

A number of additional research suggestions can be taken from the evidence provided in this review. Firstly, there is a strong need for investigations into new therapy technologies which are based on larger and more representative participant samples. The majority of studies in this chapter are centred around either small groups or case studies. This may be due to a lack of viable participants who meet the research criteria or lack of facilities. However, developments in consumer technology and the potential for strategies to be embedded in the child's natural environment, could possibly open up research to a much larger population. This could also present the opportunity to examine the effects of specific design features via randomised control group comparisons. By conducting these studies there is the opportunity to reach an optimal calibration in not only technology but also intervention protocols.

With the dynamic development of new consumer VR technology there should be continued research into any possible applications for CBT therapy for not only autistic people but also the wider profound and multiple learning difficulties population. By investigating key design features, clinicians and researchers can design and make critical suggestions for future systems and integrated frameworks. The inclusion of new technology in ASD interventions will require a level of technical or programming expertise that clinicians may not possess. With this in mind, collaborative and multidisciplinary research between engineers, programmers and clinicians is essential in developing or modifying therapy tools.

An inclusive and collaborative design process should not be limited to clinicians and engineers. A Participatory Design (PD) process that includes children with physical or mental impairments can provide a greater understanding and insight into their complex needs and preferences [247]. Commonly, children's contributions are limited to aesthetics or small non-therapeutic aspects of an intervention [248]. However, projects such as *Lands of Fog* [249] aim to provide a deeper influence on its design elements and mechanics through PD. Five sessions were conducted in which autistic children provided ideas which aided in the design of visuals, game mechanics and interactive content. This provides researchers and developers outside of the field of autism research to gain a valuable insight and make informed decisions that can strengthen the new technology as an intervention tool [250]. Furthermore, taking part in participatory design sessions provides autistic children an enriching and empowering experience as they have the opportunity to develop social and communication skills [251]. With this is mind, there is a number of researchers that believe that participatory design is essential to developing multi-disciplinary technology based projects for autistic people [250]–[255].

Frauenberger *et al* [251] involved a cohort of autistic children is the design process of *ECHOES*, a digital learning environment that targets social skills in autistic children. Following sessions with the *ECHOES* system, an annotation overlay was added to the screen that permitted the child to draw, write and place faces with different emotions on different aspects of the environment. Following this, researchers would ask high level questions regarding specific features of the virtual environment. Magkafa *et al* [256] extended participatory design to include autistic children in multiple stages of the development process of a museum based application. An initial discovery stage incorporated multiple brainstorming sessions during which participants were encouraged to provide verbal feedback, produce low-fidelity drawing of prototypes and take part in brainstorming sessions via drawing. The second PD stage centred around usability testing in order to evaluate the children's eagerness to engage with the application and its ease of use. Following time with the software a qualitative questionnaire was used which implemented a range of faces. This aided in bypassing any communication difficulties the participants may have.

However, PD workshops with autistic children can often be a challenging task as they require communication and social interactions. In addition, it is important to find the correct balance between empowering children and overwhelming them with responsibility [251]. Malinverni *et al.*[257] applied multi-modal evaluation of participatory design sessions where children were asked to perform exploration and drawing activities whilst investigators focused analysis of interactions based on gaze direction, position in the space, focus of attention and the children's drawings. Another possible challenge is to circumvent the monotropic characteristics displayed by autistic children [258]. This may manifest in participant's focusing on the minor details of a project whilst missing the relevant design decisions for that particular session [251]. In addition, this extends to autistic people having very specific interests [1] which can cause strong bias towards some design aspects of the project [251]. This matter could be avoided by focusing PD activities around highly specified design choices [251].

Finally, generalisation and maintenance of new skills acquired through computer based CBT remains under-explored. Within this review, only one study evaluated generalisation after interventions [34]. This is a difficult area in therapeutic approaches for autism, with many children showing improvements in the clinic environment but experiencing difficulty in maintaining and applying the new skills to everyday situations [259]. Although family-centred therapy is suggested to enhance generalisation [260], considerations need to be made not only when designing new intervention programs but also assessing effectiveness in the months following completion.

2.7 Summary

For autistic people who experience auditory hypersensitivity to specific environmental sounds, everyday life can often be problematic and distressing. It is therefore important that effective, engaging and affordable interventions are developed that can alleviate negative associations to particular sounds and reduce extreme adverse behaviours. This could not only improve quality of life for the individual, but also make drastic improvements to family life. This chapter has recommended a combination of serious game driven CBT and three dimensional audio. This has the potential to create an engaging, controlled and safe therapeutic environment in which users can be systematically exposed and eventually become habituated to feared auditory stimuli. This review has already presented studies that support realistic visual virtual environments employed in cognitive behavioural therapy sessions.

Consumer VR technology can provide accessible interventions that can move away from the clinic to the home, incorporating family members into therapy sessions. This allows them to gain a better understanding of their child's behaviour and specific needs, whilst also building essential social and emotional relationships. It is important that research takes advantage of new advances in interactive technology and adapts it to the requirements of autistic children. This will deliver critical knowledge into the complex workings of the autistic brain which will aid in the innovation of more efficient intervention tools and frameworks. Furthermore, it will allow autistic children the opportunity to enjoy the same enriched virtual experiences as their NT peers, creating enjoyment and empowerment.

It is important however, that collaborative research between clinicians and engineers continues in order to investigate which design aspects are necessary in developing technology driven interventions that are both useful and appealing to children on the autistic spectrum. Autism is a complex neurological condition and occurs on a broad continuum of cognitive, social and emotional capabilities. Autistic individuals may react differently to therapy on a cognitive or social level, and the effectiveness of any intervention will be dependent on the individual.

Chapter 3

The Fundamentals of Spatial Audio Rendering

3.1 Introduction

Prior to investigating how binaural based spatial audio can be used as a tool for addressing auditory hypersensitivity, it is important to first discuss the fundamentals of binaural listening and spatial audio rendering. This chapter will first introduce the core principles of how the human auditory system decodes binaural cues in order to accurately localise a free field sound source within an environment. Following this, a description of the core principles of binaural based Ambisonics will be presented, providing background knowledge of how this spatial audio rendering technique can be used to simulate three-dimensional sound fields. Finally, evidence will be provided which supports the use of spatial audio in creating plausible virtual environments.

3.2 Binaural Hearing

The success of immersive 3D audio is heavily reliant upon the capability of reproducing the human auditory system's localisation cues within the digital realm [261]. This can be accomplished via understanding the fundamental psychoacoustic principles of how a sound wave diffracts and interacts with the head and pinnae, resulting in both temporal

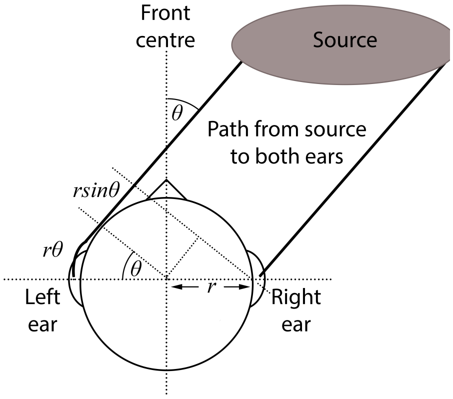


FIGURE 3.1: Image displaying sound pressure wave travelling from a source arriving at the right ear before the left. Resulting in interaural time differences. Image adapted from [265].

and amplitude changes as well as spectral cues created by pinnae filtering. It is these alterations, which vary between our two ears, that enable the brain to decode acoustic information to determine a source's locality [262]. This section will explain the binaural cues crucial to sound source localisation, and how they can be recorded as a Head Related Transfer Function (HRTF) to be used to simulate binaural hearing over headphones.

3.2.1 Interaural Time Difference

Due to the physical separation between the two ears and the shape of the human head and torso, a sound source that is not directly in-front or behind a listener will result in sound waves arriving at one ear ahead of the other. These very small delays in sound, which are at their most 0.65 msec, are known as ITD and are one of the two binaural cues fundamental to interpreting a sound's point in space along the horizontal plane (see Figure 3.1) [263], [264]. However, this hearing mechanism cannot accurately distinguish between front and rear sources, or take into account elevation.

The simplest calculation of ITD is that the ears are separated by a distance, which is usually considered approximately 18cm, and the delay is then derived by taking into account the angle of incidence. Unfortunately this is inaccurate as it overlooks the extra distance by the wave travelling around the head. Equation 3.1 is therefore the most precise method of calculation as it incorporates this extended path, although it must assume that the head is completely spherical [266].

$$ITD = \frac{r(\theta + sin(\theta))}{c} \tag{3.1}$$

r = half the distance between ears (m)

θ = angle of arrival of sound from median plane (radians)

c = speed of sound (ms^-1)

Research by Zwislocki [267] in 1956 proved that there is a maximum frequency to which ITD can be effective at source localisation (1.5 kHz), but this will decrease as the angle from the median plane increases. This frequency threshold is due to the distance a sound wave must travel in order to reach the furthest ear, causing the phase angle to differ upon arrival. The corresponding angle being both the function of the wavelength and the source positioning. For frequencies below 1.5kHz, the auditory system is sensitive to phase differences between the sound reaching each ear. This is due to the difference in path length between the sound source and the left and right ear. With frequencies below 1.5 kHz the wavelength is greater than the maximum difference in time, this phase difference will therefore provide accurate location cues. Anything above 1.5 kHz causes problems with localisation as the wavelength is similar to the temporal differences between the ears, reducing the phase difference [264]. Further investigation by [268] in 1992, showed the importance of low frequency ITD. Their experiments found that when they are present they dominate over interaural level difference and spectral shape cues present in both the low and higher frequencies. Consequently, when removed some localisation confusion occurred.

3.2.2 Interaural Level Difference

ILD are the result of the interference of sounds waves by the head, pinnae and other parts of the body such as the shoulders, these can both reflect and produce a shadowing effect. This is influenced further by the lack of symmetry of the head. Much like ITD, when a source moves away from the median plane, dissimilarities in sound intensity take place. For the ear furthest from the source, this will mean a decrease in level [264], [266].

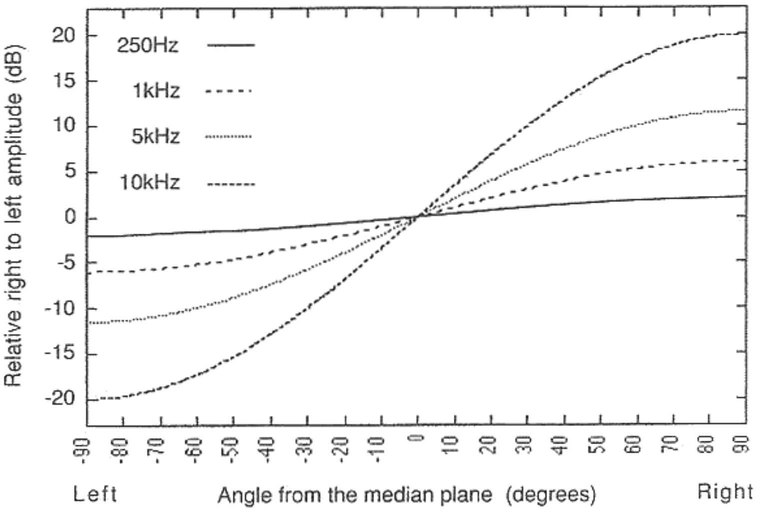


FIGURE 3.2: Interaural Level Difference - Function of angle and frequency [266].

However, unlike ITD, location cues taken from ILD are most relevant towards the higher end of the frequency spectrum, see Figure 3.2. Soundwaves with a length either equal to or shorter than the size of a human head (18cm) are most effected by the acoustic shadowing. This includes frequencies above 1.5 kHz. A lower frequency of 200 Hz has a wavelength which is roughly nine times longer than the diameter of the head, allowing the waves to travel around the head despite its azimuth. Shadowing is greater for a frequency of 10 kHz, when compared to the size of the head as its wavelength is much smaller. This increased shadowing effect therefore produces a greater difference in sound levels between the two ears [264]. It is because of this that it may be considered that ILDs occurring in individual frequency bands are more proficient at providing localisation cues rather than the overall ILD, a product of a combined effort of both the head and the filtering properties of the pinnae [263].

3.2.3 Spectral Cues

Up until the late 1960s, the pinna was not fully understood and regarded no further than an imperceptible part of the body to hold glasses and hats [264]. It is now known that it is much more important, providing localisation cues not only for the front and back, but also the elevation of a source [264]. Audio with higher frequency content (above 5 kHz), interacts with the structure of the pinna which is constructed of a complex set of multifaceted ridges. Acoustic energy is distributed and consequently reflected causing minute but significant delays. It is these delays that will deliver a comb filter effect to sound entering the inner part of the ear, representing a three dimensional function of the azimuth and elevation of the sound source [266], [269]. Another important point to consider is that the perception of elevation is also frequency dependant, direct and reflective sound waves become out of phase, this is due to the extra distance travelled by the reflections. Studies conducted by [270] showed that when a sound is located to the front there is an peak in amplitude at lower frequencies of 250-500 Hz and higher frequencies above 13 kHz. This peak moves to 8 kHz when the sound travels overhead and when the source is behind, the head peaks can be found at both 1.5 kHz and 10 kHz.

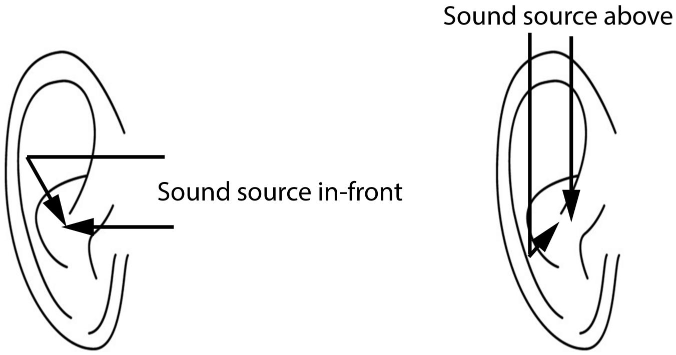


Figure 3.3: A digram to illustrate how sound from different positions interacts with the pinnae, resulting in the spectral cues essential for vertical localisation. Image adapted from [271].

A final note is that, in order to make full use of the frequency dependant elevation cues there must be a sense of familiarity between the source and the listener. An acquaintance with a sound allows them to compare the amplitudes of different frequencies with previously experienced values. Past experiments have demonstrated that unfamiliar sound sources are often perceived as coming from behind [264].

3.2.4 Head-Related Transfer Function (HRTF)

Comprised of the interaural and spectral cues previously discussed, HRTF are the functions of both frequency and position for both azimuth and elevation, containing direction-dependent information on how a sound source positioned at a point in space reaches the ear without environmental obstruction [272], [273]. From each position, a sound will be subjected to distinct frequency dependant amplitude and time delay differences providing unique binaural information that the brain can decode as a spatial cue [36]. However, spectral shaping produced by the pinnae, torso and shoulders are individualised. Therefore each HRTF will contain variances from listener to listener.

Figure 3.4: Neumann KU 100 binaural dummy head microphone, manufactured by Neumann [274].

HRTF measurements are the Fourier transform of a Head Related Impulse Response (HRIR), measuring phase shift and magnitude dependant on localisation cues mentioned in sub-sections 3.2.2, 3.2.1 and 3.2.3 [36], [272], [273](see Figure 3.6). The process of recording HRIR involves using either the human ear or a mannequin such as the Neumann KU 100 [274], [275] (see Figure 3.4). Microphones are placed in, or close to each ear canal as a means to capture the transformations of a sound as it arrives at the ear drum. From here, multiple HRIRs can be obtained by recording an output signal from a loudspeaker at a fixed radius and at multiple positions [275]. Each pair of HRIRs will then contain unique directional information associated with the elevation and azimuth of the free-field sound source. If possible, recordings should be performed within an anechoic environment in order to eliminate environmental sound reflections being encoded into the response [276]. Figure 3.5 shows a pair of HRIRs recorded for a sound source positioned at 90° azimuth and 0° elevation, displaying contrast in time and level between each ear. In addition, Figure 3.6 shows the HRTF representation of the HRIR shown in Figure 3.5. The plot

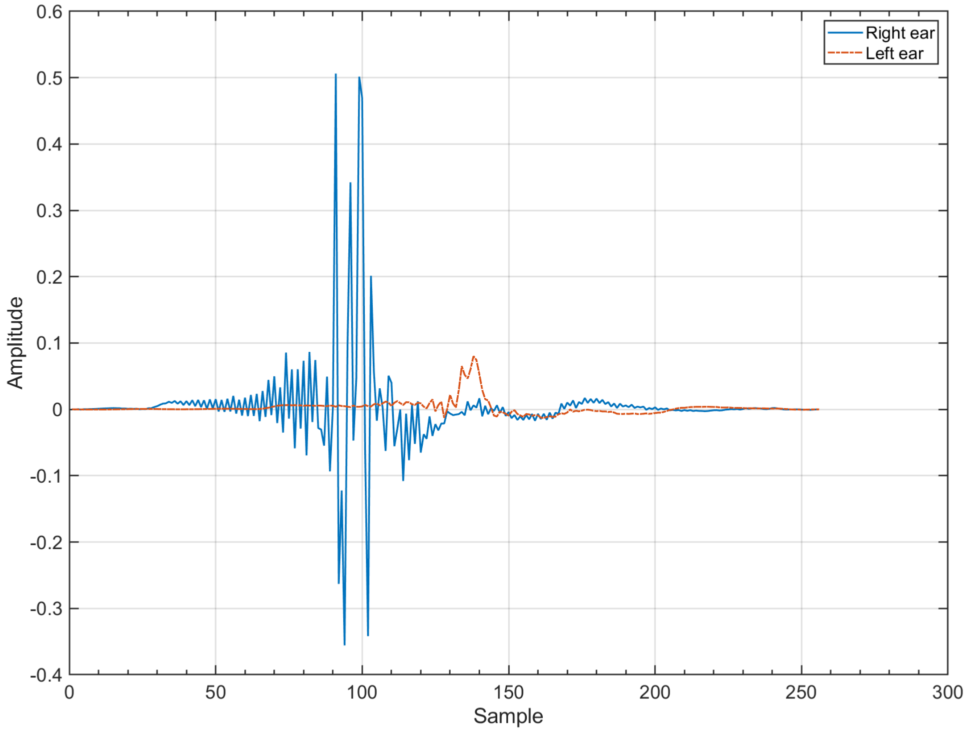


FIGURE 3.5: Time domain representations displaying the differences in time and amplitude between the left and right ear for a signal recorded at 90° azimuth and 0° elevation. HRIR taken from the Spatial Audio for Domestic Interactive Entertainment (SADIE) II database

demonstrates the spectral shaping as a result of sound diffraction by the head, torso and pinnae, in addition to the magnitude of the difference between the right and left ear spectra. Displaying the acoustic shadowing effect of the head.

HRTFs play an important role in the headphone based spatial audio reproduction systems. This is based on a concept that in order to accurately simulate spatial sound cues, a sound source must undergo the same spectral shaping virtually as it would in normal listening conditions [276]. With this in mind, perceived directionality can be achieved through the convolution of a monoaural sound source free from echo or reverberation and a pair of HRIRs recorded at the corresponding point in space [277], [278].

3.3 Room Acoustics

Perception of an environment via sound propagation not only depends on the initial sound source, but also extends to the absorption coefficients of the materials in the room and the

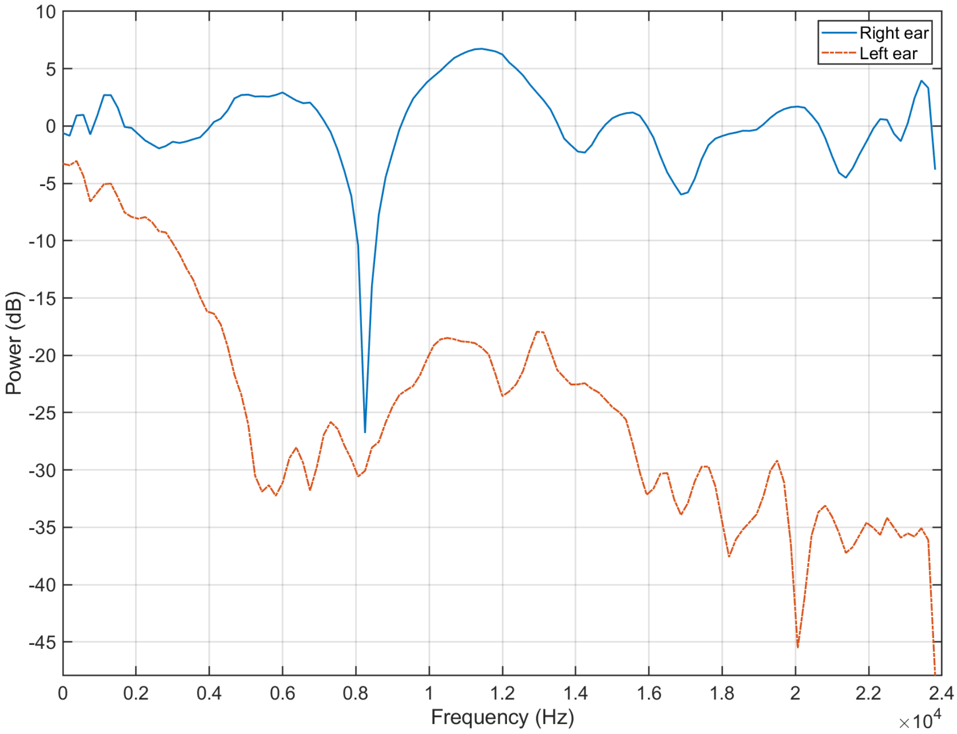


FIGURE 3.6: HRTF representation of the HRIR displayed in Figure 3.5.

geometric properties of a space. The shape of a space will have an effect particularly on the early reflections by sculpting the spatial distribution of sound wave paths to the listener. Within an enclosed environment sound energy is absorbed differently by surfaces depending upon their materials, whilst in most cases a significant amount can also be reflected back into the space. Furthermore, these reflections can produce additional reflections, a consequence of which is the creation of a diffuse sound field. Figure 3.7 exhibits the acoustic response of a location to a short impulsive signal. The initial early reflections occur as the energy first encounters surfaces, usually within the first 1-80 msec and have a significant amplitude above the noise floor. These are subsequently followed by the late reverberations which arise from further reflections between surfaces. Having considerably less energy than the first temporal section of the response, they eventually decay into silence. It is these early and late reflections which provide environmental context, by contributing to the listener's perception of size and space [36], [276]. Furthermore, any sound absorption will be frequency dependant, and therefore the timbral changes will vary between different acoustic environments.

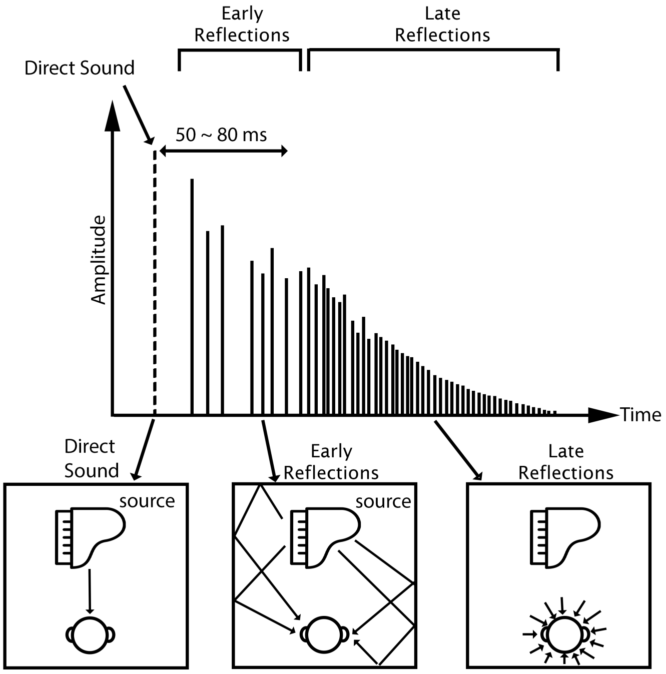


FIGURE 3.7: Image displaying direct sound followed by early and late reflections. Adapted from [279]

The acoustic characteristics of listening environments are also crucial in the localisation of a sound source within the space. The location accuracy relative to both azimuth and elevation angles can be impaired in a highly reflective environment [36]. In both real and virtual spaces, a strong reverberation can dampen the perceived attack of a direct sound. This is a process that is often exploited in orchestral pieces played in concert halls, the large reverberation time is used to blend an ensemble together [276].

Studies [276], [280] into localisation in both reverberant and dry environmental contexts, displayed an increase in errors with the presence of reverb and more precisely early reflections, with the azimuth centroids being off by an average of 12° to 23 °. Demonstrating that reverberant energy can have a negative impact upon sound source localisation. These findings were reinforced by Hartmann, affirming that the expansion of apparent source width caused by an increase in reverberation would also make location determination more difficult. The perceptual increase in area therefore presents no reliable information for azimuth and elevation discernment [281].

3.4 Distance Cues

In addition to positioning a sound source either on the horizontal or vertical plane, an ability to distinguish the distance of a sound is another fundamental component in interpreting auditory environments. Distance perception is the result of several contributing factors influenced by not only the physical distance between the listener and the sound, but also the acoustic environment in which they are located and the listener's cognitive familiarity with the sound [36], [276].

Within an environment devoid of obstacles, reflective surfaces or boundaries, the intensity of a sound is the primary consideration when interpreting distance, playing a more important role when an individual is exposed to unfamiliar sounds [276]. Within an anechoic environment the correlation between distance and sound intensity can be calculated through the inverse square law, in which the intensity of an omnidirectional sound will decrease by 6dB for each doubling of distance between the source and listener [282].

The frequency content of a sound is also a contributing parameter in the perception of distance, influenced by the sound wave absorption through the air [36]. This process is observed as a dynamic low-pass filter which is a function of the distance between source and receiver. At its most basic, the high frequency content of a far-field source will become attenuated as distance increases [283], [284]. But within the physical world there are multiple emerging conditions such as air humidity, temperature and wind, all of which will affect the spectral composition of sound. Despite the expediency of intensity and frequency formation in discriminating distance, in the absence of visual accompaniments auditory perception of distance is impeded within a non-reflective environment [36].

As mentioned in Section 3.3, the reflective properties of an acoustic environment can infer much about a listener's surroundings, this includes distance. A listener's perception of distance can also be impacted by the energy present in the late reverberations caused by reflective surfaces. Research by Sheeline in 1982 suggested that room acoustics are an important addition to intensity cues when establishing judgements in distance. They came to the conclusion that reverberation delivers the 'spatiality' that allows the listener to move from the domain corollaries of intensity to those of distance [285]. Further experimnetation by Mershon concluded that judgements in a reverberant space, when compared to those in an anechoic chamber, are more accurate at perceiving distance [286]. This is however in contrast to the findings of von Békésey, who stated that distance cues are more pronounced in anechoic spaces rather than reverberant environments [276]. When used in conjunction with the other available senses, reverberation serves as an important cue for an environment's size or boundaries. The brain can make sub-conscious decisions to place both real or virtual boundaries which present perceptual limits to a sound's source's distance from a listener [276].

3.5 Spatial Audio Reproduction through Ambisonics

Researchers within the field of spatial audio reproduction have been attempting to recreate auditory environments and deliver a sense of spatial realism to a listening audience for over 139 years. One of the first stereo transmissions was conducted by Clement Ader and took place in 1881 during the first Electrical Exhibition in Paris [287]. The experiment involved placing telephone microphones across the stage at the Paris Opera and broadcasting the audio to telephone receivers in a separate venue. Here listeners were able to experience some level of spatial acuity. From this initial experiment, spatial sound reproduction techniques have moved from two channel stereo to higher multi-channel audio playback techniques such as surround sound [288] and Ambisonics [289]. All of which can be experienced using both multi-channel speaker setups or rendered binaurally over headphones.

Due to its prevalence as an audio renderer for VR applications [290]–[293], binaural based Ambisonics will be the central audio rendering technique used as part of the research investigations presented in this thesis. Therefore it is essential to understand the fundamental concepts of ambisonic recording and reproduction techniques before they are to be used to help target auditory hypersensitivity in individuals with ASD.

Developed by Micheal Gerzon[289] in the early 1970s, Ambisonics is a complete approach to encoding, recording, transmission and reproduction of a 360 degree audio sound field using spherical harmonics [294]. Spherical harmonics are a set of mathematical functions that describe other functions located on the surface of a sphere [295]. In traditional multi-channel audio such as stereo and 5.1 surround, each audio channel has a signal that directly coincides with a fixed position within a fixed loudspeaker configuration. In contrast, each channel of an ambisonic recording contains information regarding the

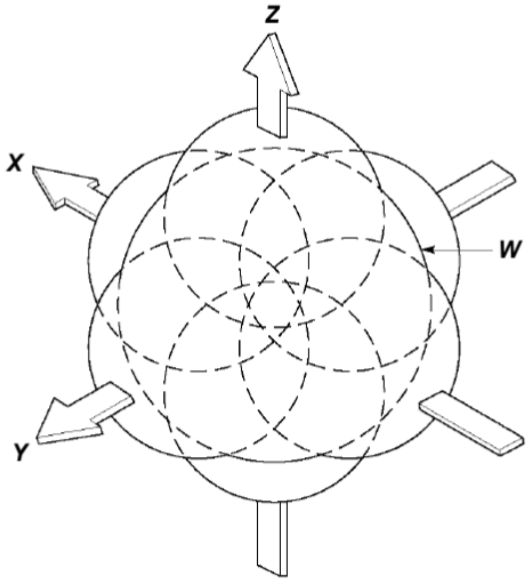


FIGURE 3.8: The components of a 1^{st} order Ambisonic B-Format signal. X,Y and Z represent the three figure of eight components, with W representing the omni-directional pressure component [36].

physical characteristic of the acoustic field which can be decoded to any number of physical or virtual loudspeakers in any configuration [36]. This creates a distinct advantage in that the recording and reproduction processes are entirely independent from each other.

First Order Ambisonics (FOA) or B-Format is the base form of Ambisonics that is capable of representing the pressure and velocity properties of a 360-degree sound field. Using four channels (see Figure 3.8), B-format is comprised of an omni-directional component (W-channel) alongside three orthogonal figure of eight components (X, Y & Z channels). From here, a monophonic sound source s can be encoded at any point across the sphere represented by azimuth θ and elevation ϕ Cartesian angles relative to the fixed front facing position using the following [296]:

$$W = s\left[\frac{1}{\sqrt{2}}\right] \tag{3.2}$$

$$X = s\,[\,cos\,\theta\,cos\,\phi\,] \tag{3.3}$$

$$Y = s\,[\,sin\,\theta\,cos\,\phi\,] \tag{3.4}$$

$$Z = s\,[\,sin\,\phi\,] \tag{3.5}$$

Encoding single or multiple sound sources surrounding a listener can be accomplished using ambisonic panner plugins for digital audio workstations, an example of which is the Facebook 360 spatial audio workstation [297]. In addition, a B-format signal can be obtained using an A-Format microphone similar to that shown in Figure 3.9. Four capsules with sub-cardioid patterns are arranged in a tetrahedral pattern array facing the A-format directions outlined in Table 3.1 [36]. From this point the A-format signal can be converted to B-format with the following [298]:

$$W = A + B + C + D \tag{3.6a}$$

$$X = A + B - C - D \tag{3.6b}$$

$$Y = A - B + C - D \tag{3.6c}$$

$$Z = A - B - C + D \tag{3.6d}$$

This is the followed by frequency dependant filtering and phase compensation [77]. Figure 3.9 displays a tetrahedral microphone array and illustrates the capsules mentioned in Equation 3.6 and Table 3.1.

Table 3.1: Capsule labelling and arrangement in a tetrahedral A-Format Soundfield microphone [36]

Capsule	Direction
A	Front-left-up
B	Front-right-down
C	Back-left-down
D	Back-right-up

In addition to positioning a sound source using spherical coordinates, the use of Ambisonics also allows for the reproduction of distance cues [299]. This is accomplished by controlling the attenuation of amplitude and frequency content of a virtual sound source. Furthermore, adjusting the ratio between direct and reverberant sound can also play a crucial role in simulating the perception of distance within a virtual sound environment.

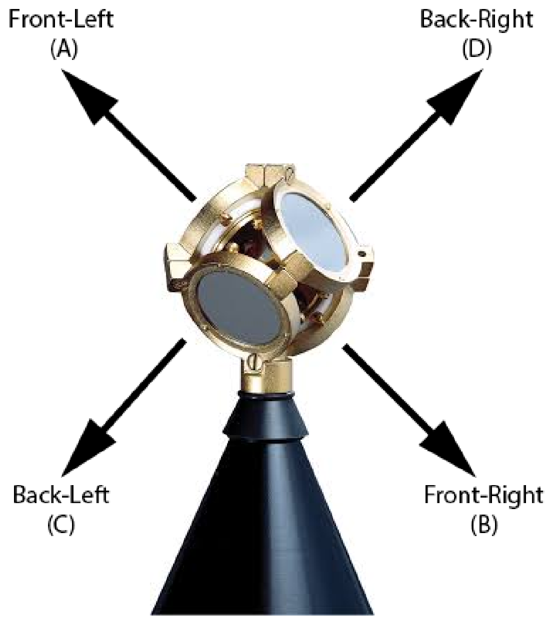


FIGURE 3.9: Illustration displaying a Soundfield tetrahedral microphone array with [300]

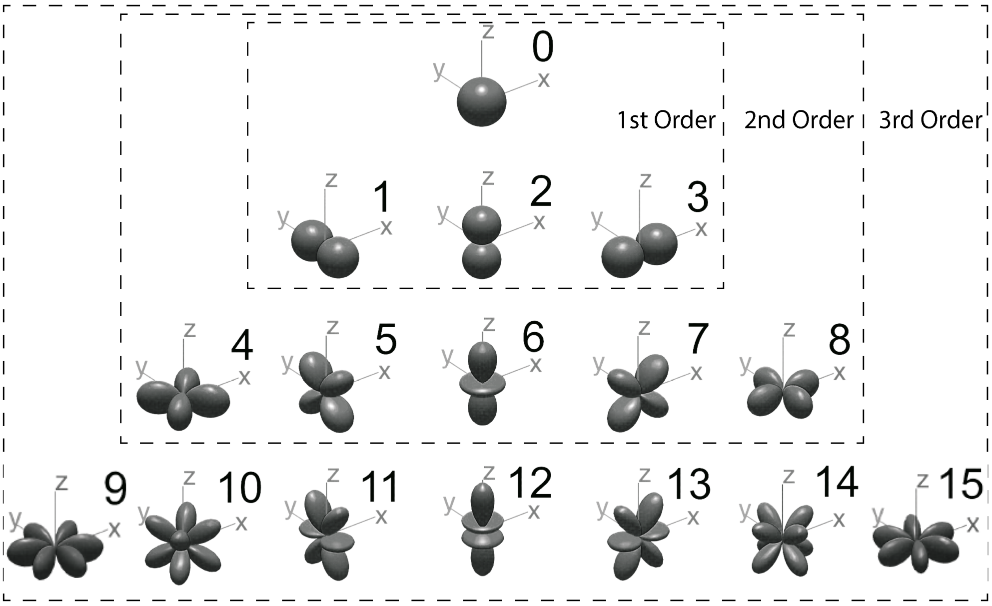


FIGURE 3.10: Ambisonics spherical harmonics for orders up to 3^{rd} order. FOA incorporates the top two lines (4 channels) and 3OA incorporates all four lines (16 channels). Reproduced from Narbutt *et al.* [301]

First order Ambisonics does however experience low spatial resolution which can negatively impact upon sound source localisation, especially in auditory stimuli with high frequency content [302]. This can be significantly improved with the implementation of higher orders which encode and decode the soundfield using a greater number of spherical harmonics (see Figure 3.11) [302]. However, increasing the ambisonic order (M) does require an increase in the amount of channels and the minimum number of loudspeakers (N), the number of which is represented using:

$$N = (M+1)^2 \tag{3.7}$$

Today, most VR systems employ a virtual ambisonic approach in order to synthesise three-dimensional soundscapes binaurally over headphones. This technique is based on decoding ambisonic information to virtual loudspeakers which places the listener in the centre of the soundfield. Binaural reproduction is achieved by convolving audio with HRTF measurements that are appropriate to the virtual loudspeaker's position in space [276], [303], [304].

One significant advantage to using Ambisonics is that an encoded soundfeld can be rotated relatively easily through the use of rotation matrices. An arbitrary rotation of a 1st order B-format soundscape characterised by a rotation matrix R would be as follows [305]:

$$W' = W \tag{3.8a}$$

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = R \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{3.8b}$$

With this in mind, any rotation matrix can be applied to an ambisonic file for pitch, roll and yaw rotation. For example, rotation across the z axis of an angle of θ can be defined as [305]:

$$W' = W \tag{3.8c}$$

$$X' = X\, coz\, \theta - Y\, sin\, \theta \tag{3.8d}$$

$$Y' = X\, sin\, \theta + Y\, cos\, \theta \tag{3.8e}$$

$$Z' = Z \tag{3.8f}$$

This creates a plausable and realistic virtual audio scene that responds dynamically to the head rotation of the listener. Current VR head devices can provide stable multi-modal interaction through accelerometers which can detect 360 orientation. Once the rendering system has received the x, y, and z rotational information, it performs a rotation of the encoded soundfield. This results in the perceived localisation of a dynamic virtual sound source with 3 DoF[277], [278]. This can then be extended further to achieve 6DoF, by the position of the HMD relative to the virtual environment. Through the use of object-based representation, each virtual sound source is also rendered using distance dependant volume attenuation and reverb based on the tracking data [299], [306]–[309].

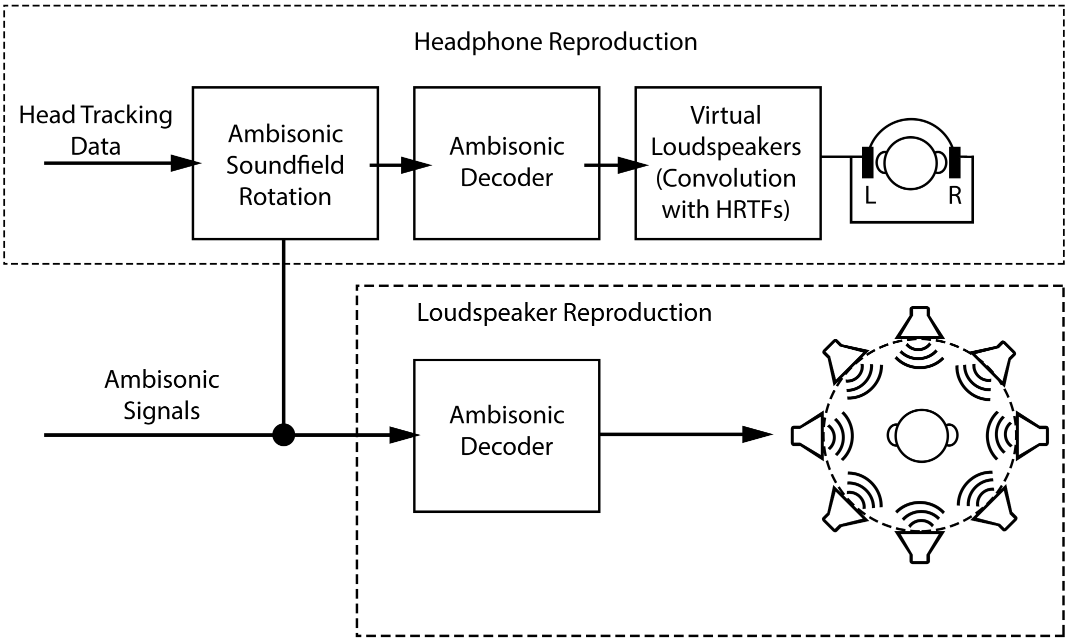


Figure 3.11: Diagram illustrating the signal chain for Ambisonic rendering over multi-loudspeaker array and a binaural headphone system. Adapted from [310].

3.6 The importance of spatial audio in virtual environments

Larson *et al.* [311] suggest that the spatial properties of virtual auditory environments have been of significant importance since the introduction of stereo sound in the 1930s. Further stating that spatial sound is used to simulate an auditory reality that gives the user the impression they are surrounded by a 3D audio environment [311]. In terms of precision, spatial acuity is somewhat inferior to vision and the somatic sensory system when perceiving the environment [312]. However, it is not unimportant. Spatial hearing is a critical way in which the user perceives space, providing information about events and objects that are far beyond the field of view [313].

To perceive the real world, the human brain must decode a constant stream of multi-modal information from the various sensory channels [314]. An example of this occurs during typical cases of face to face communication. The brain must interpret visual information such as moving lips and facial expressions, whilst also listening to localised speech sound from the mouth [246]. This cross modality synchronisation can also be echoed in VR by using 3D audio, by simulating the shared spatial qualities of both audio and visual stimuli that occupy the same time and space.

The following sections will examine how spatial audio can directly impact the presence and emotions felt by an individual within a virtual environment. As discussed in Chapter 2 these psychological mechanisms are considered crucial components of a VR exposure intervention, creating a safe and controllable environment for habituation towards a feared stimulus.

3.6.1 Presence

The concept of presence and its importance to VR applications, especially for VRET applications, has been discussed in Section 2. Often presence is looked upon as a product of user immersion [315]. Within VR, this is reliant on the extent and fidelity of the technology delivering the sensory stimuli to the user through head-mounted displays, haptic feedback and spatialised sound rendering systems, all contributing to establishing the sensation of ‘being there’ by creating a realistic and immersive environment. However, despite vision being reputed as the principle system for spatial localisation [316], Larson *et al* [241] suggest that auditory input is fundamental to realising a full sense of presence as it provides a constant stream of omidirectional sensory information, unlike its visual counterpart which is inherently directional and can be filtered by closing the eyes [317]. Therefore, spatial hearing is a critical way in which we perceive space, providing information about events and objects that are far beyond the field of view [313]. As discussed in Section 3.2 humans can localise both static and dynamically moving sound sources within a three dimensional acoustic space through a combination of direct and reflective sounds, both of which facilitate the sense of object and spatial presence [241]. This is also fundamental in administrating cross-modal consistency between the sensory information rendered through a head-mounted display and headphones, with presence being influenced by the extent to which the visual and auditory stimuli represent the same point in the virtual space [318]. With this in mind, the utilisation of binaural spatial audio reproduction systems that create 3D sound environments could have a positive impact upon an individual's sense of presence and immersion within a VE.

First, lets take into account the accuracy of localisation of a virtual sound source achieved by using binaural delivered Ambisonics. This has been well documented in research which observes increased spatial attention [319] and decreases in localisation times of virtual audio sources when using binaural Ambisonics [320]. Extending upon this, localisation accuracy is increased significantly when moving from 1st to 3$^{\text{rd}}$ order Ambisonics [78].

There has been a number of investigations that evaluate the relationship between localisation through binaural based spatial audio and presence felt within a virtual environment. A study by Larsen *et al* [321] compared presence felt in stereo to headphone delivered Ambisonics auditory environments within an immersive video game. Whilst results from participant questionnaires indicated no significant differences between the two experimental conditions, those within the 3D audio group displayed a significant increase in phasic electrodermal activity in response to sound related events. The author concluded that the physiological responses could indicate a subconsious difference in levels of immersion or presence. Ruedigier *et al.* [322] presented audio and visual environments over a HMD. In order to compare different headphone rendering techniques, audio was reproduced over headphones using either 4th order Ambisonics or stereo rendering techniques. Using eye tracking and audio source localisation to measure presence, the results indicated that the spatial audio condition experienced higher levels of localisation and presence.

Two experiments conducted by Hendrix in 1994 [323] investigated how three-dimensional audio influences the feeling of presence within a stereoscopic virtual environment. The first compared a silent VE to one with spatial sound. The second was a comparison between two environments with spatialised and non-spatialised sound. In both cases it was found that the spatial audio rendering significantly increased the feeling of presence, with the interactive sounds perceived to be originating from the visual cue. However, interestingly the use of spatial audio did not increase realism. Hendrix suggested that one possible explanation was a lack of reverberation within the auditory environment.

Reproducing the response of an acoustic space is an integral part of designing a virtual environment. Wave propagation which affects sound intensity, timings and early to late reflections, provide enough information to judge the size, shape and material make up of the surrounding space [278], [324]. The simulation of acoustic environments allows the listener to consciously or unconsciously draw upon past experiences to make sense of the new one that has been presented to them [325]. Larson *et al.* conducted two studies [245], [246] investigating the influence of simulated room acoustics on the sense of presence within a virtual auditory environment. Both studies presented participants with anechoic sound sources and audio convolved with Binaural room impulse response (BRIR)s taken from a computed generated space. The results showed a consistent positive correlation between room acoustic simulation and higher presence ratings, with an increase in auralisation quality also significantly increasing the sense of presence.

3.6.2 Spatialised Sound in Exposure Therapy

The use of VR to treat anxiety and other psychiatric disorders relies on the accurate visual representation of the feared stimuli, with sound playing an accompanying and lesser role [326], [327]. It has already been discussed in Chapter 2 that presence and fear are mutually dependant upon each other [228], [328]. Furthermore Section 3.6.1 presents empirical evidence that links the use of binaural spatial audio to increased presence. However, when removed from accompanying visual stimuli and VR, sound alone can also be used to induce an emotional response. Panksepp and Bernatzky suggest that the use of sound can have a greater neurological impact on the subcortical emotional systems than visuals [329]. Research by Ekman and Kajastila [330] observed a correlation between horizontal sound source direction and the perceived *'scariness'* of the sound. Results showed that through manipulation of the location it was possible to increase associated fear response to the stimulus. This work was extended further in a study by Hughes and Kearney [244] in which Ambisonics was used to incorporate height into sound source directionality. It was found that the location of a sound source had a significant effect on the emotional response towards the presented auditory stimuli, with elevation having a significant effect.

Spatial audio rendering techniques have been used to support visual environments in VRET interventions for combat related PTSD [331]–[333], fear of moths (mottephobia) [334], fear of the dark [335] and to desensitise autistic children to airport stressors [336]. However, despite its inclusion in VRET and research linking spatial sound to increased presence and immersion, there is limited research with mixed results investigating its capability as a primary tool for inducing the required amount of anxiety for successful exposure therapy. Brinkman *et al.* [243] compared the effects of different audio rendering techniques over headphones on the participant's self reported anxiety. The first experiment exposed individuals to the sound of a wasp flying around an indoor VR space. Results showed that compared to mono, stereo and Dolby 5.1 surround, Ambisonics driven audio generated significantly higher levels of presence and anxiety. However when the visual representation was introduced, audio did have a positive effect on anxiety levels, although there was no significant difference between stereo and 3D sound conditions. Outside of exposure therapy research, empirical evidence shows that the directional properties of a virtual sound source rendered using Ambisonics can be manipulated to invoke a more fearful emotional response [244], [330].

Finally, Argo [337] developed a soundscape exposure therapy that made use of CBT techniques and higher order Ambisonics rendered over multiple loudspeaker arrays. The program was built not to tackle a specific phobia, but rather aimed at those who experience general feelings of anxiety. The produced soundscapes did not aim to be a realistic reconstruction of an auditory environment. Instead they were designed to be hyperreal, with the use of spatial panning and anxiety-eliciting sounds such as vacuum cleaners, doors slamming and sirens. These sounds were identified from an online psychology forum [337]. During exposure sessions, participants experienced feelings of anxiety and stress. However, following the intervention program they became habituated to the physical symptoms of anxiety they experienced in everyday life.

3.7 Discussion

Chapter 2 examined the scope of intervention techniques used to address auditory hypersensitivity autistic people. Literature indicates that a negative response to specific stimuli is a psycho-emotional reaction rather than physiological differences within the auditory system. Therefore the most successful intervention techniques have aimed to habituate the individual to a problematic sound through graduated exposure. More recently, this form of exposure therapy has been embedded into serious game mechanics that use simulated auditory stimuli to aid in the reduction of auditory hypersensitivity. Though these interventions have displayed some degrees of success, dynamic and realistic virtual auditory environments reproduced by spatial audio rendering systems could be used to improve such treatments. This would adopt a similar approach to VR exposure therapy, which is grounded in over 20 years of empirical evidence and exposes the user to perceptually realistic virtual feared stimuli.

Within both physical and virtual worlds, the directional properties of sound play a critical role in how an individual perceives space, delivering essential information on the environment beyond a field of view. Through the use of spatial audio rendering techniques such as binaural based Ambisonics, it is theoretically possible to simulate the directional properties of a sound source at any lateral and vertical point across a sphere. Furthermore, static and dynamic virtual sound sources can be transformed to accommodate for the rotation of a listener's head, all of which contributes to a realistic and believable virtual auditory environment. It is this plausibility that has led to binaural based spatial audio being used to reinforce visual environments with VR applications, delivering a perceptual union between what the user sees and what they hear. This could create a distinct advantage for autistic people who can experience impaired imagination and require contextualisation of stimuli within therapy. Finally, the distance between the sound and a listener can be inferred by making adjustments to amplitude and frequency content, in addition to manipulating the ratio between direct and reverberant sound. Distance between a feared auditory stimuli and the listener has been used as an approach for gradual and safe exposure for in-vivo interventions. Consequently, distance simulation could be an important factor in delivering a graduated exposure hierarchy within a safe and controlled virtual audio environment. The perception of a virtual environment as real has a significant impact upon the immersion and presence felt within a VR experience, and it is the feeling of presence which has been identified as one of the contributing factors in the successful outcomes of VRET used in the treatment of phobias. In Chapter 2 it was discussed that VR has been used to successfully target some of the core impairments experienced by people on the autism spectrum, including anxiety. In addition, research has also shown that autistic individuals also experience feelings of presence within VR environments. Taking into consideration the enhanced perception of realism and the increased presence experienced by those listening to 3D audio environments, its application in interventions for auditory hypersensitivity in autism requires some empirical exploration.

However, the application of spatial audio rendering systems moves beyond improved acoustic realism. Today, there is a wide range of consumer technology that is capable of delivering a virtual audio environment with head-tracked 3^{rd} order Ambisonics, for example VR head-mounted displays and mobile phones. Deploying interventions using this technology would increase the accessibility of the therapy. Further to this, it will allow the intervention to be conducted within an environment in which the individual feels comfortable. A child's family plays a significant role in the intervention framework, often providing important individualised information in the design process, but also in delivering the intervention itself. Despite autistic children experiencing social and communication difficulties, parents experience increased levels of relationship closeness with their children when compared with parents in the TD population, and this strong parent-child relationship can be harnessed to facilitate early social development [338]–[340]. There are a number of advantages to providing family based interventions. Not only are they cost effective, but they also create enriched and responsive environments for parents who wish to deliver additional support. Further to this, such interventions will move away from the clinic and increase the child's time spent in their more natural home or community environment, increasing motivation and engagement with therapy [260]. Finally, by doing so the individual is placed within the setting in which they are exposed to the problematic stimuli, further contextualising the simulated stimuli and resulting in generalisation of newly acquired skills to post-therapy contexts [341].

3.8 Summary

This chapter presented the fundamental concepts of binaural hearing and how they are reproduced within ambisonic audio rendering systems. In addition, it was discussed how realistic virtual auditory environments have a significant impact upon presence and immersion experience with VR, followed by the application of spatial audio to exposure therapy. This is important to consider for the scope of this thesis as it provides empirical evidence that supports investigations into the use of spatial audio to address auditory hypersensitivity in autism. Not only for the benefits of increased realism, immersion and presence, but also in increasing accessibility to therapy.

Chapter 4

Measuring the Behavioural Response to Spatial Audio within Virtual Reality in Autistic Children and Adolescents

4.1 Introduction

As previously discussed in Chapter 3 of this thesis, spatial sound has been identified as a tool to elicit a wide array of emotions and psychological effects, including the degrees of anxiety and presence required for interventions such as VRET to be effective. The therapeutic outcomes of VR interventions often rely on the immersion of the participant through place illusion [342]. This is particularly important when designing virtual interventions for autistic people [219], a population who have difficulties in generalising new skills to real world applications. However, as discussed in Chapter 2, together with social and communication impairments and displaying repetitive interests and behaviours, autistic people often experience challenges in processing sensory information [1]. These challenges are particularly prevalent within the auditory domain, with empirical evidence liked to impaired auditory scene analysis and sound source localisation.

To date there is no study which investigates how individuals diagnosed with ASD respond to spatial audio events and specifically those within a VR environment. Within the context of this thesis, difficulties in processing sound environments, such as reduced spatial attention towards auditory stimuli, may have a negative impact upon the proposed intervention tool which is centred around rendering virtual soundscapes. This chapter therefore aims to examine the auditory spatial attention and sound localisation ability of autistic children and adolescents within a multi-modal VR game environment with spatial audio rendered binaurally over headphones. By conducting investigations within a VR environment similar to which participants may experience within an intervention setting, test parameters may be precisely controlled whilst the user interactions and behavioural responses can be robustly recorded. Furthermore, the inclusion of dynamic visuals and game mechanics maintains participant immersion, therefore examining a pattern of behaviour closer to that within a conventional VR experience removed from experimental conditions [343]. Research conducted by Newbutt *et al*[219] evaluated the feeling of immersion and presence experienced in VR by 11 autistic individuals, with participants reporting high levels of spatial presence, engagement and ecological validity. Further to this work, Wallace et al. [235] measured spatial presence within the 'blue room', an advanced CVE. The room makes use of visual images projected onto four screens located on three walls and the ceiling, with audio being delivered over a surround sound loudspeaker system. The paper reported that autistic children showed equal amounts of attention to the virtual content than NT controls and engaged with the virtual environment requirements. However, loudspeaker-based systems are often less effective at simulating the necessary ITD and ILD cues needed for accurate three-dimensional spatial audio reproduction. In contrast, head-tracked headphone-based systems are more successful at recreating dynamic soundfields which respond to the listener's position within the virtual world [278].

Initially, this chapter will present a pilot study which investigates how a small group of autistic children respond to binaurally rendered 3^{rd} order ambisonic audio events delivered in a mixed reality environment. The study uses multiple sound design techniques that test localisation and spatial attention towards complex target auditory stimuli amongst the ambient sounds of the virtual environment. Following this, the results, materials and procedure are discussed in order to update the protocol of the main larger sample size study. By examining behavioural responses towards spatialised sounds in virtual environments, this work provides empirical evidence to inform approaches to sound design for VR interventions for autistic people.

4.2 Pilot Study

A pilot study was conducted in order to evaluate how a group of autistic children and adolescents responded to the experimental virtual environment outlined in Section 4.2.5. It was crucial to examine if the group displayed any spatial attention to auditory stimuli rendered through 3^{rd} order ambisonics, and if it was affected by variable background audio levels.

4.2.1 Design

The study followed a within-subject single-arm design in which all participants were subjected to the same experimental conditions. All audio was rendered using binaural based 3^{rd}-order ambisonics. Each participant completed a two-phase experiment which tested spatial audio attention and auditory localisation. The two dependant variables were head rotation on the y axis, and time taken (seconds) to locate auditory stimulus.

4.2.2 Hypotheses

The experiment was designed to explore how the different three dimensional auditory stimuli influence the user behaviour. The hypotheses considering the design of the experiment were:

H_1– Participant head rotation data will be influenced by spatial auditory events in Phase one.

H_2–Participants will have less head orientation towards Event 4 (speech like stimuli) than Events 1, 2 and 3 (non-speech stimuli) in phase one.

H_3– The mean time taken to locate virtual characters with $\leq$ 22.6dB background sound level will be higher than those with $\leq$ 43.8dB background sound level in phase 2.

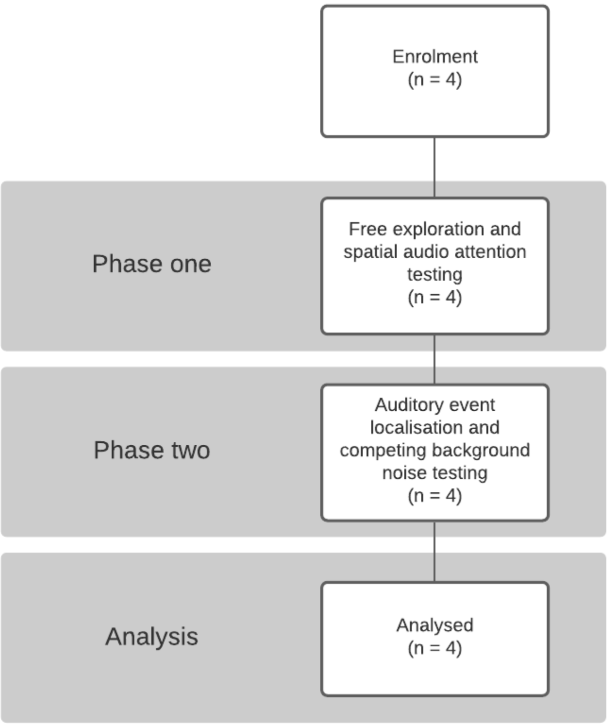


FIGURE 4.1: Flow chart displaying experimental procedure.

H_4– The mean time taken to locate virtual characters with speech like stimuli will be significantly higher than bell stimuli.

4.2.3 Participants

A small study group was comprised of four children and adolescents (all male, mean age = 12.75, Standard Deviation (SD) = 6.65, range of 7–19 years) recruited through a charity. Three of the four participants were diagnosed with autism spectrum disorder, with one not diagnosed but exhibiting autistic symptoms. All the individuals taking part in the study displayed similar social, cognitive and motor abilities.

Participants were admitted into the pilot study after informed consent and assent was obtained from their parent(s) or legal guardian(s). This can be can be found in the accompanying data outlined in Appendix B.

4.2.4 Method

4.2.5 Materials

The study was conducted in the 3sixty demonstration space, University of York, York, UK. The area features four high definition projection screens each measured $6.85m$ x $4.55m$, with a floor space of $6.85m^2$. User's head position and rotation was tracked by six OptiTrack Flex 3 motion tracking cameras, [1] using reflective markers placed on specially designed headwear. OptiTrack Motive [2] motion capture software then sent positional data with 6DoF to the game engine. Only participants were tracked, any accompanying family members or support staff were not able to interact with the VE. For safety reasons and to maximise coverage of participant tracking, an area measuring $3.50m^2$ was marked on the floor. Participants were only permitted to move within this defined section of the CVE.

Each camera has a horizontal field of view of 56°, and was arranged to secure maximum coverage of the 3.16m x 3.16m play area (see fig 4.2). This provides precise and robust coordinates of participants and physical game objects with six degrees of freedom. This data is then sent over a local area network into the game engine. To reduce unwanted noise from the head rotation data, all information was passed through a simple linear interpolation low pass filter. This eliminated sudden and confusing changes in perceived sound source location relative to the users head position.

4.2.6 Game Engine

The game engine used is Unity3D [344], which allows for software deployment across multiple platforms including Windows, Mac and console systems. Visuals for the 3D environment were constructed using pre-bought assets and the drag and drop editor functionality. All game mechanics and audio programming scripts were written in C#.

---

[1] Flex 3 Camera - https://optitrack.com/products/flex-3

[2] OptiTrack Motive - http://optitrack.com/products/motive

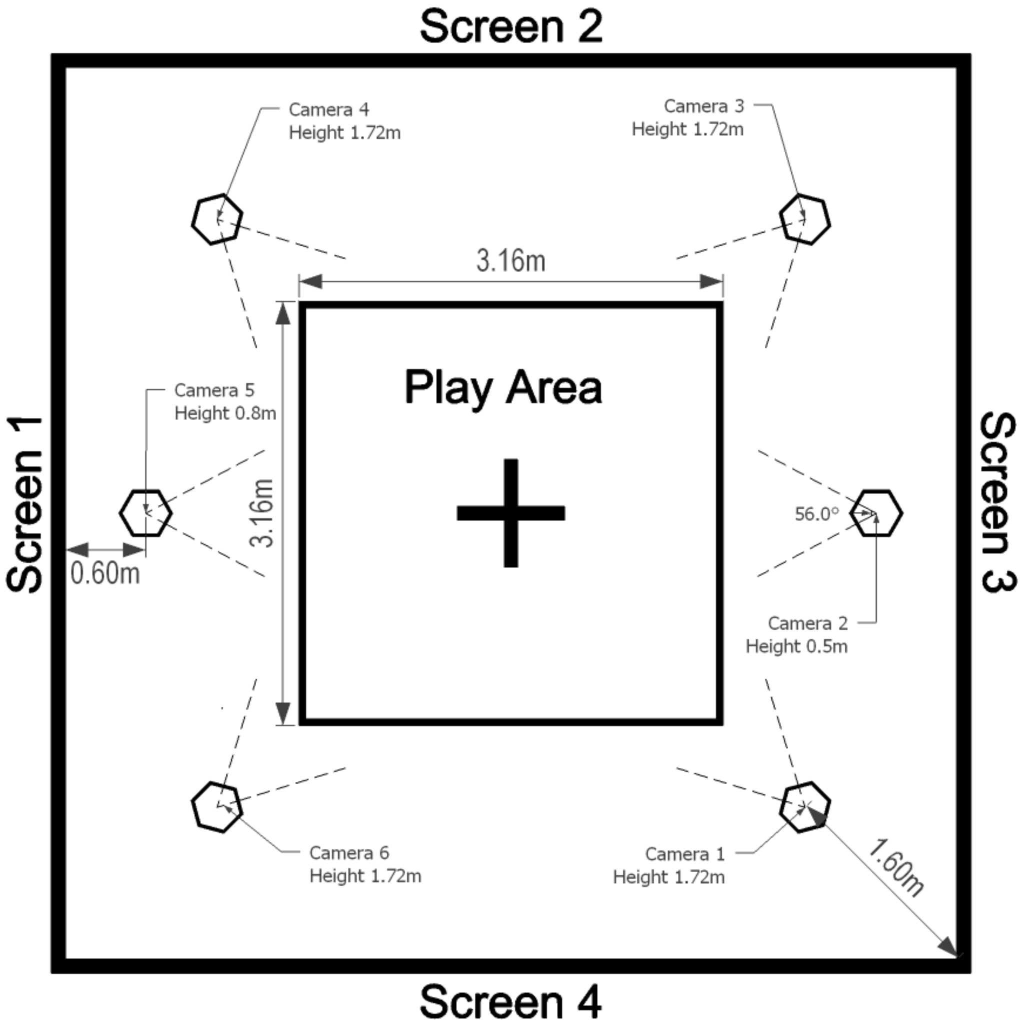


FIGURE 4.2: Floor plan of play area and tracking camera positioning.

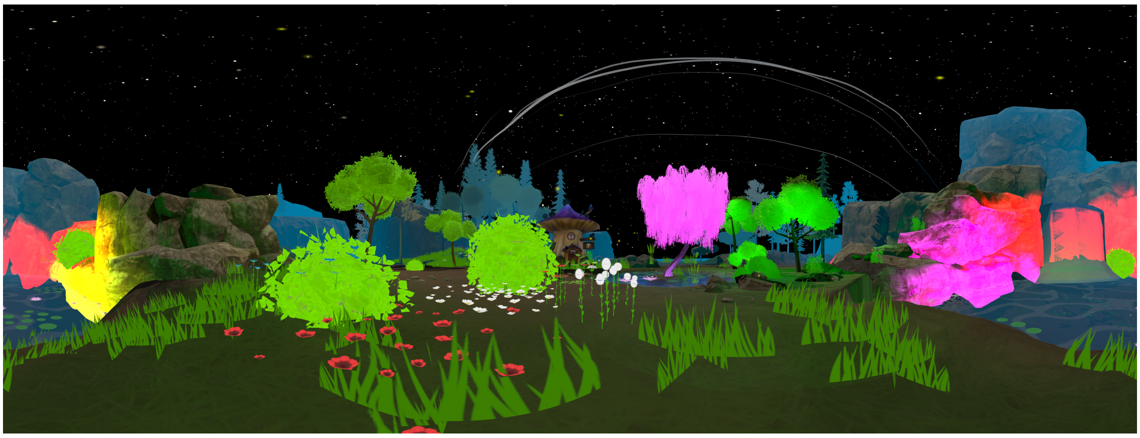


FIGURE 4.4: Visual Virtual Environment—360 degree capture from within the VE

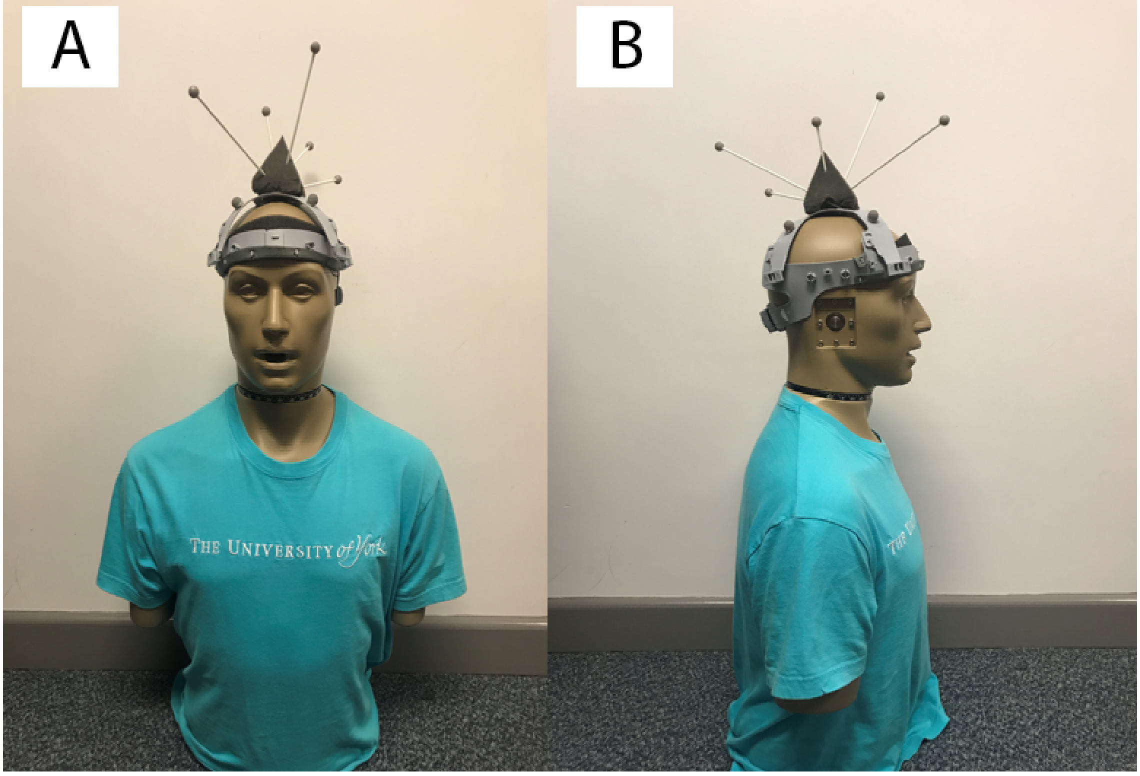


FIGURE 4.3: Infra-red head-tracking marker placement. (A) Headwear with tracking markers (front). (B) Headwear with tracking markers (side)

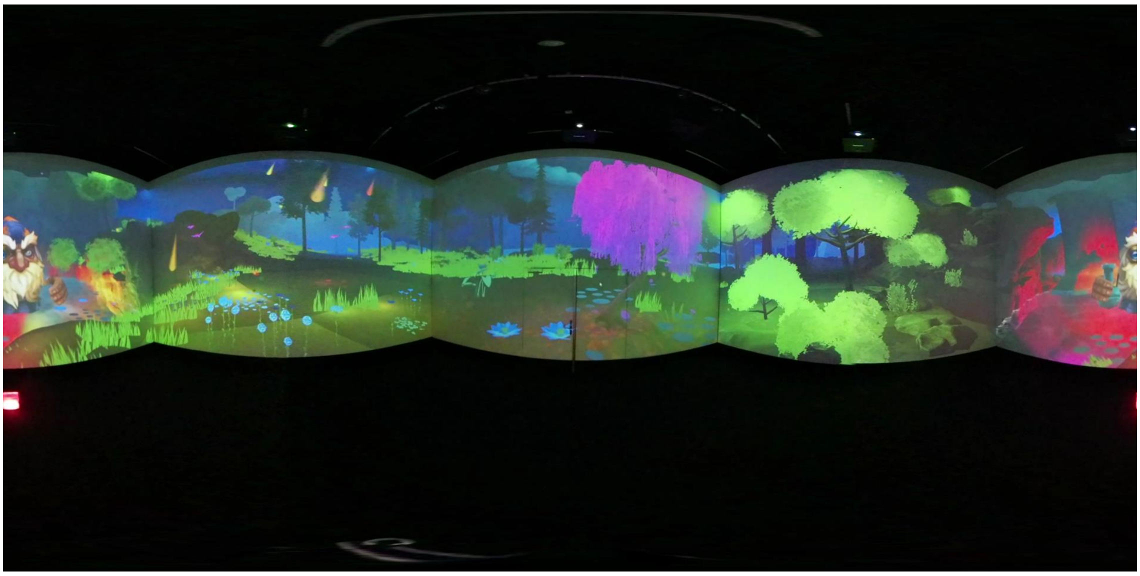


FIGURE 4.5: Visual Virtual Environment—360 degree photo taken from within the 3sixty space

4.2.7 Virtual Environment

Visual Environment: The VE used for this experiment represents an enchanted forest setting at night (see Figure 4.4). It has been noted that when designing visual aspects of VR for ASD applications , the presence of visual clutter significantly affects user performance at completing tasks [345]. Therefore, the graphical environment was designed to be engaging without presenting visual elements that may have driven user attention away from the auditory stimuli. Furthermore, autistic individuals that also exhibit abnormalities in processing visual information can often attempt to avoid visual input such as bright lights [346]. The use of a darker virtual environment reduces the possibility of anxiety related behaviours that are a result of sensory over-stimulation [347].

Auditory Stimuli: Figures 4.6 & 4.7 are overhead views of the virtual environment displaying the auditory stimuli placement and movement area, with details of each audio object in Tables 4.1 and 4.2. The variety of audio stimuli types were implemented in order to cover a broad range of the frequency spectrum. Often, investigations that evaluate localisation using Ambisonic audio utilise white noise bursts as they reproduce all frequencies at an equal amplitude [310], [348]. However, to make the experience more enjoyable for the participants and to match the visual environment, naturalistic stimuli were incorporated. A similar approach was used in a study conducted by Grani *et al* [319], in which sounds such as geese, crows, snakes and mosquitoes were used to evaluate spatial audio attention within a CVE.

All stimuli were recorded or edited using Reaper v5.79 [349] digital audio workstation and Wwise 2018.1 [350] game audio engine. Participants would listen to 3^{rd} order ambisonic audio rendered via the Google Resonance spatial audio Software Development Kit (SDK) [351]. Audio was presented at a fixed playback level of 68 dBA. This was fixed using a -20dBFS RMS value of a pink noise test signal playing at a distance of 1m from the listener. Audio was delivered using Superlux HD 681 semi-open headphones and a wearable stereo radio receiver, avoiding cable issues during free movement around the experiment space. However, only participants received binaural audio rendering, whereas accompanying support staff or family members received mono representations of the soundscape. This was implemented to reduce any influence upon the participant's behavioural response to spatial audio events.

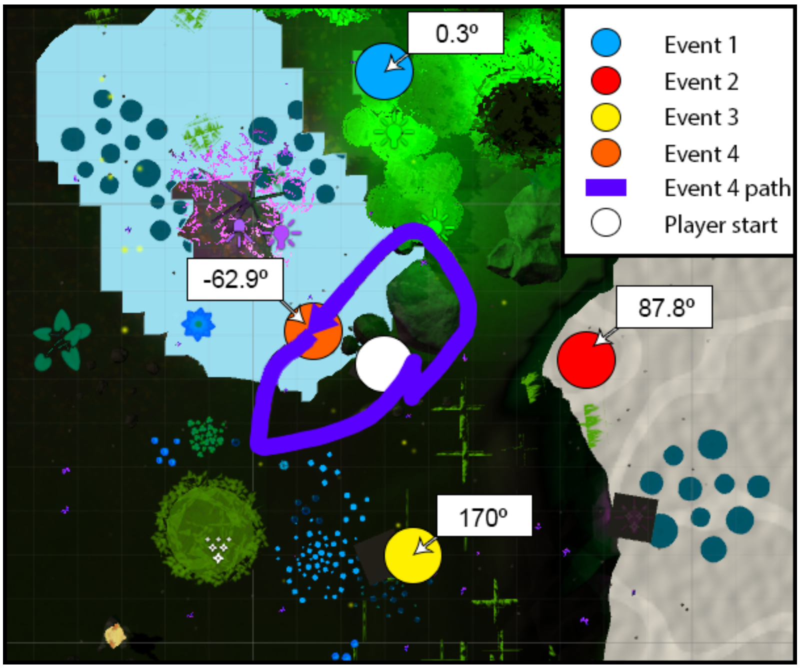


Figure 4.6: Top-down view of auditory stimuli positions within the virtual environment during Phase One.

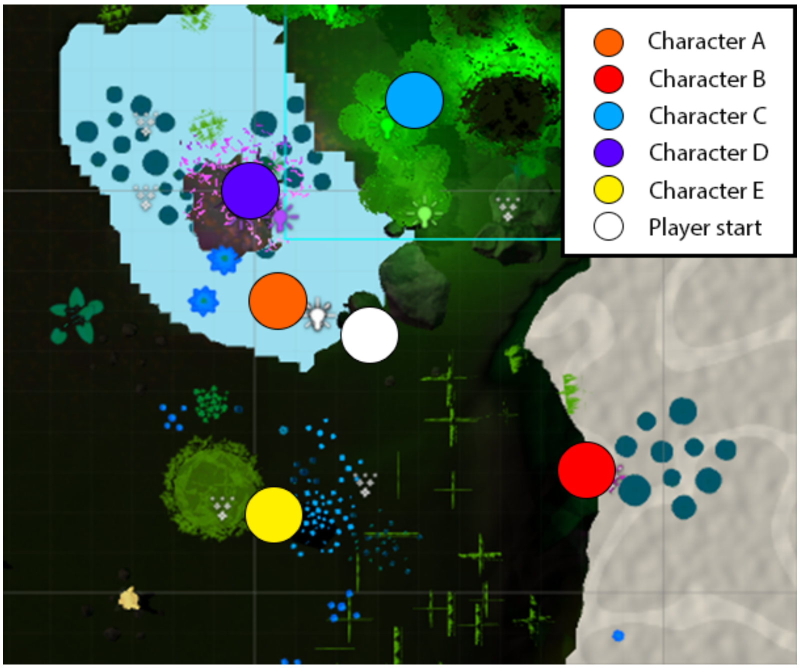


Figure 4.7: Top-down view of auditory stimuli positions within the virtual environment during Phase Two.

Table 4.1: Phase One Auditory Stimuli

Auditory Event	Features
Event 1	MIDI based music - Synth Pad
Event 2	Bubble sound FX
Event 3	MIDI based music - Vibraphone
Event 4	Explosion & Speech like stimuli

Table 4.2: Phase Two Auditory Stimuli

Auditory Event	Features	Background Sound Level (LUFS)
Character 1	Speech like stimuli and bell	$\leq$ -43
Character 2	Bell	$\leq$ -53
Character 3	Bell	$\leq$ -43
Character 4	Speech like simuli and bell	$\leq$ -53
Character 5	Speech like stimuli	$\leq$ -43

The audio events used to test responsiveness to spatial audio were included in two experiment phases.

Phase 1: Four spatial auditory events (see Table 4.1) were placed in the VE, with one appearing from each projector screen (see Figures 4.2 and 4.6 for positioning). Audio Events 1, 2 and 3 were all stationary with the exception of Event 4 which moved around the entire space over the course of 30 seconds. For each event all environment background audio was muted in order to reduce possible effects caused by complications in auditory scene analysis.

Phase 2: Five spatial audio events representing virtual characters were placed throughout the VE (see Figure 4.7 for positioning). In terms of type, audio stimuli were divided into simple and social types. Simple stimuli consisted of a bell sound effect, with the social stimuli representing a speech-like sound. All events were stationary except for movement towards the player once the character was found. These would be presented to each participant in the same order.

A spatialised composite background audio track was provided which was derived from several discrete auditory objects placed throughout the VE, matching the visual scene. This included sounds representing wind, water, frogs, birds, crickets and wind chimes.

To evaluate if associated impairments in separating target auditory objects from competing background noises would affect user experience, each character would be presented within varying levels of background audio (see Table 4.2). The level of the background sounds was either ≥ -43 LUFS or ≥ -53 LUFS. This was achieved by implementing an audio object-based hierarchical model. This was based on a simplified technique designed by Ward et al. [352] that increases the accessibility of complex audio content for those with hearing impairments. Background audio levels would be reduced by removing the bird, cricket, and wind chime ambient sound objects; the aim of this process is to maintain participant immersion.

4.2.8 Procedure

All four participants took part in a two phase experimental session which lasted no more than 30 minutes. If required, participants would be accompanied by their parent/support worker, who would be present alongside the experimenter within the 3sixty space whilst the investigation took place. If present, the parents/support workers were allowed to communicate with the participant to provide assistance if they became distressed. However, they were not permitted to deliver instructions. Prior to the start of the session, the experimenter would explain the use of the motion tracking system, the structure of the experiment and place the headphones and tracking dots on the participant in order for them to become familiarised with the soundscape and device. The participant was informed that they could terminate the experiment at any time. The investigation would also finish if there were any signs of distress or discomfort.

Phase 1: Participants were permitted to explore the virtual and physical space unaided for a period of five minutes. During this time the participant was introduced to the 3sixty and the audio/visual system in the form of spatial audio and visual events that occurred at predetermined intervals. If a parent/support worker was present, verbal communication with the participant was minimised to control any influence that may of led to any guided exploration during the experiment.

Phase 2: Participants took part in a spatial audio game which required them to find hidden characters within the virtual world based on localising sounds that they emitted and limited visual representations. By tracking head rotation within the VE via the OptiTrack system, it was possible to determine if the participant was looking at the hidden character. Once this occurred the experimental system would record the time taken to locate the sound. In order to eliminate accidental findings, correct localisation of a sound would be recorded after 2 seconds of the participant looking at the virtual sound source. This two seconds would then be deducted from the localisation time. Once found, participants were encouraged to call over the found characters who would appear to fly towards them. This phase was guided by the experimenter, who would activate each hidden creature at a time when the participant was comfortable to do so. This was determined by asking the participant if they wished to continue. Once again parent/support worker communication was restricted as to avoid an affect on the outcome of the investigation. The length of this phase would be dependant on the participant's ability to progress. Once each virtual character was found the participant would be given positive reinforcement in the form of verbal praise.

4.2.9 Outcome Measures

Phase One: An accuracy metric (a) was used to evaluate participant performance during the spatial audio attention task. This was quantified by calculating the relative difference between the participant head rotation on the azimuth axis (θ_y) and the target position of the reproduced stimulus (θ_s) at 100 ms intervals.

$$a = 1 - \left| \frac{(\theta_s - \theta_y)}{180} \right| \tag{4.1}$$

Further data recorded the target azimuth location in degrees of each spatial auditory event with the respective start and end times, and virtual environment positions. Higher values of a during auditory stimulus playback time would represent greater spatial attention towards presented auditory targets.

Phase Two: Localisation times of auditory stimuli within the virtual environment during Phase Two were recorded for each participant for all eight virtual characters. Exploratory analyses of participant reaction time also examined the relationship between the amount of background ambient noises and spatial attention. To do so, localisation times for each character were compared based on stimulus type and background audio level.

4.2.10 Pilot Study Results

All data recorded from the investigation is located in the data folder supplied with this thesis, see Appendix B.

4.2.11 Phase One

Figure 4.8 displays the accuracy scores for all participants across Phase One of the experiment. With mean accuracy levels exhibiting above 0.5 across all auditory events it therefore indicates that participant head rotation was angled towards the sound's virtual azimuth position during it's presentation, thus confirming **H1**. When comparing results between social and non-social stimuli it can be seen that accuracy scores were higher for social stimuli, leading to the rejection of H_2. Further to accuracy scores, the experimenter observed that two of the participants would try and find crickets and frogs within the virtual environment, this was based on ambient spatial auditory cues and they had no visual representation.

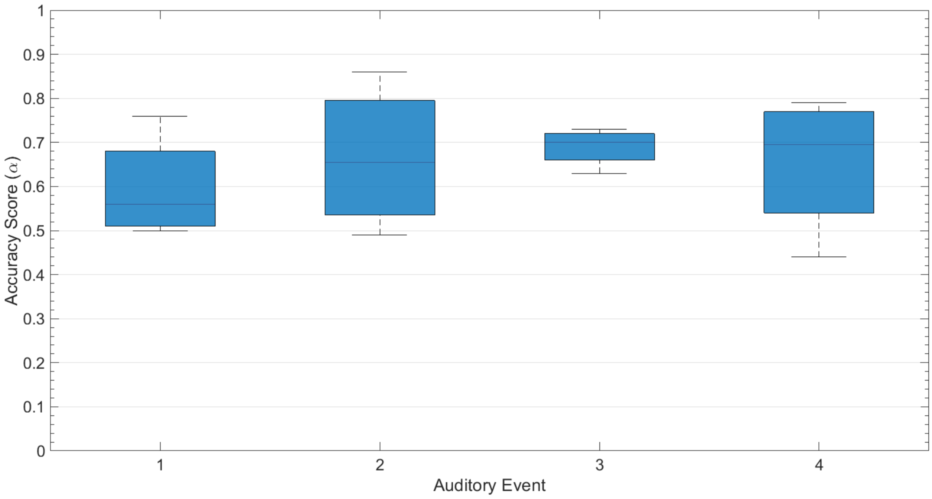


FIGURE 4.8: Box plot displaying accuracy scores recorded for all participants during Phase One.

4.2.12 Phase Two

To assess if ambient background noises may interfere with participant response to auditory cues and test H_3, the amount of time taken to locate each virtual character was recorded. Table 4.3 compares the time taken to locate each virtual character through localising their individual spatial audio cues, with Figure 4.9 displaying the mean total time. The results indicate that all participants localised the virtual character auditory cues faster when presented amongst the minimum level of ambient background sounds thus confirming H_3. These results also reject H_4 as the contributing condition to large differences in mean time is not the type of stimuli. If participants had speech sound selectivity impairments the time taken to locate characters with speech like stimuli would be higher for both background sound level settings.

Table 4.3: Time taken in seconds to locate virtual character through sound localisation

Auditory Event	**1**	**2**	**3**	**4**	**Mean -All Participants**
1	21.94	4.23	4.48	5.47	9.03
2	6.99	6.67	3.90	3.68	5.31
3	11.99	20.97	6.22	3.99	10.79
4	9.10	9.86	6.7	3.63	5.87
5	21.52	6.04	4.33	4.23	9.03

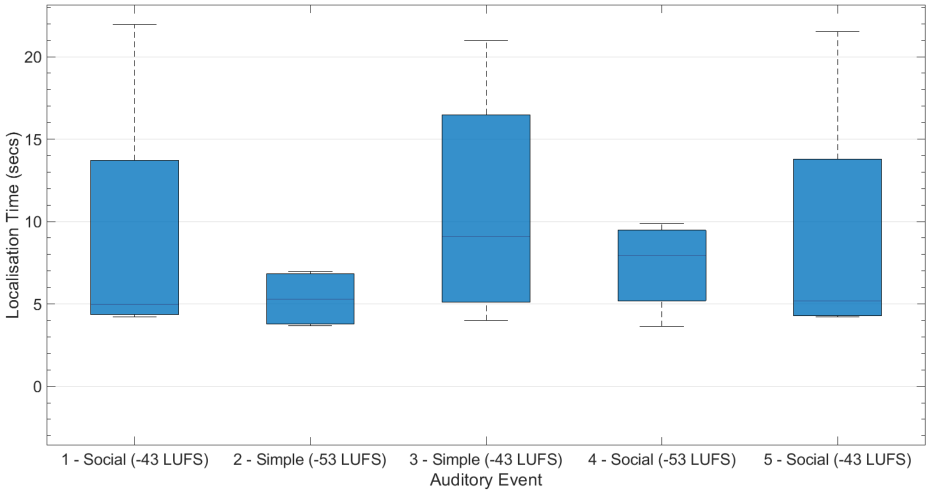


Figure 4.9: Box plot displaying localisation times recorded for all participants during Phase Two

4.2.13 Pilot Study Discussion

The results show an overall interaction with the presented spatial audio events despite varying levels of engagement between participants. Spatial audio events were able to influence user behaviour, attracting them to the appropriate areas of the virtual environment. The head rotation data collected from participants showed a general focus of attention towards auditory events, with a higher drive of user attention present during Event 4. Research has noted that autistic children display brain hyper-reactivity when exposed to novel auditory targets [353]. The combination of explosion and speech like stimuli was designed with this in mind and presents a unique audio cue that focuses attention towards it. This explanation is reinforced by the reduced attention towards Event 1, which was a musical clip (shown in Figure 4.8). However, the lower attention levels for Events 1, 2 and 3 could be explained by the fact that the visual accompaniments spanned a wider surface area which was inside the participants field of view. Therefore, they did not have to look directly at the sound source to observe the event as a whole.

Interestingly, the amount of focused head rotation in the direction of Event 4 could contradict earlier research that detected orientating deficits to speech sounds in autism [354]. This data could be explained by the moving visual stimuli that accompanied the spatial audio cue. However, in phase two there also appears to be no affect in attention to speech like stimuli with no significant differences in times to locate virtual characters B (bell) and D (speech like stimuli with bell).

Although all participants displayed an ability to determine the position of hidden virtual characters through sound localisation the results did seem to vary depending on the levels of ambient and environmental sounds. The mean time taken to find characters B and D, who had lower levels of competing audio, were shorter than those with higher. These results do support research that has surveyed difficulties in separating relevant auditory information in the presence of competing background noises [103]. These findings could have implications outside of this project when designing soundscapes for VR interventions for autism. It has been noted that when designing visual aspects of VR, the presence of visual clutter significantly affects user performance at completing tasks [345]. The same ideas should be applied to sound design. Important auditory information should be amongst a reduced level of background noise. Too much and the user may not attend to it, too little and the levels of immersion could suffer negative affects.

An interesting observation was the additional effect the 3D background environmental audio had on the user experience. Despite having no visual representation, the frog and cricket sound effects provoked verbal communication between participants 3 and 4 and the experimenter. After asking where the sound sources were coming from, both participants localised the sound sources and attempted to find the animals within the on screen VE. This finding indicates further that participants were able to engage with spatial audio within a virtual environment.

4.2.14 Pilot Study Summary

Results from this small group pilot study show that the participants displayed spatial attention towards the presented spatial auditory stimuli. Firstly, this is evident from the accuracy scores recorded during Phase one calculated using Equation 4.1, with the mean accuracy score for Event 3 being 0.7. Secondly, all participants were able to locate the virtual characters in Phase two using sound localisation. Further to this, background levels did have an impact upon localisation times as participants took longer to identify the positions of characters with a higher level of environmental audio. Finally, all participants engaged well with the virtual environment, showing interest in the visual aesthetics as well as environmental audio events. However, over the course of the study some limitations were observed altered the procedure and materials of the main study. These will be outlined in Section 4.3.

4.3 Main Experiment

As discussed in the previous sections, results from the pilot study provided evidence that autistic young people respond to spatial audio events, with this response being affected by competing background audio. Further to this, the virtual environment provides a robust experimental setting that can deliver auditory stimuli and record participant behaviour. However, there were a number of limitations to the pilot study which can be addressed in order to better inform the larger main experiment.

Firstly, conducting the study in the 3sixty space limits the amount of participants as it required parents to travel to the facility. The experiment was organised as part of an activity day in conjunction with the recruitment charity of which they have limited numbers of autistic participants. To address this issue the virtual environment was adapted into a VR application which permitted travelling to two special education schools, significantly increasing the sample size. An increased participant sample allowed for the addition of a second experimental condition group to serve as controls. The second group would listen to mono audio representations of the virtual audio environment allowing for a direct comparison with an Ambisonic rendering system, isolating the impact that spatial audio has on participant behaviour.

Another modification made based on the proceeding pilot study was the inclusion of more auditory events. As mentioned in Section 4.2.10 participants engaged with the environmental background spatial audio, which had no visual representation. With this in mind, Phase one would now introduce sounds events with no visuals. Thus increasing validity of the effect spatial audio has on the research outcomes.

4.3.1 Hypothesis

It is hypothesised that subjects exposed to binaural-based spatial audio show more spatial attention to auditory stimuli compared with the mono sound control group, indexed through more precise head orientation towards the presented audio targets. Furthermore, spatial audio stimuli will influence participant positional movement within the virtual environment. Additionally, based on previous studies which investigate difficulties in auditory scene analysis experienced by autistic people, it is predicted that localisation times will be higher for audio cues with the most amount of competing environmental background noise. Finally, by examining spatial attention towards complex speech and non-speech sounds, it is possible to investigate if differences in spatial auditory attention are the consequence of a reduced orientation towards social stimuli.

4.3.2 Main Study Design

This investigation followed a between-subjects design. Participants were randomly allocated to either a 3D spatial audio, or control group. The 3D audio group would be exposed

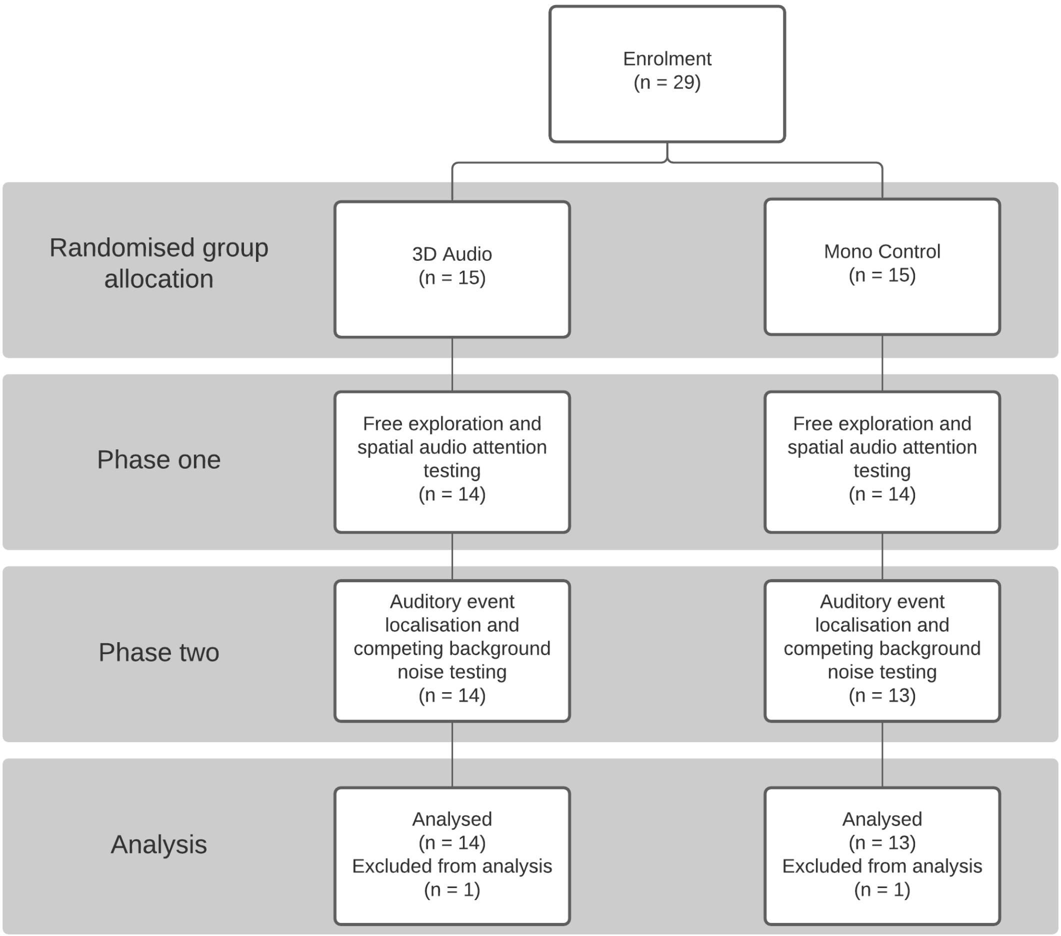


Figure 4.10: Flow chart displaying experimental procedure.

to spatial audio reproduced binaurally over headphones. The control group would listen to a monaural representation of the virtual auditory scene presented to both ears over headphones. Each participant group completed a two-phase experiment which first tested spatial audio attention, followed by a sound localisation task. Both within a multi-modal VR environment (see Figure 4.10).

4.3.3 Participants

The experimental group consisted of 29 children and adolescents (27 male and two female, mean age = 14, SD = 2.42, range of 9–19 years). Participants were recruited from two special education schools and one charity. All participants had a formal diagnosis of autism spectrum disorder obtained from their local national health trust, displaying social, cognitive, and motor abilities that allowed them to interact with the experimental software and equipment. Exclusion criteria were self-reported hearing problems and an inability to finish the task. Two participants were excluded due to not being able to complete the full experiment. This was due to them both wanting to remove the VR equipment and stop the experiment. Participants were admitted into the study after informed consent and assent was obtained from their parent(s) or legal guardian(s) which can be seen in the accompanying data outlined in Appendix B.

4.3.4 Virtual Environment

The visual virtual environment remained the same as that found in the pilot study. However, additional auditory stimuli were implemented into both phases of the experiment. These additions are outlined below.

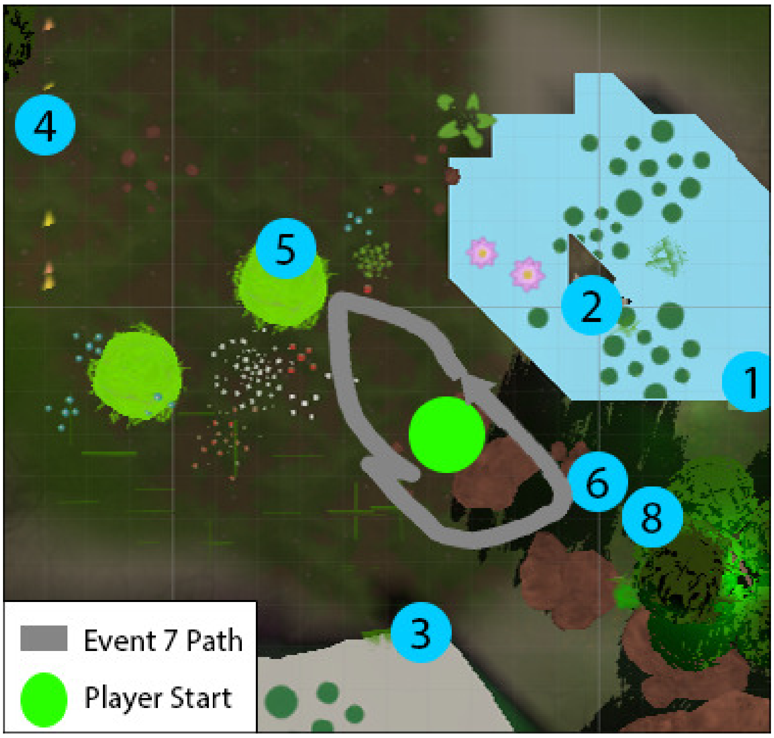


FIGURE 4.11: Top-down view of auditory stimuli positions within the virtual environment during Phase One. (1) MIDI-based music, (2) bell sound effect, (3) bubble sound effects, (4) Explosion and speech like stimuli, (5) MIDI-based music, (6) Crow sound effect, (7) Explosion and speech like stimuli, (8) Explosion and speech like stimuli.

Auditory Stimuli: Figures 4.11 & 4.12 are overhead views of the virtual environment displaying the auditory stimuli placement and movement area, with details of each audio

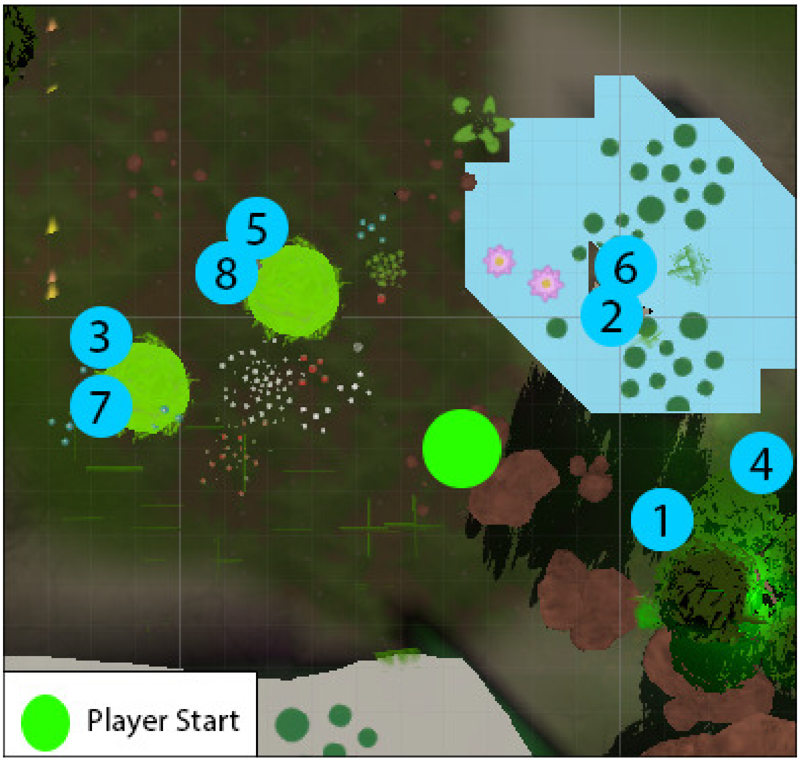


FIGURE 4.12: Top-down view of auditory stimuli positions within the virtual environment during Phase Two.

object in Tables 4.4 and 4.5. The default level for environmental background audio was −53 LUFS.

Phase One: Eight spatial auditory events (see Table 4.4) were placed in the VE. All audio events were stationary with the exception of Event 7 which moved around the entire environment over the course of 30 s. To minimise the effects of any complications in auditory scene analysis associated with ASD [103], environmental audio was reduced to −43 LUFS while each Phase One event was rendered.

Phase Two: Eight spatial audio events representing virtual characters were placed throughout the VE. To further investigate auditory detection in noise, target stimuli were played at differing signal-to-noise ratio. For auditory events with +10 SNR, the stimulus level would be 10 LUFS higher than the background noise level in order to reach +10 SNR. For auditory events with 0 SNR, the stimulus would be set at an equal level to that of the background noise.

Table 4.4: Phase One Auditory Stimuli descriptions.

Auditory Event	Stimulus	Visual Cue
1	MIDI-based music–Synth Pad	Present
2	Bell sound effects	None
3	Bubble sound effects	Present
4	Explosion & Speech-like stimuli	None
5	MIDI-based music–Synth Pad	Present
6	Crow sound effects	None
7	Explosion & Speech-like stimuli	Present
8	Explosion & Speech-like stimuli	None

Table 4.5: Phase Two Auditory Stimuli–SNR represents either the stimulus played at +10 LUFS above the background audio level, or set to a level equal to that of the background audio.

Auditory Event	Stimulus	Background Sound Level (LUFS)	SNR
1	Simple	≥ -53	+10
2	Social	≥ -43	+10
3	Social	≥ -43	0
4	Simple	≥ -43	0
5	Simple	≥ -53	0
6	Simple	≥ -43	+10
7	Social	≥ -53	0
8	Social	≥ -53	+10

4.3.5 Equipment

All audio and visual stimuli were rendered using the Oculus Rift CV1 HMD with built-in on-ear headphones. Head tracking was also achieved using the HMD, with motion tracking of participant position calculated by the Oculus Rift sensors [355].

4.3.6 Procedure

The experiments took place within the participant's recruitment school. Participants were permitted to move freely around a pre-defined tracked experimental space of 1.6 m × 1.6 m while wearing the HMD. The combination of dynamic head and positional tracking allowed movement around the virtual environment with 6DoF.

Throughout the experiment a support worker was present to communicate with the participant and provide assistance if they became distressed. However, they were not

Figure 4.13: The Oculus Rift CV1 HMD, Oculus Touch controllers and Oculus Rift Sensors

permitted to deliver instructions. Prior to the start of the session, the experimenter explained the use of the VR system, the structure of the experiment and placed the HMD on the participant in order for them to become familiarised with the device.

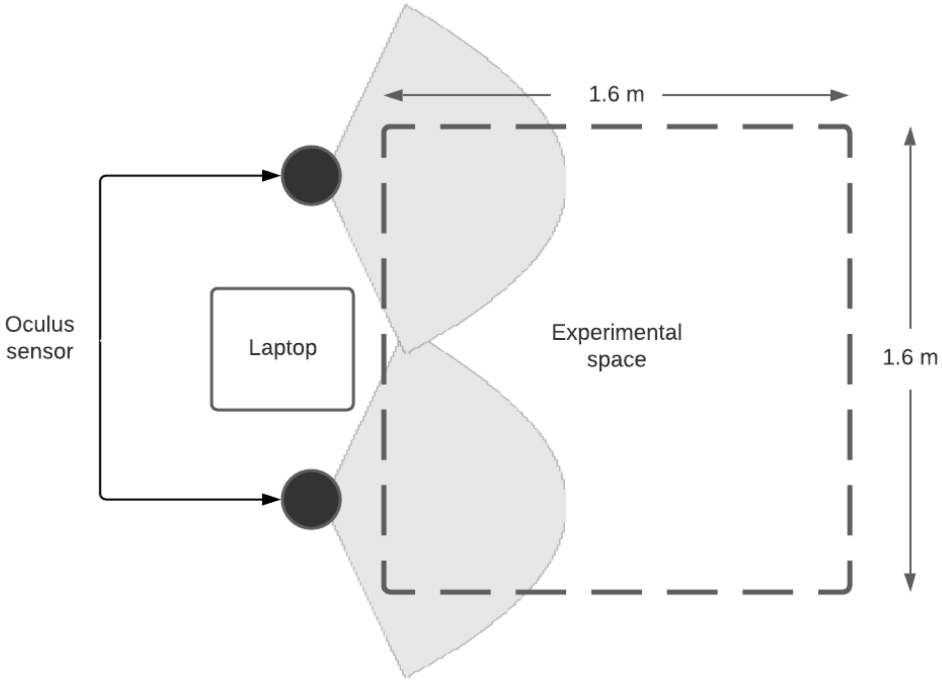


Figure 4.14: Experimental equipment layout, displaying placement of laptop and Oculus Rift sensors with 100 ° field of view.

Phase One—Free Exploration and Spatial Audio Attention Testing

Prior to the main experimental task in Phase One, subjects were introduced to the VR equipment and environment for a period of two minutes with minimal environmental audio and visuals. The experimenter explained to the participant that they were free to move around and explore the VE. This brief initial time was used to provision for any emotional excitement experienced by using VR. The participant also received support from the experimenter to become familiar with the virtual environment. Following this, participants were given a further five minutes free exploration, during which they would be exposed to eight auditory events played once at pre-determined times. Four of which had visual accompaniments and four had no visual accompaniments (see Table 4.4). This period was used to record participant head rotation and horizontal tracked positional data. Once again the experimenter told the participant that they were free to move around and explore the VE. After this, verbal communication with the participant was minimised to control any influence that may lead to any guided exploration during the experiment.

Phase Two—Spatial Audio Localisation and Background Noise Testing

Participants took part in a localisation task which required them to locate a total of eight hidden characters within the virtual world based on localising sounds that they emitted and limited visual representations. Each would play either social or simple stimuli (see Table 4.5).

The experimenter first explained the task clarifying that the participant must use their ears to find each character. To aid the participant in becoming familiarised with the target stimuli, they were presented with an example character in their field of view which plays both social and simple auditory stimuli. Following this, the experimenter activated each hidden character at a time when the participant stated they wanted to continue and repeated the explanation of the task. This followed the same data recording technique as the pilot study in which correct localisation was established after the participant looked at the virtual sound source for 2 seconds.

Once again support worker communication was restricted so as to avoid an effect on the outcome of the investigation. The length of this phase would be dependent on the participant's ability to progress. When each virtual character was found the participant was given positive reinforcement in the form of verbal praise.

4.3.7 Outcome Measurements

Phase One: Head rotational accuracy remained part of the outcome measurements for Phase one. Further to this, the use of the Oculus Rift sensor system allowed for accurate positional tracking within the virtual environment. This was implemented to better understand if spatial audio influences subject behaviour. Tracked horizontal positional data within the VE was represented by x and z and also collected at 100 ms intervals. A distance score was then calculated by measuring the distance between the participant and each of the presented auditory stimuli during playback time. Smaller distance values during auditory stimulus playback time would indicate movement towards the presented stimuli, indicating an influence of participant movement.

Phase Two: Localisation times of auditory stimuli within the virtual environment during Phase Two were recorded for each participant for all eight virtual characters. Further to group comparisons, exploratory analyses of participant reaction time also examined the relationship between the amount of background ambient noises and spatial attention. To do so, localisation times for each character were compared based on stimulus type, background audio level and SNR.

4.3.8 Main Results

All data recorded from the investigation is located on the data folder supplied with this thesis, see Appendix B.

Phase One Spatial Audio Attention Testing: A two-way mixed ANOVA was conducted to investigate the impact of 3D audio and auditory stimuli on the accuracy metric during Phase One. A comparison between test groups yielded a significant difference ($F(1, 26) = 29.43$, $p \leq 0.001$, $\eta_p^2 = 1.642$) across all auditory events, with the 3D audio test group ($M = 0.661$) scoring higher accuracy scores than the mono audio controls ($M = 0.487$). Figure 4.15 shows a comparison of group mean accuracy scores with 95% confidence intervals for each event.

In addition, there is a significant effect of stimulus type on spatial attention of both groups ($F(7, 175) = 7.657$, $p \leq 0.001$, $\eta_p^2 = 0.227$), with a significant interaction between groups and stimulus type ($F(7, 175) = 2.426$, $p = 0.021$, $\eta_p^2 = 0.072$). Finally, the use of visual

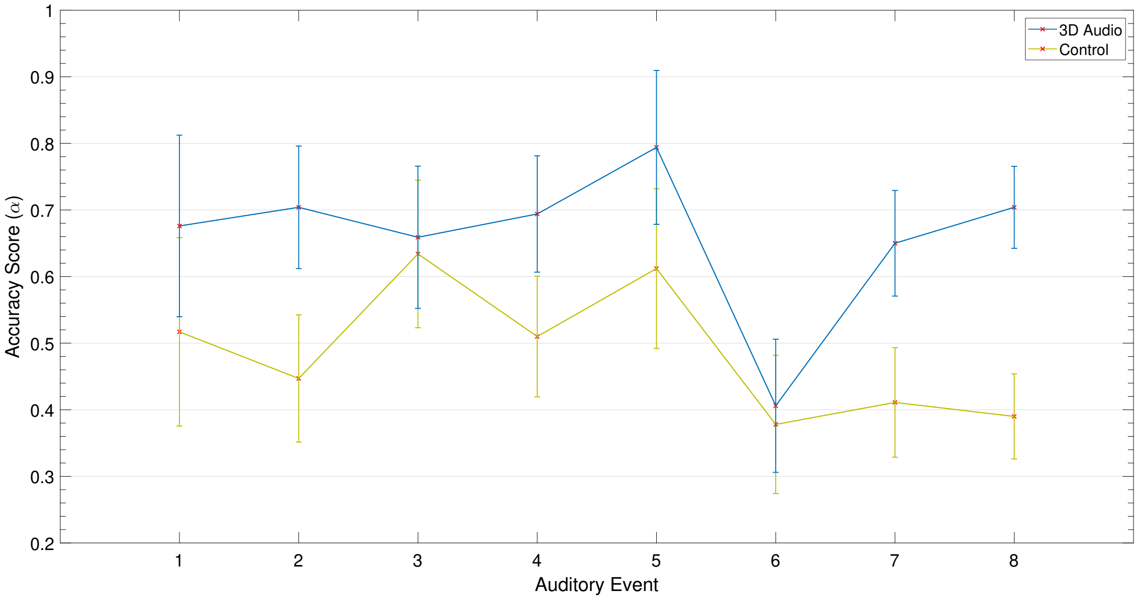
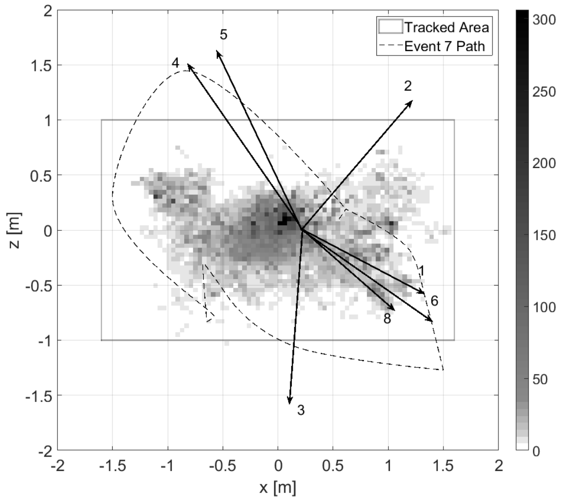


FIGURE 4.15: Mean accuracy (α) scores for each audio condition across all auditory events in Phase One.

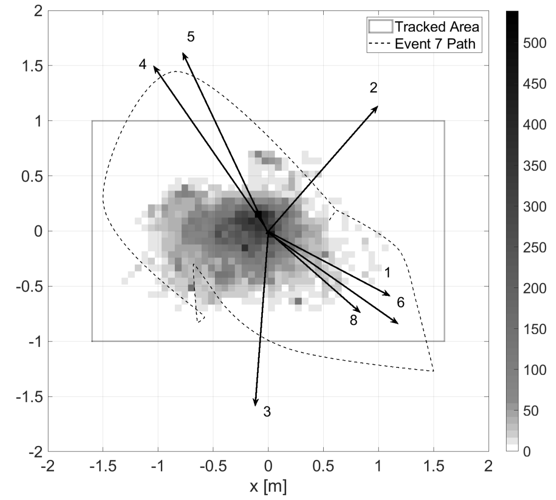
accompaniment had no significant effect on the accuracy of spatial attention in the spatial audio group yielded by visual accompaniment ($M = 0.627$ $SD = 0.187$) and no visual accompaniment ($M = 0.695$ $SD = 0.184$) paired-samples t-test; $t(55) = -1.903, p = 0.062$. However, there was a significant difference between visuals ($M = 0.543$ $SD = 0.235$) and no visuals ($M = 0.431$ $SD = 0.176$) within the mono control group; $t(51) = -2.827, p = 0.007$.

To better understand behavioural response to spatial audio within the virtual environment, tracked positional data across the two groups was also analysed. The tracked positional coordinates of all participants on the horizontal axis collected throughout the entirety of Phase One, these are visualised as 2-dimensional density plots shown in Figure 5.1. In addition, the figure also displays the directions and distance of all auditory events inside and outside of the tracked area.

The figure shows differences in the extent of exploration depending on the auditory condition. Despite both groups showing a condensation of exploration around the centre of the tracked space, the 3D audio group show a distribution of exploration in the direction of the majority of auditory targets. Further analysis of the distance between each participant and the auditory target using a two-way mixed ANOVA also yielded a significant difference between the auditory conditions ($F(7, 175) = 23.111$, $p \leq 0.001$, $\eta_p^2 = 0.308$). Participants within



(**a**) 3D audio condition



(**b**) Mono condition

Figure 4.16: Two-dimensional density plot of positional tracking data throughout Phase One for all participants. Arrows point toward the target audio objects outside of the plot area. Distance from origin to audio objects—1 (6.8 m), 2 (4.57 m), 3 (5.98 m), 4 (12.43 m), 5 (5.96 m), 6 (3.86 m), 8 (5.74 m).

the 3D audio group moved closer towards the virtual sound sources during their respective playback times.

Phase Two Spatial Audio Localisation and Background Noise Testing: A two-way mixed ANOVA showed that these variations in localisation times were also significantly different ($F(7, 175) = 7.565$, $p \leq 0.001$, $\eta_p^2 = 0.232$). Figure 4.17 shows the mean localisation times for both experiment groups across all auditory events in Phase Two, which shows shorter localisation times in the 3D audio condition ($M = 8.4$ s) than in the mono control condition ($M = 17.8$ s).

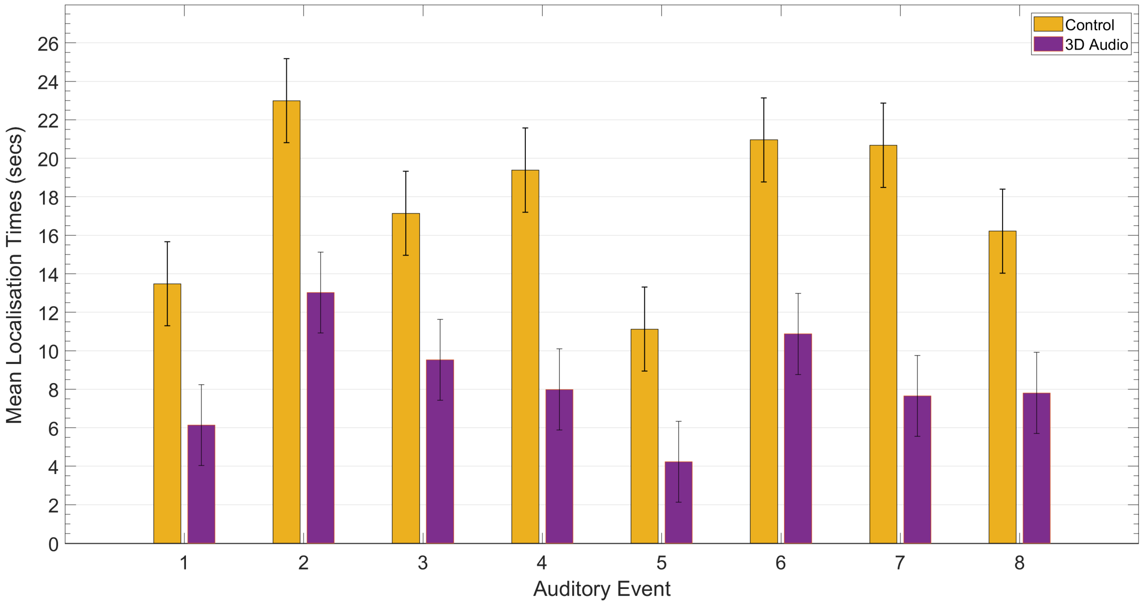


FIGURE 4.17: Mean localisation times for each audio condition across all auditory events in Phase two. The whiskers denote 95% confidence intervals.

A HLM was used to analyse the effects background audio, SNR, and type of stimulus have on localisation times within both the 3D audio and mono control conditions, the results of which can be seen in Figure 4.18. Inferential statistics reported a significant overall effect of background audio on the time taken to locate each auditory event for both groups: 3D audio ($F(1, 54.713) = 15.243$, $p < 0.001$) and mono audio ($F(1, 84) = 10.476$, $p = 0.002$). Auditory stimuli within the higher background audio level of -53 LUFS (3D audio $M = 10.4$ s and mono audio $M = 20.9$ s) displayed higher localisation times than those with -43 LUFS (3D audio $M = 6.5$ s, mono audio $M = 15.4$ s). Further to this, the effect of SNR between background sounds and auditory stimulus was only significant for the 3D audio group ($F(1, 71.029) = 5.206$, $p = 0.026$), with the time taken to correctly locate auditory stimuli $+10$ LUFS above the background audio level ($M = 7.4$ s) being lower than those with no SNR ($M = 9.5$ s). However, the background audio and SNR interaction was not significant for the 3D audio group ($F(1, 71.729) = 1.3586$, $p = 0.248$). Finally, the type of auditory stimulus yielded a significant effect on the results of localisation times for both groups: 3D audio ($F(1, 68.778) = 5.545$, $p = 0.021$) and mono audio ($F(1, 84) = 4.930$, $p = 0.029$), showing participants taking longer to locate social stimuli (3D audio $M = 9.5$ s, mono controls $M = 19.33$ s) than simple ($M = 7.3$ s, mono controls $M = 16.17$ s).

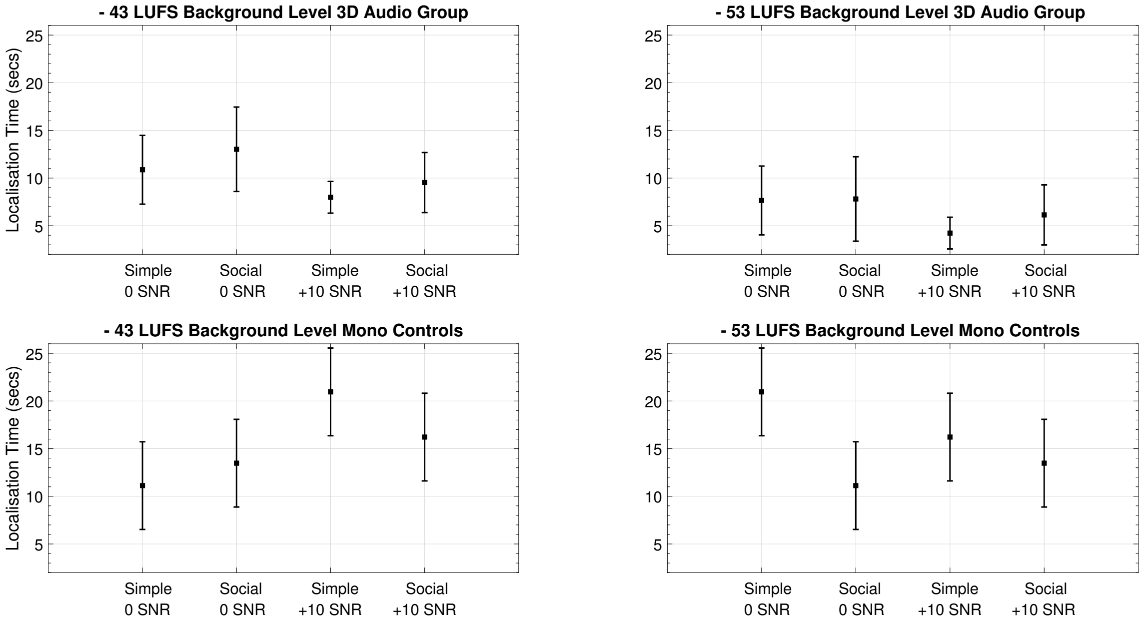


Figure 4.18: Estimated means data from HLM, comparing the effect of background audio, signal-to-noise ratio, and stimulus type for binaural spatial audio and mono control groups. The whiskers denote 95% confidence intervals.

4.4 Main Study Discussion

4.4.1 Phase One: Spatial Audio Attention Testing

Throughout Phase One, subjects in both groups displayed overall attention towards the presented auditory targets; however, those within the 3D audio group showed higher levels of head orientation in the direction of audio events. These results demonstrate that despite reported impairments in interpreting ITD and ILD binaural cues associated with autism [356], spatial audio was capable of accurately attracting participants to the areas of the virtual environment applicable to the auditory stimulus's location.

Autistic individuals have been recorded as having similar evoked brain responses to speech-like stimuli than their NT peers, but at the same time exhibit less involuntary attention towards it [354], [357]. With this in mind, it would be expected that under the test paradigms of free exploration that measures involuntary attention, accuracy scores for speech stimuli to be the lowest. However, the findings from Phase One of this experiment are contrary to this for both condition groups, with the auditory event with the lowest scores being the crow (Event 6). This auditory stimulus type would appear to be similar to the ambient environmental audio track and despite the lack of competing audio present the event was not unique enough to warrant participant attention. In the case of attention towards speech-like sounds in Phase One, these events are a combination of explosion sound effects and non-human speech-like stimuli. These were designed to create a unique audio cue that attracted attention towards them. Furthermore, studies have observed hyperactivity in the brain with response to novel auditory stimuli in children on the autistic spectrum [353].

Research by Lokki and Grohn [358] revealed that despite audio cues being less accurate for sound source localisation than accompanying visual cues, within a virtual environment audio-only cues are almost equally effective as visual only cues. In addition, when objects were placed outside of the participant's field of view or subjected to occlusion of other visual objects, audio-only cues were more successful in the preliminary phase of object localisation. This could serve as explanation for the significant effect visual accompaniments had on the overall accuracy scores for participants in the mono audio group. The visual representations would compensate for the spatial inaccuracies of the audio rendering. This would also account for the similar scores present in both groups for Event Three (e.g., bubble sound). Event Three incorporated visuals that spanned a large area of the VE which would attract user attention once it arrived into the participant field of view.

Exploration of the virtual environment is also illustrated in the positional data density plot (see Figure 5.1). For both conditions, tracked horizontal movement was concentrated close to the starting position of $(x, y = 0, 0)$. Nonetheless, participants listening to spatialised audio displayed clear movement in the direction of both the auditory only and audio-visual cues. Due to the grouping of both audio-visual and audio-only events it cannot be confirmed which events encouraged participant exploration in those areas. However, the reduced amount of tracked data in the direction of event three does suggest that the larger visual stimuli negate the need for participant exploration of the VE.

4.4.2 Phase Two: Spatial Audio Localisation and Background Noise Testing

Alongside comparing localisation times between auditory test conditions, Phase Two of this study was carried out to determine if reported impairments in decoding speech when background noise is present would have quantifiable effects on how they would respond to similar stimuli rendered via virtual spatial audio within a VR environment. Considering the heterogeneity of abilities within the sample group, it is encouraging to see that all participants were capable of performing this task without any particular difficulties. However, the differences in localisation times between participants within the experimental groups may also provide explanation for the low effect size yielded by the two-way mixed ANOVA.

Firstly, results in Phase Two showed those in the 3D audio group performing significantly better in the localisation task. These results are consistent with similar research conducted with neuro-typical participants, reporting that three-dimensional auditory displays can reduce localisation times and improve audio spatial attention in VR [320]. Furthermore, these results provide more evidence that autistic people are capable of successfully interpreting binaural information to take advantage of the spatial acuity provided by spatial audio rendering techniques.

In regards to the effect of background levels on selective hearing, results showed that localisation times were significantly longer when the level of background noise increased and/or the SNR between the auditory target and ambient noise decreased. Furthermore, the type of stimuli also had a noticeable impact on localisation times, with participants taking longer to correctly locate speech-like stimuli. This data is comparable to evidence of speech processing difficulties in autistic children in non-spatialised audio test conditions [112], [117], [357]. Although the results from the mono control group would be sufficient in evaluating the effects of background noise on the response to target audio, mono audio is rarely used with VR environments. Therefore, for this study it was important that competing background audio had a significant effect on both experiment conditions. Autistic individuals tend to demonstrate elevated speech perception thresholds, poor temporal resolution and poor frequency selectivity when compared to their NT peers [83]. The poor frequency selectivity alone could account for poor localisation along the vertical plane [112], as this is primarily attributed to the analysis of the spectral composition of the sound arriving at the ear [359]. However, a combination of all three alongside a diminished ability to correctly translate ILD and ITD sound cues would account for poor selective hearing leading to difficulties with behavioural recognition of auditory targets in background noise [83], [103].

4.4.3 Limitations

One limitation of this study is the lack of a NT control group. Comparison data may have highlighted any potential differences in performance between participants with autism and neuro-typical counterparts. Nonetheless, participants in this study still showed higher spatial attention towards binaurally rendered spatial audio, demonstrating quantifiable evidence of audio spatial interactions.

It is important to note that the balance between male and female participants is unequal. This is also the case in the investigations presented in Chapters 6 and 7. Literature documents that the prevalence of autism displays a bias towards males at a ratio of 3:1 [360]–[363]. With no medical evidence suggesting a possible cause for this disproportionality, it has been suggested that a male-centric diagnosis criteria lacks sensitivity to the female phenotype [361], [363]. The split between male and female participants in this and other studies in the thesis is a result of this. Furthermore, it could be seen as a direct representation of the population demographics within the recruitment organisations.

Another important consideration is the use of a non-individualised HRTF database within this study. The SADIE database used within Google Resonance incorporates anechoic recordings using the Neumann KU-100 dummy head which is-based upon average dimensions of an adult human head and ears [275]. HRTFs differ greatly between individuals, due to the unique shape of the pinnae, head, and torso. Therefore, the use of a generalised HRTF can sometimes result in localisation confusion and reduced externalisation of target auditory events within a virtual environment [364], [365]. In addition, recent research which compared individualised to non-individualised HRTFs did observe higher localisation errors when using a generic database [310]. Finally, this may provide some explanation for the varying accuracy and localisation times between participants within the 3D audio experiment group as well as the lower effect scores yielded in the statistical analysis. However, there are significant challenges involved in obtaining individualised recordings, with the recording of children's HRTF causing additional difficulties [366]. This warrants the use of generic dummy head recordings in the development of VR soundscapes [367].

Future investigations could also build upon the work carried out by Wallace et al. [235] and Newbutt *et al.*[219] and measure self-reported presence of autistic participants within virtual environments that use spatial audio rendering techniques. Changes in behaviour, spatial focus, and interactions when altering the audio reproduction could be compared. Furthermore, studies could be conducted to measure if removing aspects of the background audio has any negative effects on the feelings of presence felt by the participants.

4.5 Summary

Alongside realistic graphical rendering and natural approaches to computer interaction, spatialised audio can be used to significantly enhance presence and improve user interaction within a VR environment [324]. In terms of clinical applications within autism research, presenting a more realistic experience would benefit interventions such as the treatment of phobias and vocational training. By matching the visual and auditory sensory experiences of the real world, users will have a greater chance of generalising newly acquired skills into their everyday lives, therefore increasing the possible positive outcomes of VR-based therapy [259]. Within the context of this thesis, presenting a more realistic auditory experience could benefit exposure based therapy for auditory hypersensitivity.

To date, this study is the first to assess how those with autism respond to virtual spatial audio within a VR game environment. The main experiment involved 29 individuals diagnosed with autism being randomly allocated to 3D audio and mono control groups. All participants completed two experimental phases. The first compared spatial attention towards auditory stimuli by recording head rotation accuracy and horizontal movement towards presented audio. In Phase Two, participants were required to localise four social and four simple stimuli among varying levels of background audio. Results collated from head rotation and tracked positional data during the experiment do suggest that despite reported auditory processing difficulties, autistic children and adolescents can make use of auditory cues rendered by spatial audio. In addition, the amount of background audio can have significant effects on the localisation of auditory stimuli within a VE. These findings will provide valuable evidence which will inform the sound design techniques used in the forthcoming investigations presented throughout this thesis. Furthermore, the results of this research could contribute to those designing VR interventions for the autistic population. Future developers and researchers should make use of similar binaural-based spatial audio rendering approaches used in this study to increase the ecological validity of the virtual environment and deliver important information via auditory cues. However, attention should be paid to the level of background audio, as too much could have a negative impact on the experience.

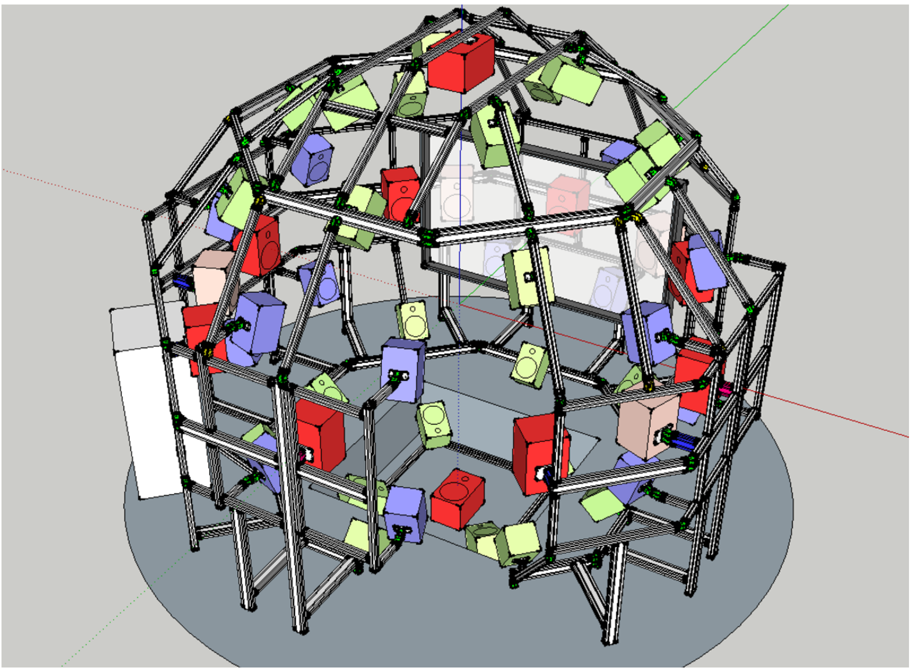
This study has shown that previously reported difficulties in auditory scene analysis associated with autism do extend into the realms of VR and binaural-based spatial audio. The amount of competing background audio can have an effect on spatial attention towards virtual sound targets and therefore this should be taken into consideration when designing three-dimensional virtual acoustic environments for ASD interventions.

Chapter 5

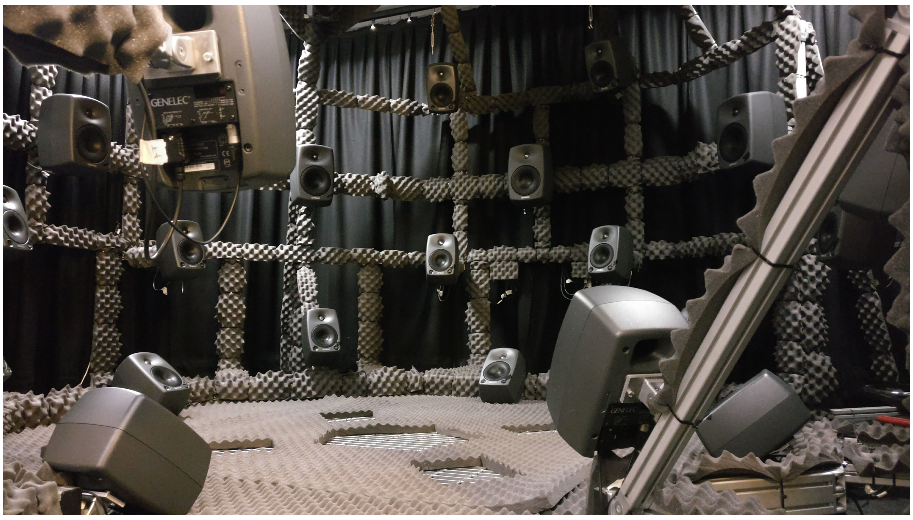
Development of SoundFields: A Serious Virtual Reality Game for Autistic Young People

5.1 Introduction

Previously in Chapter 2 it was discussed how Ambisonic audio could be embedded into exposure therapy frameworks for auditory hypersensitivity. Ambisonic spatial audio can be rendered to the listener through either loudspeaker array or binaurally via headphones, and so this presents an important consideration when designing a therapy framework for auditory hypersensitivity. Work conducted by Argo [337] has shown that immersive soundscapes presented over loudspeaker arrays are capable of eliciting the levels of anxiety needed to address mental catharsis in neurotypical adults. With this in mind, it could therefore be possible to conduct systematic desensitisation sessions within the 50 loudspeaker array located at the University of York (see Figure 5.1). This system is capable of rendering up to 5th order Ambisonics and so could simulate immersive soundscapes with the high spatial accuracy required to trigger the fear structure needed for effective exposure training. However, alongside delivering realistic auditory stimuli it is also important to motivate autistic people to engage with repeated intervention sessions if the therapeutic outcomes are to be achieved and subsequently generalised to real world applications. This could be



(**a**) Sketch up plans of 50 speaker loudspeaker array.



(**b**) Photograph taken at entrance to the loudspeaker array.

FIGURE 5.1: 50 speaker loudspeaker array located at the AudioLab, University of York.

achieved by utilising head-tracked binaural ambisonic rendering within a VR and serious game environment.

Earlier in this thesis, serious games were examined as a viable approach to providing therapy farmeworks to autistic people. For over a decade researchers have been using these interventions to improve emotional recognition, social skills and reduce anxiety including that associated with sounds [29], [30], [368]. With their effectiveness accredited to a number of factors. Firstly, autistic individuals have a tendency to enjoy video games and technology.

Secondly, the individualised VE creates a safe therapy condition to acquire and practice new skills. Finally in the context of VR based interventions, realistic models of feared stimuli can be delivered in a controlled manner at graduated levels of exposure.

The aim of this chapter is to outline a serious game targeting auditory hypersensitivity in autistic children and adolescents named SoundFields. First, a series of literature based guidelines are examined which outline the specific design and implementation requirements of a serious game for those on the autistic spectrum. This is then followed by a detailed description of the game including the therapeutic mechanisms, game-play, and sound design techniques used specifically to overcome any auditory impairments that may affect auditory scene analysis in VR. In addition, the tools used to produce the stimuli used for exposure are outlined. Finally it is discussed how the presented game complies with the design guidelines presented at the beginning of the chapter.

5.2 Design Guidelines for Serious Games and Autism

Literature suggests a series of guidelines that should be integrated into the development of serious games to increase engagement and therapeutic outcomes of an intervention. These features are founded in computer game principles which are used to entertain and engage the user [369]. These can incorporate game design features such as reward systems [66], [67], difficulty [370]–[372], narrative based goals [373]–[375], player personalisation [376] and audio [377], [378]. This section will present these guidelines in addition to examples of their implementation in previous serious game therapy applications developed for autistic young people.

5.2.1 Repeatability and predictability

One of the main characteristics of autism identified in the DMS IV [1] is that people within this population show repetitive behaviours [379]. This is noted as one of the principle reasons why they engage well with interactive technologies [380]. Moreover, repetition is crucial in providing opportunities to practice and master new educational and therapy skills. Finally, getting the player to repeat the same task delivers predictable feedback that will maintain interest and motivation in autistic individuals [178], [381]. *New*

Horizons is a serious game developed to reduce anxiety in autistic children [382]. The game permits players to repeat mini-game levels as many times as they please. Furthermore, the therapeutic goals remain the same regardless of the level and mechanics. Barajas *et al.* [204] follow a repetition mechanic in a game targeted to improve cognitive skills of autistic children. During the game, players must repeatedly place physical blocks in a pattern that mirrors an on screen display using a bespoke game controller. In a serious game designed to teach autistic people work based skills, players are required to repeatedly practice vocational activities such as planting crops and wrapping fruit within a time limit [381].

5.2.2 Reward System and Motivator

Much like real-world play, people play computer games to generate a positive feeling [66]–[69]. A human being's own subjective pleasure experiences is mediated by a set of complex neurological processes specific brain pathways known as the reward system [383]. Neurons within the reward system communicate by releasing dopamine, a chemical which plays a key role in how we feel pleasure [384]. These reward circuit responses have been linked to a diverse group of both simple and complex stimuli. At its most basic level, increased reward system activity has been observed in the reception of unconditioned stimulus such as water, fruit juice and food smells [383], [385], [386]. Further to this, studies have reported similar reward system activation in response to beautiful faces [387], social interaction [388] and music [389]. However, these stimuli are conditioned and therefore any reward status is the result of learning [390]. Similarly, research has also detected levels of dopamine release during sessions of computer game-play [68], [391].

Computer games incorporate reward mechanics aimed at producing feelings of fun by creating a sense of anticipation before the award is received [392]. Mechanics which have also been adopted by digital intervention games in order to increase motivation and engagement, resulting in the development of the target and educational therapy outcomes [393], [394].

Score systems are one of the earliest approaches to reward mechanics within games. Presenting a player with scores that directly link to in-game performance can serve as a tool for self-assessment and feelings of glory [395]. Research conducted by Malone

[396] identified that a basic scoring mechanic is one of the core features crucial to fun. Additionally, a persistent scoring system can be extended to interact with other reward mechanics such as the acquisition of in-game resources [392]. Classified as a sustenance reward, collection of resources allow the player to upgrade their avatar or virtual status [395]. Although this technique for rewarding players is a popular motivator in online multi-player games, it is also present within serious games developed for autistic children [397]. In *Caribbean Quest* [398], a game which targets improving working memory and attention, players are rewarded with ‘sand dollars’ for successfully completing in-game tasks. This in game currency can then be used to purchase trophies in a virtual game shop.

How a reward is delivered to the player is an important consideration when developing a serious game intervention for autistic children. For an individual who may have impaired cognitive abilities a quantifiable scoring system may be redundant. It has been observed that autistic children benefit from positive reinforcement in the form of audio or visual feedback, such as animations, applause or cheerful music [200]. *New Horizons* [382] which is an intervention developed to reduce anxiety in autistic players adopts an inclusive motivator approach. In addition to using a quantifiable point system, players are also rewarded positive reinforcement using visual animations and audio.

5.2.3 Customisation

The customisation of a virtual environment is a popular feature of modern day computer games. By allowing a player to customise in-game features to their own specification and preference, the game can increase the sense of immersion [399]. Computer games utilise a number of approaches to customisation. First, cosmetic customisation allows the player the change the appearance of characters and objects which does not directly impact game-play [400]. Second, customisation of character statistics and abilities can affect game-play and how challenges are overcome [401]. Finally, a player can change in-game settings such as difficulty in order to individualise the experience based on their own ability [400].

The type of customisation is very much dependant upon the style of game being played. Simulation games such as *The Sims* [402], one of the most popular simulation games in the past two decades [403], are built around requiring the player to create virtual characters and design the worlds in which they exist. Further to this, multiplayer online games such as World of Warcraft [404] and Destiny [405] allow full customisation of in-game avatars using both cosmetic and in-game abilities. How a player decides to represent themselves within a virtual environment may indicate how they expect the avatar to act as a channel of their own self-identity [406]. This personal identification with an avatar is one of the core aspects of how a player experiences a game, impacting their engagement and enjoyment [376], [407].

Personalisation of avatars has been utilised in a number of games developed for autistic children [408]–[412]. This has been shown to increase the perception of self felt within the game environment and helps in building a relationship between the autistic individual, the VE and game mechanics [380]. One notable example is *Social Stories*, a game designed to teach autistic children road safety. Photographs of the players which were placed on the character faces were observed to have a positive impact on the child's learning experience [411].

As previously mentioned, a games difficulty can be customised to adapt to the player's own abilities. This is an important consideration to make when developing serious games for those on the autistic spectrum. Autism is heterogeneous, and despite impairments being commonly identified with the condition, abilities are often unique to the individual [413]. It is therefore important that the game is capable of adapting to the player's individual skills and their own therapeutic or educational goals [380]. An example of this appears in work conducted by Chang *et al* [414], in which a serious game was developed to provide children with autism opportunities to practise everyday skills such as setting the table. Game activities were initially presented to the player at random. Upon completion, the child's performance metric informed the category and complexity of the proceeding task. This allows the experience to be customised to each child's unique skills and abilities.

5.2.4 Game Difficulty

Increasing game task difficulty is one of the foundations of computer game design. By challenging the player and implementing a form of progression, the player will maintain interest and stay motivated to perform better [370]. Furthermore, it is considered that challenge is one of the core components in delivering a fun and immersive experience [371], [415]–[417]. Feil and Scattergood [418] suggest that how a computer game challenges a player can be classified into six groups. Time based challenges allow players a limited amount of time to complete a task. In dexterity challenges, players must use physical dexterity or mental ability to make quick decisions. Memory and knowledge challenges necessitate the player to remember certain facts throughout the game. Logic challenges test player's intelligence with puzzles. Resource management challenges the player to control limited resources to complete tasks. Finally, endurance challenges test a player's ability to sustain a level of skill during a continuous course of obstacles.

Designing difficulty levels within games is important for creating the optimal experience for players. If a game is considered too easy or too hard, it will have a negative impact on a player's motivation to engage with the task [419]. Further to this, game difficulty has been seen to have a direct impact on the brain's reward system. A brain imaging study conducted by Huskey *et al* [419] observed reward system activation during difficulty balanced cognitive control tasks in a computer game. In regards to intervention based games, increasing game difficulty can also help deliver and progress the educational or therapeutic objectives of the software through challenging but achievable goals [69], [420]. This has been introduced in serious games such as *Mind Light*, a neurofeedback game to assist in the reduction of anxiety which has been used for autistic children [189], [205]. As game-play progresses, the virtual threats that are integrated into the game mechanics become increasingly more difficult to ignore and the only way to advance is to confront them.

However, due to the heterogeneity of autism, game difficulty must also be closely linked to the personalisation of the game experience to adapt to the unique cognitive, social or communication abilities of the player. This ad-hoc approach to adjusting difficulty levels had been implemented in *ECHOES* [421], a game that targets the development of social communication skills in children with ASD. Difficulty levels are not related to the individual game activities, but rather to the amount of verbal guidance they receive from the game's agent. Players who experience more difficulties in this area will be in receipt of more assistance.

5.2.5 Story Driven Goals

In addition to their fundamental game-play mechanics, computer games can also contain the components of a typical narrative such as characters, actions, events and environments [373]. Once arranged into a story-like order, it could therefore be argued that the computer game holds a narrative [375]. However, one of the key aspects that sets computer games apart from other narrative based media is interactivity [422]. In addition, a computer game narrative can be non-linear, a direct product of a game's interactivity [423]. Within a game environment, the challenges and tasks put in place by the designers help players to understand the narrative and vice-versa [423]. Furthermore, how absorbed a player is by the game story is considered to be a contribution to the sense of immersion a player feels across the entirety of the game [424].

For traditional gaming, how a narrative is experienced by a player can have a direct result of their emotional experience of the entertainment medium [371]. However, the introduction of story based goals can have a positive impact on the therapeutic outcomes of serious games. Whyte *et al* [199] suggest that one of the most powerful resources in enhancing motivation to learn within a serious game environment is the inclusion of a storyline or narrative that puts into context the educational or therapeutic goals. Furthermore, these can increase enjoyment when playing which can have positive implications on learner motivation. This is especially important when designing serious games for autistic people. Story-driven goals can give the player a reason to actively engage with the mechanics that support the game's learning outcomes. This approach was employed in the serious game *'Sinbad and the magic cure'* [29]. Here, players must actively engage with and collect problematic sound objects, delivering them to an evil witch character in return for a magic cure to save the princess.

5.2.6 Game Audio

Sound design is fundamental in the development of all computer games, including those developed for therapy, education or entertainment purposes. with that in mind, game developers are increasingly placing a strong emphasis on creating sound content for computer games [378]. Sound or music can be used as a tool to provide feedback on player actions, provide warning for in-game events, support on-screen visuals and complement visual rewards [380], [425]–[427]. All of which can significantly contribute to both immersion and engagement with the virtual environment and game mechanics [428], [429]. In addition, how sound is delivered within a VE can have a significant impact upon the emotions felt and subsequently the validity of the virtual visual and auditory stimuli (see Section 3.6.1) [430], [431]. As discussed in 2, these emotional states could be the result of neurological connections between the sense of hearing and the limbic system [432], [433].

Most modern day computer games utilise audio in well-defined categorises: music, speech and sound effects [377], [426]. Music can play an important role in creating atmosphere within a virtual environment [377]. Similar to its use in non-interactive media, music is used to emotionally engage its audience to reflect the on screen conditions [434], [435]. For example, if a player finds themselves in a combat situation, the music may become more up-tempo and aggressive in order to reflect the energetic environment [377]. Sound effects have a number of functions within a computer game's virtual environment. Firstly, sound effects give a direct notification to a player that they have performed an action within the virtual environment [377], [426]. Secondly, they can be used to orientate a player towards an area or point of interest outside of their field of view [377], [426]. Finally, sounds can be utilised in the identification of objects and to infer the object's value [377]. Regardless of how an effect is used, it is important that it is a plausible representation of the object in the virtual world [426]. In this regard, an overarching function of effects is to complement the visual stimuli in order to create a feeling of presence and immersion within the VE [377], [426], [436].

A heightened level of realism and a sense of presence are important when designing serious games for therapy or education [437]. As discussed in Chapter 2 this is important for delivering successful VRET interventions, and spatial sound can play a significant role (see Chapter 3. However, it is crucial to keep in mind the auditory processing difficulties often experienced by autistic people that were discussed earlier in this thesis, for example impaired auditory scene analysis [103]. Research conducted and reported in Chapter 4 observed that signal-to-noise ratio between target auditory stimuli and competing background noise does have a significant impact upon the localisation of spatialised auditory stimuli within a virtual environment. An impairment in analysing multiple sound sources could result in a reduction in interpreting in-game prompts and cause unnecessary stress and anxiety [380]. It is therefore essential that competing environmental in-game audio be reduced in order to compensate for any possible auditory processing impairments. This approach to delivering sound could help highlight important information via audio-visual cues and increase the ecological validity of the VE for people on the autistic spectrum.

5.3 SoundFields Description

SoundFields is the prototype of an individual interactive VR game designed for children and adolescents diagnosed with autism spectrum disorder and who also display auditory hypersensitivity towards particular sounds. It aims to help individuals become habituated to sounds that might ordinarily provoke feelings of irritation or anxiety. It achieves this by integrating principles of exposure therapy into its core game mechanics and delivering problematic sound stimuli through 3^{rd} order ambisonic rendering.

SoundFields is a VR computer game set within an enchanted forest in which the player must find and locate lost characters through four mini-games. SoundFields is divided into three smaller mini-games which are located in two separate settings within the overarching forest environment.

5.3.1 Tool used

This section presents the libraries and software tools used for the development of the serious game prototype.

Unity 3D Game Engine: The Unity3D game engine is a free to use software that allows for the design and development of 2D and 3D games including VR. The application can fully support Windows 10 and uses the C# scripting language [344].

For VR development, the SteamVR SDK was integrated into the game engine, this provides VR specific assets that communicate with major VR device manufacturers via the Open VR framework [438].

Wwise Audio Engine: Audiokenetic Wwise is a cross platform game audio engine that is free for non-commercial uses. It features audio asset editing and authoring and can be integrated into the Unity3D game engine [350]. Furthermore, Wwise supports the Google Resonance spatial audio SDK outlined below.

Google Resonance: Google resonance is an open source cross-platform spatial audio SDk used for mixed reality gaming and video that can render up to 3^{rd} order Ambisonics [351]. The SDK can be included as an audio plugin package for the Wwise game engine software and operates a binaural decoder for the audio engine with the HRTF filtering being applied to object based sound sources or multi-channel ambisonic files. Resonance is built using the SADIE HRTF database [275].

5.3.2 Therapeutic Mechanisms Used

As discussed in Chapter 2, there is convergent evidence from a number of studies that suggests auditory hypersensitivities experienced by autistic people is related to an irrational fear of the stimulus instead of physiological pain associated with the stimulus [6], [7], [126]–[128]. Therefore, the most effective treatment has been centred around CBT techniques such as systematic desensitisation [7] or exposure based training [29], [172]. However, as discussed in Chapter 2 there are also barriers to effective CBT sessions caused by the core difficulties experienced by autistic people, such as impaired communication impacting patient to therapist interactions [189] and non-contextualised stimuli via imaginal exposure [191].

The game described in this chapter integrates simplified CBT techniques based upon empirical evidence used to reduce the negative emotional reactions associated with specific sounds. Within a playful and engaging environment players will gradually and repeatedly encounter auditory stimuli which would ordinarily cause anxiety with the aim to decrease the aforementioned anxiety and increase habituation.

The primary evidence based strategy embedded into the core game mechanics is exposure training, which is empirically validated as the key CBT mechanism in reducing anxiety associated with specific fears and phobias [439]. During the course of the game, players are encouraged to interact with mechanics that will expose them to virtual representations of a feared auditory stimuli and are rewarded for doing so. This strategy of environmental sound enrichment has been applied in a number of serious games designed for auditory hypersensitivity in ASD whereby graded sound exposure follows a step-wise model that gradually increases the volume of the sound [29], [30], [368].

The effective outcome of VRET for phobias is based on the controlled delivery of sensory information, creating a realistic simulation of feared stimuli within the virtual environment. SoundFields accomplishes this by rendering feared auditory stimuli using head-tracked binaural based spatial audio. This sound reproduction technique is capable of simulating the movement of an audio source within a 360 degree space and adapt its movement to compensate for the movement of the player's head within the virtual environment. Furthermore, it is also possible to emulate the acoustic parameters of the environment the sound is most encountered in. The enhanced perceptual realism of soundscapes experienced whilst using spatial audio [440] aims to address difficulties in contextualisation of the presented stimuli encountered by autistic individuals during virtual CBT sessions. Finally, by using spatial audio, exposure hierarchies similar to those applied in previous exposure training for auditory hypersensitivity [7] can be recreated within the virtual environment. This is achieved through simulated distance attenuation properties of sound, moving the virtual sound source closer to the player in a stepped process (see Table 5.1) .

TABLE 5.1: Exposure Hierarchy. Each exposure level corresponds to the distance between the participant and the virtual sound source, distance is represented in metres .

Exposure Level	Virtual distance between player and stimulus
1	25m
2	15m
3	5m
4	2.5m

As previously mentioned in Section 5.2, rewarding a player for successfully completing game mechanics can have a positive impact upon a player's motivation to perform well in a game. Within SoundFields, the player is rewarded a larger amount of in-game currency and auditory praise via sound effects upon interaction with mechanics which render representations of target auditory stimuli. A strategy employed in the game *Sinbad and the Magic Cure* [29]. It is therefore expected that by offering the player an increased reward they will be encouraged to expose themselves to target stimuli. Therefore, positively impacting upon the proposed therapeutic outcomes of the game.

The combination of hierarchical exposure based training and positive reinforcement as an approach to reducing specific phobias has received empirical support in the treatment of autistic children. Ricciardi *et al* [441] eliminated avoidance behaviour in an autistic child with a specific fear of animatronic objects. The intervention involved a fear hierarchy that introduced the child to the objects at decreasing distances in parallel with verbal positive reinforcement. In another study, Schmidt *et al* [442] used graduated exposure combined with positive reinforcement in the form of verbal praise and edibles to introduce an autistic boy to feared rooms in his school. These, amongst other studies therefore recommend the use of incremental exposure in the presence of positive reinforcement to foster positive associations with feared stimuli [162], [443], [444]. Secondly, the previously mentioned studies also advise conducting contact desensitisation without escape prevention as an effective fear-reduction intervention for autistic children. This is the process of not allowing the individual to remove themselves from a phobic situation [441]. In regards to SoundFields, this approach may have a negative impact of player motivation to engage with therapy based mechanics. Therefore, exposure within the presented VR game is entirely voluntary, consequently permitting the player to terminate interaction with the feared stimuli at the cost of additional in-game rewards.

Finally, the repeated confrontation of a feared object is the core element of many treatments for phobia or anxiety based disorders [445]. Within SoundFields, the application of voluntary exposure through the removal of escape prevention means that an embedded reward system that positively reinforces successful interaction with a conditioned stimulus is crucial for encouraging repeated exposure. A similar design approach has been integrated into serious games targeted to reduce anxiety in autistic children including auditory hypersensitivity. *Sinbad and the Magic Cure* [29] rewards players for collecting in-game objects that emit target stimuli and require repeated collection of these objects to progress in the game. In another game designed by Morris [30], players were motivated to voluntarily engage with target sounds multiple times and at varying exposure levels through in-game positive reinforcement.

5.3.3 Game-play Description

The core game-play consists of two mini-games in which the player is given opportunities to voluntarily expose themselves to feared target auditory stimuli. Due to the heterogeneous nature of ASD, abilities can vary between each child, this therefore necessitates the need to individualise the gaming experience [199]. This is achieved in SoundFields through the inclusion of mini-games and permitting the user to choose which game they play.

Figure 5.2: Virtual shop where players can purchase cosmetic variations of their magic wand using currency earned in-game.

Games can also be repeated as many times as the player wishes. This repeatability makes serious game experiences enjoyable and motivating for those on the autistic spectrum[381]. Furthermore, it can increase generalisation by allowing repeated exposure to stimuli [191]. Mini-game mechanics are designed to provide numerous opportunities for the player to be exposed to averse auditory stimuli, however as discussed in Section 5.3.2 all exposure during game-play is voluntary and is fostered by the embedded reward system.

SoundFields features an in-game currency in the form of gems which could be used to purchase cosmetic alterations to the player's magic wand (see Figure 5.2). Shop items are organised into pricing tiers designed to encourage players to acquire more currency. The minimum amount players can be awarded for successful completion of game tasks is one gem, however they will be awarded 10 gems for successful completion of tasks which incorporate exposure to target sounds. Motivators such as this are an important element in the design of games for both educational or entertainment purposes.

Mini-Game 1: Orb Hunt: The orb hunt mini game, shown in Figure 5.5, requires the player to locate orbs hidden within the virtual game environment. Once found, the player must then capture the orb by using their magic wand which gradually moves the creature towards them. During game-play orbs will spawn randomly at pre-determined locations within the VE, these are shown in Figure 5.3 alongside the movement area for the player. Once active, the orbs will emit speech-like gibberish stimuli which aims to attract the player's spatial attention towards their position and assist them in locating it. Once it moves into the player movement area it disappears and the game score is updated.

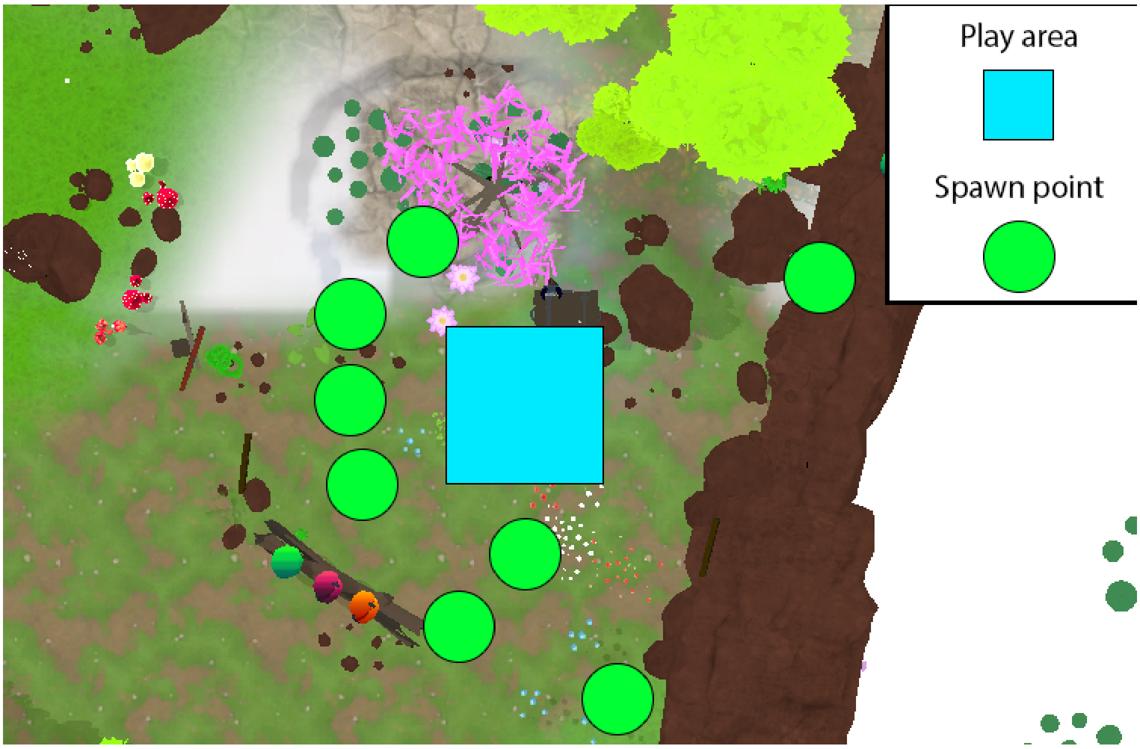


FIGURE 5.3: Top down view of the SoundFields visual environment.

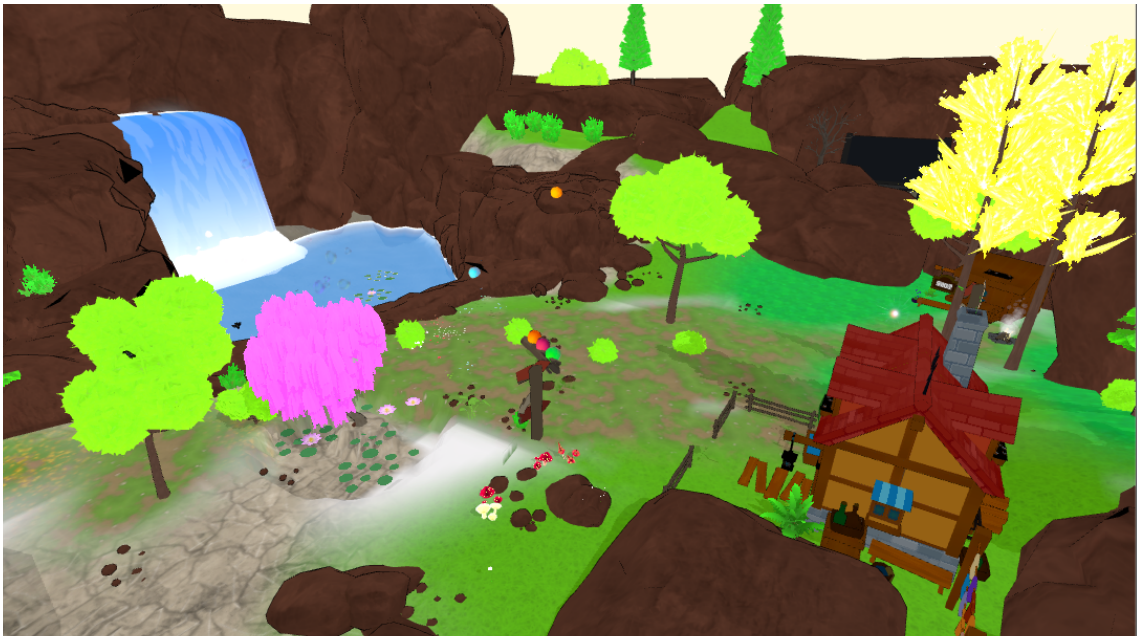


FIGURE 5.4: Top down view of the SoundFields visual environment.

Figure 5.5: Orb non-player characters used in the Orb Hunt mini game. Type A emits speech like stimuli and awards 1 gem upon capture. Type B emits target stimuli and awards 10 gems upon capture.

Two variants of the orb are used during this mini-game which are displayed in Figure 5.5. Type A will only emit speech like stimuli throughout the length of time the player interacts with it. It is important to note that the sounds used are retained from the study presented in Chapter 4, this is due to participants displaying spatial attention towards this stimuli. Once the player successfully catches the orb they are rewarded with one gem. Type B, which is denoted by its gold colour will emit speech-like stimuli until the player interacts with it. Once interaction commences the orb will play spatialised reproductions of averse sounds at the same time as speech like stimuli. In addition the orb will travel towards the player at a slower speed. Maximising the exposure time. Once the player successfully collects this orb, they are rewarded with 10 gems. As a strategy to optimise the potential therapeutic outcomes, the spawn, locate and collect game mechanic is looped until the player wishes to stop it. This creates multiple chances for the player to be exposed towards their target averse stimuli. Furthermore, it is based upon the repeated and predictable game-play that can maintain interest and motivation to engage with serious games for this population [178], [381].

Mini-Game 2: Bubble Rescue: The objective of mini-game two (see Figure 5.6) is for the player to rescue orbs trapped within floating bubbles by using their magic wand to pop them. During game-play bubbles containing the orbs will spawn randomly at predetermined locations within the VE waterfall shown in Figure 5.6. Bubbles will float upwards until it has either been popped by the player or it reaches a virtual height on 20m, at this point it will disappear and respawn.

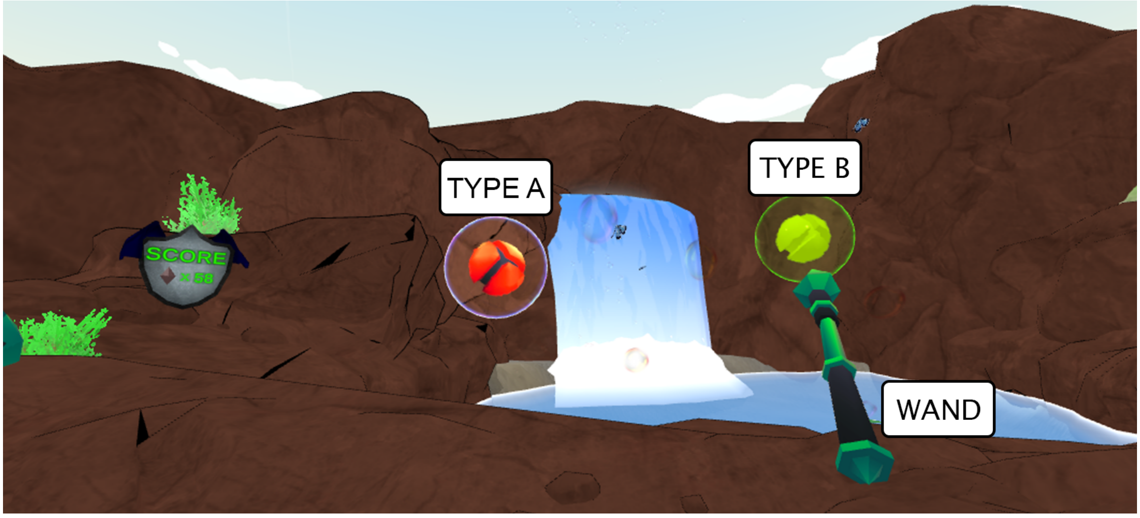


Figure 5.6: Orb non-player characters used within the Bubble Rescue mini-game. Type A emits speech like stimuli during interaction and awards 1 gem upon rescue. Type B emits target stimuli during interaction and awards 10 gems upon rescue.

Similar to the orb hunt mini-game there are two variations of characters the players can interact with, these are highlighted in Figure 5.6. Type A will only emit speech like stimuli, once the virtual bubble is popped, the orb will play a positive reinforcement auditory stimulus in the form of cheering and the player is awarded one gem. Type B is represented as a golden bubble, this bubble will emit speech like stimuli until it is interacted with. Once the player begins interaction, a reproduction of the target averse stimulus specific to that player will be played at the corresponding exposure hierarchy. This is represented by the virtual distance between the player and sound source. If the player ceases interaction the target stimuli will stop. Once the bubble has been popped by the player the target audio will stop, the positive reinforcement will be played and the player will be awarded ten gems.

5.3.4 Game Flow

The overall flow of the game is shown in Figure 5.7. Each rectangle represents an area of game-play that the player is able to interact with. The arrows indicate how the player can move freely between each of the game components. As previously mentioned in Sections 5.3.3 and 5.3.3, users can play either mini-game for as long as they please, thus this is facilitated by the freedom of movement between the four game areas. By permitting the player non-linear movement through the game environment it raises opportunity for mini-games to be repeated and for the personalisation of the game experience.

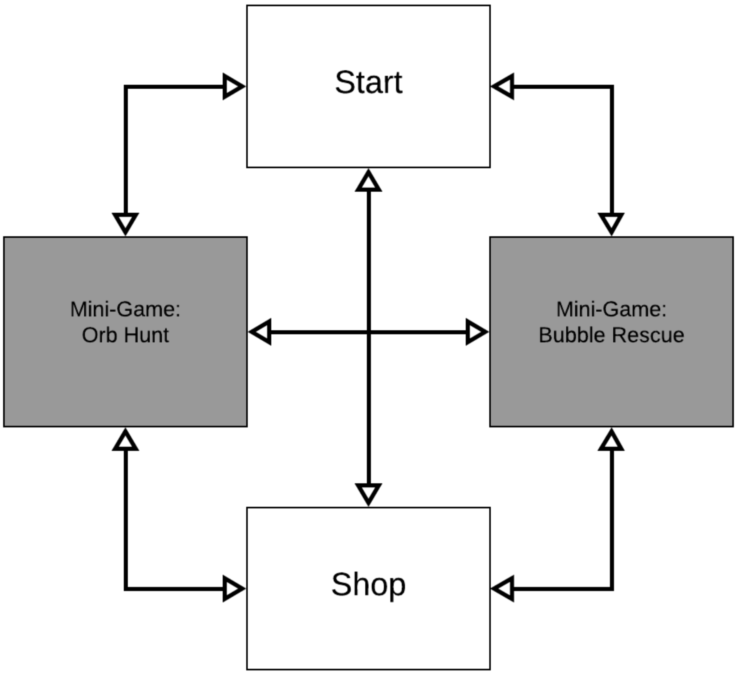


Figure 5.7: SoundFields game flow. The chart displays freedom of movement permitted between in-game areas.

5.4 Sound Design

From previous research outlined in Chapter 2 and the investigation conducted in Chapter 4, there is evidence to show that auditory processing impairments experienced by autistic children and adolescents can have a negative impact upon how they perceive a soundscape be it in the real or virtual world. With that in mind it was crucial to implement specific approaches to sound design in order to maximise the experience for the player. This will be explained in two categories, the first will discuss the implementation of environmental and non-target audio. The second category will consider the basic template of how the target audio stimuli were created.

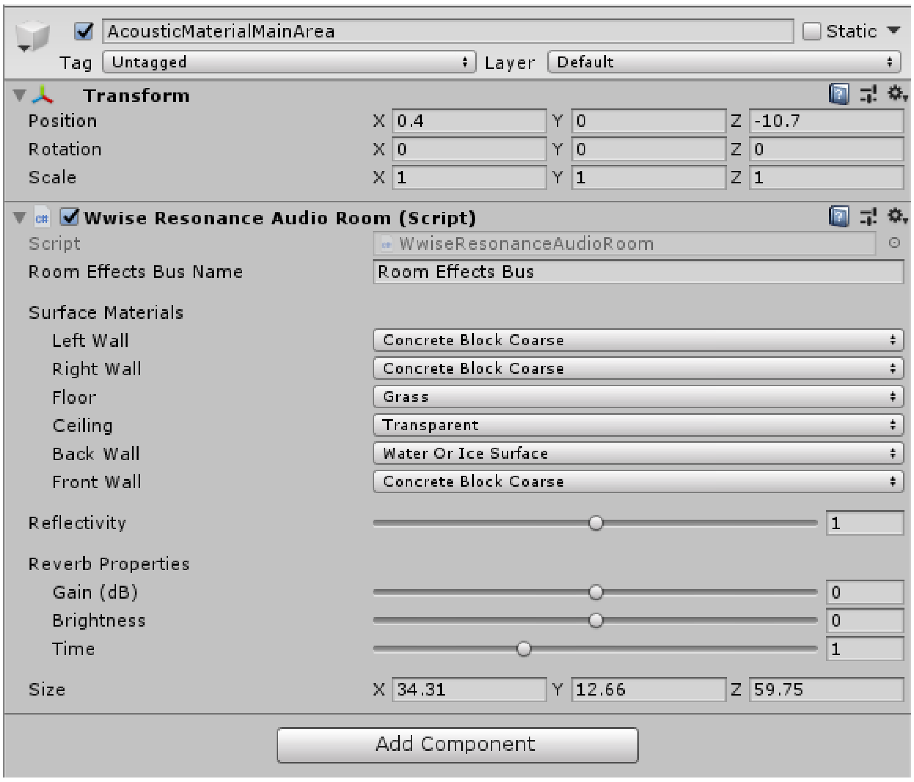
All audio stimuli were recorded or edited using the Reaper [349] digital audio workstation and WWise [350] game audio engine. Stimuli were a combination of original spatial audio recordings extracted from the Eigenscape database [446], the Vocal Interaction in an Immersive Virtual Acoustic (VIIVA) system [447] and recordings obtained from FreeSound.org [448] under the Creative Commons licence. 3D binaural-based spatial audio was rendered via the Google Resonance spatial audio SDK [351] which was included in Wwise as a plugin.

5.4.1 Environmental Audio

An ambient soundscape was created placing virtual mono sound sources throughout the unity game environment, these are displayed in Figure 5.9. A Wwise spatial audio listener is placed on the virtual camera which is positioned at the player's head. This game component will then process the audio in conjunction with the Resonance HRTF database to spatially render sounds based on the player's movement with 6DoF simulating positioning. Distance is simulated by predefined attenuation curves within Wwise that adjust amplitude and spectral content relative to the distance between the player and the sound source. This attenuation is processed before HRTF filtering. Finally, reverberation properties of the sounds are also manipulated by defining the material coefficients of the virtual 3D objects placed in the game environment (see Figure 5.8).

One of the design considerations recommended by literature and from the data collected earlier in this thesis was the use of Clear Audio, mentioned in Section 5.2.6. In the case of designing serious games for autistic young people this defines delivering key auditory objects in a manner that distinguishes them from competing background noise. As demonstrated in Chapter 4, reducing all background noise and increasing the signal-to-noise between that and the stimuli can significantly improve spatial attention towards both speech and non-speech sound objects. Furthermore, Schelinski *et al* [108] observed recently that a small difference in SNR of 1.45 dB can have a significant difference in speech recognition when presented in noise.

As discussed in Chapter 3.6.1, soundscapes that closely resemble the visual environment are essential for maintaining the presence felt by a player within VR. This has meaningful implications for the therapeutic outcomes for an intervention that is based within VR in which the concept of presence is key [219]. Therefore reducing all background noise simultaneously could create an unnatural inconsistency between the visual and auditory



☑ AcousticMaterialMainArea ☐ Static
Tag: Untagged
Layer: Default

Transform
Position: X 0.4, Y 0, Z -10.7
Rotation: X 0, Y 0, Z 0
Scale: X 1, Y 1, Z 1

☑ **Wwise Resonance Audio Room (Script)**
Script: WwiseResonanceAudioRoom
Room Effects Bus Name: Room Effects Bus

Surface Materials
Left Wall: Concrete Block Coarse
Right Wall: Concrete Block Coarse
Floor: Grass
Ceiling: Transparent
Back Wall: Water Or Ice Surface
Front Wall: Concrete Block Coarse

Reflectivity: 1

Reverb Properties
Gain (dB): 0
Brightness: 0
Time: 1

Size: X 34.31, Y 12.66, Z 59.75

FIGURE 5.8: Defining the virtual acoustic materials present within the main play area of the SoundFields VE.

environment, resulting in a breakdown of the immersion. With this considered, an object based approach was used to reduce SNR and increase the clarity of presented sounds.

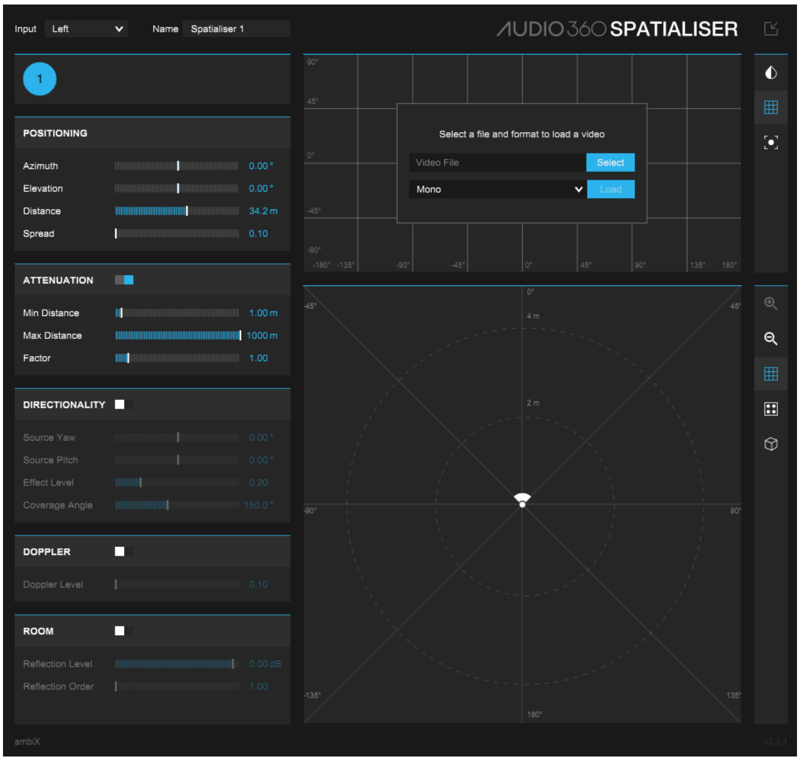
Dividing an audio scene into discrete audio objects can be implemented to deliver adaptive content in order to improve the accessibility of media for listeners with hearing impairments [352], [449], [450]. With this approach the signal-to-noise ratio of objects important to the narrative can be increased by reducing the volume of audio objects that have been categorised as being less significant to the narrative. In addition Tang et al [451] used a tracked intelligibility metric to optimise speech intelligibility. When the metric fell below the specified value, background objects were reduced to increase speech to background ratio.

SoundFields implements an automatic increase of signal-to-noise ratio for target sounds using a similar approach outlined above by adjusting background object levels. Background sounds were categorised as being environmental or diffuse sounds such as birds, waterfall or the wind in trees. This category was then further divided into two further classifications which established if the object had a visual component. When a target sound being either speech stimuli or averse target stimuli is played, the volume of background objects with no visual components are attenuated. By only adjusting the loudness of these objects the spatial attention towards target stimuli will be improved through an increased SNR and the levels of presence and immersion should be maintained by preserving the link between visual and auditory stimuli.

FIGURE 5.9: Locations of environmental audio objects placed throughout the VE. (1) Waterfall (2) Pond Waves (3) Birds (4) Crickets (5) Leaves in wind (6) Frogs.

5.4.2 Target Audio Stimuli

Target averse auditory stimuli were created within the Reaper [349] digital audio workstation using the Facebook 360 Spatial Workstation [297] virtual studio technology (VST) plugins to output 16 channel 3^{rd} order Ambisonic files. Ambisonic files are then imported into the Wwise audio engine software to be processed by the Google Resonance spatial audio plugin. During game play, the Ambisonic file is rotated along the horizontal and vertical axis based upon the head rotation data of the VR HMD. This approach to audio rendering was chosen to maximise computational efficiency by processing one file at a time. The 16 channel files contain all the encoded spatial information regarding directionality, distance and acoustic environment, therefore only requiring one audio file. An object based approach would require individual movement for each stimuli within the VR scene. In addition, this allows for a dynamic target stimulus library to be built that can be rapidly expanded without altering the application's source code and virtual environment.



Input: Left
Name: Spatialiser 1
Azimuth 0.00 °
Elevation 0.00 °
Distance 34.2 m
Spread 0.10
Min Distance 1.00 m
Max Distance 1000 m
Factor 1.00
Source Yaw 0.00 °
Source Pitch 0.00 °
Effect Level 0.20
Coverage Angle 150.0 °
Doppler Level 0.10
Reflection Level 0.00 dB
Reflection Order 1.00
Mono

FIGURE 5.10: Facebook Spatial Audio Workstation - Spatialiser Audio Plugin. Used to position an audio source using azimuth and elevation coordinates. Distance attenuation can also be simulated here.

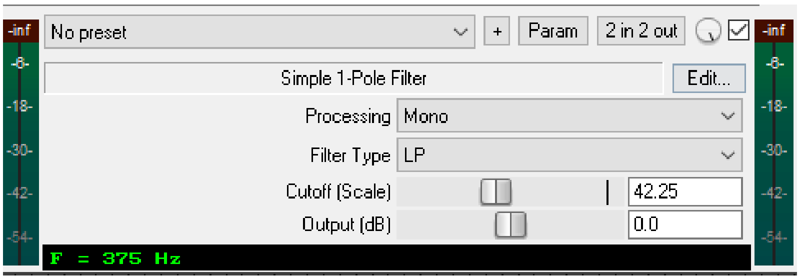
Within Reaper, the directional properties of a mono sound source are simulated by automating the azimuth and elevation properties of the Spatialiser plugin displayed in Figure 5.10. Positioning was designed to match its real-world equivalent as closely as possible. Target stimuli used in the research presented in Chapter 7 such as a fire alarm, would be set at a fixed point in the virtual space. In comparison, the ambulance stimuli used as target stimuli in Chapters 6 and 7, would appear to move/ drive from a distance and travel directly past the listener.

As discussed in Section 5.3.2, previous games that target auditory hypersensitivity in autistic children simulate the distance exposure hierarchies used in traditional interventions by gradually increasing the volume levels of the sound [29], [30], [368]. However, the perception of distance between a listener and a sound source is dependant upon the relative loudness, frequency attenuation caused by air absorption and the ratio between direct and reverberant sound energy.

TABLE 5.2: Sound absorption coefficients of brick walls. Table adapted from [453]

Material	Octave Band Frequency in Hz						
	125	250	500	1k	2k	4k	8k
Walls, hard surfaces average (brick walls, plaster, hard floors, etc)	0.02	0.02	0.03	0.03	0.04	0.05	0.05
Walls, rendered brickwork	0.01	0.02	0.02	0.03	0.03	0.04	0.04

Exposure hierarchies discussed in Section 5.3.2 simulated in this project were done so by adjusting the distance property of the Spatialiser plugin (see Figure 5.10). This simultaneously processes the audio signal's amplitude, frequency content and reverberation attributes based upon the relative distance between the sound source and the listener. Finally, additional audio filtering was implemented based upon the target stimulus and the distance. Previous studies which involve systematic and graded exposure towards auditory stimuli often initiate exposure with the stimulus situated within another room [7], [71]. Material such as brick often reflects high frequency sound waves which results in a build up of reverberant energy within a space in which the sound source resides. However, lower frequency waves will be absorbed into the walls and are transmitted through the material (see Table 5.2). Though the lower frequency sounds can still be heard, energy is absorbed by the material resulting is some extent of attenuation [452], [453]. In order to simulate this as accurately as possible, a lower pass filter is placed on the monaural sound source prior to the spatialisation plugin (see Figure 5.11). Combined with the distance attenuation reproduced by the Facebook Spatial Audio Workstation, this reproduces a sound source heard from a neighbouring room.



Simple 1-Pole Filter

Processing: Mono

Filter Type: LP

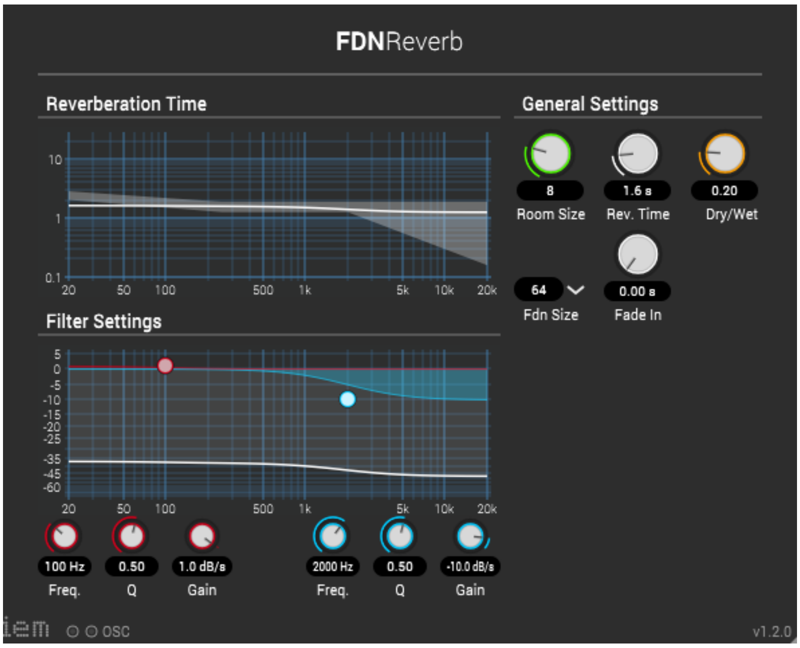
Cutoff (Scale): 42.25

Output (dB): 0.0

F = 375 Hz

FIGURE 5.11: Low pass filter simulating children playing outside of listener environment. Stimuli was taken from the library of target stimuli used in the research presented in Chapters 6 and 7.

Following two exposure levels in which the sound source is reproduced in another room, the sound source appears to the occupies the same acoustic space as the listener. For static sound sources that are often experienced within an indoor environment, reverberation is applied after the spatialisation plugin to simulate early and late reflections. An example of such a stimuli would be a hair dryer, which will be featured as a target stimuli in the research presented in Chapter 7. This was achieved using the IEM FND Reverb [454] plugin which is capable of simulating reverberation up to 7$^{\text{th}}$ order Ambisonics, therefore maintaining the sound source position within the space and simultaneously reproducing sound reflections.



Room Size: 8
Rev. Time: 1.6 s
Dry/Wet: 0.20
Fdn Size: 64
Fade In: 0.00 s
Freq.: 100 Hz
Q: 0.50
Gain: 1.0 dB/s
Freq.: 2000 Hz
Q: 0.50
Gain: -10.0 dB/s

FIGURE 5.12: IEM FND Reverb used to reproduce a hair dryer stimulus within a bathroom environment. Stimuli was taken from the library of target stimuli used in the research presented in Chapters 6 and 7.

5.5 Comparison with Design Guidelines

Section 5.2 presented design guidelines for the development of serious games for autistic young people as recommended from research literature. How these guidelines have been implemented during the development of the SoundFields game are summarised in the following:

- **Repeatability and Predictability**: Players can repeat mini-games as often as they like with the game goals and mechanics remaining the same. In addition, there is no time limit.

- **Reward System and Motivator**: The player receives points and in-game currency for successfully completing game tasks, further points are rewarded for interacting with game mechanics that include exposure to averse target auditory stimuli. Currency game can be used to purchase cosmetic alterations of the player's magic wand. Positive reinforcement is given in the form of visual animations and audio effects

- **Customisation**: Players can spend in-game currency to customise the appearance of the magic wand to their preference. The user also has full control over which mini-games they interact with and the duration of which they play them. The hierarchical structure of exposure levels and amount of game enemies cannot automatically adapt to the player's ability but they can be adjusted in a settings menu.

- **Story Driven Goals**: Each mini-game has a specific goal that can be achieved. However, the global goal is to collect and rescue as many orbs as possible within the VE. At this time there is a very limited narrative.

- **Game Difficulty**: There is one level of difficulty implemented into the core game mechanic loops of both mini-games. The player is challenged by reducing the virtual distance between themselves and the averse target stimuli.

- **Game Audio**: All audio is rendered to 3^{rd} order Ambisonics using Google Resonance spatial audio SDK. Similar to the experimental environment presented in Chapter 4, spatial sound directs players to characters that are essential to the game mechanics. In addition to this, sound provides a direct feedback to player actions within the VE. Spatial attention towards target and non-target stimuli by attenuating the intensity of background audio objects with non-visual components. This reduces signal-to-noise ratio and compensates for any impairments in auditory scene analysis that may be experienced by the player. Finally, accurate spatial representations of feared auditory stimuli are rendered using Ambisonics, which is the core component of the proposed digital intervention and research.

5.6 Summary

This chapter has outlined the essential elements for developing a serious game to target auditory hypersensitivity in autisitc young people. First, a number of design guidelines recommended by literature for developing serious games for autistic people was described. Throughout the development process of this project these design patterns have been essential in creating a game environment for autistic children. This was followed by a detailed description of the serious game SoundFields. Here the theoretical basis for the game as an exposure therapy intervention was explained and how the theory was embedded into the various game components. Each mini-game embodies mechanics that will motivate the player to actively and repeatedly engage with target averse stimuli with the overarching objective of habituation to their problematic sounds. However, the crucial component of the presented game is a theory to build upon pre-existing serious games developed for auditory hypersensitivity by reinforcing exposure therapy mechanics with realistic simulations of averse sounds rendered through 3rd order Ambisonics. It is this theory that is at the centre of the motivation for this thesis.

The information provided is a detailed description of the method and approach to how spatial audio renderings of problematic sounds will be delivered during the experimental investigations presented in the forthcoming chapters. However, the implementation of spatial audio using VR as an alternative to loudspeaker arrays extends further than the central hypothesis of the thesis. Part of the core rationale of using computer games for therapy is to increase the accessibility of an intervention protocol by allowing the user to achieve therapeutic outcomes through game play outside of clinical environments and specialist equipment. In respect to this project, using consumer grade equipment and software to test the main research question will also simultaneously evaluate the use of spatial audio as an accessible alternative or augmentation to in vivo exposure cognitive behavioural therapy techniques for auditory hypersensitivity experienced by autistic children and adolescents.

Chapter 6

Usability Study of the SoundFields Prototype

6.1 Introduction

This chapter presents a small group study evaluating the use of binaural based 3^{rd} order Ambisonics as an audio rendering technique to target auditory hypersensitivity experienced by autistic children and adolescents. As discussed in Chapter 3, spatial audio has been observed as being able to have a positive impact on psychological mechanisms such as presence felt in VR [245], [455], [456] and eliciting feelings of anxiety and fear [243], [244], [337], both of which are fundamental to VR based exposure training [228]. Furthermore, it could be considered that spatial audio can be an effective audio rendering approach for exposure based therapy for those with autism by delivering realistic virtual soundscapes. This is a similar approach to the use of realistic visual environments VRET in treating specific phobias in autistic people [34], [237]. It is these considerations that form the central novel motivation for this chapter and thesis, building upon previous achievements that use non-spatialised audio rendering techniques for this sensory processing issue [29], [30]. Therefore, the primary aim of this chapter is to examine the impact of repeated exposure towards Ambisonic rendered averse stimuli upon individuals with autism experiencing auditory hypersensitivity.

Further to examining Ambisonics as a rendering technique this investigation will also consider how the target auditory stimuli are presented. That being the VR serious game environment detailed in Chapter 5. There is already empirical evidence that supports the use of exposure based training for auditory hypersensitivity for autism embedded into computer game mechanics [29], [30], however the current application is the first to utilise VR technology. It has already been discussed in the previous chapter (see Chapter 5) that utilising technology could enable effective head tracked Ambisonics to be delivered in an engaging serious game based intervention. Firstly, this is based on over two decades of research that observes people on the autistic spectrum successfully engaging with VR based interventions. Secondly, results from Chapter 4 demonstrate that autistic young people respond to spatial audio events presented within VR environments. This is despite observed auditory impairments which effect how they translate spatial auditory information in the physical world. Finally, VR today is one of the most prevalent and accessible formats to which head-tracked spatialised sound can be experienced. With this in mind, the evaluation of the delivery of the auditory stimuli within a VR game is an important research consideration.

6.2 Hypotheses

The investigation presented in this chapter aims to answer the following research question: *Can negative emotional associations to specific sounds be reduced in autistic children and adolescents by using 3^{rd} order Ambisonics within a VR exposure based game?* Based upon this research question and the research outcomes outlined in Section 6.5.5, the following hypotheses are put forward:

- H_1: Compared to baseline measurements, participants will display significantly lower levels of self-reported negative emotional responses towards problematic target stimuli following the experimental period.
- H_2: Self-reported anxiety scores for non-target stimuli will not be effected by the experimental period.
- H_3: Participants will record increasingly longer amounts of voluntary interaction with target stimuli when comparing the first and last experimental sessions.

6.3 Experimental Design

The study followed a within-subject single-arm design in which each participant would experience the same experimental condition, listening to 3^{rd} order Ambisonic renders of problematic auditory stimuli delivered using the SoundFields VR game (see Chapter 6). All the participants completed the four experimental sessions held over the course of four weeks.

6.4 Participants

An experimental cohort consisted of a small cohort 6 adolescents (4 male and 2 female, mean age = 17.7, SD = 1.03, range of 16–19 years). All participants had a formal diagnosis of autism spectrum disorder obtained from their local national health trust. They displayed social, cognitive, and motor abilities that allowed them interact with the game intervention and VR equipment. In addition, all participants experienced auditory hypersensitivities to single or multiple auditory stimuli. Exclusion criteria were self-reported physiological hearing problems; an inability to finish the task. An information package was provided to the participants parent(s) or legal guardian(s). Participants were admitted into the study after informed consent and assent was obtained from their parent(s) or legal guardian(s).

6.5 Methods

6.5.1 Game intervention

For the duration of this study, all participants played the SoundFields VR game. For a full description of the virtual environment and game mechanics please refer to Chapter 5.

6.5.2 Auditory stimuli

For the duration of the experimental period, all audio was rendered using binaural based 3^{rd} order Ambisonics. The SoundFields game includes a library of target auditory stimuli comprised of problematic sounds identified by parents/guardians of the participants during the recruitment process. Each stimulus is replicated in four separate 16 channel 3^{rd} order Ambisonic files representing a step-wise exposure hierarchy. As discussed in Chapter 5, exposure levels correspond to a distance between the listener and the virtual sound source. For this approach distance levels were based on an adaption of those used by Koegel *et al.* [7], in which a speaker playing animal noises was moved closer to the participant during a systematic desensitisation intervention. The four exposure levels and the associated distances can be seen in Table 6.1.

TABLE 6.1: Exposure Hierarchy used within SoundFields. Each exposure level corresponds to the distance between the participant and the virtual sound source, distance is represented in metres .

Exposure Level	Virtual distance between player and stimulus
1	25m
2	15m
3	5m
4	2.5m

During each experimental session target stimuli are presented to the participant a maximum of 20 times through the exposure based mechanics embedded into each mini-game explained in Chapter 5. A C# script was used to calculate the probability of each target stimulus being made accessible within the virtual environment based upon the total amount of times the participant has successfully completed an exposure mechanic task and the remaining time of the session; Mini-Game One required a golden orb being collected, Mini-Game Two required a golden bubble to be popped. The total amount of time for a single interaction with these mechanics is a maximum of $\approx$ 5 secs. For Mini-Game One this is according to the time it takes for the orb to travel towards the participant. The time is only an approximation as participant movement can result in variations in virtual distances. For Mini-Game two, each virtual bubble takes 5 seconds to pop. However, if the bubble reaches the respawn point (see Chapter 5, Section 5.3.3) the interaction will terminate. //

Target Stimuli Design

Each target stimuli was designed to simulate its physical world counterpart as closely as possible. In regards to localisation and movement of a virtual sound source, accuracy was achieved by using 3^{rd} order Ambisonics [78]. In addition, each target stimuli was subjected to individual approaches to sound design to further extend how it may be influenced by the characteristics of an acoustic environment.

6.5.3 Experimental Procedure

All six participants underwent an intervention of 4 sessions of increasing exposure towards a particular auditory stimulus (see Table 6.2), with a duration of approximately 30 minutes. Of the six participants recruited, all were able to complete all 4 intervention sessions. During the recruitment, parents of the participants were required to complete a form which stated the particular auditory stimuli each participant displayed negative behavioural reactions towards.

Baseline Measurement

Each participant completed a baseline measurement session one week prior to the four week experimental period, during which they would be asked to rate their own emotional experiences towards presented auditory stimuli within an application developed in Unity3D (see Figure 6.1). This would be included in the primary outcomes of the study and are explained fully in Section 6.5.5. A total of twenty-two sounds were delivered, eleven representing all target sounds identified via the informed consent forms, and eleven 'relaxing' soundscapes taken from the Eigenscape database of Ambisonic recordings [446]. The database is comprised of acoustic scenes recorded using the mh Acoustics Eigenmike. The auditory scenes taken from the Eigenscape database for this study were: beach, park and woodland. All stimuli were rendered using non-head tracked binaural based 3^{rd} order Ambisonics as measurements were recorded outside of VR. In addition, target stimuli representing the highest level of the exposure hierarchy (see table 6.1). In order to minimise the likelihood of participant distress caused by sounds, stimuli would only play when the participant pressed on the space key of the laptop used during the experiment. Once the stimuli stopped the investigator would ask how the sound made them feel and the participant would provide feedback either verbally or by pointing at the corresponding face on the screen. Table 6.2 provides the demographic data for each participant alongside their target stimuli. The stimuli highlighted in this table were then included as target stimuli within the SoundFields game, played back at the corresponding exposure levels.

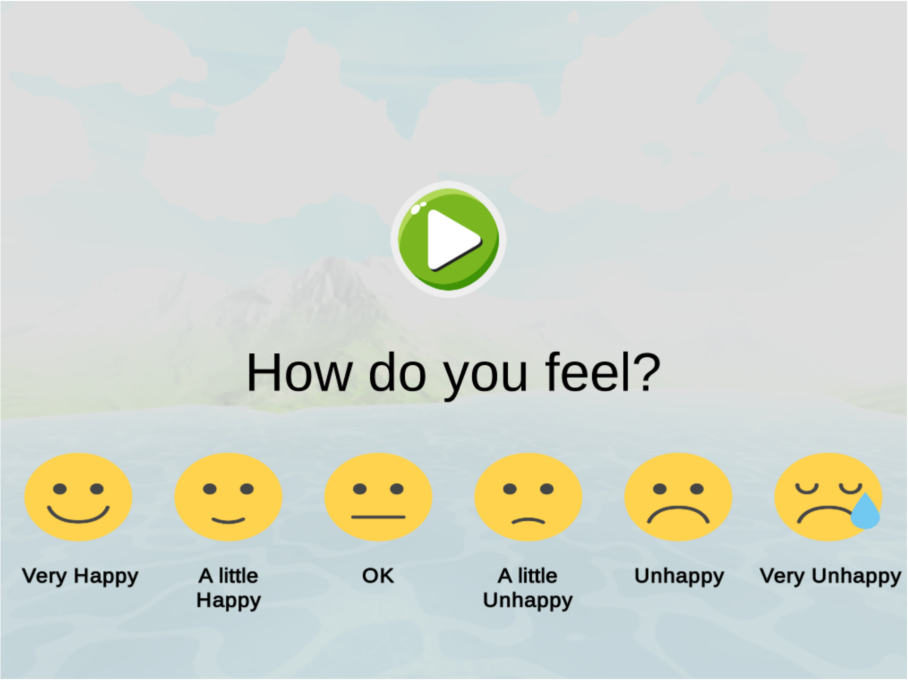


Figure 6.1: Audio Interactive Questionnaire

Table 6.2: Participant Demographics including identified problematic sounds.

Participant	Gender	Age	Stimulus
A	M	18	Group Singing, Baby Crying
B	M	18	Group Singing, Children Screaming / Shouting
C	F	17	Children Screaming / Shouting
D	M	16	Sirens, Children Screaming / Shouting
E	F	19	Children Screaming / Shouting
F	M	18	Children Screaming / Shouting, Baby Crying

Experimental Session Procedure

Each participant was given the opportunity to play the SoundFields game for a period of approximately 30 minutes, once a week over the course of four weeks. A member of staff from the recruitment school was present to communicate with participants and to provide assistance if they became distressed, however they would not be permitted to provide instructions. At the start of the session the investigator would select the target stimuli identified from the baseline measurement session and specific to the participant via an in-game menu (see Figure 6.2). The investigator would also inform the participant that they are free to terminate the session at any time. Additionally, the session would also finish if the participant became distressed or if the member of staff considered it necessary.

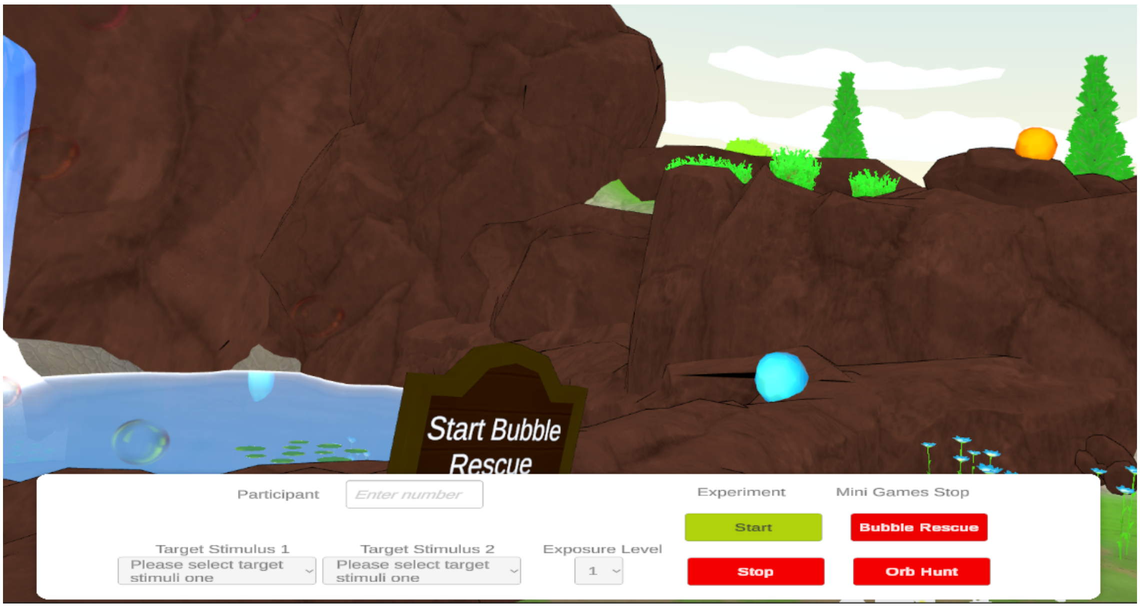


FIGURE 6.2: SoundFields in-game menu that is only visible to experimenter.

The first session began with the investigator providing a description of the basic game controls, mechanics, goals and rewards system embedded in the SoundFields mini-games to the participant. It was also explained that golden non-player characters emit sounds which the participant may not like or find annoying, however if they successfully collect them they would be rewarded with a larger amount of in-game currency. Following this, the participant was shown the virtual shop and informed how they can purchase in-game items. Throughout each session, the investigator limited communication to supplying support with game controls and to giving verbal praise when successfully interacting with target stimuli followed by asking how they felt.

For the remainder of the first and subsequent sessions, participants were free to move around the virtual environment, interacting with objects and playing the two mini-games as many times as they pleased. To simulate exposure hierarchies used in proven desensitisation approaches [7], the stimulus would be played at the corresponding distance (see Table 7.3). This began from at the furthest distance during session one. However, from the second session the virtual sound source would only be moved closer to the participant if in the previous session they displayed no signs of distress, voluntarily interacted with exposure mechanics and gave no negative feedback at the end of the session.

After playing the game each participant would repeat the same audio questionnaire described in Section 6.5.3 to record their own perceived emotional response towards the target stimuli. Each session then ended with the investigator asking how they felt about the game and if they had any negative or positive feedback.

6.5.4 Equipment Setup

All audio and visual stimuli were rendered using the Oculus Rift CV1 HMD. Head tracking was also achieved using the HMD, with motion tracking of participant position calculated by the Oculus Rift sensors. Player input was achieved using the Oculus Touch controllers.[1] Throughout each session participants were permitted to move freely around a pre-defined tracked experimental space of 1.6 m × 1.6 m while wearing the HMD. Audio was presented to the participant using the Sennheiser HD-650 headphones.

Figure 6.3: The Oculus Rift CV1 HMD, Oculus Touch controllers and Oculus Rift Sensors

---

[1] https://www.oculus.com/rift/

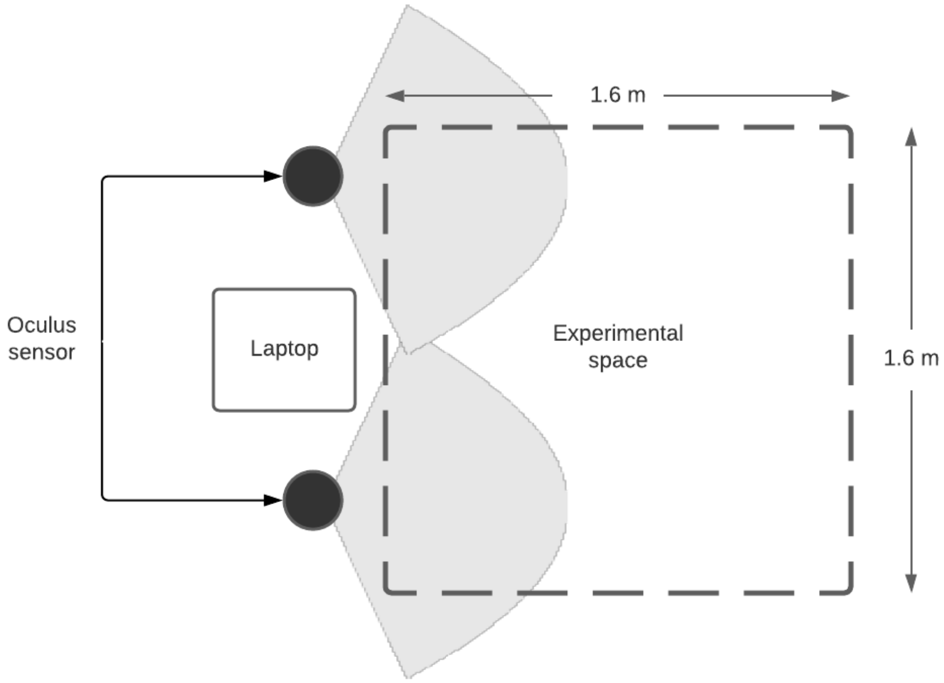


Figure 6.4: Experimental equipment layout, displaying placement of laptop and Oculus Rift sensors with 100 ° field of view.

6.5.5 Data Collection

The primary aim of this study is to evaluate if the use of binaural based Ambisonics can be used to reduce negative associations with problematic sound displayed by autistic people. Research outcomes were evaluated using quantitative measurements, both of which are outlined below.

Primary Outcome

SRER: An assessment was used to measure the emotional associations each participant experienced when presented with both target and non-target stimuli. Using a standalone application developed in Unity3D, participants were required to self-report their perceived emotional response towards the presented auditory stimuli. Once the survey is completed, the software will export the participant's response to each sound in txt file format. The application utilised an analogue Likert scale between 1-6 depicted by simple face images and accompanied by short descriptive text, whereby 1 represented 'very happy' and 6 represents 'very sad' (see Figure 6.1). Likert scales using smiley faces have been used extensively in the subjective measurement of children's emotional and physical responses in research [457]–[461]. By presenting a set of smiley face images this technique provides the child with an effective method to communicate their own subjective assessment of the situation or question, regardless of language or reading abilities [462]. This is therefore an effective tool for autistic individuals who often experience difficulties in communication and emotional recognition. Importantly, this approach has been used as an assessment tool in autism research for a VR intervention targeting the reduction of social anxiety [463] and in the investigation of the *Sinbad and the Magic Cure* project [29], a serious game developed to address auditory hypersensitivity in children with autism.

Secondary Outcome

Tracked Voluntary Participant Interaction with Target Auditory Stimuli: During each experimental session data is recorded within the SoundFields application that documents voluntary interaction with target stimuli. This is achieved by measuring the total amount of time each participant engages with the exposure based mechanics of mini-games that result in the rendering of target audio stimuli. For a full description of the mechanics please refer to section 5.3.3 in Chapter 5. As mentioned in Section 6.5.2 each target stimuli is presented to participants a maximum of 20 times, therefore the highest possible value recording is ≈ 200 seconds. Once the experimental session is completed a text file is exported that contains the total tracked voluntary interactions for each of the target stimuli.

Despite the extensive use of a smiley face Likert scale in research, there is a small amount of literature conducted with neuro-typicals that observes validity issues with this technique due to issues in communication [464] and social desirability bias [465]. With this in mind, recording voluntary interaction could bypass these challenges as well as circumventing emotional recognition problems that may occur as a result of the core symptoms associated with autism during the subject self report of emotional associations.

6.6 Results

This section presents the results from inferential and descriptive statistical analysis of participant self-reported anxiety and the recorded tracked interaction times across the four week experimental period. All the statistical analysis was performed in IBM SPSS

Statistics 26. All data recorded from the investigation is located on the data folder supplied with this thesis, see Appendix B.

6.6.1 Self-Reported Emotional Response

Table 6.3 shows each of the six participant's problematic auditory stimuli alongside the pre and post intervention self-reported anxiety scores. Baseline measurements for all participants ranged from 4 to 6, whilst post intervention measurements ranged from 3 to 5. All participants experienced a decrease in their self-reported emotional scores following the 4 week intervention. The average decrease was 1.7 points, with a minimum decrease of 1 point and a maximum of 3 points. The mean self-reported levels of anxiety across all participants for both target and non-target stimuli are displayed in Figure 6.5. Finally, a pairwise comparison of pre and post primary assessment scores was conducted using the Wilcoxon Signed-Rank test. This test was chosen due to the small sample size of the study group. Statistical analysis indicated that the score means significantly decreased between the baseline and Week 4 measurement ($Z = -2.232$, $p = 0.026$). These results therefore confirm $H1$. Finally non-target stimuli were played to each participant during each measurement session, the average decrease in emotional response scale points for these was 0.2 points. Statistical analysis using the Wilcoxon Signed Rank test revealed that there was no significant differences between pre and post test scores ($Z = -1.219$, $p = 0.223$) , thus providing support for $H2$.

Table 6.3: Particpant Self-Reported Levels of Anxiety Scores.

Participant	Stimulus	Baseline SUD score	Final Score SUD score
A	Group Singing	5	3
	Baby Crying	6	4
B	Group Singing	5	3
	Children Screaming / Shouting	4	3
C	Children Screaming / Shouting	6	5
D	Sirens	5	3
	Children Screaming / Shouting	5	3
E	Children Screaming / Shouting	5	4
F	Children Screaming / Shouting	4	3
	Baby Crying	6	3

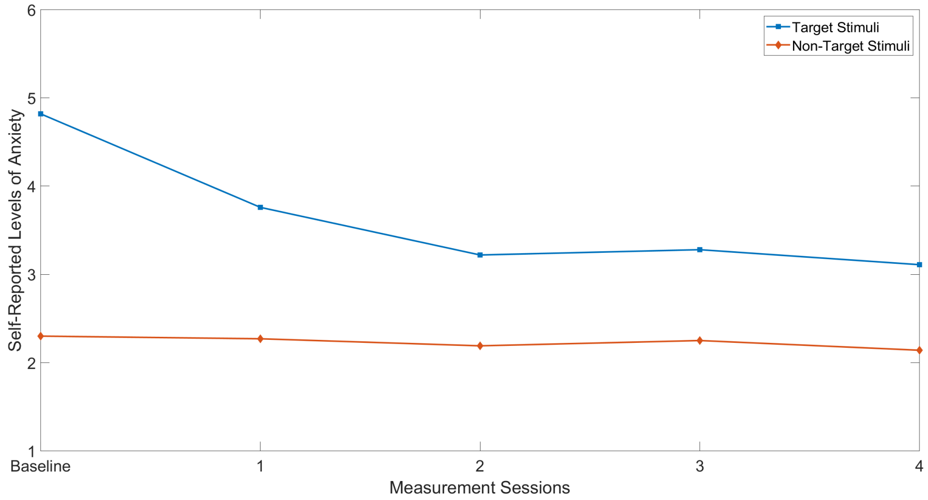


FIGURE 6.5: Calculated mean of all participants self-reported levels of anxiety for both target and non-target auditory stimuli.

TABLE 6.4: Participant Tracked Interaction Time for each target stimuli across all four experimental sessions, represented in seconds.

Participant	Stimulus	Session 1	Session 2	Session 3	Session 4
A	Group Singing	71.45	44.39	68.13	100.98
	Baby Crying	43.00	57.74	89.45	81.57
B	Group Singing	92.00	75.03	60.03	84.26
	Children Screaming / Shouting	71.45	44.39	68.13	100.98
C	Children Screaming / Shouting	65.84	61.31	23.06	71.82
D	Sirens	70.32	41.91	93.87	100.53
	Children Screaming / Shouting	89.02	63.55	84.20	100.87
E	Children Screaming / Shouting	38.46	45.04	53.19	51.93
F	Children Screaming / Shouting	100.07	73.03	99.30	99.15
	Baby Crying	82.21	78.41	99.95	92.55

6.6.2 Tracked Interaction Time

Table 6.4 displays the total amount of recorded voluntary interactions with target auditory stimuli for all participants across all four intervention sessions. Statistical comparison between the first and last sessions using the Wilcoxon pairwise test indicated a significant increase ($Z = -2.201$, $p = 0.028$) in the mean amount of tracked interaction time for each participant's target stimuli. Figure 6.6 shows the total tracked exposure times for all participants across the four week intervention period. Despite a gradual increase shown during sessions one, three and four, times are diminished during session two. With these results considered, $H3$ can be accepted.

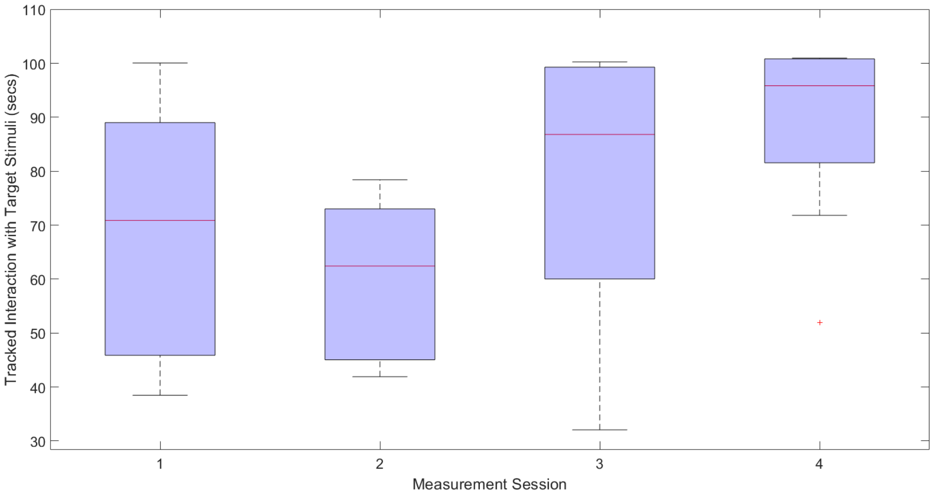


FIGURE 6.6: Box plot displaying the tracked interaction time with target stimuli for all participants across each session.

6.7 Discussion

This study evaluated the potential of using binaural based 3rd-order Ambisonics as a rendering technique to present feared auditory stimuli to the participant in a VR game targeting auditory hypersensitivity in autistic young people. Similar to the use of realistic visual three-dimensional stimuli rendered in VR for desensitisation to specific phobias and environments that may provoke anxiety in autistic individuals [237], [466], [467], SoundFields produces realistic three-dimensional auditory stimuli that can simulate the movement and acoustic environments of feared sounds. For any form of VRET to be considered successful, the virtual environment must be capable of eliciting emotions such as anxiety in order for new positive associations with the stimuli to be formed through controlled and graduated exposure [468]. The current study found that all of the target stimuli activated a level of anxiety in participants during baseline measurements which resulted in the higher scores reported in the self-reported emotional assessment scale, representing negative emotions associated with the stimulus. These results are consistent with similar research conducted with neurologically healthy participants, reporting that the use of spatialised sound can activate the anxiety provoking structure [243], [337]. Furthermore, primary outcome measurements for non-target stimuli display no significant changes across the four week experiment. In the absence of an experimental control period or group, these values can serve as further justification that the rendered stimuli can cause the negative emotional associations required for an exposure based intervention.

The presented intervention was successful in reducing the participants perceived anxiety towards target auditory stimuli over four weekly playing sessions. This was evident from the significant decrease in the self-reported emotional response scores in the pre and post study measurement sessions, which was the main outcome measure highlighted in Table 6.3 and Figure 6.5. This falls in-line with similar research that reveals a reduction in stress associated with presented audio stimuli within a computer game environment [29], [30]. Additionally, the increase in time participants voluntarily interacted with target stimuli can also be interpreted as an increase in tolerance. Results show a significant increase in tracked interaction time between sessions one and four. This is further supported by the implemented exposure hierarchy which moves the virtual sound stimuli closer to the participant. However, interestingly there is a drop in the amount of tracked interaction during session two (see Figure 6.6). This could be explained by the reduction in virtual distance between the participant and the target stimuli from 25 to 15 meters which would result in a perceptible rise in loudness after the first session.

Research has observed that much like their typically developing peers, autistic children often enjoy playing computer games during their spare time [469], [470]. This has been echoed in the reported effectiveness of serious game interventions used to improve sensory integration [471], social communication [73], emotional recognition [472] and auditory hypersensitivity [29], [30]. By integrating therapy frameworks into computer game mechanics it is possible to motivate participants to engage with the intervention repeatedly over time. This was experienced during the present study with all six participants playing the game for the 30 minute session across the entire four week period, creating an opportunity for the experimental group to have repeated graduated exposure to averse stimuli.

In addition, the mechanism by which participants were exposed to target stimuli is also an important consideration. It is possible for each player to complete each gaming session without any exposure to target stimuli. However, the system that rewards players with in-game currency motivates them to interact with the game mechanics which would expose them to perceptually abhorrent sounds, potentially developing new and positive associations with the stimulus.

Finally, another meaningful observation was the shared negative emotions associated with particular stimuli between several of the participants. Work investigating phobias in autistic children has also noted common themes such as a fear of buses, toilets, weather and social situations [34], [473]. With this in mind, future adaptations of SoundFields could include a library of sounds which can be integrated into the application. This could negate the need for bespoke variants and increasing the accessibility of this approach to therapy.

6.8 Limitations

Despite positive results, the investigation presented in this chapter does however have caveats that will need to be addressed. Firstly, it is not possible to determine if the use of binaural based spatial audio specifically had a positive impact on the outcomes of this study. For this to be achieved the following investigation requires two experimental groups with a larger sample size. This would allow a direct comparison between 3^{rd} order Ambisonics and traditional audio rendering techniques such as stereo. Furthermore as previously mentioned in the discussion, without a crossed carry-over randomised controlled design, it is not possible to distinguish if the repeated exposure to stimuli during measurement sessions had an impact on the final results. Finally, the future study will benefit from additional follow up measurements to provide data to assess the generalisation of new positive associations.

6.9 Summary

This chapter presented a study designed to answer the following research question: *Can negative emotional associations to specific sounds be reduced in autistic children and adolescents by using 3^{rd} order Ambisonics within a VR exposure based game?* This was accomplished by measuring the participant's self-reported emotional associations with the auditory stimuli, and by recording voluntary interactions with exposure based mechanics.

The experiment observed a decrease in participant self-reported levels of anxiety towards target stimuli after just four 30 minute sessions playing the game. In addition, despite the increase of exposure levels as the experiment progressed, the time each participant exposed themselves to target stimuli also increased. These results therefore indicate that binaural based Ambisonics could be an effective tool for use in a therapy framework that targets auditory hypersensitivity in autistic people. However, at this stage it is not possible to determine if Ambisonic audio has a greater impact upon increasing tolerance towards target sounds. This could only be achieved via the study in the next chapter which directly compares data from experimental conditions listening to either Ambisonic or stereo audio renders.

Another noteworthy outcome of this study is the positive effects playing a VR game designed for therapy can have on an autistic child or adolescent. Throughout the entire four week investigation, all of the participants who completed the study enjoyed the sessions and actively engaged with the game and the simple therapy centred mechanics. Although VR has been an extremely active area of research for autism for over two decades, this is the first VR application to target auditory hypersensitivity. The key reasoning behind this is the increased availability of VR technology and spatial audio production tools to the consumer market in recent years. This study could therefore hopefully enhance the support for VR as a tool for delivering engaging therapy frameworks to autistic people, and at a time when this technology is becoming evermore accessible.

Chapter 7

Evaluating the effect of Ambisonic Audio in the Reduction of Auditory Hypersensitivity in Autistic Young People

7.1 Introduction

This chapter presents a longitudinal study designed to further explore the research outcomes achieved in Chapter 6. Results indicated that spatial audio could be used within an exposure based VR game targeting auditory hypersensitivity, with all participants displaying a significant decrease in self reported anxiety following the experimental period. However despite the positive results, the lack of comparable data with other audio rendering techniques means it is not possible to truly evaluate if the increased realism achieved through spatial audio would have any significant impact on any increase in tolerance experienced by the individual. With this in mind, this chapter compares the use of binaural based 3$^{\text{rd}}$ order Ambisonic renders of problematic sounds with those delivered using standard two-channel stereo.

By adopting a similar protocol and outcome measurements to that found in the previous chapter (see Chapter 6) but with the introduction of two audio rendering conditions, a direct comparison can be made between levels of SRER and the total amount voluntary interaction time with target stimuli displayed by the two experimental conditions. By taking into account previous research investigating interventions for auditory hypersensitivity and emotional responses to spatial audio, it is expected that those listening to Ambisonic renders of averse sounds will experience higher reductions in SRER. However, at this point it is unclear if any differences displayed between the two conditions will be of statistical significance.

To the author's knowledge Chapter 6 was the first study to investigate a proposed VR intervention used for auditory hypersensitivity. As mentioned before, the results were encouraging, however the small sample size ($n = 6$) and the lack of control period raised particular issues regarding low statistical power, inflated effect size estimation and influences on results from outside factors such as training effects. To minimise the affect of these issues this study required a larger experimental population and a further scientific control. By implementing these it would produce further validation for the use of this approach for reducing negative associations with environmental sounds for autistic people.

7.2 Hypothesis

The investigation presented in this chapter aims to answer the following research question: *Compared to stereo rendering, can the use of 3^{rd} order Ambisonics as a means to deliver problematic stimuli improve the outcomes of a VR exposure based therapy for auditory hypersensitivity in autism?* Based upon this research question, the following hypotheses have been developed:

- H_1: Participants in both audio conditions will display significantly lower levels of self reported anxiety following the experimental period.
- H_2: Participants in the 3D audio experimental group will display significantly lower levels of self reported anxiety than those within the stereo rendering group following the experimental period.
- H_3: Participants in the 3D audio experimental group will display significantly lower levels of self-reported anxiety recorded at the four week follow-up than those in the stereo audio group.

- H_4: Self-reported anxiety scores for both target and non-target stimuli across both groups will not be effected by the control period.
- H_5: All participants will record significantly longer amounts of voluntary interactions with problematic auditory stimuli between the first and last experimental session.
- H_6: Participants in the 3D audio group will record significantly longer amounts of voluntary interaction with problematic auditory stimuli than those in the stereo rendering group.

7.3 Experimental Design

Participants were randomly allocated to one of either two experimental conditions, a 3D audio group or a stereo audio group. The 3D audio group would be exposed to auditory stimuli delivered via head-tracked binuaural based 3^{rd} order Ambisonics rendered over headphones. Those in the stereo audio group would listen to a head-tracked stereo representation of environmental and target averse audio rendered over headphones. The investigation followed a in-between subjects crossover study design to test the hypothesis (see Figure 7.3). Participants from both experimental groups took part in both control and experimental periods, Group 1 in an A-B, and Group 2 in a B-A sequence, however each participant would experience the experimental condition based upon their pre-assigned experimental audio rendering group. Each period consisted of four sessions (one per week) with a duration of 40 minutes. During the experimental period participants would have the opportunity to interact with game mechanics that exposed them to target stimuli. Control period sessions would involve game mechanics with no target stimuli exposure. All sessions were conducted at the school the participant was recruited at and each participant was accompanied by a member of school staff.

Primary outcome measurements were recorded at a baseline assessment stage and at the end of the final experimental and control period session. A primary outcome follow-up assessment took place 4 or 5 weeks after the experimental period (see Section 7.5.5). This difference in time is a consequence of a 1 week school holiday falling at the end of the first four week period. Those in Group A recorded at 5 weeks, those in Group B recorded at 4 weeks. Secondary outcome measurements were recorded during each experimental period session (see Section 7.5.5).

Employing a crossover study design has two key advantages for an investigation of this classification. Firstly each participant acts as their own control. Therefore when comparing experimental periods a cross over study should produce a reduced volume of standard errors by eliminating between-subject variability [474], [475]. Secondly, the more precise estimate of treatment effect consequently allows for a smaller sample size. This presented a substantial benefit during participant recruitment as the investigation was limited in terms of timescale and accommodating the schedules of the participants and their educational institutions.

7.4 Participants

The experimental group consisted of 22 children and adolescents (18 male and 4 female, mean age = 12.23, SD = 1.56, range of 8–15 years) who were recruited through three special education schools. All participants had a formal diagnosis of autism spectrum disorder obtained from their local national health trust. They displayed social, cognitive, and motor abilities that allowed them to interact with the game intervention and VR equipment. Exclusion criteria were self-reported hearing problems; an inability to finish the task. An experiment information pack was provided to the parent/guardian of each participant which included the informed consent form (see Appendix B). In addition, parents/guardians were required to provide information about any sounds which the participant may find either annoying or distressful.

Out of the 22 participants recruited, a total of 20 completed the experimental period (Group 1: $n = 10$, Group 2: $n = 10$), with two not being able to complete all experimental four sessions. A total of 14 participants successfully completed the control period (Group 1: $n = 10$, Group 2: $n = 4$). This was due to school closures put in place as a result of the COVID'19 pandemic.

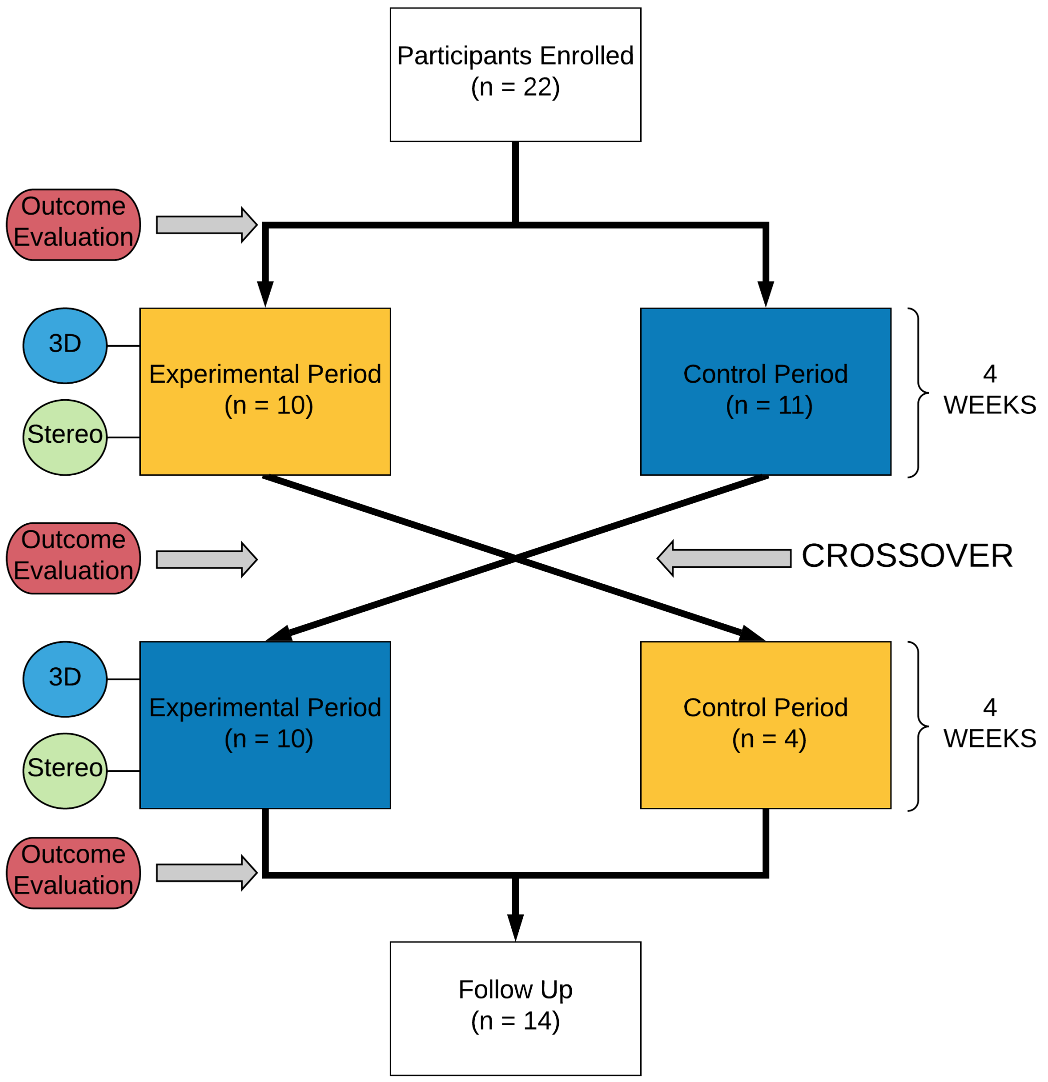


FIGURE 7.1: Crossover study design.

7.5 Method and Materials

7.5.1 Equipment

All visuals were rendered using the Oculus Rift S head mounted display using an MSI GP74 gaming laptop. Audio was delivered using Sennheiser HD 650 open back headphones. Participant head rotation and positional data within the virtual environment was tracked with 6DoF using the Oculus Rift S cameras. Participants controlled the in-game avatar using the Oculus Touch controllers. Throughout each session participants were permitted to move freely around a pre-defined experimental space of 1.6 m × 1.6 m while wearing the HMD.

FIGURE 7.2: The Oculus Rift S HMD and Oculus Touch controllers

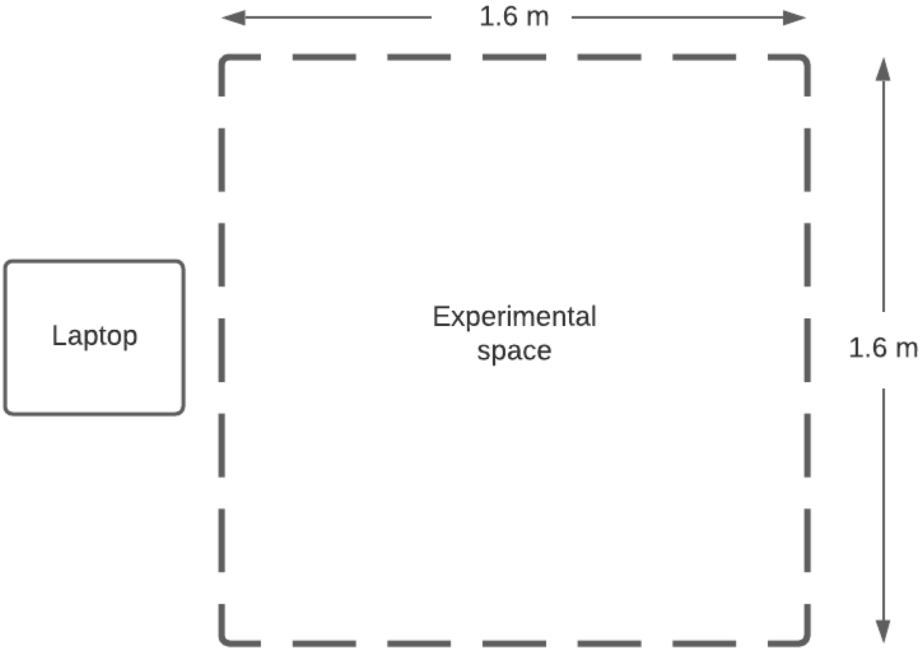


FIGURE 7.3: Experimental equipment placement

7.5.2 Game Interventions

Experimental Period

Extending the investigation conducted in Chapter 6, SoundFields was again used for the experimental condition. As mentioned in Section 2.6.1, collating feedback from autistic children through participatory design is an essential aspect of developing interactive technology based interventions for this population [250], [253], [256]. Although no formal PD sessions took place during the investigation presented in Chapter 6, some feedback regarding the virtual environment and game mechanics was collected. As mentioned in Section 6.5.3, alongside asking how they were feeling each participant was asked what they liked and disliked about the VR game. A number of the participants noted that the game would benefit from supplementary non-player characters to serve as enemies to the player. This would provide an extra level of challenge to help maintain player interest in the mini-games. Therefore, in addition to the core mechanics, during both mini-games players must defend themselves against non-player characters that appear throughout the game's environment (see Figure 7.4). When players are hit by the projectiles produced by the non-player character, one gem is deducted from the score. Although these are not part of the exposure therapy framework, the inclusion of these characters aims to scaffold the game narrative whilst also increasing the challenge and interest for the player. These characters can also be permanently deactivated should the player want them to be.

Often participants would end the session with a surplus of gems earned during game-play, resulting in feedback requesting additional items to purchase with in-game currency. To overcome this the player was given a pair of virtual gloves which accurately simulate hand movement. The addition of the virtual gloves presented another opportunity for avatar customisation, expanding on the amount of items that can be purchased in the in-game shop. These require a higher quantity of virtual currency than the wand variations and therefore aim to increase player motivation to engage with exposure game mechanics in order to receive greater rewards. Furthermore, in the previous version of SoundFields the player's magic wand was solely controlled with the right hand, this could present difficulties in terms of control for left handed players. Extending interaction to both hands increases the accessibility of the game to both left and right handed players.

Figure 7.4: New non-player character added to SoundFields. Character would spawn at random intervals and at random positions throughout the game-play session.

Control Period

An additional virtual environment was included in the SoundFields application for use throughout the control period of the investigation and included a further two mini-games. In-game control remained the same as the experimental period VE, with the player interacting with the environment and game mechanics using the virtual magic wand. Furthermore, players were free to repeat each mini-game as many times as they pleased. The system also incorporated a reward system in which in-game currency would be rewarded for completed tasks. Mini-Game One required the player to repeat presented random sequences of coloured shapes and tones. If the player is successful the series will become more complex by incrementing itself by one step. During game-play the mechanics are controlled by shooting the corresponding shape using the virtual magic wand and for each correctly repeated sequence the player is awarded with one gem. If the player incorrectly repeats the sequence they are given positive reinforcement in the form of visuals and music. The virtual environment for this game was designed to resemble a tropical beach, similar

FIGURE 7.5: An addition of gloves to the virtual shop, giving the player more opportunity for avatar customisation

to the experimental period VE this area also included simple interactive environmental elements (see Figure 7.6).

FIGURE 7.6: Control period Mini-Game One.

The second mini-game was comprised of a cartoon style woodland which is displayed in Figure 7.7. When the game begins, the player must use the virtual wand to target a group of moving non-playable characters resulting in them disappearing. When the entire group of characters has been successfully removed, the player is awarded one gem and a new cluster with increased numbers is rendered. In order to challenge the player and maintain interest and motivation, non-playable characters generate virtual projectiles that reduce speed when within 2 meters radius of the player. This allowed the player to circumvent the projectile. Failing to do so results in the loss of one of three lives. When all three lives are spent there is an opportunity to restart the game.

FIGURE 7.7: Control period Mini-Game Two and non-player characters.

7.5.3 Auditory Stimuli

Throughout the experimental period the auditory stimuli was rendered to the participant dependant upon their pre-allocated experimental group. This incorporated both environmental and averse target sounds. Two versions of SoundFields were built which either rendered sound using stereo or Ambisonics. Each participant would play the game appropriate to their assigned experimental condition. Stimuli were presented a maximum of 20 times, using the same format explained in Chapter 6.

New averse target stimuli were created to extend the built-in stimulus library used in Chapter 6. This was based upon further sounds highlighted by parents/carers during the

Table 7.1: A list of all target stimuli alongside their use within the corresponding experimental conditions

Condition	**Target Stimulus**	**Participants (n)**
3D Audio	Alarm	2
	Baby	2
	Engine	2
	Fireworks	2
	Hair Dryer	1
	Children Fighting	4
	Children Playing	7
Stereo	Alarm	4
	Baby	2
	Engine	2
	Fireworks	3
	Hair Dryer	3
	Children Fighting	2
	Children Playing	3

recruitment phase. Furthermore, to accommodate for the stereo experimental condition the stimulus library was reproduced in stereo. This can be achieved by bypassing the Resonance plugin in Wwise, however the Ambisonic files would still contain the distance attenuation and movement simulation effects produced by using the Facebook 360 Spatial Audio Workstation plugins. To replicate previous studies that have utilised stereo rendering [30], an exposure level hierarchy was realised by increasing the volume of the stimuli in a step-wise fashion. Sound levels were matched to the corresponding exposure level in the binaural spatial audio experimental condition. Finally for non-static sound sources, simple movement was replicated by stereo panning between the left and right channels of the audio file. All stimuli from both experimental conditions can be found in the accompanying data file outlined in B.

7.5.4 Experimental Procedure

Baseline Assessment

Baseline measurements were recorded one week prior to the beginning of the intervention. Each participant completed an identical audio based questionnaire in which they would rate their emotional perception of specific sounds. A series of emojis were designed to translate an analogue scale into graphical information that could be understood by the participant and bypass any possible communication impairments (see Figure 6.1 and Section 6.5.5). A total of sixteen different sounds were included, eight representing all disturbing sounds provided through parent questionnaires, and seven ‘relaxing’ soundscapes taken from the Eigenscape database [446]. All audio was presented twice in random order to collect a consistent response from particpants for each stimuli. Audio was delivered using binaural based 3$^{\text{rd}}$ order Ambisonics at the highest exposure level. Participant responses to the presented stimuli recorded in Unity3D and exported as a .txt file. In addition, all participants during this session used the VR equipment to become accustomed to the controls and head mounted display.

Participants would be randomly assigned to the experimental condition based upon the order to which they arrived at the baseline measurement session. Following this session, target stimuli would be allocated based upon the two stimuli with the highest SRER values. If more than one stimuli shared the same value the stimulus would be designated using a random number generator.

Session Procedure

Experimental Period

Each participant was given the opportunity to play the SoundFields game for a period of approximately 30 min, once a week over the course of four weeks. Throughout the experiment a support worker would be present to communicate with the participant and provide assistance if they became distressed. However, they were not permitted to deliver instructions. At the beginning of each session the investigator would select the target stimuli for the participant from a library of sounds built into the application. When the participant began the first session, they were explained that the goal of the game was to collect orbs throughout the environment and in return they would be rewarded with one unit of in-game currency. It was also explained that golden orbs would play sounds they may find annoying or not like, but if they were successful in collecting them they would be rewarded with ten units of in-game currency. The participant would then be shown the in-game shop and explained that it is possible to purchase new items using the currency they have earned. The investigator would also tell the participant that if at any time they wanted to stop then the session would end. The session would also end if the participant became distressed or if the member of staff deemed it necessary. Throughout each session, the investigator would limit interaction with the participants to giving assistance with gaming controls and to provide verbal praise when they successfully interacted with objects that emitted problematic target sounds followed by asking how they felt.

Participants were free to move around the virtual environment and play the available mini-games as many times as they pleased. During game-play each of the two target stimuli assigned to the participant could appear within the game a maximum of twenty times. To simulate exposure hierarchies used in proven desensitisation approaches [7], the stimulus would be played at the furthest distance based on the current session (see Table 7.3). In addition, target stimuli would be delivered to the participant based upon their assigned experimental condition. From session two, virtual auditory stimuli were moved closer to the participant at the beginning of each session. However this would only be done if in the previous session the participant showed no distress, voluntarily interacted with exposure based mechanics and gave no negative feedback at the end of the session.

All sessions finished with the investigator asking the participant how they felt about playing the game and if they had enjoyed themselves. They would also be asked what they liked and disliked about the session. After 4 weeks the participant would complete the audio based questionnaire.

Table 7.2: Exposure hierarchy. Each exposure level corresponds to the distance between the participant and the virtual sound source for participants in the 3D audio group; distance is represented in metres.

Session	**Virtual Distance between Player and Stimulus**
1	25 m
2	15 m
3	5 m
4	2.5m

Table 7.3: Exposure hierarchy for Stereo audio group. Each exposure level corresponds to an amplitude matched to the relative 3D audio exposure level

Exposure Level	**≈ Audio Level (LUFS)**
1	-53
2	-48
3	-26
4	-21

Control Period

Sessions during the control period would follow the same format as those in the experimental period. However, the key difference would be that the available mini-games would not contain exposure based mechanics which played target stimuli. Measurements recorded during this time provides an additional baseline to investigate if the intervention had any effect. Please see Section 7.5.2 for full game details.

Follow-Up Assessment During follow-up assessments, participants would complete the the same audio based questionnaire used to collect the primary outcome.

7.5.5 Data Collection

The purpose of this investigation was to determine if differences in audio rendering techniques used for delivering target audio stimuli to participants had an effect on any habituation towards participant specific sounds following the experimental period. To achieve this the primary and secondary outcomes used in Chapter 6 were implemented in this study.

7.6 Results

This section presents the results from inferential and descriptive statistical analysis of participant self reported anxiety and the recorded tracked interaction times for both the 3D audio and stereo experimental condition groups. All the statistical analysis was performed in IBM SPSS statistics 26 [476]. All data recorded from the investigation is located on the data folder supplied with this thesis, see Appendix B.

7.6.1 Self-Reported Emotional Response

SRER data for target stimuli collected during the investigation for both experimental conditions was found to be normally distributed as indicated by a Shapiro-Wilk test [477], [478]: 3D audio Group ($W(34) = .970$, $p = .473$) and stereo audio group ($W(34) = .946$, $p = .094$).

The mean scores of SRER levels recorded for target stimuli across both audio conditions are shown in Figure 7.8. These results display a reduction in self-reported anxiety for both experimental condition groups when comparing data collected at baseline and after the four week experimental period. A HLM with no random intercept was employed to evaluate the effect that the experimental period, control period and audio rendering techniques had on the mean self reported anxiety scores for target audio stimuli. The residual covariance structure was specified as Compound Symmetry (CS). The model investigated the effects of the experimental periods and the experimental audio conditions on the self-reported anxiety in response to target stimuli. This was based on SRER values recorded at each outcome measurement session (see Figure 7.3).

By examining Table 7.4 it can be seen that the experimental periods had a significant effect on the overall reduction in SRER values across both experimental periods. Further post-hoc pairwise comparison between baseline and post-experimental period scores also revealed a significant reduction in estimated marginal means (see Table 7.5). Therefore H_1 can be accepted. In addition, Table 7.5 reveals no significant differences between self-reported anxiety levels recorded at baseline and post-control period sessions, leading to the acceptance of H_4. Finally, 7.5 demonstrates no significant change in SRER between the post-experimental period and the 4 week follow-up measurement. This suggests that any reduction in anxiety associated with target stimuli across both conditions was maintained for at least 4 weeks. However, statistics indicate no significant differences between the two experimental conditions, therefore H_3 is rejected.

In regards to the effect of the experimental condition, HLM analysis indicates that the use of binaural based 3$^{\text{rd}}$ order Ambisonics does have a significantly positive effect upon the participants self reported anxiety towards the target auditory stimuli presented within the VR environment. In light of these results, H_2 can be accepted. Finally, results in Table 7.4 specifies there is no significant interaction between the audio condition and the experimental periods.

Figure 7.9 shows SRER levels for non-target stimuli recorded for both experimental conditions across the entire investigation. The plot shows extremely similar levels for both groups with very little difference between the values across all four outcome measurement sessions. In addition, non-target stimuli data was analysed using the same statistical tests as the target stimuli, the results of which can be seen in Tables 7.6 and 7.7 . For both

Table 7.4: HLM statistical results displaying the effect of both experimental conditions and periods on SRER values for target stimuli.

Source	Numerator df	Denominator df	F	Sig.
Experimental Condition	1	17.773	6.783	.018
Experimental Periods	3	45.907	20.547	.000
Group * Week	3	45.907	2.183	.103

Table 7.5: Post-hoc statistical analysis investigating differences in SRER values between baseline and sequential measurement sessions for target stimuli.

(I) Measurement	(J) Measurement	Mean Difference (I-J)	Std. Error	df	Sig.^{c}
Baseline	Post-Experimental	1.425*	.209	43.522	.000
	Post-Control	.221	.232	47.336	1.000
	4 Week Follow Up	1.185*	.232	47.336	.000

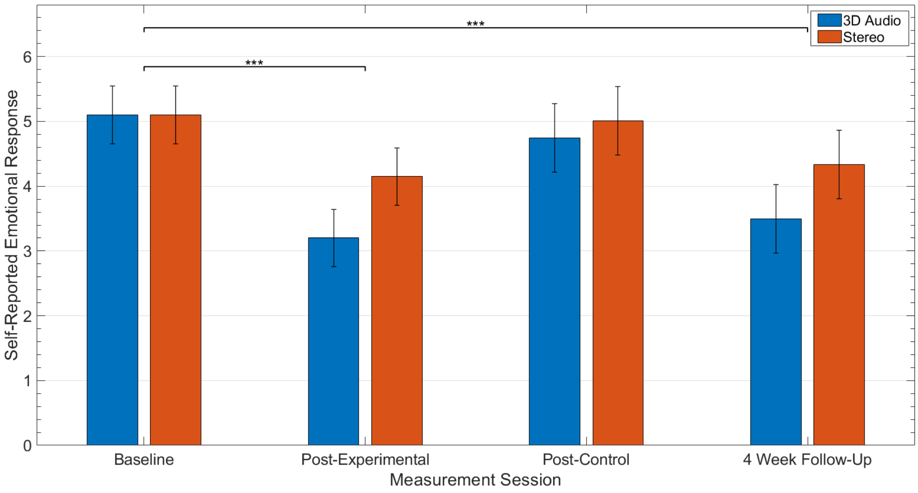


Figure 7.8: Mean self-reported emotional response to target stimuli for both experimental conditions. Whiskers denotes ±95% confidence intervals. P values (*** < 0.001) were determined from post-hoc pairwise comparison test and are indicated above the bars.

Table 7.6: HLM statistical results displaying the effect of both experimental conditions and periods on SRER values for non-target stimuli.

Source	Numerator df	Denominator df	F	Sig.
Group	1	17.027	1.269	.276
Week	3	41.670	.838	.481
Group * Week	3	41.670	1.066	.374

audio conditions there was no significant effect of the experimental periods on the SRER values. Furthermore, HLM analysis revealed that the audio rendering condition had no significant effect on the SRER values for non-target stimuli over the experimental period.

Descriptive statistics were used to examine the mean changes in SRER scores for each

Table 7.7: Post-hoc statistical analysis investigating differences in SRER values between Baseline and sequential measurement sessions for target stimuli.

(I) Measurement	(J) Measurement	Mean Difference (I-J)	Std. Error	df	Sig.^{c}
Baseline	Post-Experimental	.079	.073	40.854	1.000
	Post-Control	-.031	.083	42.177	1.000
	4 Week Follow Up	.065	.03	42.177	1.000

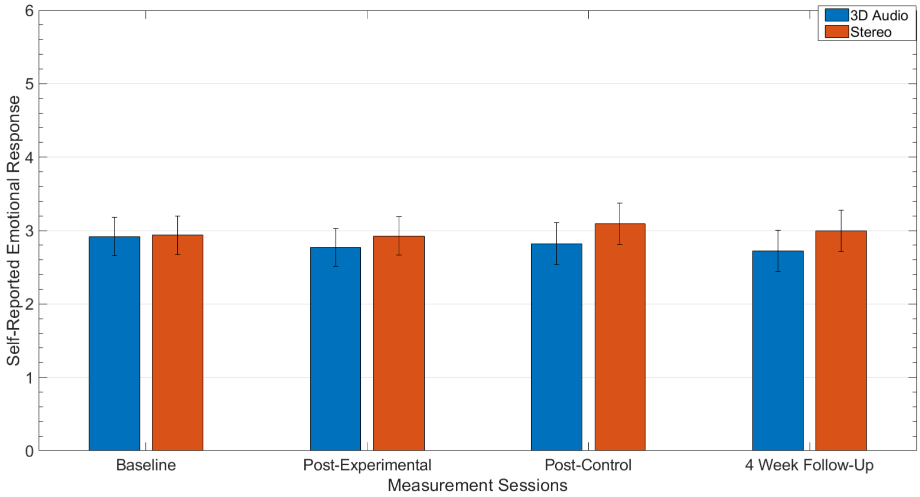


Figure 7.9: Mean self-reported emotional response to non-target stimuli for both experimental conditions. Whiskers denotes ±95% confidence intervals.

individual target stimuli between the pre and post experimental period measurements. It is important to note that although the sirens stimuli was used as a target stimuli it was done so for only one participant in the stereo condition, therefore it has been omitted from this analysis.

In reference to Figure 7.10, it can be seen that across all stimuli, with the exception of 'fireworks', the 3D audio group experienced the greater increase in changes between self-reported anxiety values between the baseline and post experimental assessments. Furthermore, Table 7.8 shows the mean baseline, and post-experimental self reported anxiety figures for each target stimulus alongside the corresponding percentage difference in SRER scores across each audio condition. In support of the data shown in Figure 7.10, it can be seen that the experimental period had a greater positive impact on those participants exposed to the 'children playing' stimulus in the 3D audio group with a decrease of SRER of 50%, this was followed by the 'children fighting' stimulus which showed a 33.18% decrease. In comparison, participants in the stereo condition experienced some of the lowest decreases in anxiety, 'children playing' experiencing a decrease of 3.77% and 'children fighting' anxiety levels reducing by 9.52%.

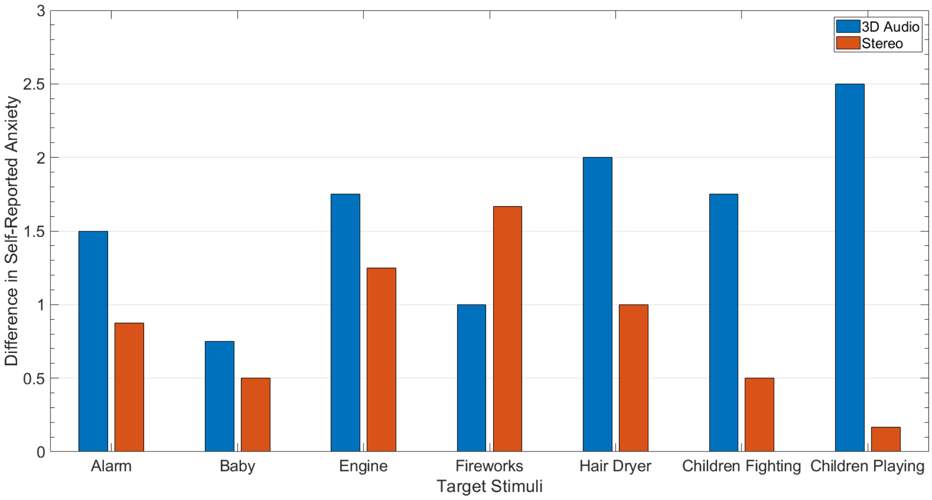


FIGURE 7.10: Decreases in negative self emotional responses for each stimuli across both experimental conditions.

TABLE 7.8: Changes in mean self-reported emotional response scores for target stimuli across both experimental conditions, showing pre and post measurement scores with percentage decrease.

Condition	**Target Stimulus**	**Participants (n)**	**Pre-Test (M)**	**Post-Test (M)**	**% Decrease**
3D Audio	Alarm	2	4.75	3.25	31.59
	Baby	2	5	4.25	10.53
	Engine	2	5	3.25	35
	Fireworks	2	5.5	4.5	18.18
	Hair Dryer	1	6	4	33.33
	Children Fighting	4	5.12	3.37	34.18
	Children Playing	7	5	2.5	50
Stereo	Alarm	4	5.25	4.37	16.76
	Baby	2	5.5	5	9.09
	Engine	2	5.25	4	23.81
	Fireworks	3	5.33	3.67	31.14
	Hair Dryer	3	5	5	20
	Children Fighting	2	5.25	4.75	9.52
	Children Playing	3	4.5	4.33	3.77

7.6.2 Tracked Interaction Times

Throughout each experimental period session, the total amount of time each participant spent voluntarily interacting with game mechanics which exposed them to target stimuli was recorded. Data from the 3D audio group was found to be normally distributed using the Shapiro-Wilk Test [477], [478] ($W(40) = .950$, $p = .0.77$), however the stereo condition failed the same test for normality ($W(40) = .937$, $p = .0.28$). With this in mind, a HLM with no random intercept was used due to its robustness with non-normal data distribution [479]. Statistical tests investigated the effect of the experimental period and audio rendering techniques on the total amount of time participants voluntarily interacted within-game mechanics that delivered target auditory stimuli. This was compared between both experimental groups and across the four week experimental period. The covariance structure for this model was specified as CS.

In Figure 7.11 it can be seen that the amount of tracked interaction time for both the 3D audio and Stereo groups increases across the four week experimental period. Moreover this is expressed in statistical testing, showing that the subsequent experimental sessions had a significant effect upon tracked interaction times ($F(58) = 46.361$, $p = 0.001$), leading to a support of H_5. In addition, this plot shows that participants in the 3D audio group did voluntarily interact more with target audio stimuli. Statistical analysis showed there was a significant interaction between the audio rendering technique and the experimental period ($F(58) = 3.825$, $p = 0.05$).

However, further analysis was conducted on each individual experimental session. When examining Figure 7.11 it is clear that there is very little difference between the two audio conditions, which is reflected in the similarity between the median lines and the distribution of data between the highest and lowest values. This is supported by an independent samples t-test that indicated no significant difference between the 3D audio ($M = 90.33$, $SD = 36.39$) and stereo groups ($M = 84.89$ $SD = 40.13$); ($t(18) = 0.317$, $p < 0.755$). Similar results were also identified for session two in which no significant effect was recorded ($t(18) = 1.523$, $p < 0.145$) despite the increased differences between the 3D audio ($M = 116.66$ $SD = 32.59$) and stereo ($M = 92.33$ $SD = 12.20$) groups. It is from session three that the contrast between the groups becomes more apparent with the experimental condition having a significant effect upon the recorded times ($t(18) = 2.218$, $p < 0.040$). The significance

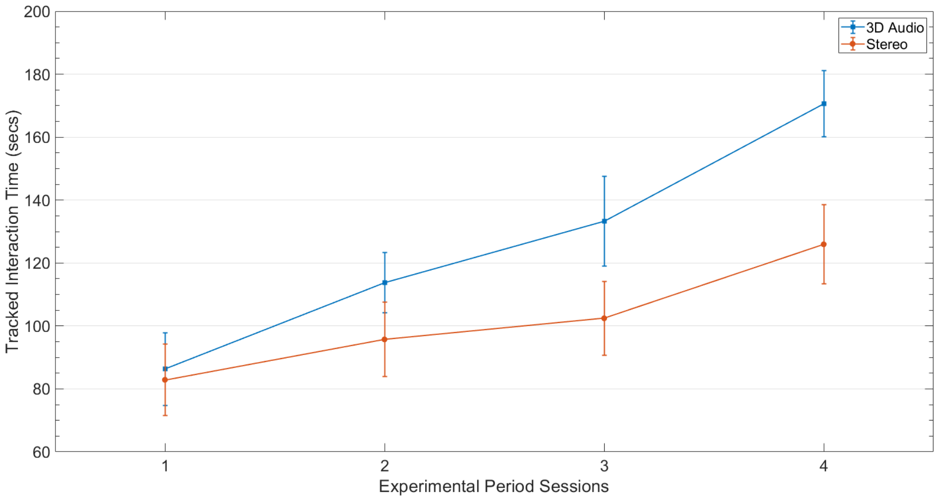


FIGURE 7.11: Mean tracked interactions times across all experimental sessions. The whiskers denote standard-error.

is then sustained with data recorded in session four ($t(18) = 2.138$, $p < 0.047$). This is reflected in Figure 7.11, where the 3D audio group ($M = 163.05$ $SD = 35.51$) can be seen interacting more with target stimuli than the stereo condition ($M = 126.95$ $SD = 39.88$) during session four.

7.7 Discussion

The main objective of this study was to compare binaural based 3$^{\text{rd}}$ order Ambisonics to stereo rendering as an approach to delivering audio stimuli in an exposure training VR game targeting auditory hypersensitivity autistic young people. In line with the outcomes of Chapter 6, results from this investigation support the use of exposure based training within a serious game environment in reducing self reported anxiety associated with specific environmental sounds for autistic people. This is also in keeping with previous research in which serious games have shown success in targeting this specific sensory processing issue [29], [30]. What further reinforces these findings is the implementation of the four week control period into the experimental design. Across both audio conditions there was no significant changes in SRER values between measurements collected before and after the completion of the control sessions. Moreover, these narrow changes in values were reproduced for non-target stimuli. Similar to Chapter 6, how the participants felt about non-target sounds remained reasonably consistent across the entirety of the eight week program. Taking into account the results from both the control period and non-target sounds, it should be considered that any impact on self-reported anxiety is specific to the exposure based game-mechanics and not a general training effect.

The comparison of audio rendering techniques was the primary objective of this chapter and focused on the core research question of this thesis. According to the statistical analysis of SRER scores, those participants listening to averse stimuli delivered via binaural based 3^{rd} order Ambisonics showed a significantly increased improvement in self-reported anxiety than those in the stereo rendering condition following the four weekly experimental sessions. Furthermore, the appreciable improvements of the participants in the 3D audio group cannot be attributed to differences in baseline measurements. Statistical analysis showed that the baseline anxiety scores for these two groups were not significantly different.

When examining these results it is important to consider again the use of VR therapy to reduce specific phobias in individuals with and without autism spectrum disorder [34], [237]. Firstly as reported in Chapter 2, this technology has the ability to accurately and safely simulate a controlled three-dimensional environment to create a bespoke exposure therapy experience according to the user's needs [480]. However, what separates VR from flat-screen solutions is the sense of presence felt by the user, and it is presence that is considered an important component in activating the relevant fear structures to achieve habituation [228]. Also, it has already been observed that autistic children experience similar levels of presence as their NT peers within virtual environments [234]. Secondly, but more crucial to the application of this research, is the role of spatial audio in influencing the sense of presence in VR. Realistic auditory environments rendered through spatial audio have been shown to increase levels of presence [245], [278], [455], [456]. Furthermore as noted in previous chapters, similar audio rendering techniques to those used in this study have been observed to elicit feelings of fear and anxiety [243], [244], [337]. With this in mind the results of this study could be expected. Those with autism often experience difficulties with imagination [481], [482] which leads to a need for stimuli to be contextualised. Therefore, participants in the 3D audio condition who experience a sound closer to what they hear in the real world in terms of localisation, dynamic movement and environmental acoustic characteristics should feel increased similarity between the real and virtual stimulus over those in the stereo audio group. This familiar interpretation of virtual stimuli rendered within virtual environments by autistic young people has also been attributed to the successful outcomes of past VR based interventions [222]–[224]. Further to this, White *et al.* [191] notes that autistic people require frequent practice with contextualised exposure in order to increase the chances of reducing anxiety associated symptoms. This could serve as a justification for the larger decreases in SRER displayed by participants listening to the *'Children Fighting'* and *'Children Playing'* target stimuli compared to those listening to *'Fireworks'* or *'Alarms'* in the 3D audio condition. This study was carried out within the school each participant was recruited during their school day. Consequently in addition to the repeated exposure to these stimuli with the VR game, the opportunity to real-world exposure of these environmental soundscapes is much higher within the school environment than the other target stimuli such as the *'Fireworks'*.

By recording the voluntary exposure to target stimuli during each experimental session it is possible to measure any increase in tolerance towards a participant's target sounds to support measured values of self-reported anxiety. When examining the results collected across the completed experimental period, Figure 7.11 shows that the 3D audio group spent longer times interacting with stimuli than those in the Stereo group. However, these were only statistically different between Sessions 3 & 4. This is most likely explained by the similar times recorded by both groups during the first experimental period session. The lack of significant differences measurements recorded in session one in terms of the mean and distribution of values could be interpreted in two ways. First, the participants would be experiencing a new virtual environment with new mechanics and so lower times could correspond to periods exploring and experimenting with the VR environment and game, subsequently spending less time interacting with exposure based mechanics. Secondly, this would be the first time each participant heard their target sounds outside of the outcome measurement sessions, and therefore are less likely to want to engage. However, comparing tracked interaction times between the two groups on a session by session basis does reveal some statistically significant differences from the third session, with the 3D audio group interacting more with the game mechanics that deliver problematic sounds. It could therefore be accepted that these results mirror the larger decreases in self-reported anxiety values between the baseline and post-experimental session measurements. Finally, one crucial feature of these results is that each session represents an increase in the exposure hierarchy with the virtual sound source being moved closer to the participant, resulting in an increase in perceptual loudness. With this in mind, an increase in voluntary interaction with a sound that the participant has personally reported a negative emotional association with makes these results more meaningful. As increases in tracked interaction times and decreases in self-reported anxiety values were recorded across both audio conditions, it is important to consider the impact game-play and the virtual environment had on the outcomes of this study.

In parallel to the advantages of using 3D audio as a technique for simulating realistic sound environments that have been identified in this study, the context in which spatial sound is delivered is also a key advantage of the format. Computer games have been successfully used in the past to target auditory hypersensitivity in autistic young people [29], [30]. However as stated in Chapter 6, this is the first game in this context that uses VR. The levels of engagement with the SoundFields game displayed by participants is consistent with those in the study presented in Chapter 6, however these results could be considered more meaningful due to the larger sample size of this investigation. Across both audio conditions all participants with the exception of two completed the four weekly experimental sessions, with six not finishing the control period due to school closures. Alongside the quantifiable results measuring engagement with exposure based mechanics described above, statements were also provided by staff from the recruitment schools (full statements can be seen in Appendix A.1 & A.2). It was noted that participants were 'highly motivated' and 'excited' to begin their weekly VR sessions. Also staff observed that participants maintained interest during the 8 week study despite some being expected to become disinterested as the study progressed. These comments and the continued interaction with exposure mechanics recorded in-game during this study, are corroborated by literature acknowledging that computer game based interventions are enjoyed by autistic people [29], [30], [74], [186]–[188], [483]. Both this and the previous study (see Chapter6) provide evidence that exposure based therapy embedded into VR based computer game mechanics can improve self-reported anxiety levels associated with problematic sounds. The voluntary interaction with target stimuli supported by the in-game reward system aimed to develop the participant's confidence and create positive associations with the stimuli (see Chapter 5.3.2). Both the primary and secondary outcomes of this investigation indicate that this applied to participants across both experimental groups. Adding to this, comments received by one recruitment school witnessed a change in tolerance to sounds with participants disregarding target stimuli as they were 'too focused on the task in hand

and having too much fun'.

7.8 Limitations

Despite the positive results from this study there are a small number of limitations that must be considered. Firstly, the small sample size is an important limitation to consider. This was not only affected by the amount of suitable candidates identified by the recruitment institutions, but also school closures implemented during the COVID'19 pandemic. Leading to 30% ($n = 6$) of the participants being unable to complete the control period. This reduces the statistical power of the results that validate the experimental sessions as the primary influencing factor of the investigation. Additionally, the sample size creates difficulties in generalising these results to the larger autistic population. Another consideration of a smaller group is a lack of representation across the broader autism phenotype. The primary data collection method for measuring emotional response to auditory stimuli accounted for participants with varying communication impairments. However, there was insufficient number of participants with cognitive impairments or co-morbid learning difficulties. This did not provide enough insight into how autistic children with varying abilities would interact with both the game itself and the exposure therapy intervention. Nonetheless, the technology employed throughout this and the preceding experiments could deliver a suitable approach to improving upon this sample size. The limited equipment needed allows for the experiments to be carried out at the recruitment schools. With a small amount of training, staff would be capable of conducting the experiment over a longer period, allowing more opportunity for participants to be recruited. This is made further possible by the automatic data collection of participant interactions within the virtual reality environment.

Despite the results identifying a significant increase in tolerance towards target auditory stimuli across both experimental conditions, there is no data which examines how the experiment impacted the participants' life outside of time spent playing SoundFields. The use of questionnaires, such as the the Parent Stress Index [484] or Short Sensory Profile [105], could provide evidence of any generalisation of tolerance towards real-world auditory stimuli which cause emotional distress. This would contribute to the evaluation computer game based CBT interventions for auditory hypersensitivity, but also aid in measuring the effectiveness of simulating these sounds using spatial audio rendering. In addition to parental questionnaires, this and preceding studies would have benefited from qualitative data collection with the aim of evaluating how the participants themselves felt about the game environment, game-mechanics and the intervention itself. Chapter 1 outlined the importance of participatory design sessions when developing game centred interventions for autistic children. The SoundFields environment and game elements are very different from the first iteration which was used as part of the study in Chapter 4, changes were made based upon the opinions and suggestions of the participants. However, these were not collected during any official participatory design sessions. For this and any future game based intervention to be successful and benefit the autistic population, it is important to listen to those with the most valuable insights. The people who will be using and benefiting from the game itself.

7.9 Summary

Expanding upon the results gathered in Chapter 6, this chapter presented an experiment designed to evaluate the use of binaural based 3$^{\text{rd}}$ order Ambisonics to address auditory hypersensitivity compared against a stereo audio rendering approach. Self-reported emotional responses were once again used as the primary outcomes to quantify any changes in tolerance towards target and non-target audio stimuli following the four week experimental period. In addition, data was also gathered during each experimental period session which recorded the total amount of time each participant voluntarily interacted with in-game mechanics that delivered target auditory stimuli.

Results from this investigation fall in line with those from the previous chapter, suggesting that a serious game based in VR can be used to reduce a player's self reported anxiety associated with specific sounds. This is based upon a decrease in SRER values following four sessions playing the game from a larger experimental sample size and with the inclusion of a control period. However, importantly the larger decrease in SRER experienced by participants in the 3D audio group supports the use of binaural based 3$^{\text{rd}}$ order Ambisonics as an approach to rendering target auditory stimuli within an exposure based training framework for auditory hypersensitivity. This apparent advantage to using spatialised sound is further corroborated by the additional time those in the 3D audio group spent interacting with exposure based mechanics over the stereo audio group. These results could be explained by the enhanced realism and presence experienced by an individual when listening to Ambisonic renders of a soundscape.

Chapter 8

Summary and Conclusions

This chapter will conclude this thesis by presenting a summary of the research contributions presented, followed by a re-statement of the core hypothesis. Finally, recommendations will be made for future research conducted within the context of spatial audio based therapy for autistic people, with some final remarks from the author.

8.1 Summary

Following the introductory chapter, Chapters 2 and 3 presented the background literature required to contextualise the novel research within the thesis. Chapter 2 introduced a number of auditory processing difficulties that can be experienced by autistic people such as impaired auditory scene analysis and interpretation of the binaural cues essential for sound localisation, all of which could have a negative impact on how autistic people decode a virtual auditory environment within VR. Following this, the issue of auditory hypersensitivity was discussed including the possible causes and the range of interventions that have been developed to reduce the negative psycho-emotional responses to environmental sounds. Keeping in mind the successful therapy frameworks used to address auditory hypersensitivity such as systematic desensitisation, the chapter ended with a review exploring how cognitive behavioural therapy has been embedded into VR and serious game technology, creating an engaging and accessible approach to therapy that is capable of utilising binaural based Ambisonics as an audio rendering system.

Chapter 3 presented the principles of binaural listening that are fundamental in sound source localisation including, inter-aural differences in time and intensity between the left and right ear, and the spectral shaping of sound caused by the pinnae, head and torso. This was followed with a description of HRTFs, which capture binaural sound cues at the point of the ear canal and can be filtered with a monaural sound source to simulate spatial sound from a specific point in space using binaural technology. Ambisonics was then introduced, an audio rendering technique capable of decomposing a three-dimensional soundfield using spherical harmonics, resulting in accurate sound source localisation with the introduction of higher-order Ambisonics. This was extended to binaural based Ambisonics which simulates virtual loudspeakers utilising HRTFs, creating an immersive soundfield which can dynamically rotate to the movement of the listener's head. Chapter 3 then discussed the important role binaural based spatial audio plays in delivering plausible virtual environments, enhancing emotional response, immersion and presence which are essential psychological mechanisms for VR exposure therapy. Finally, examples are given of when spatial audio has been at the centre of exposure therapy research.

Considering the auditory processing deficits experienced by autistic individuals, Chapter 4 presented a novel study that aimed to investigate if these difficulties in processing audio information would directly impact how individuals with autism interact with a presented virtual spatial audio environment. Two experiments were conducted with participants diagnosed with ASD ($n = 29$) that compared: (1) behavioural reaction between spatialized and non-spatialized audio; and (2) the effect of background noise on participant interaction. Participants listening to binaural-based spatial audio showed higher spatial attention towards target auditory events. In addition, the amount of competing background audio was reported to influence spatial attention and interaction. These findings suggest that despite associated sensory processing difficulties, autistic young people can correctly decode the auditory cues simulated in current spatial audio rendering techniques. Furthermore, the outcomes of this study can be used to inform approaches to sound design for future VR ASD interventions, including the VR game developed for the subsequent chapters of the thesis.

Chapter 5 described in detail the development of SoundFields: a VR game designed to address auditory hypersensitivity in autistic children and adolescents. First the guidelines for designing serious games for autistic individuals were presented. These were integrated into the development of SoundFields in order to maximise motivation and engagement with the game environment and mechanics. From here, the chapter presented an explanation of the core exposure therapy concepts and how they are embedded into the in-game mechanics. This adopts a similar approach to preceding serious games that target auditory hypersensitivity in autism by rewarding voluntarily interaction with adverse stimuli via an in-game narrative. However, the novel aspect of the game is the utilisation of binaural 3^{rd} order Ambisonics to render target stimuli. Problematic sounds are simulated in a way that would closely simulate the real world equivalent, through movement within a three-dimensional acoustic environment that responds to the head rotation of the listener. Further to this, exposure hierarchies that adjust the distance between the listener and the virtual sound source can be simulated by manipulating the frequency and amplitude content alongside adjustments to reverberant energy.

The feasibility of both SoundFields and binaural 3^{rd} order Ambisonics in reducing levels of negative emotions associated with identified problematic sounds was evaluated in a study presented in Chapter 6. The experiment was conducted with six participants diagnosed with ASD who displayed hypersensitivity to specific sounds. During the course of the investigation participants played the game weekly over four weeks and all participants actively engaged with the VR environment and enjoyed playing the game. Following this period, a comparison of pre- and post-study measurements showed a significant decrease in negative emotions linked to target auditory stimuli. Further to this, secondary outcome measurements indicated a significant increase in voluntary interaction with problematic sounds within the game environment, despite an increase in exposure hierarchy. First, these findings suggest that the SoundField environment and game mechanics are a viable intervention approach for auditory hypersensitivity. Secondly, the results validate 3^{rd} order Ambisonics as a rendering technique for target stimuli.

Finally, Chapter 7 presented a longitudinal study with a larger experimental population ($n = 20$) designed to expand upon the research outcomes achieved in Chapter 6. Although the results from the previous study were promising, a direct comparison with stereo audio rendering was required to determine if increased realism achieved through spatial audio would have any significant impact on any increase in tolerance. Results were consistent with the previous study, with a significant decrease in negative emotions associated with target stimuli displayed across both experimental conditions. However, those listening to

3^{rd} order Ambisonic reproductions of target stimuli exhibited a significantly larger decrease. These results were supported by participants within the 3D audio condition recording significantly more voluntary interactions with target stimuli.

8.2 Conclusions

The hypothesis which was specified in Section 1.2 and forms the central motivation for the research presented in this thesis is as follows:

> *The use of binaural ambisonic rendering to deliver auditory stimuli within an exposure based serious game will have a positive impact on reducing associated negative emotions of autistic children and adolescents experiencing auditory hypersensitivity.*

The investigations presented in this thesis support and confirm this hypothesis. Firstly Chapter 4 aimed to investigate the behavioural response to spatial audio events within a VR environment with varying levels of competing background audio. As described in Chapter 2, autistic people can experience particular auditory impairments which could have a negative impact on how they respond to spatial audio events within a virtual acoustic environment, these being impaired decoding of binaural cues and difficulties in auditory scene analysis. Initially, results from Chapter 4 showed that participants displayed accurate localisation and spontaneous attention towards spatial audio events. This supports the use of binaural based Ambisonics within VR applications developed for ASD. However, the study also demonstrated that SNR between target sounds and competing background audio can have a significant effect on virtual sound source localisation. This suggests that impaired auditory scene analysis observed in the physical world is maintained in virtual auditory environments.

Approaches to reproducing the SoundFields auditory environments were established in the outcomes of Chapter 4, focusing on controlling SNR between competing background audio and target sounds. By doing so, this maximised capability of binaural based Ambisonics at delivering realistic stimuli for autistic people, whilst enabling the acknowledgement of key auditory stimuli. However, results from Chapter 4 also provide further contributions outside of the context of this thesis. Literature recognises VR as an important format for delivering therapy, education and training for those with ASD. In addition, it is recognised that spatial audio plays an important role in scaffolding realistic 3D visual environments by creating a plausible and immersive VR environment. Therefore, similar sound design considerations could be implemented in future VR based interventions as a means to optimise immersion and interaction with the VE and therapy mechanics for autistic individuals.

Chapters 6 and 7 demonstrated that exposure to adverse auditory stimuli rendered using 3^{rd} order Ambisonics can significantly reduce self-reported negative emotions associated with the sound. Furthermore, a comparison to stereo rendering showed spatial audio rendering in Chapter 7 demonstrated 3D Signal-to-noise can significantly improve this reduction. These results are also supported by the significant increase in voluntary interaction with target stimuli within the VR game-mechanics, with participants listening to spatial audio recording a greater amount of interaction in Chapter 7. Taking into account literature that recognises auditory hypersensitivity in autism as a psycho-emotional response to sound, these results therefore validate the use of binaural based Ambisonics as an audio rendering tool for exposure based therapies.

However, the use of spatial audio extends past the reproduction of realistic simulations of adverse auditory stimuli. Virtual reality is the most prevalent format in which head-tracked binaural based Ambisonics can be presented to the listener. This creates opportunities for therapy to be amalgamated into serious games, creating an engaging intervention platform. Although serious games and VR have been utilised extensively in ASD interventions for over two decades, the game developed as part of this thesis is the first VR application developed to target auditory hypersensitivity. Results from Chapters 6 and 7 provide empirical evidence that supports previous research recommending serious game interventions for auditory hypersensitivity [29], [30]. Participants across both experimental conditions displayed a significant decrease in self-reported emotional responses to target stimuli, with no significant changes to non-target stimuli scores. Further to this, no significant changes in self-reported emotional responses for non-target sounds and target stimuli were recorded following the control period. Within SoundFields and other similar applications, players voluntarily repeatedly exposed themselves to adverse auditory stimuli throughout gameplay. This creates opportunities for the player to become habituated to the stimuli whilst simultaneously establishing a positive association with it. One important component of

SoundFields is the in-game motivator system that rewards players who successfully interact with mechanics that deliver adverse stimuli with an increasing exposure hierarchy. The impact of this system is primarily highlighted by the tracked interaction times displayed for both experimental conditions. This outcome measurement recorded a significant increase in voluntary interactions despite the increasing amplitude intensity of the target stimuli. Furthermore as mentioned in the statements provided by the recruitment schools, the participants enjoyed playing the game.

Computer-based interventions also increase accessibility to treatment for a number of impairments faced by autistic individuals, allowing for therapy to be implemented within a family or community environment and with consumer available equipment. Today, VR can be implemented using PC based and mobile phone technology. All of which are capable of rendering 3^{rd} order Ambisonics over headphones. The investigations presented throughout the thesis were conducted within schools, environments in which many of the participants experienced their target stimuli. In terms of an exposure based therapy this would serve as an advantage, repeating the exposure within real-world contexts. This was particularly evident for those participants who experienced negative emotional reactions to school sounds, with the 3D audio group demonstrating a greater percentage decrease than other stimuli.

8.3 Future Work

The investigations presented within this thesis offer some interesting avenues for future examinations within the field of VR based interventions for autistic people and the treatment of auditory hypersensitivity. This section will provide some considerations for further research.

Compare behavioural response to spatial audio within VR between ASD and TD children:

The investigation presented in Chapter 4 could be further extended to include NT participants, allowing for a direct comparison in behavioural responses towards spatial audio events and sound interpretation amongst competing background audio. Comparison data may highlight any potential differences in performance between participants with autism and NT counterparts. In regards to background stimuli, the presented study was limited to the specific group of sounds that represented the immediate visual environment. Future studies would benefit with the inclusion of multiple environments, thereby permitting a wider range of competing auditory stimuli with contrasting frequency content. Additional SNR levels could also be included to expand on the results gained from Chapter 4. Finally, the study lacked a form of hearing screening to detect physiological impairments that may affect auditory processing. By including this into future procedures, any experimental disparity between ASD and NT participants would most likely be the result of neurological differences between the two groups.

The effect of binaural based Ambisonics on presence within Virtual Reality experienced by autistic individuals:

The early chapters of this thesis discussed the concept of presence, demonstrating how a psychological immersion within a VR environment is crucial for the successful outcomes of VR exposure therapies. Research has shown that spatial audio rendering techniques that allow for accurate localisation of dynamic virtual sound sources also play a significant role in achieving presence, though these studies have only included NT participants. There have been investigations that measure self-reported presence within VE experienced by autistic individuals, however the effect of the auditory modality has not yet been explored. Although it has been observed that autistic and NT participants experience similar levels of presence and ecological validity within VR, any differences in auditory processing may effect this. Coupled with the previous recommendations for future research, this would lead to a greater understanding of how autistic individuals respond to virtual spatialised sound. Therefore creating a more stable foundation for developing subsequent VR based interventions, maximising engagement with virtual environments and potential therapeutic outcomes.

Longitudinal study which includes parental questionnaires and physiological measurements

Despite the results of Chapters 6 and 7 supporting the use of binaural based Ambisonics to help reduce auditory hypersensitivity, the following proposals would provide further robustness. Firstly, the inclusion of outcome measurements collected from parents and school staff would provide information regarding participant emotional response towards real-world auditory stimuli. This could be evaluated comparing pre- and post-intervention measurements using the Short Sensory Profile [105] and parental questionnaires such as the Parent Stress Index [484]. Furthermore, these measurements would provide evidence of any generalisation of tolerance following the experimental period. Secondly, physiological measurements would scaffold the primary and secondary measurements used in Chapters 6 and 7, bypassing any social and communication impairments. The measurement of skin conductance, blood volume pulse, electrodermal activity and skin temperature have all been used in the assessment of anxiety symptoms in autistic people [485], [486].

Comparison between clinician based in-vivo treatments for auditory hypersensitivity and spatial audio framework:

There are multiple studies which compare VR with in-vivo exposure therapy techniques, with results suggesting that VR simulations could be as effective as their real world counterparts. A study comparing clincian based in-vivo treatment with one driven by Ambisonics could present additional robust support in the use binaural based spatial audio in the treatment of auditory hypersensitivity in autism.

Explore the implementation of an augmented reality based game for auditory hypersensitivity:

The implementation of binaural based ambisonic rendering is not limited to virtual reality applications. One interesting avenue for future research could therefore integrate spatial audio simulations into exposure based games deployed to augmented reality (AR) technology. AR overlays three-dimensional computer generated objects over the real world via a see-through display [487]. Devices such as the Magic Leap [488] and Microsoft HoloLens [489] also permit user head and positional tracking with 6DoF via spatial computing [490]. This presents a distinct advantage over VR based systems as the user could experience the simulated stimuli within the environment they are commonly exposed to, thus further increasing the context of the intervention. Another key advantage of AR is that it is not restricted to a single user. This would allow family members or staff to participate with the intervention mechanics, providing support but also aiding in the communication and social development of the user [491]–[493]

8.4 Final Remarks

The core motivation behind this thesis was to investigate the use of binaural based Ambisonics in the reduction of auditory hypersensitivity experienced by autistic young people. The results from this research have demonstrated that exposure to Ambisonic rendered stimuli within a VR game can aid in the reduction of negative emotions associated with specific sounds. Furthermore, this reduction is significantly improved when listening to spatialised sound over stereo. This provides a strong foundation of support for integration of spatial audio into future interventions developed for this particular sensory processing issue.

This thesis has shown that spatialised sound is an important component in the delivery of plausible multi-modal environments within VR. Chapter 4 observed organic behavioural responses towards spatial audio stimuli within VR in addition to impaired localisation of virtual sound sources amongst competing background audio. It is therefore hoped that these outcomes can be applied to the wider area of VR interventions developed for ASD, informing approaches to reproducing virtual acoustic environments that maintain immersion whilst achieving the objective of the target sounds.

Virtual reality has been proven to be an important tool for delivering therapy to those with autism for over two decades, and advancements in consumer equipment are allowing the technology to become increasingly accessible. This allows therapy frameworks to be delivered via a fun and motivating format. Similar to the projects that inspired this research [29], [30], SoundFields achieved comparable results by motivating and rewarding players who interact with problematic auditory stimuli. Both experimental observations and those from staff show that participants enjoyed their time in the experiment and maintained interest throughout. Combined with the reduction in associated negative emotions, this is encouraging to see in terms of research but also a rewarding experience.

It is hoped that the research presented in this thesis will provide encouragement for inter-disciplinary research that designs interactive interventions for autistic young people. The development of consumer interactive technology presents an exciting time for the creation of realistic and engaging virtual environments. However, these can be utilised for purposes outside of entertainment. They can deliver enriching and engaging environments for education and therapy that provide meaningful benefits to autistic people.

Appendix A

Evaluating the effect of Ambisonic Audio in the Reduction of Auditory Hypersensitivity in young people with ASD - School statements

This appendix includes statements provided by staff from two of the recruitment schools, giving feedback regarding the experiment from the perspective of the teachers.

A.1 School A

Parental consent was given for a wide range of pupils with a diagnosis of ASC to take part in the project. The pupils ranged from non-verbal pupils with lower cognitive abilities working within a semi-formal curriculum pathway and with low attention to task, to pupils working within a higher formal curriculum pathway demonstrating much higher cognitive skills.

As a school, we set high expectations, but the response to the simulating games which were presented during the project exceeded all our expectations. The fact that some of our pupils tolerated the head- set was an unexpected response, and then to see the pupils following instructions and spending substantial amounts of time on the games was again a response which we didn't predict. It was a surprise for all the staff involved, to see all the pupils being highly motivated by this weekly therapy, and some pupils who we felt would become disinterested in the games as the weeks progressed, were highly motivated when they saw Daniel, who very quickly became an object of reference!

A.2 School B

Our pupils that participated in this study initially did so with some trepidation along with the excitement of being able to 'play' with technology not normally accessible to them. From a staff point of view, I was fascinated to see how quickly the children adapted to the game - wearing the head and hands equipment and being thrust into a surreal landscape that was very alien to them. The speed in which they were able to co-ordinate the controls and begin to build up achievements within the game was astounding! Every Monday, all involved would come into school so excited knowing that it was VR day! As the weeks went on we witnessed their confidence in the game, and their ability to play it, build and build. We saw their reactions to some sounds, that in the first instance were too uncomfortable to listen to, change into the pupil having tolerance and sometimes even barely noticing they were occurring! They were too focused on the task in hand and having too much fun! The opportunity to participate in this study has been gratefully received and was carried out in the most respectful and professional manner. We look forward to hearing the outcome.'

Appendix B

Supporting Data Index

The appendix provides content information and organisation of the attached data file. The data is divided into folders based upon the relevant chapters presented in this thesis.

B.1 Chapter 4 Data and VR environment

Contains VR experimental environments and operating instructions. In addition, experiment informed consent forms and Video captures of VR environment are also included.

B.2 Chapter 6 Data and VR environment

Contains experiment informed consent forms.

B.3 Chapter 7 Data and VR environment

Contains experiment informed consent forms and Video captures of VR environment.

B.4 Target Auditory Stimuli

Contains document with links to Ambisonic target stimuli files used for both experiments in Chapters 6 & 7.

B.5 SoundFields VR

Contains VR builds of both audio conditions for experimental and control periods.

B.6 Audio Questionnaire

Contains unity build of audio questionnaire software used to measure emotional responses to target stimuli.

Glossary

Ambisonics A technique for recording and rendering three-dimensional audio by decomposing a 360 degree sound-field.

Anechoic Free from echo and reverberation energyhttps://www.overleaf.com/project/5e18648465cc290

Avatar The player representation of themselves within a computer game's virtual environment.

dBHL A measurement of sound intensity used by clinicians, often in Audiogram tests.

Degrees of Freedom A 3D space has six degrees in total. Three represent movement along the rotational axis (yaw, pitch and roll). The other three represent translational movement (elevate, straff, and surge).

Free Field Environment An environment in which no sound reflection can occur.

In-Virtuo Exposure Directly facing a virtual representation of a feared object, situation or activity within virtual reality.

In-Vivo exposure Directly facing a feared object, situation or activity in real life.

Loudness Unit Full Scale A precise measurement of loudness to full scale.

Mono Audio A technique for rendering audio over one channel.

Non-Player Character A character within a computer game that is not controlled by the player but controlled by the software itself.

Serious Game A computer game developed for purposes outside the field of entertainment including education, training and therapy.

Signal-to-Noise ratio The ratio between the target audio signal and the level of background noise.

Stereo Audio A technique for rendering audio over two channels.

List of Acronyms

Notation	**Description**
ABM	Attention Bias Modification
AIT	Auditory Integration Training
ASA	Auditory Scene Analysis
ASD	Autism Spectrum Disorder
BRIR	Binaural room impulse response
CAI	Computer Assisted Interventions
CBT	Cognitive Behavioural Therapy
CS	Compound Symmetry
CVE	Collaborative Virtual Environment
FOA	First Order Ambisonics
GAS	Goal Attainment Scaling
GUI	Graphical User Interface
HLM	Hierarchical Linear Model
HMD	Head Mounted Display
HRTF	Head Related Transfer Function

Notation	**Description**
IFL	Identity First Language
ILD	Interaural Level Difference
ITD	Interaural Time Difference
NT	Neurotypical
PD	Participatory Design
PFL	Person First Language
PTSD	Post-Traumatic Stress Disorder
SADIE	Spatial Audio for Domestic Interactive Entertainment
SD	Standard Deviation
SDK	Software Development Kit
SIT	Sensory Integration Therapy
SNR	Signal-to-noise ratio
SRER	Self-Reported Emotional Response
TUI	Tangible User Interface
VE	Virtual Environment
VR	Virtual Reality
VRET	Virtual Reality Exposure Therapy

References

[1] A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[2] T. Watson, *Music therapy with adults with learning disabilities*. Routledge, 2007.

[3] E. J. Marco, L. B. N. Hinkley, S. S. Hill, and S. S. Nagarajan, "Sensory processing in autism: A review of neurophysiologic findings," *Pediatric research*, vol. 69, no. 5 Pt 2, 48R, 2011.

[4] B. A. Pfeiffer, K. Koenig, M. Kinnealey, M. Sheppard, and L. Henderson, "Effectiveness of sensory integration interventions in children with autism spectrum disorders: A pilot study," *American Journal of Occupational Therapy*, vol. 65, no. 1, pp. 76–85, 2011.

[5] B. Rimland and S. M. Edelson, "Brief report: A pilot study of auditory integration training in autism," *Journal of autism and developmental disorders*, vol. 25, no. 1, pp. 61–70, 1995.

[6] L. N. Stiegler and R. Davis, "Understanding sound sensitivity in individuals with autism spectrum disorders," *Focus on Autism and Other Developmental Disabilities*, vol. 25, no. 2, pp. 67–75, 2010.

[7] R. L. Koegel, D. Openden, and L. K. Koegel, "A systematic desensitization paradigm to treat hypersensitivity to auditory stimuli in children with autism in family contexts," *Research and Practice for Persons with Severe Disabilities*, vol. 29, no. 2, pp. 122–134, 2004.

[8] O. T. Leyfer, S. E. Folstein, S. Bacalman, N. O. Davis, E. Dinh, J. Morgan, H. Tager-Flusberg, and J. E. Lainhart, “Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders,” *Journal of autism and developmental disorders*, vol. 36, no. 7, pp. 849–861, 2006.

[9] F. J. Van Steensel, S. M. Bögels, and S. Perrin, “Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis,” *Clinical child and family psychology review*, vol. 14, no. 3, p. 302, 2011.

[10] T. H. Ollendick, N. J. King, and P. Muris, “Fears and phobias in children: Phenomenology, epidemiology, and aetiology,” *Child and Adolescent Mental Health*, vol. 7, no. 3, pp. 98–106, 2002.

[11] A. Narzisi, C. Costanza, B. Umberto, and M. Filippo, “Non-pharmacological treatments in autism spectrum disorders: An overview on early interventions for pre-schoolers,” *Current clinical pharmacology*, vol. 9, no. 1, pp. 17–26, 2014.

[12] S. J. Webb, E. J. Jones, J. Kelly, and G. Dawson, “The motivation for very early intervention for infants at high risk for autism spectrum disorders,” *International journal of speech-language pathology*, vol. 16, no. 1, pp. 36–42, 2014.

[13] K. Sullivan, W. L. Stone, and G. Dawson, “Potential neural mechanisms underlying the effectiveness of early intervention for children with autism spectrum disorder,” *Research in developmental disabilities*, vol. 35, no. 11, pp. 2921–2932, 2014.

[14] S. J. Rogers, “Brief report: Early intervention in autism,” *Journal of autism and developmental disorders*, vol. 26, no. 2, pp. 243–246, 1996.

[15] P. J. Santosh and J. Singh, “Drug treatment of autism spectrum disorder and its comorbidities in children and adolescents,” *BJPsych Advances*, vol. 22, no. 3, pp. 151–161, 2016.

[16] N. Sathe, J. C. Andrews, M. L. McPheeters, and Z. E. Warren, “Nutritional and dietary interventions for autism spectrum disorder: A systematic review,” *Pediatrics*, e20170346, 2017.

[17] K. Chenausky, A. Norton, H. Tager-Flusberg, and G. Schlaug, “Auditory-motor mapping training: Comparing the effects of a novel speech treatment to a control treatment for minimally verbal children with autism,” *PloS one*, vol. 11, no. 11, e0164930, 2016.

[18] G. Bumin, M. Huri, S. Salar, and H. Kayihan, "Occupational therapy in autism," in *Autism Spectrum Disorder-Recent Advances*, InTech, 2015.

[19] J. Case-Smith, L. L. Weaver, and M. A. Fristad, "A systematic review of sensory processing interventions for children with autism spectrum disorders," *Autism*, vol. 19, no. 2, pp. 133–148, 2015.

[20] K. Durkin, "Videogames and young people with developmental disorders.," *Review of General Psychology*, vol. 14, no. 2, p. 122, 2010.

[21] A. McCoy, J. Holloway, O. Healy, M. Rispoli, and L. Neely, "A systematic review and evaluation of video modeling, role-play and computer-based instruction as social skills interventions for children and adolescents with high-functioning autism," *Review Journal of Autism and Developmental Disorders*, vol. 3, no. 1, pp. 48–67, 2016.

[22] F. Van Steensel and S. Bögels, "Cognitive-behavioral therapy for anxiety disorders in children with and without asd," *J Consult Clin Psychol*, vol. 83, no. 3, pp. 512–23, 2015.

[23] N. Miller, J. Wyatt, L. B. Casey, and J. B. Smith, "Using computer-assisted instruction to increase the eye gaze of children with autism," *Behavioral Interventions*, vol. 33, no. 1, pp. 3–12, 2018.

[24] R. Rosen, P. L. Weiss, M. Zancanaro, and E. Gal, "Usability of a video modeling computer application for the vocational training of adolescents with autism spectrum disorder," *British Journal of Occupational Therapy*, vol. 80, no. 4, pp. 208–215, 2017.

[25] D. K. Murray, "Autism and information technology: Therapy with computers," in *Autism and Learning (Classic Edition)*, Routledge, 2011, pp. 98–113.

[26] D. Moore, P. McGrath, and J. Thorpe, "Computer-aided learning for people with autism–a framework for research and development," *Innovations in Education and Training International*, vol. 37, no. 3, pp. 218–228, 2000.

[27] H. A. M. Noor, F. Shahbodin, and N. C. Pee, "Serious game for autism children: Review of literature," *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, vol. 6, no. 4, pp. 554–559, 2012.

[28] R. Fino, M. J. Lin, A. Caballero, and F. F. Balahadia, “Disaster awareness simulation for children with autism spectrum disorder using android virtual reality,” *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 9, no. 2-6, pp. 59–62, 2017.

[29] H. M. Zakari, M. Poyade, and D. Simmons, “Sinbad and the magic cure: A serious game for children with asd and auditory hypersensitivity,” in *International Conference on Games and Learning Alliance*, Springer, 2017, pp. 54–63.

[30] R. Morris *et al.*, “Managing sound sensitivity in autism spectrum disorder: New technologies for customized intervention,” PhD thesis, Massachusetts Institute of Technology, 2009.

[31] D. Freeman, P. Haselton, J. Freeman, B. Spanlang, S. Kishore, E. Albery, M. Denne, P. Brown, M. Slater, and A. Nickless, “Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial,” *The Lancet Psychiatry*, 2018.

[32] B. M. Choo, A. M. Talicol, E. J. Panganiban, and E. C. J. Ong, “Use of tethered virtual reality headsets in exposure therapy for agoraphobia and social phobia.,” in *KHD@ IJCAI*, 2017, pp. 20–24.

[33] D. C. Beidel, B. C. Frueh, S. M. Neer, C. A. Bowers, B. Trachik, T. W. Uhde, and A. Grubaugh, “Trauma management therapy with virtual-reality augmented exposure therapy for combat-related ptsd: A randomized controlled trial,” *Journal of anxiety disorders*, 2017.

[34] M. Maskey, J. Lowry, J. Rodgers, H. McConachie, and J. R. Parr, “Reducing specific phobia/fear in young people with autism spectrum disorders (asds) through a virtual reality environment intervention,” *PloS one*, vol. 9, no. 7, e100374, 2014.

[35] N. C. Burbules, “Rethinking the virtual,” in *The international handbook of virtual learning environments*, Springer, 2006, pp. 37–58.

[36] F. Rumsey, “Spatial audio,” in. Focal Press, 2013, pp. 5–26.

[37] F. Ortolani, “Introduction to ambisonics,” *Ironbridge Electronics*, 2015.

[38] G. Vivanti, “Ask the editor: What is the most appropriate way to talk about individuals with a diagnosis of autism?” *Journal of autism and developmental disorders*, vol. 50, no. 2, pp. 691–693, 2020.

[39] L. Kenny, C. Hattersley, B. Molins, C. Buckley, C. Povey, and E. Pellicano, “Which terms should be used to describe autism? perspectives from the uk autism community,” *Autism*, vol. 20, no. 4, pp. 442–462, 2016.

[40] R. Cooper, K. Cooper, A. J. Russell, and L. G. Smith, ““i’m proud to be a little bit different”: The effects of autistic individuals’ perceptions of autism and autism social identity on their collective self-esteem,” *Journal of Autism and Developmental Disorders*, vol. 51, no. 2, pp. 704–714, 2021.

[41] M. Anderson-Chavarria, “The autism predicament: Models of autism and their impact on autistic identity,” *Disability & Society*, pp. 1–21, 2021.

[42] A. F. Crocker and S. N. Smith, “Person-first language: Are we practicing what we preach?” *Journal of multidisciplinary healthcare*, vol. 12, p. 125, 2019.

[43] D. Marks, “Models of disability,” *Disability and rehabilitation*, vol. 19, no. 3, pp. 85–91, 1997.

[44] J. Piven, J. T. Elison, and M. J. Zylka, “Toward a conceptual framework for early brain and behavior development in autism,” *Molecular psychiatry*, vol. 22, no. 10, pp. 1385–1394, 2017.

[45] M. Dapretto, M. S. Davies, J. H. Pfeifer, A. A. Scott, M. Sigman, S. Y. Bookheimer, and M. Iacoboni, “Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders,” *Nature neuroscience*, vol. 9, no. 1, pp. 28–30, 2006.

[46] D. S. Dunn and E. E. Andrews, “Person-first and identity-first language: Developing psychologists’ cultural competence using disability language.,” *American Psychologist*, vol. 70, no. 3, p. 255, 2015.

[47] M. A. Gernsbacher, *Editorial perspective: The use of person-first language in scholarly writing may accentuate stigma*, 2017.

[48] J. Sinclair, “Why i dislike “person first” language,” *Autonomy, the Critical Journal of Interdisciplinary Autism Studies*, vol. 1, no. 2, 2013.

[49] P. M. Levisohn, “The autism-epilepsy connection,” *Epilepsia*, vol. 48, pp. 33–35, 2007.

[50] R. Tuchman and I. Rapin, “Epilepsy in autism,” *The Lancet Neurology*, vol. 1, no. 6, pp. 352–358, 2002.

[51] T. Armstrong, *The Power of Neurodiversity: Unleashing the Advantages of Your Differently Wired Brain (published in Hardcover as Neurodiversity)*. Da Capo Lifelong Books, 2011.

[52] L. Mottron, M. Dawson, I. Soulieres, B. Hubert, and J. Burack, "Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception," *Journal of autism and developmental disorders*, vol. 36, no. 1, pp. 27–43, 2006.

[53] N. Bagatell, "Orchestrating voices: Autism, identity and the power of discourse," *Disability & Society*, vol. 22, no. 4, pp. 413–426, 2007.

[54] L. Arnold, "A brief history of "neurodiversity" as a concept and perhaps a movement," *Autonomy, the Critical Journal of Interdisciplinary Autism Studies*, vol. 1, no. 5, 2017.

[55] R. Chapman, "Neurodiversity theory and its discontents: Autism, schizophrenia, and the social model of disability," *The Bloomsbury Companion to Philosophy of Psychiatry, edited by Şerife Tekin and Robyn Bluhm*, pp. 371–90, 2019.

[56] S. M. Bury, R. Jellett, J. R. Spoor, and D. Hedley, ""it defines who i am" or "it's something i have": What language do [autistic] australian adults [on the autism spectrum] prefer?" *Journal of autism and developmental disorders*, pp. 1–11, 2020.

[57] S. M. Robertson and A. D. Ne'eman, "Autistic acceptance, the college campus, and technology: Growth of neurodiversity in society and academia," *Disability Studies Quarterly*, vol. 28, no. 4, 2008.

[58] B. Sanchez, "Autism spectrum disorder and bilingualism: An integration of neurological, language, and relational dynamics in increasingly multicultural and globalized systems," PhD thesis, Chicago School of Professional Psychology, 2007.

[59] D. Baker, "Particular polyglots: Multilingual students with autism," *Disability & Society*, vol. 28, no. 4, pp. 527–541, 2013.

[60] S. K. Kapp, K. Gillespie-Lynch, L. E. Sherman, and T. Hutman, "Deficit, difference, or both? autism and neurodiversity.," *Developmental psychology*, vol. 49, no. 1, p. 59, 2013.

[61] H. Thomas and T. Boellstorff, "Beyond the spectrum: Rethinking autism," *Disability Studies Quarterly*, vol. 37, no. 1, 2017.

[62] M. Botha, J. Hanlon, and G. L. Williams, “Does language matter? identity-first versus person-first language use in autism research: A response to vivanti,” *Journal of Autism and Developmental Disorders*, pp. 1–9, 2021.

[63] T. Shakespeare, “Disability: Suffering, social oppression, or complex predicament?” In *The contingent nature of life*, Springer, 2008, pp. 235–246.

[64] J. Park, R. Fitzgerald, and M. Legge, “The predicament of d/deaf: Towards an anthropology of not-disability,” *Human Organization*, pp. 154–163, 2015.

[65] J. Steigner and S. U. Ruhlin, “Systematic desensitization of hyperacusis and vocal pitch disorder treatment in a patient with autism,” *Internet Journal of Allied Health Sciences and Practice*, vol. 12, no. 1, p. 6, 2014.

[66] M. Balkaya and G. Catak, “Why games are fun? the reward system in the human brain,” in *Gamer Psychology and Behavior*, Springer, 2016, pp. 21–38.

[67] B. James, B. Fletcher, and N. Wearn, “Three corners of reward in computer games,” 2013.

[68] M. J. Koepp, R. N. Gunn, A. D. Lawrence, V. J. Cunningham, A. Dagher, T. Jones, D. J. Brooks, C. J. Bench, and P. Grasby, “Evidence for striatal dopamine release during a video game,” *Nature*, vol. 393, no. 6682, pp. 266–268, 1998.

[69] A. K. Przybylski, C. S. Rigby, and R. M. Ryan, “A motivational model of video game engagement,” *Review of general psychology*, vol. 14, no. 2, pp. 154–166, 2010.

[70] K. Porayska-Pomsta, C. Frauenberger, H. Pain, G. Rajendran, T. Smith, R. Menzies, M. E. Foster, A. Alcorn, S. Wass, S. Bernadini, *et al.*, “Developing technology for autism: An interdisciplinary approach,” *Personal and Ubiquitous Computing*, vol. 16, no. 2, pp. 117–127, 2012.

[71] H. J. Jackson and N. J. King, “The therapeutic management of an autistic child’s phobia using laughter as the anxiety inhibitor,” *Behavioural and Cognitive Psychotherapy*, vol. 10, no. 4, pp. 364–369, 1982.

[72] K. M. Lee, “Presence, explicated,” *Communication theory*, vol. 14, no. 1, pp. 27–50, 2004.

[73] S. Bernardini, K. Porayska-Pomsta, and T. J. Smith, “Echoes: An intelligent serious game for fostering social communication in children with autism,” *Information Sciences*, vol. 264, pp. 41–60, 2014.

[74] H. M. Zakari, M. Ma, and D. Simmons, “A review of serious games for children with autism spectrum disorders (asd),” in *International Conference on Serious Games Development and Applications*, Springer, 2014, pp. 93–106.

[75] B. H. Sørensen and B. Meyer, “Serious games in language learning and teaching-a theoretical perspective.,” in *DiGRA Conference*, 2007.

[76] P. M. Kato, S. W. Cole, A. S. Bradlyn, and B. H. Pollock, “A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial,” *Pediatrics*, vol. 122, no. 2, e305–e317, 2008.

[77] M. A. Gerzon, “The design of precisely coincident microphone arrays for stereo and surround sound,” in *Audio Engineering Society Convention 50*, Audio Engineering Society, 1975.

[78] L. Thresh, C. Armstrong, and G. Kearney, “A direct comparison of localization performance when using first, third, and fifth ambisonics order for real loudspeaker and virtual loudspeaker rendering,” in *Audio Engineering Society Convention 143*, Audio Engineering Society, 2017.

[79] T. McKenzie, D. Murphy, and G. Kearney, “Assessing the authenticity of the kemar mouth simulator as a repeatable speech source. (preprint),” *Audio Engineering Society Convention Paper, October 2017 18-21. 143rd Convention*, 0ADAD.

[80] D. Strickland, L. M. Marcus, G. B. Mesibov, and K. Hogan, “Brief report: Two case studies using virtual reality as a learning tool for autistic children,” *Journal of Autism and Developmental Disorders*, vol. 26, no. 6, pp. 651–659, 1996.

[81] T. Ouimet, N. E. Foster, A. Tryfon, and K. L. Hyde, “Auditory-musical processing in autism spectrum disorders: A review of behavioral and brain imaging studies,” *Annals of the New York Academy of Sciences*, vol. 1252, no. 1, pp. 325–331, 2012.

[82] E. Applebaum, A. L. Egel, R. L. Koegel, and B. Imhoff, “Measuring musical abilities of autistic children,” *Journal of autism and developmental disorders*, vol. 9, no. 3, pp. 279–285, 1979.

[83] J. I. Alcántara, E. J. Weisblatt, B. C. Moore, and P. F. Bolton, “Speech-in-noise perception in high-functioning individuals with autism or asperger’s syndrome,” *Journal of Child Psychology and Psychiatry*, vol. 45, no. 6, pp. 1107–1114, 2004.

[84] D. A. Treffert and G. L. Wallace, “Islands of genius,” *Scientific American*, vol. 286, no. 6, pp. 76–85, 2002.

[85] B Hermelin and N O’connor, “Idiot savant calendrical calculators: Rules and regularities,” *Psychological medicine*, vol. 16, no. 4, pp. 885–893, 1986.

[86] L. Mottron and S. Belleville, “A study of perceptual analysis in a high-level autistic subject with exceptional graphic abilities,” *Brain and cognition*, vol. 23, no. 2, pp. 279–309, 1993.

[87] B. Hermelin and N. O’Connor, “Art and accuracy: The drawing ability of idiot-savants,” *Journal of Child Psychology and Psychiatry*, vol. 31, no. 2, pp. 217–228, 1990.

[88] B. Hermelin, N O’connor, and S Lee, “Musical inventiveness of five idiots-savants,” *Psychological medicine*, vol. 17, no. 3, pp. 685–694, 1987.

[89] R. L. Young and T Nettelbeck, “The abilities of a musical savant and his family,” *Journal of autism and developmental disorders*, vol. 25, no. 3, pp. 231–248, 1995.

[90] A. L. Hill, “Idiot savants: Rate of incidence,” *Perceptual and Motor Skills*, vol. 44, no. 1, pp. 161–162, 1977.

[91] P. Heaton, B. Hermelin, and L. Pring, “Autism and pitch processing: A precursor for savant musical ability?” *Music Perception: An Interdisciplinary Journal*, vol. 15, no. 3, pp. 291–305, 1998.

[92] P. Heaton, K. Hudry, A. Ludlow, and E. Hill, “Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders,” *Cognitive neuropsychology*, vol. 25, no. 6, pp. 771–782, 2008.

[93] A. Järvinen-Pasley, J. Pasley, and P. Heaton, “Is the linguistic content of speech less salient than its perceptual features in autism?” *Journal of autism and developmental disorders*, vol. 38, no. 2, pp. 239–248, 2008.

[94] C. R. Jones, F. Happé, G. Baird, E. Simonoff, A. J. Marsden, J. Tregay, R. J. Phillips, U. Goswami, J. M. Thomson, and T. Charman, “Auditory discrimination and auditory sensory behaviours in autism spectrum disorders,” *Neuropsychologia*, vol. 47, no. 13, pp. 2850–2858, 2009.

[95] P. Heaton, K. Williams, O. Cummins, and F. Happé, “Autism and pitch processing splinter skills: A group and subgroup analysis,” *Autism*, vol. 12, no. 2, pp. 203–219, 2008.

[96] P. K. Kuhl, S. Coffey-Corina, D. Padden, and G. Dawson, “Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures,” *Developmental science*, vol. 8, no. 1, 2005.

[97] A. Bonnel, S. McAdams, B. Smith, C. Berthiaume, A. Bertone, V. Ciocca, J. A. Burack, and L. Mottron, “Enhanced pure-tone pitch discrimination among persons with autism but not asperger syndrome,” *Neuropsychologia*, vol. 48, no. 9, pp. 2465–2475, 2010.

[98] T. Lepistö, T. Kujala, R. Vanhala, P. Alku, M. Huotilainen, and R. Näätänen, “The discrimination of and orienting to speech and non-speech sounds in children with autism,” *Brain research*, vol. 1066, no. 1, pp. 147–157, 2005.

[99] E. J. Flagg, J. E. O. Cardy, W. Roberts, and T. P. Roberts, “Language lateralization development in children with autism: Insights from the late field magnetoencephalogram,” *Neuroscience letters*, vol. 386, no. 2, pp. 82–87, 2005.

[100] B. Boets, J. Verhoeven, J. Wouters, and J. Steyaert, “Fragile spectral and temporal auditory processing in adolescents with autism spectrum disorder and early language delay,” *Journal of autism and developmental disorders*, vol. 45, no. 6, pp. 1845–1857, 2015.

[101] G. J. Brown and D. Wang, “Separation of speech by computational auditory scene analysis,” in *Speech enhancement*, Springer, 2005, pp. 371–402.

[102] A. S. Bregman, *Auditory scene analysis: The perceptual organization of sound.* MIT press, 1994.

[103] I.-F. Lin, A. Shirama, N. Kato, and M. Kashino, “The singular nature of auditory and visual scene analysis in autism,” *Phil. Trans. R. Soc. B*, vol. 372, no. 1714, p. 20 160 115, 2017.

[104] G. Dawson, K. Toth, R. Abbott, J. Osterling, J. Munson, A. Estes, and J. Liaw, “Early social attention impairments in autism: Social orienting, joint attention, and attention to distress.,” *Developmental psychology*, vol. 40, no. 2, p. 271, 2004.

[105] S. D. Tomchek and W. Dunn, “Sensory processing in children with and without autism: A comparative study using the short sensory profile,” *American Journal of occupational therapy*, vol. 61, no. 2, pp. 190–200, 2007.

[106] R. W. Peters, B. C. Moore, and T. Baer, “Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people,” *The Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 577–587, 1998.

[107] W. B. Groen, L. van Orsouw, N. Ter Huurne, S. Swinkels, R.-J. van der Gaag, J. K. Buitelaar, and M. P. Zwiers, “Intact spectral but abnormal temporal processing of auditory stimuli in autism,” *Journal of autism and developmental disorders*, vol. 39, no. 5, pp. 742–750, 2009.

[108] S. Schelinski and K. von Kriegstein, “Brief report: Speech-in-noise recognition and the relation to vocal pitch perception in adults with autism spectrum disorder and typical development,” *Journal of Autism and Developmental Disorders*, vol. 50, no. 1, pp. 356–363, 2020.

[109] W. A. Teder-Sälejärvi, K. L. Pierce, E. Courchesne, and S. A. Hillyard, “Auditory spatial localization and attention deficits in autistic adults,” in, 2-3, vol. 23, Elsevier, 2005, pp. 221–234.

[110] R. J. Kulesza and K. Mangunay, “Morphological features of the medial superior olive in autism,” in, vol. 1200, Elsevier, 2008, pp. 132–137.

[111] R. J. Kulesza Jr, R. Lukose, and L. V. Stevens, “Malformation of the human superior olive in autistic spectrum disorders,” in, vol. 1367, Elsevier, 2011, pp. 360–371.

[112] E. Visser, M. P. Zwiers, C. C. Kan, L. Hoekstra, A. J. van Opstal, and J. K. Buitelaar, “Atypical vertical sound localization and sound-onset sensitivity in people with autism spectrum disorders,” in, 6, vol. 38, Canadian Medical Association, 2013, p. 398.

[113] P. Rochat and T. Striano, “Social-cognitive development in the first year,” in, 1999, pp. 3–34.

[114] J. Morton and M. H. Johnson, “Conspec and conlern: A two-process theory of infant face recognition.,” in, 2, vol. 98, American Psychological Association, 1991, p. 164.

[115] N. Boddaert, P. Belin, N. Chabane, J.-B. Poline, C. Barthélémy, M.-C. Mouren-Simeoni, F. Brunelle, Y. Samson, and M. Zilbovicius, "Perception of complex sounds: Abnormal pattern of cortical activation in autism," in, 11, vol. 160, Am Psychiatric Assoc, 2003, pp. 2057–2060.

[116] G. Dawson, A. N. Meltzoff, J. Osterling, J. Rinaldi, and E. Brown, "Children with autism fail to orient to naturally occurring social stimuli," in, 6, vol. 28, Springer, 1998, pp. 479–485.

[117] L. N. Soskey, P. D. Allen, and L. Bennetto, "Auditory spatial attention to speech and complex non-speech sounds in children with autism spectrum disorder," in, 8, vol. 10, Wiley Online Library, 2017, pp. 1405–1416.

[118] L. Kanner *et al.*, "Autistic disturbances of affective contact," *Nervous child*, vol. 2, no. 3, pp. 217–250, 1943.

[119] U. Frith and M. Mira, "Autism and asperger syndrome," *Focus on Autistic Behavior*, vol. 7, no. 3, pp. 13–15, 1992.

[120] S. Bettison, "The long-term effects of auditory training on children with autism," *Journal of autism and developmental disorders*, vol. 26, no. 3, pp. 361–374, 1996.

[121] S. M. Edelson, D. Arin, M. Bauman, S. E. Lukas, J. H. Rudy, M. Sholar, and B. Rimland, "Auditory integration training: A double-blind study of behavioral and electrophysiological effects in people with autism," *Focus on Autism and Other Developmental Disabilities*, vol. 14, no. 2, pp. 73–81, 1999.

[122] S. J. Rogers, S. Hepburn, and E. Wehner, "Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders," *Journal of autism and developmental disorders*, vol. 33, no. 6, pp. 631–642, 2003.

[123] G. T. Baranek, B. A. Boyd, M. D. Poe, F. J. David, and L. R. Watson, "Hyperresponsive sensory patterns in young children with autism, developmental delay, and typical development," *American Journal on Mental Retardation*, vol. 112, no. 4, pp. 233–245, 2007.

[124] A. M. Tharpe, F. H. Bess, D. P. Sladen, H. Schissel, S. Couch, and T. Schery, "Auditory characteristics of children with autism," *Ear and hearing*, vol. 27, no. 4, pp. 430–441, 2006.

[125] J. S. Gravel, M. Dunn, W. W. Lee, and M. A. Ellis, "Peripheral audition of children on the autistic spectrum," *Ear and Hearing*, vol. 27, no. 3, pp. 299–312, 2006.

[126] E. Gomes, N. T. Rotta, F. S. Pedroso, P. Sleifer, and M. C. Danesi, "Auditory hypersensitivity in children and teenagers with autistic spectrum disorder," *Arquivos de Neuro-psiquiatria*, vol. 62, no. 3B, pp. 797–801, 2004.

[127] J. R. Lucker, "Auditory hypersensitivity in children with autism spectrum disorders," *Focus on Autism and Other Developmental Disabilities*, vol. 28, no. 3, pp. 184–191, 2013.

[128] J. R. LUCKER and A Doman, "Auditory hypersensitivity and autism spectrum disorders: An emotional response," *Autism Science Digest: The Journal of Autismone*, no. 04, pp. 103–108, 2012.

[129] P. J. Jastreboff, M. M. Jastreboff, *et al.*, "Tinnitus retraining therapy (trt) as a method for treatment of tinnitus and hyperacusis patients," *Journal of the American Academy of Audiology*, vol. 11, no. 3, pp. 162–177, 2000.

[130] M. M. Jastreboff and P. J. Jastreboff, "Components of decreased sound tolerance: Hyperacusis, misophonia, phonophobia," *ITHS News Lett*, vol. 2, no. 5-7, 2001.

[131] C. Herráiz and I. Diges, "Tinnitus and hyperacusis/phonophobia," in *Textbook of tinnitus*, Springer, 2011, pp. 455–461.

[132] Z. A. Asha'ari, N. M. Zain, and A. Razali, "Phonophobia and hyperacusis: Practical points from a case report," *The Malaysian journal of medical sciences: MJMS*, vol. 17, no. 1, p. 49, 2010.

[133] A. R. Møller, J. K. Kern, and B. Grannemann, "Are the non-classical auditory pathways involved in autism and pdd?" *Neurological Research*, vol. 27, no. 6, pp. 625–629, 2005.

[134] J. R. Lucker and A. Doman, "Neural mechanisms involved in hypersensitive hearing: Helping children with asd who are overly sensitive to sounds," *Autism research and treatment*, vol. 2015, 2015.

[135] F. Musiek, A. Mohanani, E. Wierzbinski, G. Kilgore, J. Hunter, and J. Marotto, "The non-classical pathway: Too great to be ignored," *The Hearing Journal*, vol. 64, no. 10, pp. 6–8, 2011.

[136] A. R. Moller and P. R. Rollins, “The non-classical auditory pathways are involved in hearing in children but not in adults,” *Neuroscience Letters*, vol. 319, no. 1, pp. 41–44, 2002.

[137] M. D. Shen, D. D. Li, C. L. Keown, A. Lee, R. T. Johnson, K. Angkustsiri, S. J. Rogers, R.-A. Müller, D. G. Amaral, and C. W. Nordahl, “Functional connectivity of the amygdala is disrupted in preschool-aged children with autism spectrum disorder,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 55, no. 9, pp. 817–824, 2016.

[138] A. Etkin and T. D. Wager, “Functional neuroimaging of anxiety: A meta-analysis of emotional processing in ptsd, social anxiety disorder, and specific phobia,” *American Journal of Psychiatry*, vol. 164, no. 10, pp. 1476–1488, 2007.

[139] E. Courchesne and K. Pierce, “Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection,” *Current opinion in neurobiology*, vol. 15, no. 2, pp. 225–230, 2005.

[140] R. L. Moseley and F. Pulvermüller, “What can autism teach us about the role of sensorimotor systems in higher cognition? new clues from studies on language, action semantics, and abstract emotional concept processing,” *Cortex*, 2017.

[141] R. Stoner, M. L. Chow, M. P. Boyle, S. M. Sunkin, P. R. Mouton, S. Roy, A. Wynshaw-Boris, S. A. Colamarino, E. S. Lein, and E. Courchesne, “Patches of disorganization in the neocortex of children with autism,” *New England Journal of Medicine*, vol. 370, no. 13, pp. 1209–1219, 2014.

[142] J. Bachevalier and K. A. Loveland, “The orbitofrontal–amygdala circuit and self-regulation of social–emotional behavior in autism,” *Neuroscience & Biobehavioral Reviews*, vol. 30, no. 1, pp. 97–117, 2006.

[143] A. C. Laurent and E. Rubin, “Challenges in emotional regulation in asperger syndrome and high-functioning autism,” *Topics in Language Disorders*, vol. 24, no. 4, pp. 286–297, 2004.

[144] A. C. Samson, O. Huber, and J. J. Gross, “Emotion regulation in asperger’s syndrome and high-functioning autism.,” *Emotion*, vol. 12, no. 4, p. 659, 2012.

[145] C. A. Mazefsky, J. Herrington, M. Siegel, A. Scarpa, B. B. Maddox, L. Scahill, and S. W. White, "The role of emotion regulation in autism spectrum disorder," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 52, no. 7, pp. 679–688, 2013.

[146] C. A. Mazefsky, K. A. Pelphrey, and R. E. Dahl, "The need for a broader approach to emotion regulation research in autism," *Child development perspectives*, vol. 6, no. 1, pp. 92–97, 2012.

[147] L. Berkovits, A. Eisenhower, and J. Blacher, "Emotion regulation in young children with autism spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 47, no. 1, pp. 68–79, 2017.

[148] M. Edelstein, D. Brang, R. Rouw, and V. S. Ramachandran, "Misophonia: Physiological investigations and case descriptions," *Frontiers in Human Neuroscience*, vol. 7, p. 296, 2013.

[149] T. Attwood *et al.*, *The complete guide to asperger's syndrome*, 2007.

[150] N. Ikuta, R. Iwanaga, A. Tokunaga, H. Nakane, K. Tanaka, and G. Tanaka, "Effectiveness of earmuffs and noise-cancelling headphones for coping with hyper-reactivity to auditory stimuli in children with autism spectrum disorder: A preliminary study," *Hong Kong Journal of Occupational Therapy*, vol. 28, pp. 24–32, 2016.

[151] T. J. Kiresuk and R. E. Sherman, "Goal attainment scaling: A general method for evaluating comprehensive community mental health programs," *Community mental health journal*, vol. 4, no. 6, pp. 443–453, 1968.

[152] L. Turner-Stokes, "Goal attainment scaling (gas) in rehabilitation: A practical guide," *Clinical rehabilitation*, vol. 23, no. 4, pp. 362–370, 2009.

[153] B. Pfeiffer, L. Stein, C. Shui, and A. Murphy, "Effectiveness of noise-attenuating headphones on physiological responses for children with autism spectrum disorders," *Frontiers in Integrative Neuroscience*, vol. 13, p. 65, 2019.

[154] G. Berard, *Hearing equals behavior*. Keats Publ., 1993.

[155] H. M. Link, "Auditory integration training (ait): Sound therapy? case studies of three boys with autism who received ait," *British Journal of Learning Disabilities*, vol. 25, no. 3, pp. 106–110, 1997.

[156] A. Stehli, *The sound of a miracle*. Avon Books, 1992.

[157] O. C. Mudford, B. A. Cross, S. Breen, C. Cullen, D. Reeves, J. Gould, and J. Douglas, “Auditory integration training for children with autism: No behavioral benefits detected,” *American Journal on Mental Retardation*, vol. 105, no. 2, pp. 118–129, 2000.

[158] W. Zollweg, D. Palm, and V. Vance, “The efficacy of auditory integration training: A double blind study,” *American Journal of Audiology*, vol. 6, no. 3, pp. 39–47, 1997.

[159] W. T. ODonohue, “Cognitive behavior therapy: Core principles for practice,” in. Wiley, 2012, pp. 75–97.

[160] J. Marshall, R. J. Hill, and P. Dodrill, “A survey of practice for clinicians working with children with autism spectrum disorders and feeding difficulties,” *International journal of speech-language pathology*, vol. 15, no. 3, pp. 279–285, 2013.

[161] J Marshall, R Ware, J Ziviani, R. Hill, and P Dodrill, “Efficacy of interventions to improve feeding difficulties in children with autism spectrum disorders: A systematic review and meta-analysis,” *Child: care, health and development*, vol. 41, no. 2, pp. 278–302, 2015.

[162] S. Tyner, A. Brewer, M. Helman, Y. Leon, J. Pritchard, and M. Schlund, “Nice doggie! contact desensitization plus reinforcement decreases dog phobias for children with autism,” *Behavior analysis in practice*, vol. 9, no. 1, pp. 54–57, 2016.

[163] V. McMullen, S. L. Mahfood, G. L. Francis, and J. Bubenik, “Using prediction and desensitization techniques to treat dental anxiety: A case example,” *Behavioral Interventions*, vol. 32, no. 1, pp. 91–100, 2017.

[164] *Get started with the listening program® for optimal human performance*, `https://advancedbrain.com/getstarted/`, (Accessed on 02/13/2018).

[165] L. Hall and J. Case-Smith, “The effect of sound-based intervention on children with sensory processing disorders and visual–motor delays,” *American Journal of Occupational Therapy*, vol. 61, no. 2, pp. 209–215, 2007.

[166] B. M. Gee, K. Thompson, and H. St John, “Efficacy of a sound-based intervention with a child with an autism spectrum disorder and auditory sensory over-responsivity,” *Occupational therapy international*, vol. 21, no. 1, pp. 12–20, 2014.

[167] B. Gee, K. Thompson, A. Pierce, M. Toupin, and J. Holst, "The effectiveness of auditory stimulation in children with autism spectrum disorders: A case–control study," *International Journal of Therapy and Rehabilitation*, vol. 22, no. 1, pp. 36–46, 2015.

[168] S. W. Porges, O. V. Bazhenova, E. Bal, N. Carlson, Y. Sorokin, K. J. Heilman, E. H. Cook, and G. F. Lewis, "Reducing auditory hypersensitivities in autistic spectrum disorder: Preliminary findings evaluating the listening project protocol," *Frontiers in Pediatrics*, vol. 2, p. 80, 2014.

[169] S. W. Porges, M. Macellaio, S. D. Stanfill, K. McCue, G. F. Lewis, E. R. Harden, M. Handelman, J. Denver, O. V. Bazhenova, and K. J. Heilman, "Respiratory sinus arrhythmia and auditory processing in autism: Modifiable deficits of an integrated social engagement system?" *International Journal of Psychophysiology*, vol. 88, no. 3, pp. 261–270, 2013.

[170] S. A. Schoen, L. J. Miller, and K. E. Green, "Pilot study of the sensory over-responsivity scales: Assessment and inventory," *American Journal of Occupational Therapy*, vol. 62, no. 4, pp. 393–406, 2008.

[171] A. J. Ayres, *Sensory integration and learning disorders*. Western Psychological Services, 1972.

[172] T. Smith, D. W. Mruzek, and D. Mozingo, "Sensory integration therapy," *Controversial therapies for autism and intellectual disabilities: Fad, fashion, and science in professional practice*, pp. 247–269, 2015.

[173] H Kaplan, M Clopton, M Kaplan, L Messbauer, and K McPherson, "Snoezelen multi-sensory environments: Task engagement and generalization," *Research in developmental disabilities*, vol. 27, no. 4, pp. 443–455, 2006.

[174] W. Farr, N. Yuill, and H. Raffle, "Social benefits of a tangible user interface for children with autistic spectrum conditions," *Autism*, vol. 14, no. 3, pp. 237–252, 2010.

[175] K. E. Ringland, R. Zalapa, M. Neal, L. Escobedo, M. Tentori, and G. R. Hayes, "Sensorypaint: A multimodal sensory intervention for children with neurodevelopmental disorders," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 2014, pp. 873–884.

[176] N. Newbutt, M. Donegan, *et al.*, “A brief review: Assistive technology and autism, a proposal for virtual tools for improved communication and emotional recognition,” 2010.

[177] B. O. Ploog, A. Scharf, D. Nelson, and P. J. Brooks, “Use of computer-assisted technologies (cat) to enhance social, communicative, and language development in children with autism spectrum disorders,” *Journal of autism and developmental disorders*, vol. 43, no. 2, pp. 301–322, 2013.

[178] R. Iovannone, G. Dunlap, H. Huber, and D. Kincaid, “Effective educational practices for students with autism spectrum disorders,” *Focus on autism and other developmental disabilities*, vol. 18, no. 3, pp. 150–165, 2003.

[179] F. J. Sansosti, K. A. Powell-Smith, and R. J. Cowan, *High-functioning autism/Asperger syndrome in schools: Assessment and intervention.* Guilford Press, 2010.

[180] W. Keay-Bright, “Can computers create relaxation? designing reactickles© software with children on the autistic spectrum,” *CoDesign*, vol. 3, no. 2, pp. 97–110, 2007.

[181] S. E. Lindley, J. Le Couteur, and N. L. Berthouze, “Stirring up experience through movement in game play: Effects on engagement and social behaviour,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2008, pp. 511–514.

[182] A. M. Piper, E. O’Brien, M. R. Morris, and T. Winograd, “Sides: A cooperative tabletop computer game for social skills development,” in *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, ACM, 2006, pp. 1–10.

[183] I. Giannopulu and G. Pradel, “Multimodal interactions in free game play of children with autism and a mobile toy robot,” *NeuroRehabilitation*, vol. 27, no. 4, pp. 305–311, 2010.

[184] L. Bartoli, C. Corradi, F. Garzotto, and M. Valoriani, “Exploring motion-based touchless games for autistic children’s learning,” in *Proceedings of the 12th International Conference on Interaction Design and Children*, ACM, 2013, pp. 102–111.

[185] M. Wilson, “Six views of embodied cognition,” *Psychonomic bulletin & review*, vol. 9, no. 4, pp. 625–636, 2002.

[186] N. Uzuegbunam, W.-H. Wong, S.-c. S. Cheung, and L. Ruble, “Mebook: Kinect-based self-modeling intervention for children with autism,” in *Multimedia and Expo (ICME), 2015 IEEE International Conference on*, IEEE, 2015, pp. 1–6.

[187] M. Saiano, L. Pellegrino, M. Casadio, S. Summa, E. Garbarino, V. Rossi, D. Dall’Agata, and V. Sanguineti, “Natural interfaces and virtual environments for the acquisition of street crossing and path following skills in adults with autism spectrum disorders: A feasibility study,” *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 17, 2015.

[188] G. Altanis, M. Boloudakis, S. Retalis, and N. Nikou, “Children with motor impairments play a kinect learning game: First findings from a pilot case in an authentic classroom environment,” *J Interact Design Architect*, vol. 19, pp. 91–104, 2013.

[189] L. A. Wijnhoven, D. H. Creemers, R. C. Engels, and I. Granic, “The effect of the video game mindlight on anxiety symptoms in children with an autism spectrum disorder,” *BMC psychiatry*, vol. 15, no. 1, p. 138, 2015.

[190] M. Silver and P. Oakes, “Evaluation of a new computer intervention to teach people with autism or asperger syndrome to recognize and predict emotions in others,” *Autism*, vol. 5, no. 3, pp. 299–316, 2001.

[191] S. W. White, D. Oswald, T. Ollendick, and L. Scahill, “Anxiety in children and adolescents with autism spectrum disorders,” *Clinical psychology review*, vol. 29, no. 3, pp. 216–229, 2009.

[192] P. Ince, G. Haddock, and S. Tai, “A systematic review of the implementation of recommended psychological interventions for schizophrenia: Rates, barriers, and improvement strategies,” *Psychology and Psychotherapy: Theory, Research and Practice*, vol. 89, no. 3, pp. 324–350, 2016.

[193] L. Van Der Krieke, L. Wunderink, A. C. Emerencia, P. De Jonge, and S. Sytema, “E–mental health self-management for psychotic disorders: State of the art and future perspectives,” *Psychiatric Services*, vol. 65, no. 1, pp. 33–49, 2014.

[194] J. H. Wright and A. S. Wright, “Computer-assisted psychotherapy,” *The Journal of psychotherapy practice and research*, vol. 6, no. 4, p. 315, 1997.

[195] S. Arnab, T. Lim, M. B. Carvalho, F. Bellotti, S. Freitas, S. Louchart, N. Suttie, R. Berta, and A. De Gloria, "Mapping learning and game mechanics for serious games analysis," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 391–411, 2015.

[196] L. A. Annetta, "The "i's" have it: A framework for serious educational game design.," *Review of General Psychology*, vol. 14, no. 2, p. 105, 2010.

[197] C. E. Catalano, A. M. Luccini, and M. Mortara, "Best practices for an effective design and evaluation of serious games," 2014.

[198] T. Susi, M. Johannesson, and P. Backlund, "Serious games: An overview," 2007.

[199] E. M. Whyte, J. M. Smyth, and K. S. Scherf, "Designing serious game interventions for individuals with autism," *Journal of autism and developmental disorders*, vol. 45, no. 12, pp. 3820–3831, 2015.

[200] A. Constantin, H. Johnson, E. Smith, D. Lengyel, and M. Brosnan, "Designing computer-based rewards with and for children with autism spectrum disorder and/or intellectual disability," *Computers in Human Behavior*, 2017.

[201] S Fletcher-Watson, H Pain, S Hammond, A Humphry, and H McConachie, "Designing for young children with autism spectrum disorder: A case study of an ipad app," *International Journal of Child-Computer Interaction*, vol. 7, pp. 1–14, 2016.

[202] K. M. Kapp, *The gamification of learning and instruction: game-based methods and strategies for training and education.* John Wiley & Sons, 2012.

[203] A. Z. Hassan, B. T. Zahed, F. T. Zohora, J. M. Moosa, T. Salam, M. M. Rahman, H. S. Ferdous, and S. I. Ahmed, "Developing the concept of money by interactive computer games for autistic children," in *Multimedia (ISM), 2011 IEEE International Symposium on*, IEEE, 2011, pp. 559–564.

[204] A. O. Barajas, H. Al Osman, and S. Shirmohammadi, "A serious game for children with autism spectrum disorder as a tool for play therapy," in *Serious Games and Applications for Health (SeGAH), 2017 IEEE 5th International Conference on*, IEEE, 2017, pp. 1–7.

[205] E. A. Schoneveld, M. Malmberg, A. Lichtwarck-Aschoff, G. P. Verheijen, R. C. Engels, and I. Granic, “A neurofeedback video game (mindlight) to prevent anxiety in children: A randomized controlled trial,” *Computers in Human Behavior*, vol. 63, pp. 321–333, 2016.

[206] Y. Bar-Haim, D. Lamy, L. Pergamin, M. J. Bakermans-Kranenburg, and M. H. Van Ijzendoorn, “Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study.,” *Psychological bulletin*, vol. 133, no. 1, p. 1, 2007.

[207] D. S. Pine, S. M. Helfinstein, Y. Bar-Haim, E. Nelson, and N. A. Fox, “Challenges in developing novel treatments for childhood disorders: Lessons from research on anxiety,” *Neuropsychopharmacology*, vol. 34, no. 1, p. 213, 2009.

[208] P. Muris and A. P. Field, “Distorted cognition and pathological anxiety in children and adolescents,” *Cognition and emotion*, vol. 22, no. 3, pp. 395–421, 2008.

[209] T. Shechner, A. Rimon-Chakir, J. C. Britton, D. Lotan, A. Apter, P. D. Bliese, D. S. Pine, and Y. Bar-Haim, “Attention bias modification treatment augmenting effects on cognitive behavioral therapy in children with anxiety: Randomized controlled trial,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 53, no. 1, pp. 61–71, 2014.

[210] D. M. Romano, G. Sheppard, J. Hall, A. Miller, and Z. Ma, “Basic: A believable, adaptable socially intelligent character for social presence,” in *Proceedings of the 8th Annual International Workshop on Presence*, Citeseer, 2005, pp. 287–290.

[211] G. Riva, M. Bacchetta, M. Baruffi, S. Rinaldi, F. Vincelli, and E. Molinari, “Virtual reality-based experiential cognitive treatment of obesity and binge-eating disorders,” *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, vol. 7, no. 3, pp. 209–219, 2000.

[212] S. Shema-Shiratzky, M. Brozgol, P. Cornejo-Thumm, K. Geva-Dayan, M. Rotstein, Y. Leitner, J. M. Hausdorff, and A. Mirelman, “Virtual reality training to enhance behavior and cognitive function among children with attention-deficit/hyperactivity disorder: Brief report,” *Developmental neurorehabilitation*, vol. 22, no. 6, pp. 431–436, 2019.

[213] T. D. Parsons and A. A. Rizzo, “Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis,” *Journal of behavior therapy and experimental psychiatry*, vol. 39, no. 3, pp. 250–261, 2008.

[214] G Rivaetal, “Virtual reality: A new clinical setting lab,” *Virtual Environments in Clinical Psychology and Neuroscience: Methods and Techniques in Advanced Patient-therapist Interaction*, vol. 58, p. 73, 1998.

[215] M. R. Kandalaft, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman, “Virtual reality social cognition training for young adults with high-functioning autism,” in, 1, vol. 43, Springer, 2013, pp. 34–44.

[216] M. L. Max and J. C. Burke, “Virtual reality for autism communication and education, with lessons for medical training simulators.,” in, vol. 39, 1997, pp. 46–53.

[217] D. J. Cox, T. Brown, V. Ross, M. Moncrief, R. Schmitt, G. Gaffney, and R. Reeve, “Can youth with autism spectrum disorder use virtual reality driving simulation training to evaluate and improve driving performance? an exploratory study,” *Journal of autism and developmental disorders*, vol. 47, no. 8, pp. 2544–2555, 2017.

[218] L. Lamash, E. Klinger, and N. Josman, “Using a virtual supermarket to promote independent functioning among adolescents with autism spectrum disorder,” in *Virtual Rehabilitation (ICVR), 2017 International Conference on*, IEEE, 2017, pp. 1–7.

[219] N. Newbutt, C. Sung, H.-J. Kuo, M. J. Leahy, C.-C. Lin, and B. Tong, “Brief report: A pilot study of the use of a virtual reality headset in autism populations,” in, 9, vol. 46, Springer, 2016, pp. 3166–3176.

[220] A. Leonard, P Mitchell, and S Parsons, “Finding a place to sit: A preliminary investigation into the effectiveness of virtual environments for social skills training for people with autistic spectrum disorders,” *Virtual Reality and Associated Technologies, Veszprem, Hungary, University of Reading*, 2002.

[221] S. Parsons and S. Cobb, “State-of-the-art of virtual reality technologies for children on the autism spectrum,” *European Journal of Special Needs Education*, vol. 26, no. 3, pp. 355–366, 2011.

[222] D. C. Strickland, D. McAllister, C. D. Coles, and S. Osborne, "An evolution of virtual reality training designs for children with autism and fetal alcohol spectrum disorders," *Topics in language disorders*, vol. 27, no. 3, p. 226, 2007.

[223] D. Charitos, G. Karadanos, E. Sereti, S. Triantafillou, S. Koukouvinou, and D. Martakos, "Employing virtual reality for aiding the organisation of autistic children behaviour in everyday tasks," in *Proceedings of ICDVRAT*, 2000, pp. 147–152.

[224] M. Wang and E. Anagnostou, "Virtual reality as treatment tool for children with autism," in *Comprehensive guide to autism*, Springer, 2014, pp. 2125–2141.

[225] M. Krijn, P. M. Emmelkamp, R. Biemond, C. d. W. de Ligny, M. J. Schuemie, and C. A. van der Mast, "Treatment of acrophobia in virtual reality: The role of immersion and presence," *Behaviour research and therapy*, vol. 42, no. 2, pp. 229–239, 2004.

[226] M. B. Powers and P. M. Emmelkamp, "Virtual reality exposure therapy for anxiety disorders: A meta-analysis," *Journal of anxiety disorders*, vol. 22, no. 3, pp. 561–569, 2008.

[227] M. Price, N. Mehta, E. B. Tone, and P. L. Anderson, "Does engagement with exposure yield better outcomes? components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia," *Journal of anxiety disorders*, vol. 25, no. 6, pp. 763–770, 2011.

[228] M. Price and P. Anderson, "The role of presence in virtual reality exposure therapy," *Journal of anxiety disorders*, vol. 21, no. 5, pp. 742–751, 2007.

[229] M. Lombard, "Resources for the study of presence: Presence explication," *Retrieved September*, vol. 3, p. 2000, 2000.

[230] M. Schuemie, M Bruynzeel, L Drost, M Brinckman, G De Haan, P. Emmelkamp, and C. Van der Mast, "Treatment of acrophobia in virtual reality: A pilot study," in *Conference Proceedings Euromedia*, 2000, pp. 271–275.

[231] Y. Ling, W.-P. Brinkman, H. T. Nefs, C. Qu, and I. Heynderickx, "Effects of stereoscopic viewing on presence, anxiety, and cybersickness in a virtual reality environment for public speaking," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 3, pp. 254–267, 2012.

[232] D. Villani, C. Repetto, P. Cipresso, and G. Riva, “May i experience more presence in doing the same thing in virtual reality than in reality? an answer from a simulated job interview,” *Interacting with Computers*, vol. 24, no. 4, pp. 265–272, 2012.

[233] Y. Ling, H. T. Nefs, N. Morina, I. Heynderickx, and W.-P. Brinkman, “A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders,” *PloS one*, vol. 9, no. 5, e96144, 2014.

[234] S. Wallace, S. Parsons, A. Westbury, K. White, K. White, and A. Bailey, “Sense of presence and atypical social judgments in immersive virtual environments: Responses of adolescents with autism spectrum disorders,” *Autism*, vol. 14, no. 3, pp. 199–213, 2010.

[235] S. Wallace, S. Parsons, and A. Bailey, “Self-reported sense of presence and responses to social stimuli by adolescents with autism spectrum disorder in a collaborative virtual reality environment,” *Journal of Intellectual & Developmental Disability*, vol. 42, no. 2, pp. 131–141, 2017.

[236] S. H. Spence, “A measure of anxiety symptoms among children,” *Behaviour research and therapy*, vol. 36, no. 5, pp. 545–566, 1998.

[237] M. Maskey, J. Rodgers, V. Grahame, M. Glod, E. Honey, J. Kinnear, M. Labus, J. Milne, D. Minos, H. McConachie, *et al.*, “A randomised controlled feasibility trial of immersive virtual reality treatment with cognitive behaviour therapy for specific phobias in young people with autism spectrum disorder,” *Journal of autism and developmental disorders*, vol. 49, no. 5, pp. 1912–1927, 2019.

[238] J. N. Meindl, S. Saba, M. Gray, L. Stuebing, and A. Jarvis, “Reducing blood draw phobia in an adult with autism spectrum disorder using low-cost virtual reality exposure therapy,” *Journal of Applied Research in Intellectual Disabilities*, vol. 32, no. 6, pp. 1446–1452, 2019.

[239] I. T. Miller, B. K. Wiederhold, C. S. Miller, and M. D. Wiederhold, “Virtual reality air travel training with children on the autism spectrum: A preliminary report,” *Cyberpsychology, Behavior, and Social Networking*, vol. 23, no. 1, pp. 10–15, 2020.

[240] E. Gomes, F. S. Pedroso, and M. B. Wagner, “Auditory hypersensitivity in the autistic spectrum disorder,” *Pró-Fono Revista de Atualização Científica*, vol. 20, no. 4, pp. 279–284, 2008.

[241] P. Larsson, A. Väljamäe, D. Västfjäll, A. Tajadura-Jiménez, and M. Kleiner, "Auditory-induced presence in mixed reality environments and related technology," in *The Engineering of Mixed Reality Systems*, Springer, 2010, pp. 143–163.

[242] B. Shinn-Cunningham and R. D. Shilling, "Virtual auditory displays," in *Handbook of Virtual Environments*, CRC Press, 2002, pp. 105–132.

[243] W.-P. Brinkman, A. R. Hoekstra, and R. van EGMOND, "The effect of 3d audio and other audio techniques on virtual reality experience," *ANNUAL REVIEW OF CYBERTHERAPY AND TELEMEDICINE 2015*, p. 44, 2015.

[244] S. Hughes and G. Kearney, "Fear and localisation: Emotional fine-tuning utlising multiple source directions," in *Audio Engineering Society Conference: 56th International Conference: Audio for Games*, Audio Engineering Society, 2015.

[245] P. Larsson, D. Västfjäll, and M. Kleiner, "Effects of auditory information consistency and room acoustic cues on presence in virtual environments," *Acoustical Science and Technology*, vol. 29, no. 2, pp. 191–194, 2008.

[246] P. Larsson, D. Vastfjall, and M. Kleiner, "Better presence and performance in virtual environments by improved binaural sound rendering," in *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*, Audio Engineering Society, 2002.

[247] C. Frauenberger, J. Good, and W. Keay-Bright, "Designing technology for children with special needs: Bridging perspectives through participatory design," *CoDesign*, vol. 7, no. 1, pp. 1–28, 2011.

[248] L. Malinverni, J. Mora-Guiard, V. Padillo, L. Valero, A. Hervás, and N. Pares, "An inclusive design approach for developing video games for children with autism spectrum disorder," *Computers in Human Behavior*, vol. 71, pp. 535–549, 2017.

[249] J. Mora-Guiard, C. Crowell, N. Pares, and P. Heaton, "Lands of fog: Helping children with autism in social interaction through a full-body interactive experience," in *Proceedings of the The 15th International Conference on Interaction Design and Children*, ACM, 2016, pp. 262–274.

[250] L Millen, S. Cobb, H Patel, *et al.*, "Participatory design with children with autism," in *Proceedings 8th Intl. Conference on Disability, VR and Associated Technologies*, 2010, pp. 93–101.

[251] C. Frauenberger, J. Good, A. Alcorn, and H. Pain, “Supporting the design contributions of children with autism spectrum conditions,” in *Proceedings of the 11th International Conference on Interaction Design and Children*, 2012, pp. 134–143.

[252] S. Parsons, L. Millen, S. Garib-Penna, and S. Cobb, “Participatory design in the development of innovative technologies for children and young people on the autism spectrum: The cospatial project,” *Journal of Assistive Technologies*, 2011.

[253] L. Millen, S. Cobb, and H. Patel, “Participatory design approach with children with autism,” 2011.

[254] M. Madsen, R. El Kaliouby, M. Eckhardt, M. E. Hoque, M. S. Goodwin, and R. Picard, “Lessons from participatory design with adolescents on the autism spectrum,” in *CHI’09 Extended Abstracts on Human Factors in Computing Systems*, 2009, pp. 3835–3840.

[255] S. Merter and D. Hasırcı, “A participatory product design process with children with autism spectrum disorder,” *CoDesign*, vol. 14, no. 3, pp. 170–187, 2018.

[256] D. Magkafa, N. Newbutt, and M. Palmer, “Implementing co-design practices for the development of a museum interface for autistic children,” *Recent Advances in Technologies for Inclusive Well-Being: Virtual Patients, Gamification and Simulation*, pp. 421–443, 2021.

[257] L. Malinverni, J. Mora-Guiard, and N. Pares, “Towards methods for evaluating and communicating participatory design: A multimodal approach,” *International Journal of Human-Computer Studies*, vol. 94, pp. 53–63, 2016.

[258] D. Murray, M. Lesser, and W. Lawson, “Attention, monotropism and the diagnostic criteria for autism,” *Autism*, vol. 9, no. 2, pp. 139–156, 2005.

[259] K. Dautenhahn, “Design issues on interactive environments for children with autism,” in *In: Procs of ICDVRAT 2000, the 3rd Int Conf on Disability, Virtual Reality and Associated Technologies*, University of Reading, 2000.

[260] L. Schreibman and R. L. Koegel, “Fostering self-management: Parent-delivered pivotal response training for children with autistic disorder.,” 1996.

[261] F. Baumgarte and C. Faller, “Binaural cue coding-part i: Psychoacoustic fundamentals and design principles,” in, 6, vol. 11, IEEE, 2003, pp. 509–519.

[262] J. C. Middlebrooks and D. M. Green, "Sound localization by human listeners," in, 1, vol. 42, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, 1991, pp. 135–159.

[263] T. R. Anderson and R. H. Gilkey, "Binaural and spatial hearing in real and virtual environments:" in. Lawrence Erlbaum Associates, 1997, pp. 1–23.

[264] S. Handel, "Listening: An introduction to the perception of auditory events," in. MIT Pr., 1993, pp. 99–112.

[265] Glackin, "A spiking neural network model of the medial superior olive using spike timing dependent plasticity for sound localization," *Frontiers in Computational Neuroscience*, 2010. DOI: `10.3389/fncom.2010.00018`.

[266] D. M. Howard, J. A. S. Angus, and J. J. Wells, "Acoustics and psychoacoustics," in. Focal Press, 2013, pp. 107–119.

[267] J. Zwislocki and R. Feldman, "Just noticeable differences in dichotic phase," in, 5, vol. 28, ASA, 1956, pp. 860–864.

[268] F. L. Wightman and D. J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1648–1661, 1992.

[269] A. D. Musicant and R. A. Butler, "The influence of pinnae-based spectral cues on sound localization," *The Journal of the Acoustical Society of America*, vol. 75, no. 4, pp. 1195–1200, 1984.

[270] D. Wright, J. H. Hebrank, and B. Wilson, "Pinna reflections as cues for localization," *The Journal of the Acoustical Society of America*, vol. 56, no. 3, pp. 957–962, 1974.

[271] B. Xie, *Head-related transfer function and virtual auditory display.* J. Ross Publishing, 2013.

[272] T. Potisk, "Head-related transfer function," in *Seminar Ia, Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia*, 2015.

[273] X.-L. Zhong and B.-S. Xie, "Head-related transfer functions and virtual auditory display," *Soundscape Semiotics - Localisation and Categorisation*, 2014. DOI: `10.5772/56907`.

[274] Neumann, *Ku 100.* [Online]. Available: `https://en-de.neumann.com/ku-100`.

[275] G. Kearney and T. Doyle, “An hrtf database for virtual loudspeaker rendering,” in *Audio Engineering Society Convention 139*, Audio Engineering Society, 2015.

[276] D. R. Begault and L. J. Trejo, “3-d sound for virtual reality and multimedia,” 2000.

[277] S. T. Hayes and J. A. Adams, “Device motion via head tracking for mobile interaction,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, vol. 60, 2016, pp. 774–778.

[278] D. N. Zotkin, R. Duraiswami, and L. S. Davis, “Rendering localized spatial audio in a virtual auditory space,” *IEEE Transactions on multimedia*, vol. 6, no. 4, pp. 553–564, 2004.

[279] A. C. Isbert, *Diseño acústico de espacios arquitectónicos.* Univ. Politèc. de Catalunya, 1998, vol. 4.

[280] D. R. Begault, “Perceptual effects of synthetic reverberation on three-dimensional audio systems,” *Journal of the Audio Engineering Society*, vol. 40, no. 11, pp. 895–904, 1992.

[281] W. M. Hartmann, “Localization of sound in rooms,” *The Journal of the Acoustical Society of America*, vol. 74, no. 5, pp. 1380–1391, 1983.

[282] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, “Auditory distance perception in humans: A summary of past and present research,” *ACTA Acustica united with Acustica*, vol. 91, no. 3, pp. 409–420, 2005.

[283] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, “Creating interactive virtual acoustic environments,” in, 9, vol. 47, Audio Engineering Society, 1999, pp. 675–705.

[284] P. Zahorik and F. L. Wightman, “Loudness constancy with varying sound source distance,” *Nature neuroscience*, vol. 4, no. 1, pp. 78–83, 2001.

[285] C. W. Sheeline, “An investigation of the effects of direct and reverberant signal interaction on auditory distance perception,” PhD thesis, Stanford University, 1982.

[286] D. H. Mershon and L. E. King, “Intensity and reverberation as factors in the auditory perception of egocentric distance,” *Attention, Perception, & Psychophysics*, vol. 18, no. 6, pp. 409–415, 1975.

[287] B. F. Hertz, “100 years with stereo!: The beginning,” in *Audio Engineering Society Convention 68*, Audio Engineering Society, 1981.

[288] M. Kerins, *Beyond Dolby (stereo): cinema in the digital sound age.* Indiana University Press, 2010.

[289] M. A. Gerzon, “Periphony: With-height sound reproduction,” *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.

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

[291] D. G. Malham and A. Myatt, “3-d sound spatialization using ambisonic techniques,” *Computer music journal*, vol. 19, no. 4, pp. 58–70, 1995.

[292] S.-N. Yao, “Headphone-based immersive audio for virtual reality headsets,” *IEEE Transactions on Consumer Electronics*, vol. 63, no. 3, pp. 300–308, 2017.

[293] R. Faria and J. Zuffo, “An auralization engine adapting a 3d image source acoustic model to an ambisonics coder for immersive virtual reality,” in *Audio Engineering Society Conference: 28th International Conference: The Future of Audio Technology–Surround and Beyond*, Audio Engineering Society, 2006.

[294] M. Frank, F. Zotter, and A. Sontacchi, “Producing 3d audio in ambisonics,” in *Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology–Cinema, Television and the Internet*, Audio Engineering Society, 2015.

[295] *Spherical harmonics.* [Online]. Available: `https://brilliant.org/wiki/spherical-harmonics/`.

[296] B. Wiggins, “An investigation into the real-time manipulation and control of three-dimensional sound fields,” PhD thesis, University of Derby, 2004.

[297] *Facebook 360 spatial workstation*, `https://facebook360.fb.com/spatial-workstation/`, Accessed: 2020-04-17.

[298] P. G. Craven and M. A. Gerzon, *Coincident microphone simulation covering three dimensional space and yielding various directional outputs*, US Patent 4,042,779, 1977.

[299] G. Kearney, M. Gorzel, H. Rice, and F. Boland, “Distance perception in interactive virtual acoustic environments using first and higher order ambisonic sound fields,” *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 61–71, 2012.

[300] *Soundfield microphone array.* [Online]. Available: `https://www.creativefieldrecording.com/wp-content/uploads/2017/03/Soundfield-Microphone-Tetrahedral-Microphone-Array-149x300.jpg`.

[301] M. Narbutt, J. Skoglund, A. Allen, M. Chinen, D. Barry, and A. Hines, "Ambiqual: Towards a quality metric for headphone rendered compressed ambisonic spatial audio," *Applied Sciences*, vol. 10, no. 9, p. 3188, 2020.

[302] S. Moreau, J. Daniel, and S. Bertet, "3d sound field recording with higher order ambisonics–objective measurements and validation of a 4th order spherical microphone," in *120th Convention of the AES*, 2006, pp. 20–23.

[303] M. Noisternig, A. Sontacchi, T. Musil, and R. Holdrich, "A 3d ambisonic based binaural sound reproduction system," in *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, Audio Engineering Society, 2003.

[304] M. Noisternig, T. Musil, A. Sontacchi, and R. Holdrich, "3d binaural sound reproduction using a virtual ambisonic approach," in *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003*, IEEE, 2003, pp. 174–178.

[305] D. Arteaga, *Introduction to ambisonics*, 2015.

[306] M. Taylor, A. Chandak, Q. Mo, C. Lauterbach, C. Schissler, and D. Manocha, "Guided multiview ray tracing for fast auralization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 11, pp. 1797–1810, 2012.

[307] A. Rungta, C. Schissler, N. Rewkowski, R. Mehra, and D. Manocha, "Diffraction kernels for interactive sound propagation in dynamic environments," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 4, pp. 1613–1622, 2018.

[308] A. Plinge, S. J. Schlecht, O. Thiergart, T. Robotham, O. Rummukainen, and E. A. Habets, "Six-degrees-of-freedom binaural audio reproduction of first-order ambisonics with distance information," in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2018.

[309] D. Rivas Méndez, C. Armstrong, J. Stubbs, M. Stiles, and G. Kearney, "Practical recording techniques for music production with six-degrees of freedom virtual reality," in *Audio Engineering Society Convention 145*, Audio Engineering Society, 2018.

[310] T. Rudzki, I. Gomez-Lanzaco, J. Stubbs, J. Skoglund, D. T. Murphy, and G. Kearney, "Auditory localization in low-bitrate compressed ambisonic scenes," *Applied Sciences*, vol. 9, no. 13, p. 2618, 2019.

[311] P. Larsson, A. Väljamäe, D. Västfjäll, A. Tajadura-Jiménez, and M. Kleiner, "Auditory-induced presence in mixed reality environments and related technology," *The Engineering of Mixed Reality Systems, Human-Computer Interaction Series*, vol. 10, pp. 143–163, 2010.

[312] K. M. Stanney and K. S. Hale, *Handbook of virtual environments: Design, implementation, and applications.* CRC Press, 2014.

[313] T. Hermann and H. Ritter, "Sound and meaning in auditory data display," *Proceedings of the IEEE*, vol. 92, no. 4, pp. 730–741, 2004.

[314] A. Kohlrausch and S. van de Par, "Auditory-visual interaction: From fundamental research in cognitive psychology to (possible) applications," in *Human Vision and Electronic Imaging IV*, International Society for Optics and Photonics, vol. 3644, 1999, pp. 34–45.

[315] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *Journal of computer-mediated communication*, vol. 3, no. 2, JCMC321, 1997.

[316] M. Radeau, "Auditory-visual spatial interaction and modularity.," *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 1994.

[317] R. H. Gilkey and J. M. Weisenberger, "The sense of presence for the suddenly deafened adult: Implications for virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 4, no. 4, pp. 357–363, 1995.

[318] K. Ozawa, Y. Chujo, Y. Suzuki, and T. Sone, "Psychological factors involved in auditory presence," *Acoustical Science and Technology*, vol. 24, no. 1, pp. 42–44, 2003.

[319] F. Grani, F. Argelaguet, V. Gouranton, M. Badawi, R. Gaugne, S. Serafin, and A. Lecuyer, "Design and evaluation of binaural auditory rendering for caves," in *2014 IEEE Virtual Reality (VR)*, IEEE, 2014, pp. 73–74.

[320] E. R. Hoeg, L. J. Gerry, L. Thomsen, N. C. Nilsson, and S. Serafin, "Binaural sound reduces reaction time in a virtual reality search task," in *2017 IEEE 3rd VR workshop on sonic interactions for virtual environments (SIVE)*, 2017, pp. 1–4.

[321] P. M. Larsen JJ, "The effect of spatial audio on immersion, presence, and physiological," Master's thesis, Aalborg University, 2015.

[322] P. Ruediger, J. Spilski, N. Kartal, S. Gsuck, N. O. Beese, S. J. Schlittmeier, T. Lachmann, and A. Ebert, "Cognitive indicators for acoustic source localization and presence in a vivid 3d scene,"

[323] C. M. Hendrix, "Exploratory studies on the sense of presence in virtual environments as a function of visual and auditory display parameters," Master's thesis, Citeseer, 1994.

[324] M. Naef, O. Staadt, and M. Gross, "Spatialized audio rendering for immersive virtual environments," in *Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, 2002, pp. 65–72.

[325] R. Brøvig-Hanssen and A. Danielsen, "The naturalised and the surreal: Changes in the perception of popular music sound," *Organised Sound*, vol. 18, no. 1, pp. 71–80, 2013.

[326] J. L. Maples-Keller, B. E. Bunnell, S.-J. Kim, and B. O. Rothbaum, "The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders.," *Harvard review of psychiatry*, vol. 25, no. 3, pp. 103–113, 2017.

[327] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, "The impact of perception and presence on emotional reactions: A review of research in virtual reality," *Frontiers in psychology*, vol. 6, p. 26, 2015.

[328] G. Robillard, S. Bouchard, T. Fournier, and P. Renaud, "Anxiety and presence during vr immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games," *CyberPsychology & Behavior*, vol. 6, no. 5, pp. 467–476, 2003.

[329] J. Panksepp and G. Bernatzky, "Emotional sounds and the brain: The neuro-affective foundations of musical appreciation," *Behavioural processes*, vol. 60, no. 2, pp. 133–155, 2002.

[330] I. Ekman and R. Kajastila, "Localization cues affect emotional judgments–results from a user study on scary sound," in *Audio Engineering Society Conference: 35th International Conference: Audio for Games*, Audio Engineering Society, 2009.

[331] B. O. Rothbaum, L. Hodges, R. Alarcon, D. Ready, F. Shahar, K. Graap, J. Pair, P. Hebert, D. Gotz, B. Wills, *et al.*, "Virtual reality exposure therapy for ptsd vietnam veterans: A case study," *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, vol. 12, no. 2, pp. 263–271, 1999.

[332] D. P. Wood, J. Murphy, K. Center, R. McLay, D. Reeves, J. Pyne, R. Shilling, and B. K. Wiederhold, "Combat-related post-traumatic stress disorder: A case report using virtual reality exposure therapy with physiological monitoring," *Cyberpsychology & Behavior*, vol. 10, no. 2, pp. 309–315, 2006.

[333] D. P. Wood, J. Webb-Murphy, K. Center, R. McLay, R. Koffman, S. Johnston, J. Spira, J. M. Pyne, and B. K. Wiederhold, "Combat-related post-traumatic stress disorder: A case report using virtual reality graded exposure therapy with physiological monitoring with a female seabee," *Military medicine*, vol. 174, no. 11, pp. 1215–1222, 2009.

[334] A. F. Abate, M. Nappi, and S. Ricciardi, "Ar based environment for exposure therapy to mottephobia," in *International Conference on Virtual and Mixed Reality*, Springer, 2011, pp. 3–11.

[335] V. D. H. Thanh, O. Pui, and M. Constable, "Room vr: A vr therapy game for children who fear the dark," in *SIGGRAPH Asia 2017 Posters*, 2017, pp. 1–2.

[336] M. Poyade, G. Morris, I. Taylor, and V. Portela, "Designing a virtual reality exposure therapy to familiarise and desensitise to environmental stressors in airports," 2017.

[337] J. Argo, "Immersive soundscapes to elicit anxiety in exposure therapy: Physical desensitization and mental catharsis," PhD thesis, University of Glasgow, 2017.

[338] J. S. Karst and A. V. Van Hecke, "Parent and family impact of autism spectrum disorders: A review and proposed model for intervention evaluation," *Clinical child and family psychology review*, vol. 15, no. 3, pp. 247–277, 2012.

[339] G. Montes and J. S. Halterman, “Psychological functioning and coping among mothers of children with autism: A population-based study,” *Pediatrics*, vol. 119, no. 5, e1040–e1046, 2007.

[340] G. Mahoney and B. Wiggers, “The role of parents in early intervention: Implications for social work,” *Children & Schools*, vol. 29, no. 1, pp. 7–15, 2007.

[341] L. K. Koegel, R. L. Koegel, J. K. Harrower, and C. M. Carter, “Pivotal response intervention i: Overview of approach,” *Journal of the Association for Persons with Severe Handicaps*, vol. 24, no. 3, pp. 174–185, 1999.

[342] M. Slater, “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,” in, 1535, vol. 364, The Royal Society, 2009, pp. 3549–3557.

[343] O. Rummukainen, T. Robotham, S. J. Schlecht, A. Plinge, J. Herre, and E. A. Habels, “Audio quality evaluation in virtual reality: Multiple stimulus ranking with behavior tracking,” in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2018.

[344] *Unity3d - game engine*, `https://unity3d.com`, Accessed: 2018-09-10.

[345] L. Bozgeyikli, “Virtual reality serious games for individuals with autism spectrum disorder: Design considerations,” in, 2016, pp. 180 –183.

[346] S. R. Leekam, C. Nieto, S. J. Libby, L. Wing, and J. Gould, “Describing the sensory abnormalities of children and adults with autism,” in, 5, vol. 37, Springer, 2007, pp. 894–910.

[347] M. Liss, J. Mailloux, and M. J. Erchull, “The relationships between sensory processing sensitivity, alexithymia, autism, depression, and anxiety,” in, 3, vol. 45, Elsevier, 2008, pp. 255–259.

[348] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel, “Investigation on localisation accuracy for first and higher order ambisonics reproduced sound sources,” *Acta Acustica united with Acustica*, vol. 99, no. 4, pp. 642–657, 2013.

[349] *Reaper- digital audio workstation*, `https://www.reaper.fm/`, Accessed: 2020-04-17.

[350] *Audiokinetic wwise*, `https://www.audiokinetic.com/products/wwise`, Accessed: 2018-09-10.

[351] *Google resonance - spatial audio sdk*, `https://developers.google.com/resonance-audio/`, year = Accessed: 2018-06-30.

[352] L. Ward, B. Shirley, and J. Francombe, "Accessible object-based audio using hierarchical narrative importance metadata," in *Audio Engineering Society Convention 145*, Audio Engineering Society, 2018.

[353] M. Gomot, M. K. Belmonte, E. T. Bullmore, F. A. Bernard, and S. Baron-Cohen, "Brain hyper-reactivity to auditory novel targets in children with high-functioning autism," in, 9, vol. 131, Oxford University Press, 2008, pp. 2479–2488.

[354] R Čeponienė, T. Lepistö, A. Shestakova, R. Vanhala, P Alku, R. Näätänen, and K Yaguchi, "Speech–sound-selective auditory impairment in children with autism: They can perceive but do not attend," in, 9, vol. 100, National Acad Sciences, 2003, pp. 5567–5572.

[355] *Oculus rift head mounted display*, `https://www.oculus.com/rift/`, Accessed: 2019-01-02.

[356] M Kashino, S Furukawa, T Nakano, S Washizawa, S Yamagishi, A Ochi, A Nagaike, S Kitazawa, and N Kato, "Specific deficits of basic auditory processing in high-functioning pervasive developmental disorders," in *36th ARO (Association for Research in Otolaryngology) MidWinter Meeting, Baltimore, MD, USA*, 2013.

[357] A. J. Whitehouse and D. V. Bishop, "Do children with autism 'switch off'to speech sounds? an investigation using event-related potentials," in, 4, vol. 11, Wiley Online Library, 2008, pp. 516–524.

[358] T. Lokki and M. Grohn, "Navigation with auditory cues in a virtual environment," in, 2, vol. 12, IEEE, 2005, pp. 80–86.

[359] D. McAlpine, "Creating a sense of auditory space," in, 1, vol. 566, Wiley Online Library, 2005, pp. 21–28.

[360] S. Baron-Cohen, M. V. Lombardo, B. Auyeung, E. Ashwin, B. Chakrabarti, and R. Knickmeyer, "Why are autism spectrum conditions more prevalent in males?" *PLoS Biol*, vol. 9, no. 6, e1001081, 2011.

[361] M.-C. Lai, M. V. Lombardo, B. Auyeung, B. Chakrabarti, and S. Baron-Cohen, “Sex/gender differences and autism: Setting the scene for future research,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 54, no. 1, pp. 11–24, 2015.

[362] A. K. Halladay, S. Bishop, J. N. Constantino, A. M. Daniels, K. Koenig, K. Palmer, D. Messinger, K. Pelphrey, S. J. Sanders, A. T. Singer, *et al.*, “Sex and gender differences in autism spectrum disorder: Summarizing evidence gaps and identifying emerging areas of priority,” *Molecular autism*, vol. 6, no. 1, pp. 1–5, 2015.

[363] R. Loomes, L. Hull, and W. P. L. Mandy, “What is the male-to-female ratio in autism spectrum disorder? a systematic review and meta-analysis,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 56, no. 6, pp. 466–474, 2017.

[364] B. U. Seeber and H. Fastl, “Subjective selection of non-individual head-related transfer functions,” Georgia Institute of Technology, 2003.

[365] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, “Localization using nonindividualized head-related transfer functions,” *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993.

[366] J Fels, P Buthmann, and M Vorländer, “Head-related transfer functions of children,” *Acta Acustica united with Acustica*, vol. 90, no. 5, pp. 918–927, 2004.

[367] Y.-M. Hur, Y.-C. Park, S.-P. Lee, and D.-h. Youn, “Efficient individualization method of hrtfs using critical-band based spectral cue control,” in, 4, vol. 30, The Acoustical Society of Korea, 2011, pp. 167–180.

[368] D. Line, C. Loureiro, and R. O. Prates, “Musical app in hypersensitivity to sounds and neurodevelopmental disorders: Applicable strategies,” in *Proceedings of the 11th International Conference of Students of Systematic Musicology*, 2018.

[369] H. Desurvire, M. Caplan, and J. A. Toth, “Using heuristics to evaluate the playability of games,” in *CHI’04 extended abstracts on Human factors in computing systems*, 2004, pp. 1509–1512.

[370] B. Bostan and S. Öğüt, “Game challenges and difficulty levels: Lessons learned from rpgs,” in *International simulation and gaming association conference*, 2009, pp. 1–11.

[371] H. Qin, P.-L. P. Rau, and G. Salvendy, “Effects of different scenarios of game difficulty on player immersion,” *Interacting with Computers*, vol. 22, no. 3, pp. 230–239, 2010.

[372] M.-V. Aponte, G. Levieux, and S. Natkin, “Difficulty in videogames: An experimental validation of a formal definition,” in *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, 2011, pp. 1–8.

[373] K. M. Lee, N. Park, and S.-A. Jin, “Narrative and interactivity in computer games.,” 2006.

[374] C. A. Lindley, “Story and narrative structures in computer games,” *Bushoff, Brunhild. ed*, 2005.

[375] B. Neitzel, “Narrativity of computer games,” in *Handbook of Narratology*, De Gruyter, 2014, pp. 608–622.

[376] S. Turkay and C. K. Kinzer, “The effects of avatar-based customization on player identification,” in *Gamification: Concepts, methodologies, tools, and applications*, IGI Global, 2015, pp. 247–272.

[377] K. Jørgensen, “On the functional aspects of computer game audio,” 2006.

[378] J. Friberg and D. Gärdenfors, “Audio games: New perspectives on game audio,” in *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, 2004, pp. 148–154.

[379] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[380] L. Bartoli, F. Garzotto, M. Gelsomini, L. Oliveto, and M. Valoriani, “Designing and evaluating touchless playful interaction for asd children,” in *Proceedings of the 2014 conference on Interaction design and children*, 2014, pp. 17–26.

[381] J. Kwon and Y. Lee, “Serious games for the job training of persons with developmental disabilities,” *Computers & Education*, vol. 95, pp. 328–339, 2016.

[382] S. Carlier, S. Van der Paelt, F. Ongenae, F. De Backere, and F. De Turck, “Using a serious game to reduce stress and anxiety in children with autism spectrum disorder,” in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2019, pp. 452–461.

[383] S. N. Haber and B. Knutson, "The reward circuit: Linking primate anatomy and human imaging," *Neuropsychopharmacology*, vol. 35, no. 1, pp. 4–26, 2010.

[384] R. A. Bressan and J. A. Crippa, "The role of dopamine in reward and pleasure behaviour–review of data from preclinical research," *Acta Psychiatrica Scandinavica*, vol. 111, pp. 14–21, 2005.

[385] G. S. Berns, S. M. McClure, G. Pagnoni, and P. R. Montague, "Predictability modulates human brain response to reward," *Journal of neuroscience*, vol. 21, no. 8, pp. 2793–2798, 2001.

[386] J. O'Doherty, M. L. Kringelbach, E. T. Rolls, J. Hornak, and C. Andrews, "Abstract reward and punishment representations in the human orbitofrontal cortex," *Nature neuroscience*, vol. 4, no. 1, pp. 95–102, 2001.

[387] S. Hamann and H. Mao, "Positive and negative emotional verbal stimuli elicit activity in the left amygdala," *Neuroreport*, vol. 13, no. 1, pp. 15–19, 2002.

[388] J. K. Rilling, D. A. Gutman, T. R. Zeh, G. Pagnoni, G. S. Berns, and C. D. Kilts, "A neural basis for social cooperation," *Neuron*, vol. 35, no. 2, pp. 395–405, 2002.

[389] K. Blum, T. J. Chen, A. L. Chen, M. Madigan, B. W. Downs, R. L. Waite, E. R. Braverman, M. Kerner, A. Bowirrat, J. Giordano, *et al.*, "Do dopaminergic gene polymorphisms affect mesolimbic reward activation of music listening response? therapeutic impact on reward deficiency syndrome (rds)," *Medical hypotheses*, vol. 74, no. 3, pp. 513–520, 2010.

[390] B. Knutson, A. Westdorp, E. Kaiser, and D. Hommer, "Fmri visualization of brain activity during a monetary incentive delay task," *Neuroimage*, vol. 12, no. 1, pp. 20–27, 2000.

[391] K. A. Mathiak, M. Klasen, R. Weber, H. Ackermann, S. S. Shergill, and K. Mathiak, "Reward system and temporal pole contributions to affective evaluation during a first person shooter video game," *BMC neuroscience*, vol. 12, no. 1, pp. 1–11, 2011.

[392] H. Wang and C.-T. Sun, "Game reward systems: Gaming experiences and social meanings.," in *DiGRA conference*, Citeseer, vol. 114, 2011.

[393] F. F.-H. Nah, V. R. Telaprolu, S. Rallapalli, and P. R. Venkata, "Gamification of education using computer games," in *International Conference on Human Interface and the Management of Information*, Springer, 2013, pp. 99–107.

[394] P. Howard-Jones, S. Demetriou, R. Bogacz, J. H. Yoo, and U. Leonards, "Toward a science of learning games," *Mind, Brain, and Education*, vol. 5, no. 1, pp. 33–41, 2011.

[395] N. Hallford and J. Hallford, *Swords and circuitry: A designer's guide to computer role-playing games*. Premier Press, 2001.

[396] T. W. Malone, "Toward a theory of intrinsically motivating instruction," *Cognitive science*, vol. 5, no. 4, pp. 333–369, 1981.

[397] N. Yee, "Motivations for play in online games," *CyberPsychology & behavior*, vol. 9, no. 6, pp. 772–775, 2006.

[398] K. A. Kerns, S. Macoun, J. MacSween, J. Pei, and M. Hutchison, "Attention and working memory training: A feasibility study in children with neurodevelopmental disorders," *Applied Neuropsychology: Child*, vol. 6, no. 2, pp. 120–137, 2017.

[399] K. Khowaja and S. S. Salim, "Serious game for children with autism to learn vocabulary: An experimental evaluation," *International Journal of Human–Computer Interaction*, vol. 35, no. 1, pp. 1–26, 2019.

[400] R. Cuthbert, S. Turkay, and R. Brown, "The effects of customisation on player experiences and motivation in a virtual reality game," in *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 2019, pp. 221–232.

[401] S. Adinolf and S. Turkay, "Controlling your game controls: Interface and customization," in *Proceedings of the 7th international conference on Games+ Learning+ Society Conference*, 2011, pp. 13–22.

[402] [Online]. Available: `https://www.ea.com/en-gb/games/the-sims`.

[403] O. b. K. Townsend, *The best life simulation games of all time*. [Online]. Available: `https://www.ranker.com/list/the-best-life-simulation-games/kyle-townsend`.

[404] B. Entertainment, *Return to outland on june 1, 2021*. [Online]. Available: `https://worldofwarcraft.com/en-gb/`.

[405] [Online]. Available: `https://www.bungie.net/7/en/Destiny/NewLight`.

[406] S. Turkle, "Constructions and reconstructions of self in virtual reality: Playing in the muds," *Mind, Culture, and Activity*, vol. 1, no. 3, pp. 158–167, 1994.

[407] C. Klimmt, D. Hefner, and P. Vorderer, “The video game experience as “true” identification: A theory of enjoyable alterations of players’ self-perception,” *Communication theory*, vol. 19, no. 4, pp. 351–373, 2009.

[408] A. Alcorn, H. Pain, G. Rajendran, T. Smith, O. Lemon, K. Porayska-Pomsta, M. E. Foster, K. Avramides, C. Frauenberger, and S. Bernardini, “Social communication between virtual characters and children with autism,” in *international conference on artificial intelligence in education*, Springer, 2011, pp. 7–14.

[409] S. Tsikinas and S. Xinogalos, “Design guidelines for serious games targeted to people with autism,” in *Smart Education and e-Learning 2019*, Springer, 2019, pp. 489–499.

[410] D. Deriso, J. Susskind, L. Krieger, and M. Bartlett, “Emotion mirror: A novel intervention for autism based on real-time expression recognition,” in *European Conference on Computer Vision*, Springer, 2012, pp. 671–674.

[411] K. T. Ying, S. B. M. Sah, and M. H. L. Abdullah, “Personalised avatar on social stories and digital storytelling: Fostering positive behavioural skills for children with autism spectrum disorder,” in *2016 4th International Conference on User Science and Engineering (i-USEr)*, IEEE, 2016, pp. 253–258.

[412] S. Alves, A. Marques, C. Queirós, and V. Orvalho, “Lifeisgame prototype: A serious game about emotions for children with autism spectrum disorders.,” *PsychNology Journal*, vol. 11, no. 3, 2013.

[413] C. Lord, E. H. Cook, B. L. Leventhal, and D. G. Amaral, “Autism spectrum disorders,” in, 2, vol. 28, Elsevier, 2000, pp. 355–363.

[414] M. Chang, R. Kuo, C.-W. Lyu, and J.-S. Heh, “A situated game for autistic children learning activities of daily living,” in *2012 IEEE Fourth International Conference On Digital Game And Intelligent Toy Enhanced Learning*, IEEE, 2012, pp. 217–220.

[415] R. Rouse III, *Game design: Theory and practice.* Jones & Bartlett Publishers, 2004.

[416] C. Fabricatore, M. Nussbaum, and R. Rosas, “Playability in action videogames: A qualitative design model,” *Human-Computer Interaction*, vol. 17, no. 4, pp. 311–368, 2002.

[417] S. H. Hsu, F.-L. Lee, M.-C. Wu, *et al.*, “Designing action games for appealing to buyers,” *Cyberpsychology & behavior*, vol. 8, no. 6, pp. 585–591, 2005.

[418] J. Feil and M. Scattergood, *Beginning game level design.* Thomson Course Technology, 2005.

[419] R. Huskey, B. Craighead, M. B. Miller, and R. Weber, “Does intrinsic reward motivate cognitive control? a naturalistic-fmri study based on the synchronization theory of flow,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 18, no. 5, pp. 902–924, 2018.

[420] M. J. Habgood and S. E. Ainsworth, “Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games,” *The Journal of the Learning Sciences*, vol. 20, no. 2, pp. 169–206, 2011.

[421] K. Porayska-Pomsta, K. Anderson, S. Bernardini, K. Guldberg, T. Smith, L. Kossivaki, S. Hodgins, and I. Lowe, “Building an intelligent, authorable serious game for autistic children and their carers,” in *International Conference on Advances in Computer Entertainment Technology*, Springer, 2013, pp. 456–475.

[422] H. Qin, P.-L. P. Rau, and G. Salvendy, “Player immersion in the computer game narrative,” in *International Conference on Entertainment Computing*, Springer, 2007, pp. 458–461.

[423] H. Qin, P.-L. Patrick Rau, and G. Salvendy, “Measuring player immersion in the computer game narrative,” *Intl. Journal of Human–Computer Interaction*, vol. 25, no. 2, pp. 107–133, 2009.

[424] G. Christou, “The interplay between immersion and appeal in video games,” *Computers in human behavior*, vol. 32, pp. 92–100, 2014.

[425] K. Jørgensen, *Left in the dark: playing computer games with the sound turned off.* Routledge, 2017.

[426] J. R. Parker and J. Heerema, “Audio interaction in computer mediated games,” *International Journal of Computer Games Technology*, vol. 2008, 2008.

[427] C. Johanson and R. L. Mandryk, “Scaffolding player location awareness through audio cues in first-person shooters,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 3450–3461.

[428] L. E. Nacke and M. Grimshaw, “Player-game interaction through affective sound,” in *Game sound technology and player interaction: Concepts and developments*, IGI global, 2011, pp. 264–285.

[429] I. Ekman, "On the desire to not kill your players: Rethinking sound in pervasive and mixed reality games.," in *FDG*, 2013, pp. 142–149.

[430] A. Berndt and K. Hartmann, "The functions of music in interactive media," in *Joint International Conference on Interactive Digital Storytelling*, Springer, 2008, pp. 126–131.

[431] I. Ekman, "Psychologically motivated techniques for emotional sound in computer games," *Proceedings of AudioMostly*, pp. 20–26, 2008.

[432] S. Brown, M. J. Martinez, and L. M. Parsons, "Passive music listening spontaneously engages limbic and paralimbic systems," *Neuroreport*, vol. 15, no. 13, pp. 2033–2037, 2004.

[433] J. A. Winer, "Decoding the auditory corticofugal systems," *Hearing research*, vol. 207, no. 1-2, pp. 1–9, 2005.

[434] C. Gorbman, *Unheard melodies: Narrative film music*. Indiana University Press, 1987.

[435] A. Aljanaki, F. Wiering, and R. C. Veltkamp, "Studying emotion induced by music through a crowdsourcing game," *Information Processing & Management*, vol. 52, no. 1, pp. 115–128, 2016.

[436] S. Gormanley, "Audio immersion in games—a case study using an online game with background music and sound effects," *The Computer Games Journal*, vol. 2, no. 2, pp. 103–124, 2013.

[437] A. Chalmers and K. Debattista, "Level of realism for serious games," in *2009 Conference in Games and Virtual Worlds for Serious Applications*, IEEE, 2009, pp. 225–232.

[438] *Openvr (steamworks documentation)*. [Online]. Available: `https://partner.steamgames.com/doc/features/steamvr/openvr`.

[439] U. Feske and D. L. Chambless, "Cognitive behavioral versus exposure only treatment for social phobia: A meta-analysis," *Behavior therapy*, vol. 26, no. 4, pp. 695–720, 1995.

[440] V. R. Algazi and R. O. Duda, "Headphone-based spatial sound," *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 33–42, 2011.

[441] J. N. Ricciardi, J. K. Luiselli, and M. Camare, "Shaping approach responses as intervention for specific phobia in a child with autism," *Journal of Applied Behavior Analysis*, vol. 39, no. 4, pp. 445–448, 2006.

[442] J. D. Schmidt, J. K. Luiselli, H. Rue, and K. Whalley, "Graduated exposure and positive reinforcement to overcome setting and activity avoidance in an adolescent with autism," *Behavior modification*, vol. 37, no. 1, pp. 128–142, 2013.

[443] B. Grider, J. K. Luiselli, and W. Turcotte-Shamski, "Graduated exposure, positive reinforcement, and stimulus distraction in a compliance-with-blood-draw intervention for an adult with autism," *Clinical Case Studies*, vol. 11, no. 3, pp. 253–260, 2012.

[444] L. Carter, J. M. Harper, and J. K. Luiselli, "Dental desensitization for students with autism spectrum disorder through graduated exposure, reinforcement, and reinforcement-fading," *Journal of Developmental and Physical Disabilities*, vol. 31, no. 2, pp. 161–170, 2019.

[445] P. Eelen, D. Hermans, and F. Baeyens, "Learning perspectives on anxiety disorders," *An Introduction to Clinical Management and Research*, p. 249, 2001.

[446] M. C. Green and D. Murphy, "Eigenscape: A database of spatial acoustic scene recordings," *Applied Sciences*, vol. 7, no. 11, p. 1204, 2017.

[447] H. Daffern, D. Camlin, H Egermann, A. J. Gully, G Kearney, C Neale, and J. Rees-Jones, "Exploring the potential of virtual reality technology to investigate the health and well being benefits of group singing," *International journal of performance arts and digital media*, vol. 15, no. 1, pp. 1–22, 2019.

[448] *Freesound - audio database*, `https://freesound.org/home/`, Accessed: 2020-04-17.

[449] L. A. Ward and B. G. Shirley, "Personalization in object-based audio for accessibility: A review of advancements for hearing impaired listeners," *Journal of the Audio Engineering Society*, vol. 67, no. 7/8, pp. 584–597, 2019.

[450] B. Shirley, L. Ward, and E. Chourdakis, "Personalization of object-based audio for accessibility using narrative importance," CEUR Workshop Proceedings, 2019.

[451] Y. Tang, B. M. Fazenda, and T. J. Cox, "Automatic speech-to-background ratio selection to maintain speech intelligibility in broadcasts using an objective intelligibility metric," *Applied Sciences*, vol. 8, no. 1, p. 59, 2018.

[452] M. Vorländer, *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer Science & Business Media, 2007.

[453] M. Vorlander, "Auralization,"

[454] D. Rudrich, 2018. [Online]. Available: `https://plugins.iem.at/`.

[455] S. Poeschl, K. Wall, and N. Doering, "Integration of spatial sound in immersive virtual environments an experimental study on effects of spatial sound on presence," in *2013 IEEE Virtual Reality (VR)*, IEEE, 2013, pp. 129–130.

[456] C. Hendrix and W. Barfield, "The sense of presence within auditory virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 5, no. 3, pp. 290–301, 1996.

[457] S. K. Tatla, "The development of the pediatric motivation scale for children in rehabilitation: A pilot study," PhD thesis, University of British Columbia, 2014.

[458] C. A. Roberto, J. Baik, J. L. Harris, and K. D. Brownell, "Influence of licensed characters on children's taste and snack preferences," *Pediatrics*, vol. 126, no. 1, pp. 88–93, 2010.

[459] G. Salvador-Herranz, D. Perez-Lopez, M. Ortega, E. Soto, M. Alcaniz, and M. Contero, "Manipulating virtual objects with your hands: A case study on applying desktop augmented reality at the primary school," in *2013 46th Hawaii International Conference on System Sciences*, IEEE, 2013, pp. 31–39.

[460] L. Nijs and M. Leman, "Interactive technologies in the instrumental music classroom: A longitudinal study with the music paint machine," *Computers & Education*, vol. 73, pp. 40–59, 2014.

[461] J. C. Read, S. MacFarlane, and C. Casey, "Endurability, engagement and expectations: Measuring children's fun," in *Interaction design and children*, Shaker Publishing Eindhoven, vol. 2, 2002, pp. 1–23.

[462] D. E. Hall, "Musical acoustics," in. Brooks/Cole Pub. Co, 2002, pp. 54–340.

[463] A. Adjorlu, N. B. B. Barriga, and S. Serafin, "Virtual reality music intervention to reduce social anxiety in adolescents diagnosed with autism spectrum disorder," in *16th Sound and music computing conference*, 2019.

[464] L. Reynolds-Keefer, R. Johnson, T. Dickenson, and L. McFadden, “Validity issues in the use of pictorial likert scales,” *Studies in Learning, Evaluation, Innovation and Development*, vol. 6, no. 3, pp. 15–24, 2009.

[465] L. Hall, C. Hume, and S. Tazzyman, “Five degrees of happiness: Effective smiley face likert scales for evaluating with children,” in *Proceedings of the The 15th International Conference on Interaction Design and Children*, 2016, pp. 311–321.

[466] M. Simões, M. Bernardes, F. Barros, and M. Castelo-Branco, “Virtual travel training for autism spectrum disorder: Proof-of-concept interventional study,” *JMIR serious games*, vol. 6, no. 1, e5, 2018.

[467] M. Poyade, G. Morris, I. Taylor, and V. Portela, “Using mobile virtual reality to empower people with hidden disabilities to overcome their barriers,” in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017, pp. 504–505.

[468] M. Krijn, P. M. Emmelkamp, R. P. Olafsson, and R. Biemond, “Virtual reality exposure therapy of anxiety disorders: A review,” *Clinical psychology review*, vol. 24, no. 3, pp. 259–281, 2004.

[469] M. H. Kuo, G. I. Orsmond, E. S. Cohn, and W. J. Coster, “Friendship characteristics and activity patterns of adolescents with an autism spectrum disorder,” *Autism*, vol. 17, no. 4, pp. 481–500, 2013.

[470] G. I. Orsmond and H.-Y. Kuo, “The daily lives of adolescents with an autism spectrum disorder: Discretionary time use and activity partners,” *Autism*, vol. 15, no. 5, pp. 579–599, 2011.

[471] K.-H. Li, S.-J. Lou, H.-Y. Tsai, and R.-C. Shih, “The effects of applying game-based learning to webcam motion sensor games for autistic students’ sensory integration training.,” *Turkish Online Journal of Educational Technology-TOJET*, vol. 11, no. 4, pp. 451–459, 2012.

[472] L. M. Almeida, D. P. d. Silva, D. P. Theodório, W. W. Silva, S. C. M. Rodrigues, T. A. Scardovelli, A. P. d. Silva, and M. A. S. Bissaco, “Altriras: A computer game for training children with autism spectrum disorder in the recognition of basic emotions,” *International Journal of Computer Games Technology*, vol. 2019, 2019.

[473] S. D. Mayes, S. L. Calhoun, R. Aggarwal, C. Baker, S. Mathapati, S. Molitoris, and R. D. Mayes, "Unusual fears in children with autism," *Research in Autism Spectrum Disorders*, vol. 7, no. 1, pp. 151–158, 2013.

[474] D. E. Johnson, "Crossover experiments," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 5, pp. 620–625, 2010.

[475] E. J. Mills, A.-W. Chan, P. Wu, A. Vail, G. H. Guyatt, and D. G. Altman, "Design, analysis, and presentation of crossover trials," *Trials*, vol. 10, no. 1, p. 27, 2009.

[476] *Spss software*. [Online]. Available: `https://www.ibm.com/analytics/spss-statistics-software?lnk=STW_US_STESCH`.

[477] N. M. Razali, Y. B. Wah, *et al.*, "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests," *Journal of statistical modeling and analytics*, vol. 2, no. 1, pp. 21–33, 2011.

[478] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.

[479] J. Arnau, R. Bono, M. J. Blanca, and R. Bendayan, "Using the linear mixed model to analyze nonnormal data distributions in longitudinal designs," *Behavior research methods*, vol. 44, no. 4, pp. 1224–1238, 2012.

[480] P. M. Emmelkamp, "Technological innovations in clinical assessment and psychotherapy," *Psychotherapy and psychosomatics*, vol. 74, no. 6, pp. 336–343, 2005.

[481] J. Low, E. Goddard, and J. Melser, "Generativity and imagination in autism spectrum disorder: Evidence from individual differences in children's impossible entity drawings," *British Journal of Developmental Psychology*, vol. 27, no. 2, pp. 425–444, 2009.

[482] S. E. Lind, D. M. Williams, D. M. Bowler, and A. Peel, "Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection?" *Neuropsychology*, vol. 28, no. 1, p. 55, 2014.

[483] L. A. Wijnhoven, D. H. Creemers, A. A. Vermulst, R. J. Lindauer, R. Otten, R. C. Engels, and I. Granic, "Effects of the video game 'mindlight'on anxiety of children with an autism spectrum disorder: A randomized controlled trial," *Journal of Behavior Therapy and Experimental Psychiatry*, p. 101 548, 2020.

[484] J. O. Berry and W. H. Jones, "The parental stress scale: Initial psychometric evidence," *Journal of Social and Personal Relationships*, vol. 12, no. 3, pp. 463–472, 1995.

[485] A. Kushki, E. Drumm, M. P. Mobarak, N. Tanel, A. Dupuis, T. Chau, and E. Anagnostou, "Investigating the autonomic nervous system response to anxiety in children with autism spectrum disorders," *PLoS one*, vol. 8, no. 4, e59730, 2013.

[486] S. Kuriakose and U. Lahiri, "Understanding the psycho-physiological implications of interaction with a virtual reality-based system in adolescents with autism: A feasibility study," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 665–675, 2015.

[487] L. Rosenblum, "Virtual and augmented reality 2020," *IEEE Computer Graphics and Applications*, vol. 20, no. 1, pp. 38–39, 2000.

[488] *Magic leap: Reality is just beginning.* [Online]. Available: `https://www.magicleap.com/en-us`.

[489] *Microsoft hololens: Mixed reality technology for business.* [Online]. Available: `https://www.microsoft.com/en-us/hololens`.

[490] S. Rocha and A. Lopes, "Navigation based application with augmented reality and accessibility," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–9.

[491] R. O. Kellems, C. Charlton, K. S. Kversøy, and M. Győri, "Exploring the use of virtual characters (avatars), live animation, and augmented reality to teach social skills to individuals with autism," *Multimodal Technologies and Interaction*, vol. 4, no. 3, p. 48, 2020.

[492] Y.-C. Huang and I.-J. Lee, "Using augmented reality and concept mapping to improve ability to master social relationships and social reciprocity for children with autism spectrum disorder," in *International Conference on Human-Computer Interaction*, Springer, 2020, pp. 19–37.

[493] M. M. Soares and A. d. S. O. Neves, "Augmented reality and the use of alternative communication for children with autism spectrum disorder: A literature review," in *International Conference on Applied Human Factors and Ergonomics*, Springer, 2020, pp. 324–330.