Keeping your ears on the virtual ground: Expert insights on virtual reality audio techniques and challenges

Jacob M. Hedges
Faculty of Arts and Social Sciences, University of Technology Sydney, Australia
Jacob.M.Hedges@student.uts.edu.au

Robert Sazdov
Faculty of Arts and Social Sciences, University of Technology Sydney, Australia
Robert.Sazdov@.uts.edu.au

Andrew Johnston
Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
Andrew.Johnston@.uts.edu.au

Abstract

Virtual reality (VR) audio practitioners navigate interdisciplinary spatial design decisions alongside complex interactive audio systems, resulting in novel solutions to creating immersive experiences in VR environments. Previous work has attempted to uncover these unique practitioner challenges through case studies, however direct insights from a range of VR audio professionals may provide a new understanding of how these challenges are being solved. This study employs semi-structured interviews with twelve participants identified as leading practitioners in the field to reveal several key findings on VR audio creative practice, with a focus on Head Mounted Display (HMD) content. The results aim to serve as a valuable resource for audio professionals and researchers to understand the current challenges faced by leading experts in the field. Key findings highlight the necessity for flexible design approaches due to the diverse range of VR applications, the importance of strategic audio placement and scale for immersion, and the need for highly dynamic sound object behavior to create realistic soundscapes. The difficulty of mixing audio in VR and the limitations of current tools emphasize the need for integrated and comprehensive solutions, as well as efficient auralization tools for enhancing the quality and immersive potential of VR audio production.

CCS Concepts

• **Human-centered computing** → Human computer interaction (HCI); Interaction techniques; Auditory feedback; Interaction design; Interaction design theory, concepts and paradigms; Human computer interaction (HCI); Interaction paradigms; Virtual reality; Human computer interaction (HCI); Interaction devices; Sound-based input / output.

Keywords

Virtual Reality, Sound Design, Audio, Creative Practice, Immersion, Presence

ACM Reference Format:
Jacob M. Hedges, Robert Sazdov, and Andrew Johnston. 2024. Keeping your ears on the virtual ground: Expert insights on virtual reality audio techniques and challenges. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3678299.3678337

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678337

1 INTRODUCTION

The advancement of Virtual Reality (VR) continues to present novel opportunities and challenges for audio designers, significantly expanding the landscape of audio creative practice. Well established audio practices borrowed and adapted from established media such as films and games have shaped the foundation of VR audio design. However, the unique capabilities of VR Head Mounted Display (HMD) technology invites a reevaluation of existing design strategies, fostering the potential for new audio techniques that are only just beginning to be more deeply explored within academic research [1, 2]. VR audio requires a deep understanding of spatial design decisions that impact audio processes and communication, emphasizing the interdisciplinary nature of VR system design [3]. Additionally, the challenge of accurate sound localization is crucial for believability in VR, highlighting the commitment needed to advance audio quality in VR and other immersive technologies [4]. Real-time auralization in VR must meet stringent requirements for head tracked rendering while supporting dynamic sound sources within limited computational budgets, showcasing the complex design challenges present in VR audio production [5]. This study aims to contribute to and extend the current discourse regarding the emerging sonic culture of VR audio practice through interviews conducted with leading VR audio expert practitioners. As audio practice can vary dramatically across VR, Augmented Reality (AR) and Mixed Reality (MR) fields, this study focuses solely on VR audio practice for HMD.

2 BACKGROUND AND RELATED WORKS

2.1 Unique audio design considerations for VR

The spectrum of complexity within VR experiences begins with 3 Degrees of Freedom (3DoF) content, which allows for rotational viewpoint navigation but lacks translational movement. Experiences that utilize these capabilities in conjunction with linear media are often referred to as Cinematic VR or 360 Videos [6]. On the other hand, 6 Degrees of Freedom (6DoF) VR enables users to move both rotationally and translationally within the virtual environment, resulting in the need for more complex dynamic spatial audio design strategies. VR is arguably the most relevant medium for the application of spatial audio as a functional technology [3]. This is due to the fact that our perception of the VR experience is significantly influenced by our physical orientation, particularly the orientation of our head and, consequently, our ears. Most traditional media

forms are not designed to be dependent on this point of view, and so VR offers a unique advantage to spatial design considerations due to its egocentric nature [7]. This forces a shift in focus in the application of audio design strategies towards enhancing movement, transparency, and engagement within virtual environments.

2.2 Designing for immersion

Design decisions in VR audio also shift in focus towards impacting user immersion by enhancing the sense of presence and emotional engagement within virtual environments. Immersion in VR is a critical yet ambiguously defined concept, especially concerning the role of audio in influencing immersive experiences [8, 9]. Research highlights that spatial audio attributes and continuous background auditory information are essential for replicating real-world scenes, while congruent sensory feedback and high-quality sound further bolster the sense of presence [10]. Alongside presence, another key factor of immersive experiences is involvement, which is defined as being cognitively absorbed in an activity. This facet of immersive experience is especially important in the context of digital game user experience [11]. Research on how audio impacts involvement in the context of virtual environments points to the importance of music's emotional impact, musical literacy, and interactive responsiveness [12]. Despite these insights, further research is needed to understand how sound specifically influences involvement in VR, especially in the context of VR audio practice.

2.3 Tool and workflow challenges in VR audio

The current landscape of VR audio production tools and workflows exhibits several limitations that impact creators across various stages of production. One primary concern is the underdevelopment of user interfaces for authoring 3D audio parameters, which forces users to resort to ad hoc solutions or external programming languages, highlighting a gap in functionality and workflow integration within existing tools [13, 14]. This is further complicated by the growing community of VR content creators who require immersive audio authoring tools that cater to both expert and novice users, indicating a need for more inclusive and versatile design considerations [15]. The process of editing 3D content for VR experiences is still predominantly conducted in traditional 2D desktop applications, which do not offer intuitive capabilities for previewing content in VR headsets or manipulating audio parameters directly, thereby interrupting the creative process.

2.4 Related studies

Whilst this is not a completely new field in research, there are still only a small number of studies that have helped shape an understanding of VR audio practice from a research perspective. A recent study involving interviews with audio practitioners in immersive media experiences (IMEs) revealed the importance of spatial audio, user interaction, narrative quality, and visual content in achieving immersion [1]. Key challenges identified include the difficulty in creating auditory distance and the necessity for high-quality, multisensory stimuli to enhance immersion. Another study analyzed XR audio case studies, leading to the identification of challenges and strategies across the core areas of technologies, tools, techniques, and perception (3TP) [2]. The findings emphasize the role of Ambisonics, binaural synthesis, and reverberation in spatial audio production for XR. The rapid obsolescence of tools and the need for bespoke software configurations are noted as ongoing issues, while advanced techniques such as dynamic head-tracking and 360° audio integration demonstrate experiential benefits. Both studies point to the critical need for innovation in tools and workflows, as well as a deeper understanding of perceptual factors, to overcome current limitations and enhance the immersive quality of VR audio experiences. These findings provide a comprehensive backdrop for examining the novel solutions and strategies employed by leading VR audio practitioners in navigating the challenges of immersive audio production.

Considering the identified knowledge gaps and challenges presented in this section, the research questions presented below were formulated. These were also used to guide the development of the semi-structured interview questions for the current study, which can be found in Appendix A. The goal of these questions is to both identify new strategies and challenges, as well as to further validate those found in related studies, in the hope of developing a deeper understanding of the emerging sonic culture of VR audio practice.

- How do the production challenges and techniques in VR audio differ fundamentally from those in conventional media, and what unique solutions could potentially address these differences?
- In what ways do the design decisions in VR audio influence user immersion?
- What are the current trends and limitations in the tools and workflows used for VR audio production?

3 METHODOLOGY

The purpose of this study is to identify current design strategies and challenges that are specific to the sonic culture of VR audio creative practice. As mentioned in the related studies section, a similar research approach was taken by a recent study, which explores spatial audio design challenges in IMEs, which were classified as a more general term for XR [1]. This exploratory study involved five participants being interviewed with a further three completing a survey. The goal of the current study is to complement and further expand on these findings by interviewing a larger number of participants and focusing specifically on VR HMD audio practice. To achieve this, we conducted one-on-one semi-structured interviews with VR audio expert practitioners to gain deep and relevant insights into these research areas. Semi-structured interviews are a qualitative research method where interviewers engage with participants in a conversation-like manner to gather insights into their motivations and attitudes towards the research topic [16]. These interviews blend pre-determined, structured questions with the flexibility to explore new topics that arise during the conversation, allowing for a comprehensive understanding of the subject matter.

3.1 Participant selection

Participants were selected using a non-probabilistic sampling approach. Research was conducted to identify potential leading experts in VR audio through publications, online communities and professional networks. The selection criteria of 'leading expert' was defined as having a significant and notable body of professional VR

Table 1: Thematic framework for strategies and challenges in VR audio practice

Themes	Definition	Sub themes
Design context and objectives	How the approach to VR audio design changes conceptually compared to traditional media and across different contexts of VR	Similarities with traditional media Content Dependence Perspective based considerations
Audio techniques for VR	Specific audio techniques and strategies described by participants that are crucial for effective audio design in VR environments	Audio guided focus Importance of scale Dynamic sound emitter techniques Environmental and interaction detail Spatial audio and psychoacoustic considerations
Designing audio for immersion	Audio approaches that are aimed towards improving the immersive experience as a key goal of VR	Designing for immersion, realism and plausibility Multimodal congruency and plausible reverb
Workflow and technological challenges	Tool and workflow challenges identified by experts as issues in their creative practice	Preferred workflow approaches Workflow challenges Limitations of current tools Indispensable and missing tools

practical works in the field, as well as through other contributions such as technological developments and publications. A total of 14 interviews were conducted, however two participants were later excluded from the study due to a lack in practitioner experience to meet the selection criteria of 'leading expert'. This resulted in 12 participants being included in the analysis, which is sufficient considering the relatively small diversity of the population being studied [17]. These participants' professional experience in VR audio ranged from five to 20 years and included roles in prominent VR production studios (both indie and AAA), leading VR related tech companies, and research institutions. Furthermore, additional focus was given to including a range of participants with specific expertise across the spectrum of technical and creative experience, as well as including experts in both 3DoF cinematic and 6DoF interactive VR. Quotes from participants in this paper have been anonymized.

3.2 Data analysis

This study employed thematic analysis to systematically analyze interview data through an iterative process that included initial data familiarization, coding, and ultimately identifying and reviewing core themes related to the research focus [18, 19]. After transcribing the interview recordings and reading several times to become familiar with the data, the analysis began using the NVIVO 12 qualitative analysis software. Firstly, detailed and descriptive categories (known as 'codes') were generated for the data to organize them into manageable segments. This process was iterated to develop more focused codes in order to refine the relationships within the data. The subsequent phase involved searching for and defining broader patterns of meaning or themes from these codes, assessing their relevance to ensure they accurately represented the dataset. The final stages involved critically reviewing these themes, including merging, splitting, or discarding them based on their relevance to the data, and then defining and naming the themes to clearly articulate their significance in understanding current design strategies and challenges that are specific to VR audio creative practice.

4 RESULTS AND DISCUSSION

During the thematic analysis process, 1230 initial raw codes were identified through a line-by-line coding process. These were then refined into 26 focused codes, including one titled 'Miscellaneous' for any individual lines that discussed points the researchers considered irrelevant to the study. These focused codes were analyzed and grouped into the final thematic framework illustrated in Figure 1 and further detailed in Table 1 below. The following sections discuss these themes along with key findings and excerpts from the interviews identified by the researchers as relevant to the study goal.

4.1 Theme 1 – Design context and objectives

4.1.1 Similarities with traditional media. This sub theme explores the ways that VR audio practitioners adapt their design approach compared to traditional media. As VR creative practice borrows many of its audio design approaches from these established paradigms, such as screen-based games and films, certain design strategies such as the focus of supporting the narrative remain a key objective.

> *"What I'm trying to do is I'm trying to bring the world to life. My primary task as an audio designer is to support the narrative as far as I'm concerned." (P1)*

This sentiment is echoed through P3, who explains "You just have to get good sounds and understand how to craft stuff well and do your best to make it fit the genre or the style". This sentiment suggests that the creative approach of VR and traditional media in regard to audio production are the same at a foundational level, which is to create and edit audio content that complements

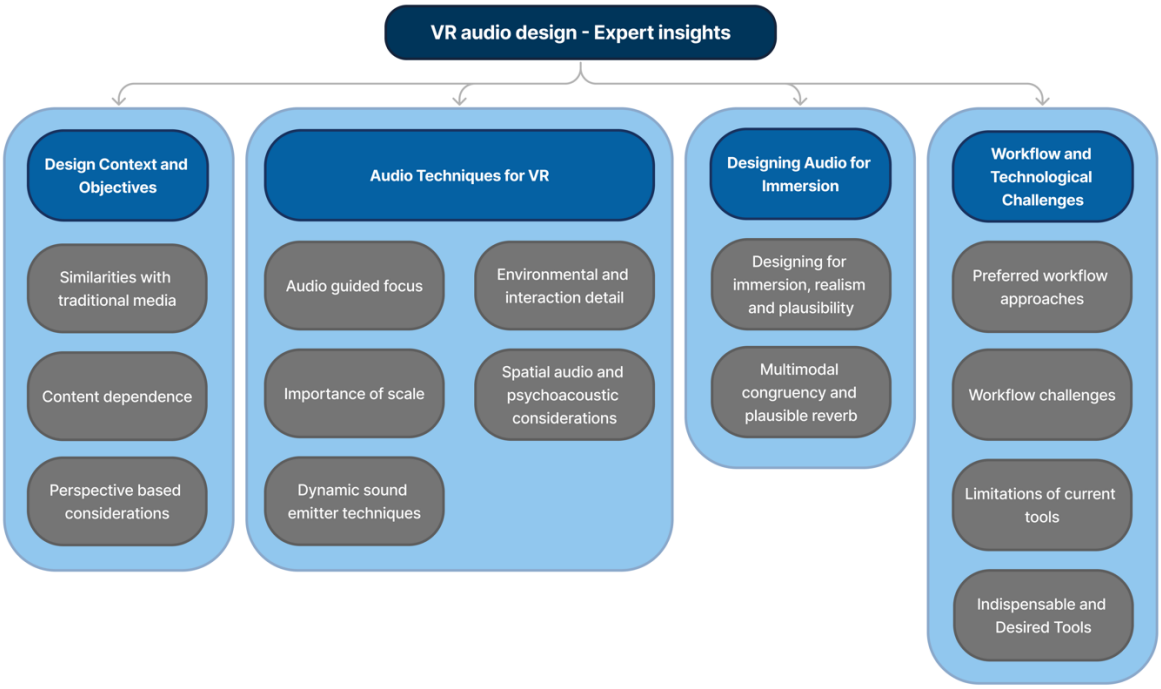

Figure 1: Illustration of overall thematic framework identified from VR audio expert insights.

the style and narrative of the overall experience. Alongside these philosophies, the similarities for VR audio production approaches to those specifically of game audio production were highlighted by several experts. For example, P7 states "Every game that we work on has its challenges. Every game that we work on has different perspectives, different aesthetics, different ways that players interact. So [VR experiences are] just another game", which points to the similar challenges faced across both mediums. Adding to this, P8 identifies that "A VR gaming project is similar, very similar to a traditional gaming project in the early stages". This observation that similarities exist in the 'early stages' suggests that the need for specific audio design considerations for VR begin later in the development process. These examples highlight the importance of having foundational knowledge in audio production techniques for traditional media in approaching VR audio, further verifying findings from similar research [1, 2].

4.1.2 Content dependence. Another key identification to be made in terms of the design context for VR audio is how practices and challenges may differ across the different applications of VR. This became an important distinction to make for many experts, as design approaches can change dramatically between linear and nonlinear contexts.

> "*In my experience, mixing for linear 3DoF involves a set of tools and a pipeline that are essentially an extension of traditional post-production mixing, while 6DoF operates within a completely different paradigm.*" (P5)

This distinction between certain design contexts is also a feature within sub contexts of traditional media, as P5 further highlights "*Even in conventional media, mixing for theatrical or TV involves a slightly different approach*". On a more philosophical level, P3 discusses how their adaptive approach to game and VR audio design guides their process "*Understanding how something's going to work informs how I need to build the content. That's been my method of building*". Whilst this points to a universal approach independent of context, the specific techniques required across different VR applications can still be very different, as P8 explains "*Depending on the product you're making, whether it could be a game or a VR cinematic experience, the approach varies significantly*".

These findings point to the fact that VR audio designers need to be flexible in their approach due to the complex range of experiences that the potential of VR technologies can create. An example from [4] serves to exemplify this sentiment using the example of a car sound across different media contexts. For instance, in conventional media like movies, a sound asset such as a car passing from left to right can be a single recording panned to match the screen movement. However, this expands in 3DoF VR, as the audience's ability to face their head (the camera) in any orientation towards the car requires a much more complex sound model to support these interactions. P12 mentions this specifically: "*With traditional media like 16:9, you always have the onscreen and offscreen sounds right? Except for that, VR does not make a difference with onscreen or offscreen. It's always onscreen*". Expanding this to 6DoF VR, the audience can walk around, lean close to, kneel in front of it or even stop the car, and VR HMD allows the users physical movements to control this. This level of detail and interactivity demands a more sophisticated approach to sound design, far beyond what is needed for linear or even basic interactive media [4].

4.1.3 Perspective based considerations. Perhaps the most crucial difference in designing audio for VR compared to traditional media is that the perspective of the virtual content is controlled by the user's physical movements, particularly that of the head [7]. The challenge of designing audio for this complex presentation is conceptually described by one participant:

> "*How do we convey that? And so the difficulty with VR is all of a sudden we've got this space that we are now in. We're not on the side watching other people in the swimming pool. We've now jumped into the swimming pool. But how do we communicate that we can go underwater and interact underwater?*" (P1)

This shift into fully dynamic interaction with a virtual environment fundamentally changes how audio must be approached, with a shift towards an egocentric design philosophy. P3 elaborates on the increased complexity of designing audio for multiple perspectives in certain VR interactive experiences. This complexity arises particularly in scenarios where the user controls a third-person character and views the experience from the 'cameras' perspective which is controlled by the HMD, like looking into a diorama: "*So there's content perspective shifts that need to happen sometimes, and we're constantly thinking about whether I am trying to tell a story from their perspective right now, or your perspective*". The change in perception granted by controlling the view of a virtual environment with head movements is explored further by P7, who states "*When you've got the headset on, you just sense things differently. The sense of depth is different*". This sentiment was shared in findings from a similar study on immersive audio practice in [1], and further aligns with research by [20] that points to the critical role of head movements in providing additional cues for evaluating spatial audio attributes, which may influence this altered 'sense of depth'.

The transition from a passive observer to an active participant within the virtual environment complicates the audio design process, demanding more dynamic and context-sensitive solutions. Furthermore, the loss of a central focus point in VR introduces additional challenges:

> "*The opportunity for variance and complexity increases in interactive audio because the content is often much less deterministic whether the end user is in control of the camera or not. You inherit all the linear problems and more because the content itself is changing dynamically in real time. The underlying challenge is not different than Live Events or Game Audio in that way, but the end user environment evolved into wearable technology that is simultaneously the camera, the viewer, and the listener in motion.*" (P9)

This sentiment encapsulates the layers of complexity in audio design approaches that compound due to the unpredictable camera focus in VR HMD environments. Multiple participants echoed this challenge, with P4 stating that "*The main difference [from traditional media] is, you are not supposed to watch towards the front*", and P11 confirming that "*When you go into VR, you don't have any center*". Finally, P12 points towards the creative freedom that this new way of interacting with media allows for both the audience and audio designers: "*And now we let the audience decide what they want to see in that sphere. And we, sound designers, are now in the lead*". These insights suggest that the challenge here lies in guiding the user's auditory experience without a predefined focal point, fundamentally altering the approach to audio design in VR compared to traditional media. This challenge is also shared with screenwriters in the VR medium [21], where the term 'Immersography' has been coined to describe the new ways of crafting narratives that are intrinsically linked to the user's spatial and interactive experiences. As the culture of VR audio practice grows alongside this, a more sophisticated and thoughtful approach is needed to ensure that the sound environment is always coherent and responsive to the user's actions.

4.2 Theme 2 - Audio Techniques for VR

4.2.1 Audio guided focus. In terms of audio techniques that evolve beyond their value from traditional media to VR, perhaps the most essential is guiding the user's spatial orientation through sound cues. This is often achieved by carefully tuning audio elements to draw the user's attention to important visual elements. This technique ensures a congruence between audio and visual elements, which is a key factor towards creating a sense of immersion in virtual environments [10].

> "*So, if you do a VR cinematic experience, or even in games, you need the audio to convey information subtly. For instance, a beeping light or a flashing object in the environment might be crucial for progressing the narrative.*" (P8)

Whilst the approach of guiding the user's attention is not new as P8 mentions above, in VR it has become a crucial tool in conveying key narrative points without a central focus point, as we can hear things in the virtual environment that are outside of our field of view. P11 further confirms that "*This is basically the most important thing that audio needs to do in VR. To give focus on things*". Extending this, P3 highlights how this technique can also be implemented subtly in an experience by taking advantage of the full 3D space allowed in VR "*We would really tune that stuff using the elevation so that if you were looking at it, it just put more focus on the thing you were looking at*". These insights collectively underscore the importance of strategic audio placement to guide the user in VR and its ability to enhance user engagement and narrative progression.

4.2.2 Importance of scale. Moving beyond traditional media techniques, a unique VR audio design approach revealed in the analysis was the importance of representing physical scale and size, as it can significantly impact the user's sense of immersion and physical presence. By taking advantage of VR's unique affordance of sensorimotor contingencies (i.e., that we are placed within a virtual world with our real body controlling our virtual spatial orientation [22]), we can further emphasize our perception of scale through strategic audio techniques.

This sentiment was echoed by multiple participants. P1 pointed out that "*Scale is the most important thing*" when asked what the key difference is between producing audio for VR compared to traditional media. P3 supported this point by explaining that when designing the audio, "*Everything needs to feel larger, more lush and rich, to give you that sensation of being small*". P3 went on further to discuss the details of this process: "*Sometimes we build things at a larger scale, so maybe they sound bigger with more detail*

than you would typically expect to hear from your own perspective". This hyper-attention to emphasizing scale emerges as a key focal point in audio techniques specific to VR. This is likely due to the heightened awareness of bodily movements that control the virtual environment - leading to an increase sense of physical presence in the virtual environment, which is a key factor that can influence the immersive experience [8]. This finding points to a focus on creating 'larger than life' sounds as opposed to aiming for replicating physical reality to take advantage of the unique opportunity of an egocentric perspective in VR HMD experiences.

4.2.3 Dynamic sound emitter techniques. Another novel audio technique for VR that emerged in the findings was the importance of placing sound emitters around the VR camera that are locked to the camera position but not to its rotation. This approach essentially allows you to create a 3DoF audio environment that follows the user within a 6DoF virtual space, which can be highly effective for producing realistic and dynamic ambiences that respond to user movements.

> "*So what I do is I set up an array. We have an emitter there, an emitter there, an emitter there, and an emitter there. Those emitters are each playing mono rain. They track where the camera is. They move with the camera.*" (P1)

Different applications of this approach were outlined across the participant, for instance P3 their approach: "*We opted to create height-based point-sourced quad emitters, three sets of them based on elevation*" which points out that these emitters can take advantage of the full 3D virtual space, essentially mimicking real-world loudspeaker setups with multiple elevation layers. The impact that this dynamic interaction between sound sources and the virtual space can have compared to simply head-locked audio was highlighted by P11 "*But as soon as you detach your audio, and so the audio is no longer connected to your head, it's now connected to the scene. So when you look around, you can hear that the sound is moving which completely changes your sense of presence*". Collectively, these techniques illustrate the intricate methods VR audio designers employ to create immersive and realistic auditory experiences, ensuring that soundscapes are coherent and dynamically responsive to user movements.

4.2.4 Environmental and interaction detail. The level of environmental detail and strategic layering was also identified as a key approach for best practice. The importance of treating environments as characters was emphasized, with this personification of environments helping to craft a dynamic and engaging auditory experience that resonates with the user on a deeper level.

> "*I try to do that with the environments that I make. Like an impressionist painting, I think, is that the technique where they do those like kind of thick, broad distances, it looks like an impression. But when you zoom in you can actually see every single one.*" (P1)

The example above comparing VR soundscape design to impressionist painting alludes to the fact that VR allows you to have a dynamic relationship with the sonic environment that is not as prevalent with traditional media. When the user is focused on the 'big picture' in VR, they hear the soundscape as a homogeneous feature. However, when they pay closer attention, they are presented with a more nuanced sonic experience. P8 confirms this by stating that "*Those small details are what the listener, the user, the person experiencing it, might not consciously recognize, but those are the details that completely sell the illusion*". This idea of the importance of detail is also prevalent in considering interactive sound elements that respond to user actions, which can add layers of realism to the VR experience. The analysis highlighted the complexity of designing interactive audio that accurately reflects the diversity of possible user actions.

> "*In real life, if they're holding this virtual sword in their real hands, there's no way for them to swing it the same way every time, there's absolutely no way. So there's an infinite number of ways that person can swing a sword, and you must create a sound for all of them.*" (P10)

This feedback loop between user actions and auditory responses is crucial for maintaining immersion in VR. In traditional video games, sound designers work with a finite set of sounds tied to specific animations, triggered by button presses. In VR, however, players interact directly with virtual objects, such as swinging a sword using a controller. This direct interaction results in an infinite variety of movements, making it challenging for sound designers to create adaptive audio that responds accurately to the nuances of each unique action. The goal is to ensure that every movement feels natural and enhances the sense of presence in the virtual world, as P10 further elaborates ""*There has to be some interaction, and sound can help give that feeling that they are there with the feedback and all the stuff*". The insights discussed in this subtheme reveal the multifaceted approach required in VR audio design, where both environmental and interaction details play crucial roles in crafting a convincing and engaging virtual experience.

4.2.5 Spatial audio and psychoacoustic considerations. Designing spatial audio for VR presents unique challenges and demands meticulous attention to detail. Due to the reliance on sensorimotor contingencies for VR to effectively translate perceptual information in a virtual environment, audio practitioners need to understand how spatial audio and psychoacoustic technologies can be leveraged to enhance the experience, without compromising the overall audio quality or emotional intention.

> "*Crafting and delivering a sound with intentional emotional power that suspends our belief and excites the brain is a delicate art. Our brains are highly sensitive and constantly filtering stimuli in our environment. Between the content creation, software rendering, and hardware rendering, a lot can go wrong to skew information and break that immersion. HRTF's and headphones are difficult to deal with because they introduce so much variance in the creation and delivery system that ultimately affects the way the sound is perceived.*" (P9)

This quote points to the issue that introducing too many spatial audio technologies can cause detrimental effects to the experience, and how understanding how the final delivery will sound is crucial. Speaking of specific approaches, P3 pointed out how "*You can't fake*

it in VR; everything really needs to be point-sourced in the engine" and that using technology such as HRTF can be flawed, as sometimes you do not want to mimic real-world environments "*And so for us, we just found it kind of homogenized things. We lost detail. We lost some of the richness, so we explored other ways to try and give you that same sensation of elevation and spatialization*" (P3).

This challenge was echoed by multiple participants, with P1 confirming "*We are inherently flawed. We are not good at triangulating. And so how the hell do I compensate for that in a video game where it's actually a natural human weakness?*". P1 continued this point to identify a specific audio technique to deal with this "*So if literally, whether it's a jet fighter or whatever, I am somehow going to make anything that needs to feel like it's moving to be pulsing. So if you've got this thing going, 'woo, woo,' it helps us triangulate a lot better. So it gives us the sense of movement*". This phenomenon is backed by research in [23] which shows that 'pulsing', or amplitude modulation (AM), significantly enhances human sound localization as the auditory system utilizes interaural time differences in the AM waveform to effectively determine the location of the sound source, particularly at higher frequencies where pure tones alone are less effective due to phase ambiguity.

Furthermore, whilst some participants championed the use of psychoacoustic based reproduction systems "*Sound localization basically can be improved a lot using binaural technologies.*" (P11), others pointed out how these can have different results depending on their implementation "*One big issue is the HRTF filters. If they are not personalized, they will work for some, but not for others.*" (P12). This points to a contention within the VR audio practice community on the effective implementation of these technologies. It is crucial to integrate these technologies carefully to ensure the final delivery leverages them to enhance the overall experience. These considerations underline the complex interplay between spatial audio accuracy and psychoacoustic realism in VR audio design.

4.3 Theme 3 - Designing Audio for Immersion

4.3.1 Designing for immersion, realism and plausibility. The focus on designing towards immersion as a metric of user experience in VR audio involves not just recreating reality but also manipulating perceptions to create an engaging and believable experience. Whilst immersion is an important factor in traditional media, due to the nature of HMD content completely blocking out your senses from the real world, sustaining immersion becomes a crucial design consideration.

> *"When I think of immersion in audio in VR, I think of something that has at least some relation to the way that sound sources behave in space. I have to say 'some' relation because you can warp that and still get the same effect."* (P2)

This quote aligns with research by [24, 25] which suggests that sounds do not need to be realistic or possible in the real world; instead, their plausibility and alignment with the user's expectations are crucial for creating a sense of presence. This is echoed by several participants, with P12 stating "*I'm trying to convince my audience of being in that real or virtual environment. So the sounds have to be convincing, but they also have to fit in the expectations of the audience.*", and P1 further elaborating "*You have to play with their perceptions. Not trying to recreate reality. We are playing with something that's not hard and fast. We're playing with perceptions*". Whilst some participants pointed to realism being a key factor in VR audio "*It's much more critical to not break the immersive environment. Things have to be realistic.*" (P8), others highlighted how the concept of immersion begins to blur when considering Mixed Reality sonic environments *"But what we do want to do is create something that is believable and plausible, like we want you to believe that this toy ball is actually in your house, and you're picking it up, and you're throwing it, and it's bouncing around. And that's really not immersion, in my opinion."* (P10). These perspectives reinforce the need for an understanding of psychoacoustics principles, so that audio designers can create both realistic and surreal virtual worlds and create plausible soundscapes for them to sell the illusion.

One key aspect of VR that requires careful attention when designing for immersion is headtracking, as outlined by P7 "*VR is unique in that you've got head tracking. So when your ambience is fixed and you've got head tracking, then I think that is a huge part of the feeling of immersion because it's like being in the real world*". Whilst this is a new challenge presented to audio designers when working in VR, there are some benefits. For example, research by [26] showed that front-back confusions in identifying sound sources decreased when listeners either moved their heads or the sound source themselves. P11 points out how this can negatively impact the experience when it is not treated in VR "*But when you go around, the fact that audio is stitched to your head, it doesn't work basically, you don't have that sense of immersion, that sense of presence in the scene*". Alongside the importance of headtracking, representation of the virtual self and user agency, P8 states that "*The next step is actually making those representations of the user interact with that environment. And that's a very easy way to amplify the immersion*". This statement supports the fact that designing for immersion is a crucial aspect of VR audio practice, which confers with findings in similar studies on VR audio practice [1, 2].

4.3.2 Multimodal congruency and plausible reverb. Designing VR audio also necessitates a careful balance of sensory inputs alongside audio for enabling both an immersive and functional user experience. Multimodal congruence, which ensures that auditory and visual stimuli align seamlessly, was seen as crucial for creating a cohesive experience across the participants' responses. Research by [10] states that this alignment is vital for maintaining immersion and preventing cognitive dissonance, as mismatched sensory inputs can negatively impact the user experience.

> *"Because now what happens is your eyes and your ears don't match and your brain's having to work really hard to try to put this whole thing together. What you're hearing and what you're seeing don't match, and the experience, for me, it just gave me a headache."* (P8)

This cognitive dissonance highlighted in the above quote suggests that audio creation for VR content cannot exist in a vacuum from the visual elements, as accurate synchronization is crucial for an immersive experience. Additionally, the integration of haptics plays a crucial role in achieving multimodal congruency, as P1 confirms "*Yeah, it's the haptics, isn't it? The actual air moving.*

That's sort of part of it". This underscores the importance of synchronizing not only auditory and visual content, but also tactile feedback to create a believable virtual environment. In addition to congruency, creating reverb that is plausible given the perceived virtual environment is essential for creating a sense of presence and immersion. P2 outlines this design consideration "*Do you want to modify the acoustic space that the person sees? Does it need to be exactly the same?*", extending on how the idea of plausibility can be applied to acoustic representations. Perhaps the most critical way to manipulate this is through reverb, as P1 states "*If the entire game occurs in an anechoic chamber, don't expect the audio to be very impactful*", pointing to the importance of reverb in creating an immersive sonic environment, as highlighted below:

> *"You're going to expect some sound that conveys the fact that you now have all these reflective surfaces around you, and if it doesn't, if the sound doesn't do that to at least a decent degree, if your reverbs are all broken that particular day, it's really hard to be in that space."* (P8)

Furthermore, the detail applied to creating plausible reverberation within VR has the potential to enhance the immersive experience, as P11 elaborates "*When you are in a cave and you hear real reflections, real diffractions of sounds, also occlusions*". Being able to integrate this level of detail is challenging, but the effort can pay off as P3 describes "*We did spend a lot of time crafting our reverbs and building custom IRs and using convolution reverbs for all our spaces. And you know, that really helps create a sense of immersion and make a space feel believable*". The observations detailed here emphasize the need for sophisticated reverb techniques and congruent multimodal feedback as key pillars in VR audio practice, further supporting previous research [1, 2, 10]. By addressing these aspects, VR audio designers can enhance user engagement and ensure a seamless integration of auditory and visual elements in virtual environments.

4.4 Theme 4 - Workflow and Technological Challenges

4.4.1 Preferred workflow approaches. VR audio production involves various tools and methods tailored to different project requirements and team dynamics. Our analysis of the preferred workflows used by the experts point to the necessity of aligning with the development team's tools, emphasizing the importance of compatibility and flexibility. In terms of content creation and 3DoF linear production, Digital Audio Workstations (DAWs) such as Reaper and Pro Tools were most commonly used, with the flexibility of Reaper being a notable benefit for VR audio production.

> *"For me, it's always building your content first. We use Reaper and we really like the extensibility of Reaper, and how much you can automate. As a small person team that also makes me happy, because it's sped up our workflow a bit and automated the process of making variations for sounds and things like that."* (P3)

> *"I use Pro Tools primarily because I'm proficient with it. While I've seen colleagues demonstrate advanced automation techniques in Reaper, I'm not sure I could achieve the same results in Pro Tools, even though it has to be said that it has recently gained scripting capabilities."* (P5)

> *"I love Pro Tools for its overview and structured approach, and Reaper is so versatile in handling Ambisonic tracks."* (P12)

> *"I'm primarily using Pro Tools and Soundminer for asset creation. Soundminer is a library tool that handles a lot of my catalog and sound design functionality, but Pro Tools is my preference for editing and mixing."* (P9)

> *"I would be doing my sound design in something like Reaper."* (P1)

In terms of audio implementation tools, the use of audio middleware was desired over native audio systems. Both Wwise and FMOD were mentioned as frequently used depending on the project requirements, with a slight preference towards Wwise. This preference reflects trends observed within the game development industry as a whole [27]. Similarly, The choice between Unity and Unreal engines often depends on the specific requirements of the project, underscoring a holistic approach to tool integration in VR audio production.

> *"I always preferred working directly with audio middleware like FMOD or Wwise rather than just going straight to Unity, because it just offers more tools in terms of audio implementation."* (P2)

> *"Unreal Engine handles native audio far better than Unity does. Unity is quite bare-bones in this regard. I would even say that if we're presented with a project developed in Unity and the production lacks the budget for audio middleware, we'll simply pass on it. Achieving the level of quality we seek is not feasible without it."* (P5)

> *"If I've got the choice, I'll be working in Unreal. By choice, I would be working with Wwise."* (P1)

> *"We're usually using FMOD for most of our implementations, and I've used a little bit of Wwise on some projects, but I prefer FMOD." (P7)*

> *"In terms of game engines, it's Unity for me. But I have done stuff in Unreal."* (P10)

> *"Both engines have their pros and cons. My preferred workflow, put shortly, is having everything as closely integrated within the same environment as possible, so that the design process is not thinking separately about VR, audio, and the interactive aspects of it."* (P6)

This flexibility identified between preferred workflows for VR was further underscored by a number of participants, leading to a consensus towards not relying on a specific workflow or set of tools. This aligns with the challenge of applying a blanket approach to VR as the use cases and context can vary widely between projects [1, 2].

> *"There are different developers with different tools that will plug into different software, and you need to know what your developers or the people producing the piece are working with so that you can deliver something that works for them."* (P2)

> *"So there's a lot of consideration around what is my development team doing, where are we publishing to, what is going to be compatible so I can achieve my level of quality, my level of taste to that thing as well."* (P2)
>
> *"Any tool can be replaced. To be honest, there are a million DAWs out there. I don't really care which one; I don't care how you get to the top of the mountain as long as we get there together."* (P3)
>
> *"My answer is evolving, but recently I've been prioritizing platforms that work in a more closely integrated way."* (P6)

Ultimately, the choice of tools and workflows in VR audio production is guided by the specific needs of the project, the preferences of the team, and the goal of creating a seamless and immersive auditory experience.

4.4.2 Workflow challenges. VR audio production presents unique workflow challenges that can complicate the development process. The primary issue mentioned by many participants was the difficulty of mixing audio within a VR environment, which often requires cumbersome hardware. This is compounded by the challenge of understanding audio content without experiencing it within the context of the HMD environment.

> *"Mixing in VR is hard. To mix in VR, we have hardware strapped to our face."* (P3)
>
> *"So to understand and really hear what your content is doing in a session, I'm just listening to it, and I don't understand what's going to happen until I throw it into the game and actually start moving my head around."* (P3)
>
> *"But if we're thinking in terms of VR, there's definitely that stepping outside of your development mode and going into the experience mode, and going back and forth, which is very disruptive."* (P6)
>
> *"How do you hear in context? You have to put the headset on, and then you have to take it off... How can you look at the computer screen and adjust your middleware, your audio implementation tools, if you've got this headset on?"* (P8)

The iterative process of creating and experiencing VR content introduces further workflow friction, with frequent adjustments needed while wearing and removing the headset. This workflow challenge of in-context VR mixing appears to be a known issue amongst tool developers, with companies such as Dear Reality addressing this issue with their 'Spatial Connect for Wwise' tool, which allows access to certain Wwise GUI windows inside VR environments for Unity development [28]. Another identified challenge involves transitioning between different versions of game engines and retooling work for different platforms and engines, which further complicates production workflows. Consistent communication and collaboration among team members is crucial to ensure smooth workflow integration, and the collaborative nature of VR audio production requires audio professionals to influence and understand decisions made by other departments. Additionally, system design-oriented thinking is a significant challenge in interactive VR audio development.

> *"It's impossible to retool the work in one engine to another engine. You can carry over assets, you can carry models, but the way you tie everything together is actually a massive part of the work. We don't have a format or a standard across these engines like you might have with digital audio workstations."* (P6)
>
> *"If you've been developing in Unreal 4.27 and suddenly decide to switch to Unreal 5.x, you might be in for an unpleasant surprise!"* (P5)
>
> *"There really needs to be a very, very, very clear and deep communication between the game designers, the artists, the directors, the animators, and the sound people."* (P1)
>
> *"As an audio person, you're very infrequently developing the whole piece, whatever it is, by yourself. So you have to work with your team in a lot more detail."* (P2)
>
> *"The challenge with interactive development is that it requires a significant amount of time to work on systems. The majority of your effort goes into designing and coding these systems."* (P5)

Overall, these insights reveal the multifaceted challenges faced by VR audio professionals. From the technical difficulties of mixing within a VR environment to the need for seamless collaboration and iterative testing, these challenges necessitate a flexible and adaptive workflow to achieve high-quality VR audio experiences.

4.4.3 Limitations of current tools. The VR audio production landscape is marked by significant limitations in the tools available, reflecting a disparity between the needs of immersive audio design and the capabilities of current technologies. A recurrent theme among participants is the insufficiency of tools compared to those available for traditional media, often necessitating compromises in achieving the desired level of immersion and quality.

> *"Once you pick the format or get prescribed to the format, then you have a choice of different tools, and they're somewhat limited. There's not as many tools as traditional media."* (P2)

This sentiment encapsulates a broader frustration with the current state of VR audio tools, where the scarcity of robust options forces audio designers to work within constrained parameters, often hindering their creative aspirations. A notable challenge identified by participants is the computational intensity required for realistic sound propagation and spatial audio rendering, which is further compounded through hardware limitations.

> *"Achieving things like sound diffraction, obstruction and occlusion can be challenging and computationally expensive."* (P5)
>
> *"You need to start to trace rays around to calculate real reflection, using the real geometry of the space. And this is very heavy from a computational point of view."* (P11)
>
> *"We're trying to recreate the full dynamic range and frequency spectrum of reality with limited hardware."* (P10)
>
> *"We still rely heavily on impulse response reverbs because achieving hyper-realistic sound in real-time on*

> *a VR device, which has processing power closer to that of a phone than a workstation, requires compromises."* (P5)

Such limitations often necessitate a balance between computational feasibility and audio fidelity, leading to less immersive experiences. Ambisonics, a popular method for spatial audio in VR [1, 2], also faces scrutiny for its limitations. While First Order Ambisonics can provide a basic level of immersion, it often falls short in delivering the nuanced audio experiences required for high-end VR projects, as noted. Higher Order Ambisonics can solve this through greater representation of the spatial scene, however this can come at the cost of computational efficiency.

> *"First order Ambisonics is something I mean, I hate it, it's very bad."* (P11)

> *"But if you try to make an ambient bed in Ambisonics, it has to be baked at some point; it has to be mixed down. And the playback of those channels becomes prohibitive on a VR system" (P7)*

> *"A first-order Ambisonic mix might tick a lot of boxes for many immersive media projects. But it may not achieve the sense of taste, the quality, or the level of immersion, the depth of experience, the 'je ne sais quoi' that you really have to go over the top of something basic." (P2)*

> *"The channel counts can get astronomical as you go into like second order and third order." (P7)*

The lack of integrated and comprehensive software solutions also emerged as a critical barrier.

> *"Honestly, I really hate to talk about software and tools because there isn't the one-stop solution that you're looking for. Everybody tries to have the shortcut, or the software that does everything for you. But this is not how it works."* (P4)

This highlights a fragmented tool ecosystem, where designers must often stitch together multiple software solutions, each with its own limitations and learning curves, to achieve their desired outcomes. Addressing these issues is crucial for advancing the field and enabling VR audio practitioners the best opportunity to craft immersive experiences.

4.4.4 Indispensable and Desired Tools for VR Audio Production. When asked if there were any tools that were indispensable in their current VR audio workflows, several experts pointed to philosophical rejection of any particular tool being indispensable. For example, P10 simply answered *"No"*, and P3 stated that their indispensable tool was "*My ears. Yeah, that's really the only answer I have",* and went on to further reinforce that *"A tool is just a tool. It's like in the same way that a game engine is just a game engine. If you can learn one DAW, you can work in other DAWs."* (P3). However, some participants highlighted various tools as essential, indicating a reliance on established software and hardware to meet current production needs.

> *"If Unity and Unreal were to disappear today, a lot of my projects would disappear with them."* (P6)

> *"Reaper."* (P7)

> *"I love fancy gear, high end studios, and am always chasing after the next big thing, but I really have learned to appreciate the staples in my toolbox that withstand the test of time and get the job done. A solid set of mics, recorder, speakers, headphones, DAW, sound library, and just a few great plugins is all I really need."* (P9)

> *"Interactive development involves a lot of heavy lifting, so having robust audio middleware is essential."* (P5)

Participants also identified several desired tools that would address existing gaps and enhance the production process. One unique finding presented the concept of a 'volumetric' sound source, taking a similar conceptual approach to 3D graphics rendering, and using this approach towards expanding real-time audio beyond simple point source emitters:

> *"I think that volumetric spatial audio would be an exciting feature to explore. The current mix of point source, speaker panning, and Ambisonics speaker agnostic workflows have been a lot of fun but challenging and not fully satisfying in many ways. Larger assets like dragons and tanks that you can really get close to would benefit from having those unique acoustic details."* (P9)

Alongside this, the desire for a tool to solve the main workflow challenge of mixing audio was also a common request among the experts interviewed.

> *"If you're dealing with something that's in 3D space, shouldn't you have 3D tools to be able to do that as well?"* (P2)

> *"I saw someone working on recently, like a way to pull up Wwise information in VR. And I'm like, that would be amazing. Anything that makes my life easier."* (P3)

> *"I think the workflow of designing these experiences in VR is the biggest missing piece."* (P6)

Further mentions of tools that would be desired included highly efficient and accurate auralization tools, such as:

> *"Native high order Ambisonics file compression and renderers in Unity and Unreal that are efficient for mobile devices."* (P9)

> *"Physics-based real-time audio generation that's also lightweight and can run on mobile with high performance"* (P10)

The importance of enhanced reverb capabilities was also noted, as well as a desire for better integration of real-world audio into VR environments:

> *"If there was like a twelfth order reverb which you could easily adapt to the physics of your room, not only the dimension but also the physics"* (P12)

> *"The ability to take actual impulse responses recorded in the real world and relatively easily get those into your development environment"* (P8)

Finally, a 'unifying' tool that could potentially reduce the complexity of managing multiple software and workflows, allowing for more efficient production, was mentioned by several experts.

> *"Something that would improve the interoperability between linear and interactive tools and workflow."* (P5)
>
> *"I always feel like it would be great if there were a tool that does everything for you, right? Where you can spatialize sound, it does have some reverb, you can go to Ambisonics, but you also can go to Dolby Atmos, maybe you can also go to 6 degrees of freedom."* (P4)
>
> *"You could think of plugins that make it easy with one push of a button, and it's being spatialized like that."* (P12)

In summary, the indispensable tools for VR audio production encompass a range of software and hardware that form the backbone of current workflows. However, the desire for more integrated, intuitive, and comprehensive tools indicates a clear direction for future advancements. By addressing these gaps, the industry can move towards more seamless and efficient VR audio production processes, ultimately enhancing the quality and immersion of VR experiences.

5 CONCLUSION

This study explores the complexities of designing audio for VR, revealing both similarities and unique challenges compared to traditional media. Through semi-structured interviews with leading VR audio expert practitioners, the research aimed to address key questions about the production challenges, techniques, and design decisions in VR audio. Thematic analysis was employed to systematically analyze data from one-on-one semi-structured interviews with 12 experts, providing deep and relevant insights into these areas. The findings reveal that VR audio design requires flexible approaches due to the diverse range of applications, emphasizing the importance of strategic audio placement and representing scale to emphasize the unique egocentric perspective of VR. The necessity of dynamic sound emitter techniques to create realistic soundscapes and effective psychoacoustic considerations are crucial for achieving immersion, realism, and plausibility, which were identified as unique and important design goals. Challenges faced by VR audio designers mainly concern complex workflow issues and tool limitations. Mixing audio in VR is particularly challenging due to hardware constraints and the need for iterative testing within the VR environment. The limitations of current tools underscore the need for integrated and comprehensive solutions. The potential for advancements in tools and techniques, such as volumetric sound sources and computationally efficient auralization tools, were identified as highly desired by the experts involved in the study. As the field of VR audio continues to grow, complementing technical research and advancements with insights and feedback from expert practitioners is crucial in developing new tools and refining existing ones to meet the evolving needs of the field.

Acknowledgments

This research was supported and funded by Dolby Australia and the Australian Government Research Training Program (RTP) Scholarship. The authors would also like to thank all the participants for their giving up their time and valuable insights for the study.

References

[1] Daniel Turner, Damian Murphy, Chris Pike, and Chris Baume. 2021. Spatial audio production for immersive media experiences: Perspectives on practice-led approaches to designing immersive audio content. *The Soundtrack*, *13*(1), 73-94.
[2] Justin Paterson and Oliver Kadel. 2023. Audio for extended realities: A case study informed exposition. *Convergence*, 13548565231169723.
[3] Thomas Deacon and Mathieu Barthet. 2022. Spatial Design Considerations for Interactive Audio in Virtual Reality. In *Sonic Interactions in Virtual Environments* (pp. 181-217). Cham: Springer International Publishing.
[4] Stephan Schütze and Anna Irwin-Schütze. 2018. *New Realities in Audio: A Practical Guide for VR, AR, MR and 360 Video*. CRC Press.
[5] Nikunj Raghuvanshi and Hannes Gamper. 2022. Interactive and Immersive Auralization. In *Sonic Interactions in Virtual Environments* (pp. 77-113). Cham: Springer International Publishing.
[6] Jon Swords and Nina Willment. 2024. The emergence of virtual production–a research agenda. *Convergence*, 13548565241253903.
[7] Michele Geronazzo and Stefania Serafin. 2023. Sonic interactions in virtual environments: The egocentric audio perspective of the digital twin. In *Sonic Interactions in Virtual Environments* (pp. 3-45) Cham: Springer International Publishing.
[8] Hyunkook Lee. 2022. A conceptual model of immersive experience in extended reality. PsyArXiv. https://doi.org/10.31234/osf.io/sefkh
[9] Dai-In Danny Han, Frans Melissen, and Mata Haggis-Burridge. 2024. Immersive experience framework: a Delphi approach. *Behaviour & Information Technology*, *43*(4), 623-639.
[10] Rolf Nordahl and Niels C. Nilsson. 2014. The sound of being there: presence and interactive audio in immersive virtual reality. In *The Oxford Handbook of Interactive Audio*, Oxford Handbooks.
[11] Gordon Calleja 2007. Digital game involvement: a conceptual model. *Games and Culture*, 2, 236-260.
[12] Isabella Van Elferen. 2016. Analyzing game musical immersion: the ALI model.
[13] Justin D. Mathew Stéphane Huot, and Brian F. Katz. 2017. Survey and implications for the design of new 3D audio production and authoring tools. *Journal on Multimodal User Interfaces*, *11*, 277-287.
[14] Anil Çamci. 2024. Modern Workflows for Procedural Audio at the Intersection of Gaming and Music Performance in Virtual Reality. In *Audio Engineering Society Conference: AES 2024 International Audio for Games Conference*. Audio Engineering Society.
[15] Anil Çamci. 2019. Some considerations on creativity support for VR audio. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 1500-1502). IEEE.
[16] William C. Adams. 2015. Conducting semi-structured interviews. *Handbook of practical program evaluation*, 492-505.
[17] Sarah E. Baker and Rosalind Edwards. 2012. How many qualitative interviews is enough.
[18] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.
[19] Victoria Clarke and Virginia Braun. 2017. Thematic analysis. *The journal of positive psychology, 12*(3), 297-298.
[20] Changjin Kim, Robert Mason, and Tim Brookes. 2013. Head movements made by listeners in experimental and real-life listening activities. *Journal of the Audio Engineering Society, 61*(6), 425-438.
[21] Maria C. Reyes. 2024. Towards Immersography: Considerations for an Integrated Understanding of Immersive Narrative Experiences. In Keith Dooley and Amanda Munt (eds.) *Screenwriting for Virtual Reality*. Palgrave Studies in Screenwriting. Palgrave Macmillan, Cham.
[22] J. Kevin O'Regan and Alva Noë. 2001. What it is like to see: A sensorimotor theory of perceptual experience. *Synthese 129*: 79-103
[23] R. H. Kay. 1982. Hearing of modulation in sounds. *Physiological Reviews*, 62(3): 894–975.
[24] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535): 3549–3557.
[25] Mathew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2).
[26] Frederick L. Wightman and Doris J. Kistler. 1999. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America, 105*(5), 2841-2853.
[27] Brian Schmidt. 2023 *Game Audio Industry Survey 2023*. GameSoundCon. Retrieved May 10, 2024, from https://www.gamesoundcon.com/post/game-audio-industry-survey-2023
[28] Dear Reality. (n.d.). SPATIAL CONNECT for Wwise. Dear Reality. Retrieved May 10, 2024, from https://www.dear-reality.com/pages/spatial-connect-wwise

A A APPENDICES

A.1 Semi-structed interview questions

(1) How would you describe the key differences between producing audio for conventional media compared to virtual reality experiences?
(2) What do you see as some of the biggest challenges in working with spatial audio in VR?
(3) What would you say are the key factors for making users feel immersed through audio in VR?
(4) Do you consciously design towards these factors, or do you think it comes naturally through wanting to design a good user experience?
(5) Could you briefly talk us through your preferred workflow when working on audio for a VR project?
(6) Are there specific software or tools that have become indispensable in your VR audio production process?
(7) Is there a missing piece or tool that doesn't exist yet that would make your workflow significantly better?