Audio Engineering Society

Convention Paper 10260

Presented at the AES 160th Convention
2026 May 28–30, Copenhagen, Denmark

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

The Missing Next Step: Sound, Agency, and Plausibility in Virtual Reality — A Narrative Review

Peter Kurucz[1], Neil Hillman[2], Stefania Serafin[3], Eve Klein[4], and Nilufar Baghaei[1]

[1]*The University of Queensland, School of Electrical Engineering and Computer Science, Australia*
[2]*The Audio Suite, Australia*
[3]*Technical University of Denmark, Department of Engineering Technology and Didactics, Denmark*
[4]*The University of Queensland, School of Music, Australia*

Correspondence should be addressed to Peter Kurucz (p.kurucz@uq.edu.au)

ABSTRACT

Sound plays a critical role in virtual reality (VR), shaping attention, narrative comprehension, emotional engagement, and experiential plausibility under conditions of embodiment and user agency. Although a growing body of research addresses VR audio techniques, perceptual effects, and sound taxonomies, existing approaches remain fragmented and largely descriptive. In particular, they do not provide a unifying, VR-specific account of how sound meaning and emotional intent are operationally linked to user agency and non-linear narrative progression. This paper presents a narrative review of selected literature spanning game audio frameworks, immersive sound design, narrative theory, and plausibility-related research in games and VR. Through synthesis of these perspectives, the review identifies a conceptual gap in current research, namely the absence of a VR-specific, agency-coupled sound design framework for structuring sound meaning and emotional intent in support of experiential plausibility as users actively shape events in interactive VR environments.

1 Introduction

Despite a growing body of work on virtual reality (VR) audio techniques, taxonomies, and perceptual effects, existing approaches do not yet provide a unifying, VR-specific sound design framework that operationally links sound meaning and emotional intent to user agency, narrative progression, and experiential plausibility. In practice, sound designers working in interactive VR must make design decisions under conditions of embodiment, mobile perspective, and non-linear interaction, where sound plays a critical role in guiding attention, conveying causality, and maintaining a coherent sense of "what is happening" as users act within the environment.

While prior frameworks from film and game audio offer valuable insights, their underlying assumptions about fixed viewpoints, linear progression, or externally authored pacing do not fully account for the experiential consequences of user agency in VR. As a result, sound design knowledge in this domain remains fragmented across technical guidelines, descriptive taxonomies, and perceptual studies.

This paper presents a narrative review of selected literature identified through targeted database searches and reference chaining, focusing on sound design, interactivity, narrative, and plausibility in games and VR,

with the aim of synthesising current approaches and identifying conceptual and practical gaps. The paper articulates the need and implications for future sound design frameworks capable of supporting coherent and believable experiences in authored interactive VR environments.

2 Scope, Definitions, and Assumptions

To establish a clear foundation for this manuscript, it is necessary to define key terms commonly used in the field, as many remain subject to varying interpretations and persistent misconceptions. In this paper, the term *virtual reality* refers to the creation of a fully virtual world that replaces the user's physical environment. This entails a complete separation from the physical environment, typically achieved through a head-mounted display (HMD) and headphones, resulting in a fully immersive experience [1].

When referring to *immersive systems*, we adopt a system-oriented perspective that encompasses characteristics such as field of view, tactile feedback, and spatial sound. In this sense, immersion is treated as an objective property of the system, shaped by the degree to which these components are coherently integrated. Contemporary HMD-based VR systems are generally capable of supporting high levels of immersion by default [2].

For the purposes of this paper, we assume controlled playback conditions in which virtual audio dominates physical-world cues, such as through the use of closed-back or high-quality headphones. Under these conditions, sound within the virtual environment is treated as perceptually authoritative. Accordingly, the experiential boundaries of a VR session can be operationally defined as the moment the user puts on an HMD and the moment it is removed. Within these boundaries, the virtual experience is treated as a self-contained perceptual system.

2.1 Included Interaction Paradigms and User Roles

This paper focuses on *interactive VR* environments, defined as virtual spaces in which users can actively interact with their surroundings. User interactivity in VR is commonly described in terms of degrees of freedom. Three degrees of freedom (3DoF) support rotational head movement without positional translation, while six degrees of freedom (6DoF) additionally enable translational movement within the virtual space [3]. Beyond locomotion, interactive systems may support object manipulation through virtual hands or controller-based interaction. Together, these interaction modalities allow users to influence events within the environment, resulting in individualised experiential and narrative trajectories shaped by user actions and in-situ decisions [4][5].

In narrative contexts, such environments may be linear or non-linear, and user actions may or may not influence narrative progression. Following Dooley's classification, user roles within VR narratives are understood to range from passive observers to active explorers, with the latter possessing agency and the capacity to influence events within the virtual space [6]. The present discussion is restricted to environments in which users assume the role of the explorer and exercise agency over their experience within non-linear narrative progression (see Figure 1).

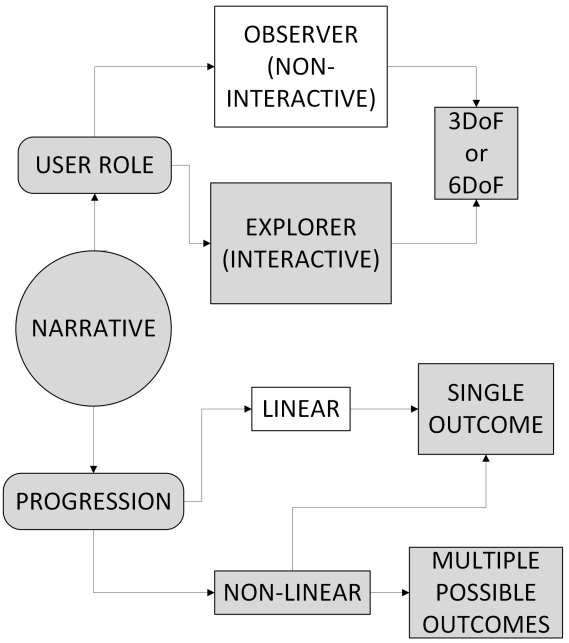


Fig. 1: Narrative progression and user roles adapted from Dooley [6] and Crawford [5]; grey items indicate this manuscript's focus.

2.2 Review Scope and Methodological Framing

We examine relevant approaches and taxonomies of sound and sound design, with a focus on their contributions to interactive virtual environments. This review is not intended to be comprehensive or systematic; rather, it highlights selected works most relevant to the scope of this paper. Sources were identified through targeted searches on Google Scholar using representative keywords including terms such as “VR sound design”, “immersive audio”, “plausibility illusion”, “presence”, “game audio frameworks”, “sound taxonomies”" and “narrative sound”, and were supplemented by reference chaining from relevant publications. Only peer-reviewed, English-language sources were included. As this is a narrative review [7], no PRISMA protocol was applied, and additional relevant studies, particularly within grey literature, may not be captured.

2.3 Exclusions: Procedural and Generative Systems

Procedurally generated audio systems, where sound content or structure is created algorithmically in response to user behaviour, have been explored in prior research [8]. Although such systems are highly relevant to future VR sound design, they are excluded from the current scope. This limitation is intentional: before extending sound design approaches to generative contexts, it is necessary to establish a shared descriptive language for sound meaning, agency, and experiential coherence within authored interactive VR environments. Such a foundation may later support the systematic integration of procedural audio techniques. The present discussion is therefore limited to non-generative VR environments in which system behaviours, narrative structures, and sound responses are authored rather than dynamically generated.

3 Sound, Narrative, and Listener Perspective in Interactive VR

In practice, individualisation of the auditory experience in interactive environments is most commonly addressed through user-exposed audio personalisation settings, such as independent control over music, dialogue, and sound-effect levels. For example, in the popular role-playing game (RPG) Baldur’s Gate 3 [9], an extensive range of audio personalisation options is available, including separate amplitude controls for Sound Effects, User Interface, Ambient sounds, Music, Voice, and Chatter, as well as adjustments to the overall dynamic range of the sound mix. While such settings allow experiences to be tailored to individual listening preferences and accessibility needs, they primarily operate at the level of mix balance and do not engage with the spatial, narrative, or affective roles of sound, limitations that become particularly salient in VR environments where listener movement and perspective are central to sound perception. The complex responsibility of addressing this falls back onto the sound development team.

When a user is able to move around in a VR space either 3DoF or 6DoF, developers must also consider how sound behaves from the user’s perspective. Sound designers typically use either head-locked stereo, spatialised sound or a hybrid combination of the two [10]. In head-locked stereo, the sounds move with the user’s head rotation, while spatialised sound is translated from Ambisonics to binaural stereo. The latter allows the sound field to rotate independently as the user moves, enabling accurate positioning of sounds within the surrounding space. Spatialised sound can create the impression of localisation of sounds in these spaces better than head-locked stereo [11].

In VR environments supporting at least 3DoF, users can actively orient toward off-screen sounds through head movement, allowing auditory cues to guide attention and acquire visual context in ways not possible in fixed-viewpoint media. As the listener moves, spatial audio rendering systems shape the sound field around the user using head-related transfer function (HRTF) data, enabling changes in localisation and timbre that reflect listener perspective [12]. When head-locked stereo playback is used, this coupling between sound and movement is absent, and sounds do not respond to the user’s orientation. By contrast, traditional 2D media rely on externally controlled camera movement and fixed listening positions, with all spatial relationships determined during post-production rather than through user action.

3.1 Diegesis, Embodiment, and Transdiegetic Sound

One of the distinct features of interactive VR, compared to linear film or even some non-linear game sound design, is the user’s ability to experience the environment through a first-person avatar. By embodying an avatar

and having the ability to influence the flow of events, the line between diegetic and extradiegetic sound elements becomes less defined. In traditional media, such as films, diegetic sounds are known to be the sounds that exist within the world of the story, such as environmental, foley sounds and general dialogue by the characters. Whereas extradiegetic (often referred to as non-diegetic) sounds are the opposite of these, sounds that do not exist in this world. In the world of traditional film sound production, it is often the narration or musical scoring [13].

When it comes to sound in computer games, the rules shift because the listener's role changes from that of a passive observer to an active participant. A new term, transdiegetic sounds have been defined by Jørgensen [14], which contains sounds that cannot be clearly specified as diegetic or non-diegetic. Jørgensen explores this hybrid space in relation to the musical score in computer games, which is technically non-diegetic since it is not part of the game world, yet it is audible to the player who controls the game character. The score is often used to foreshadow danger or highlight various narrative elements, thereby influencing the player and, by extension, the character's actions within the game world. In this way, the score could be considered part of the game world. However, it can also be argued that when a user's physical world is fully replaced by a virtual one, and they completely embody an avatar, the virtual environment becomes their new reality. In such cases, all sounds become diegetic to the user, and non-diegetic sounds could effectively cease to exist within fully interactive and immersive VR environments using HMDs.

One may ask whether sounds from the real world interfering with the VR environment could impact the user's perception of that sonic environment. In short, the answer is yes. However, much like film sound designers, we strive for precision and work under the assumption and hopes that the listener will experience the environment under controlled conditions, such as using high-quality headphones or in the case of films a properly calibrated cinema sound system. While this is a significant issue, it falls outside the scope of this paper; we assume playback via good-quality headphones that minimise external interference with the virtual soundscape.

3.2 Narrative Theory and User Agency in Interactive Media

Branigan [15, p. 36] notes: *"Narrative is a way of comprehending space, time, and causality."*. While this definition originates in the context of film, it remains relevant in interactive VR, where narrative does not necessarily follow a traditional linear structure with predefined plots and characters. Instead, it can also emerge organically through spatial design, user interaction, sensory cues, and the emotional arc of the experience. Some VR experiences retain a more conventional, story-driven approach, for example, the popular game Half-Life: Alyx [16], which closely follows a structured narrative arc. However, VR spaces also support other types of narratives. These can be shaped by the design of the environment and spatial elements rather than a straightforward storyline. For example, the popular sandbox game The Sims features an open world and open-ended structure that drives user interactions [17]. Another category is procedural narratives, which are not scripted but instead evolve based on the player's interactions within the rule-based systems of the environment like Façade [18]. Lastly, embodied narratives can emerge through the users physical and sensory experiences, often shaped by movement, perspective, and interaction in the space, such as certain theme park attractions [19].

In their book, Hamlet on the Holodeck, Murray [19] explores cyberspace, describing the key properties of digital environments in relation to users, with a focus on narratives. Although Murray does not investigate virtual scenario plausibility directly, they argue that the higher the level of procedural and participatory design (interactivity), the greater the user's sense of agency within these new virtual environments. Additionally, an increased encyclopaedic extent (the details of the content) and coherent spatial navigation (moving around in the space) can enhance overall immersion. Murray also describes the phenomenon of "Active Creation of Belief", where interactivity and immersion reinforce each other. A user being inside a coherent immersive space can be motivating for them to engage within and therefore contributing to their sensation of agency. This feeling of agency could then feed back into enhancing their immersive experience.

User expectations for realistic or near-realistic representations of a space, including the soundscape, in VR can be high in certain scenarios, which may lead to

breaks in scenario plausibility if those expectations are not met [20]. However, if the environment is well designed, users can accept this new reality. In this context we use the same outline of realism as described by Puronas [21, p. 184]: *"..realism refers to believability of sounds made by characters, gadgets and objects located within their environment."*. However, it has been found that physical realism is not a necessary requirement of the plausibility illusion (Psi) [2]. Hyper realism, in the context of sound design, means the hypertextualisation of sounds to evoke the desired impact on the listener [22]. Film and games sound designers have been using many tricks, including hyper realistic sound design approaches, to convey scenario plausibility and the desired emotional outcome in support of the narrative.

Taken together, these perspectives reveal that sound in interactive environments operates simultaneously at narrative, perceptual, and affective levels, motivating the development of sound design frameworks and taxonomies that can systematically describe these roles within games and VR.

3.3 Existing Game Audio Frameworks and Taxonomies

In the context of video games, Grimshaw-Aagaard [23] analyses the notions of fidelity, authenticity, and verisimilitude (believability) in sound design, highlighting how realism has been repeatedly mobilised as a marketing construct rather than a perceptual necessity. The author argues that designers should focus on verisimilitude rather than the realistic replication of sounds to enhance player experience.

Drawing on embodied cognition and mirror neuron theory, Karen Collins discusses how virtual characters in games are often perceived as extensions of ourselves in the digital world [24]. Collins examines how sounds produced by the digital self can influence our empathy toward virtual characters. The discussion includes not only sounds resulting from character interactions but also music and how sound goes beyond creating ambiences or just setting a mood. Collins also touches on voice, noting that a mismatch between the player's expected voice and the character's voice can create a sense of disconnection. Collins concludes by emphasising that designers have a responsibility in how they use sound to influence users' emotions, pointing out that if

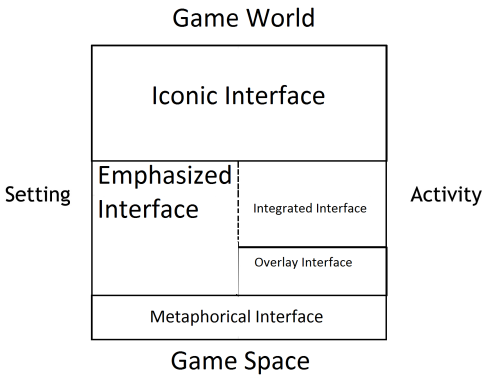


Fig. 2: The Flexible Game Audio Model for Categorization (FGAMC) model developed by Westerberg et al. [25] based on the works of Jørgensen [26] and Huiberts et al. [27].

video games can make us cry, they may also be capable of eliciting a wide range of emotional responses.

Westerberg et al. [25] developed the Flexible Game Audio Model for Categorization (FGAMC), which aims to provide a categorisation framework for game audio using two axes (see Figure 2). The first axis places sounds along a spectrum of diegesis, ranging from those that exist within the "Game World" (the world experienced by the player-character) to those within the broader "Game Space", which includes all system and interface sounds relevant to gameplay. The second axis considers the functionality of each sound, evaluating whether it serves the purpose of "Setting", sounds that define the world and narrative context, or "Activity", sounds generated by the interactions of the player's character.

When it comes to the user gaining access to narrative information, Branigan [15] discusses the concept of focalization, where information is either filtered through a character's perspective or objectively by the viewer themselves. In our case in VR, the user becomes the character by embodying them, and all information within the space is perceived and processed directly through their experience. Therefore, focalization happens through the embodied character. Any framework applied to these new reality environments must account for nuances that may cause such conflicts when translating ideas from traditional, linear narratives to non-linear ones [28].

It has often been discussed how sound can be essential for conveying presence, guiding attention, enabling

feedback, and affecting emotions in virtual environments [29, 30]. While a comprehensive sound design framework is lacking, several studies have focused on specific areas, such as the assessment of photorealistic VR environments and how the application of sound design could enhance the sense of presence for the user [31]. Others showed how the correctly synched responsiveness of sounds triggered by the user can enhance the users' presence in a VR environment [32]. However, literature on this topic is often framed more as a practical and technical guide to gain a better understanding of overall audio production for games. One example is the comprehensive work of Jean-Luc Sinclair [33], used by many, and which inspired even the sound design approach of a co-located narrative VR experience [34].

A hybrid framework which provides an approach bridging technical and theoretical is the works of Rob Bridget [35]. The "big four food groups" of sound in games, identified by the author, include sound (effects), music, voice, and the mix itself. Each categorised by their origin and function within the space. The fourth category, the mix, encompasses not only amplitude adjustments but also the prioritisation and balancing of audio elements [36]. They also recommend that due to the subjective nature of the mix within games, the overall authority of the mix should be still guided by the sound developer team, but more comprehensive controls should be handed over to the user [37]. Bridget explores game sound design not only from the technical, but psychological perspective too [38]. They state that sound can be approached objectively and from an imaginative perspective, where sound design led by storytelling can be the cause of exaggerated sounds. They also categorise sounds from a higher-level approach, based on psychological perception, whether they are objective (world), emotionally enhanced (heart) and thus modified by the character whose eyes/ears we perceive them through. Lastly, they describe intellectual sounds (mind), which are defined by the intellectual perception and interpretation of the game's character and often, by extension, the player. Their work is founded in traditional film post-production strategies and emphasise that sound can shape a narrative meaning in virtual environments.

Sound design for VR presents many new challenges for developers; however, it has been demonstrated that game sound design approaches can be adapted to these new realities. For example, in the sound design for the popular Zero Latency games, which are multiplayer VR first-person shooters, developed specifically for virtual reality [39]. Some of these techniques included the use of audio zones within the game environment, enabling the triggering and progressive mixing of sounds in the space. Another technique was the creative use of the Proximity Effect, where the lower frequency range of sounds is boosted when they are closer to the player, thereby enhancing their perceptual impact. The developers also concluded that simplifying certain aspects of the game sound design, such as spatialising only the most important sounds, improved the playability of the project for the users. This aligns with the research findings that emphasise simplification over complexity by Walther-Hansen et al. [40]. They argue that, while modern sound development tools allow us to expose the users to many sensory information it is also arguable that the strategic reduction of these, including sound design, can be beneficial to VR environments.

This approach of favouring simplification over complexity was not a new one; it was also adopted in linear media. When cinema sound design was at the same experimental stage as VR sound design is now, the auteur New Wave film director Robert Bresson said:

> *"Because the ear is profound, whereas the eye is frivolous, too easily satisfied. The ear is active, imaginative, whereas the eye is passive. When you hear a noise at night, instantly you imagine its cause. The sound of a train whistle conjures up the whole station. The eye can perceive only what is presented to it."* [41]

And Bresson went on to say:

> *"There are two kinds of sound in my films: sounds which occur during shooting and those I add later. What I add is more important because I treat these sounds as if they were actors. For example, when you go into the street and hear a hundred cars passing, what you think you hear is not what you hear, because if you recorded it by means of a magnetophone, you would find that the sound was a mere jumble. So, when I have to record the sound of cars, I go to the country and record every single car in pure silence. Then I mix all these sounds in a way that creates not what I hear in the street, but what I think I hear."* [41]

3.4 Holistic and Inclusive Approaches to VR Sound

More current research has also explored the importance of an audio-focused VR game, SoundTrek, in making VR more inclusive for visually impaired users, and in countering the field's strong visual orientation [42]. The authors developed this game to explore an approach in which allowing the user to switch between first-person and third person listening perspectives can help players navigate the virtual obstacles in an environment. Their goal was to use the game to further explore sound design considerations for creating interactive VR environments. Their approach provided a unique interaction approach, based on audio perspective, that offers insights into similar inclusive VR environment developments.

Tom Garner [43] emphasises that sound perception in virtual environments should be approached holistically, rather than being viewed solely through its individual components. Garner recommends focusing on both the sounds in the space, the receiver, physics and psychology all together from a listener-centric perspective. Garner re-developed the Sonic Virtuality model to illustrate both the physical and psychological contents of how sound is being actualised within VR environments (see Figure 3). Due to the unique nature of how perception works in VR, the author also concludes that sound design for these spaces must be carefully assessed and somewhat developed differently than in films or even in games. User interactions are key contributors to how sound is perceived in VR environments and spatial audio design requires an interdisciplinary approach, that incorporates both sound design, interaction design, psychology, and artistic practices [44].

Jain et al. [45] have collected and analysed interview data from VR sound designers to create a two-dimensional taxonomy for VR sound environments. The motivation behind the study was to better understand accessibility for hearing impaired users. Their findings were validated by Human Computer Interaction (HCI) researchers. The first dimension specified is based on the source of the sound and included: localised speech, non-localised speech, animate objects, inanimate objects, notification sounds, interaction sounds, surrounding ambience, point ambience, and music. The second dimension is based on the intent of the sound and included: conveying critical information, increasing realism, supporting rhythm or movement,

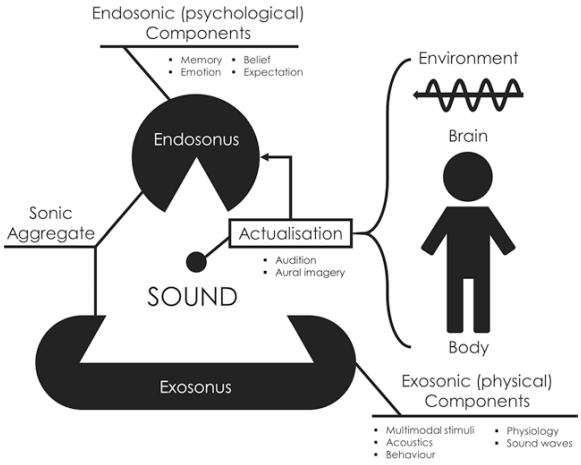


Fig. 3: The Sonic Virtuality Model, by Tom Garner [43, p. 71], illustrating physical and psychological components of sound being actualised in VR environments.

expressing affective state, providing aesthetic or decorative value, and facilitating non-critical interactions. They also found that while there are some standards for general technical sound design approaches in VR, there are no real creative approaches developed just yet. Most sound designers working in this space rely on their knowledge gained from other fields.

Furthermore, interviews conducted with professional VR sound designers by Hedges et al. [10] point out that the field is still in its infancy, and also mention that sound professionals continue to rely heavily on their expertise from linear media and computer game development projects. Their findings also showed that designing approaches are highly content-dependent, perspective is a major influencing factor, and that no established workflow or design framework is currently available to developers.

4 Synthesis and Identified Gap

Taken together, the reviewed literature demonstrates sustained interest in understanding the role of sound in interactive and immersive environments, particularly within the contexts of games and head-mounted VR. Previous work has produced a range of taxonomies and conceptual frameworks that describe sound in terms of function, perception, diegesis, intent, and technical implementation. These contributions have advanced understanding of how sound supports presence, guides

attention, conveys information, and shapes affective experience in interactive media.

However, these approaches largely remain descriptive or domain-specific and do not jointly provide an operational account of how sound meaning and emotional intent are coupled to user agency in VR. In particular, existing frameworks do not adequately address how sound design adapts as users actively shape non-linear narrative progression, nor how sound contributes to the ongoing maintenance of experiential plausibility under conditions of embodied interaction and perspective change. As a result, developers and researchers lack a shared, VR-specific conceptual basis for reasoning about how sound meaning and intent should be structured to support coherent and believable experiences in authored interactive VR environments.

One framework that is particularly relevant in this context is the Four Sound Areas (FSA), which was originally developed for linear media, particularly film sound design [46]. The FSA framework moves beyond functional classifications by incorporating emotional intent and narrative meaning, while acknowledging the dynamic nature of soundtracks. We propose that a new framework can be conceptualised by extending this model into interactive VR contexts. This structured conceptual framework can be used to map sound design dimensions to experiential outcomes, accounting not only for sound categorisation but also for its contribution to experiential and narrative impact within interactive environments. Such a framework could address the gap identified in this paper by guiding sound design in scenarios where users, acting through embodied avatars, are afforded agency over the narrative. Its formalisation and validation, however, are beyond the scope of this narrative review and remain for future work.

5 Implications for Practice and Research

The synthesis presented in this review has implications for both sound design practice and research in interactive VR. From a practitioner perspective, the absence of a shared, operational framework for reasoning about sound meaning and emotional intent under conditions of user agency places a significant burden on individual designers. In the absence of such a framework, sound design decisions in VR are often guided by personal experience, conventions inherited from linear media or games, or iterative trial-and-error processes. While these approaches can produce effective outcomes, they limit the transferability of design knowledge, complicate collaboration across teams, and make it difficult to articulate or justify design decisions in relation to experiential plausibility.

For researchers, the lack of an agency-coupled, VR-specific conceptual framework constrains the systematic study of how sound contributes to plausibility, narrative comprehension, and user experience over time. Existing taxonomies and models provide valuable descriptive accounts of sound function and perception but offer limited support for formulating testable hypotheses about how sound meaning and emotional intent should adapt as users actively shape non-linear narrative trajectories. This gap complicates efforts to compare findings across studies and to relate empirical results back to concrete design principles.

The absence of a shared, VR-specific sound design framework also has implications for emerging algorithmic and AI-driven audio systems. As adaptive, data-driven, and generative techniques become increasingly integrated into interactive VR pipelines, these systems will necessarily rely on underlying conceptual models to determine what should change, when, and why. Establishing a coherent, agency-coupled framework is therefore not only relevant for human designers, but also a prerequisite for meaningful automation and intelligent sound behaviour in future interactive VR environments.

Taken together, these implications point to the need for future research aimed at developing and validating integrative sound design frameworks that explicitly account for embodiment, agency, and narrative progression in VR. Such frameworks would support more intentional sound design practice, enable clearer communication between designers and researchers, and provide a common basis for evaluating how sound contributes to experiential plausibility in interactive virtual environments.

References

[1] Schütze, S. and Irwin-Schütze, A., "definitions and core concepts," in *New Realities in Audio: A Practical Guide for VR, AR, MR and 360 Video*, chapter 2, pp. 15–49, CRC Press, Boca Raton, 2018.

[2] Slater, M., "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), pp. 3549–3557, 2009.

[3] Lim, C. H. and Lee, S. C., "The effects of degrees of freedom and field of view on motion sickness in a virtual reality context," *International Journal of Human–Computer Interaction*, 40(19), pp. 5884–5896, 2024.

[4] Schütze, S. and Irwin-Schütze, A., "new realities," in *New Realities in Audio: A Practical Guide for VR, AR, MR and 360 Video*, chapter 1, pp. 1–14, CRC Press, Boca Raton, 2018.

[5] Crawford, C., *Chris Crawford on interactive storytelling*, Pearson Education, 2004.

[6] Dooley, K., "Virtual Reality Narration: Listening to and Reliving Stories Through First-Person Testimony," in *Virtual Reality Narratives: Embodied Encounters in Space*, chapter 3, pp. 69–93, Palgrave Macmillan, Cham, 2024.

[7] Agarwal, S., Charlesworth, M., and Elrakhawy, M., "How to write a narrative review." *Anaesthesia*, 78(9), 2023.

[8] Menexopoulos, D., Pestana, P., and Reiss, J., "The state of the art in procedural audio," *Journal of the Audio Engineering Society*, 71(12), pp. 826–848, 2023.

[9] Larian, "Baldur's Gate 3," 2023, video game.

[10] Hedges, J., Sazdov, R., and Johnston, A., "Keeping your ears on the virtual ground: Expert insights on virtual reality audio techniques and challenges," volume 61, pp. 373–384, ACM, New York, NY, USA, 2024.

[11] Schütze, S. and Irwin-Schütze, A., "why isn't this easy? The many challanges of the new realities," in *New Realities in Audio: A Practical Guide for VR, AR, MR and 360 Video*, chapter 3, pp. 51–96, CRC Press, Boca Raton, 2018.

[12] Meta (formerly Facebook) Developer Blog, "Improve Spatial Audio Quality with the Universal HRTF," `https://developers.meta.com/horizon/blog/improve-spatial-audio-universal-hrtf-meta-quest/`, 2023, accessed: 2025-08-25.

[13] Bordwell, D., Thompson, K., and Smith, J., "Narrative Form," in *Film art: An introduction*, volume 7, chapter 3, pp. 72–111, McGraw-Hill New York, New York, 2004.

[14] Jørgensen, K., "On transdiegetic sounds in computer games," *Northern Lights: Film & Media Studies Yearbook*, 5(1), pp. 105–117, 2007.

[15] Branigan, E., *Narrative comprehension and film*, Routledge, London, 1992.

[16] Valve, "Half-Life: Alyx," 2020, video game.

[17] Jenkins, H., "Game Design as Narrative Architecture," in N. Wardrip-Fruin and P. Harrigan, editors, *First person: New media as story, performance, and game*, pp. 670–689, Mit Press, 2004.

[18] Mateas, M. and Stern, A., "Build it to understand it: Ludology meets narratology in game design space," in S. de Castell and J. Jenson, editors, *Proceedings of DiGRA 2005 Conference: Changing Views: Worlds in Play*, 2005.

[19] Murray, J. H., *Hamlet on the Holodeck, updated edition*, Cambridge, MA & London: MIT Press, London, England, 2017.

[20] Kurucz, P., Serafin, S., Chitale, V., Klein, E., and Baghaei, N., "Investigating the Impact of Avatar Action Sounds on the Plausibility Illusion in Virtual Reality," in L. O'Connor, editor, *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 636–642, IEEE, Saint Malo, France, 2025.

[21] Puronas, V., "Sonic hyperrealism: illusions of a non-existent aural reality," *The New Soundtrack*, 4(2), pp. 181–194, 2014.

[22] Holman, T., "Introduction," in *Sound for Film and Television*, pp. xi–xiii, Routledge, New York, 2012.

[23] Grimshaw-Aagaard, M., "Video Game Sound Design and the Fetish of Realism," in W. Gibbons and M. Grimshaw-Aagaard, editors, *The Oxford Handbook of Video Game Music and Sound*, Oxford University Press, 2024.

[24] Collins, K., "Making gamers cry: mirror neurons and embodied interaction with game sound," in L. Roque, editor, *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, pp. 39–46, ACM, Coimbra, Portugal, 2011.

[25] Westerberg, A. and Schoenau-Fog, H., "Categorizing video game audio: An exploration of auditory-psychological effects," in M. Turunen, editor, *Proceedings of the 19th International Academic Mindtrek Conference*, pp. 47–54, 2015.

[26] Jørgensen, K., "Time for new terminology?: Diegetic and non-diegetic sounds in computer games revisited," in M. Grimshaw, editor, *Game sound technology and player interaction: Concepts and developments*, pp. 78–97, IGI Global Scientific Publishing, 2011.

[27] Huiberts, S. and van Tol, R., "IEZA: A Framework for Game Audio," *Game Developer* (formerly Gamasutra), 2008, [Online]. Available: https://www.gamedeveloper.com/audio/ieza-a-framework-for-game-audio. Accessed: 2025-08-13.

[28] Garner, T. A., "Cinematic sound design for players," in M. Filimowicz, editor, *Foundations in Sound Design for Interactive Media*, chapter 1, pp. 1–23, Routledge, New York, 2019.

[29] Kenwright, B., "There's more to sound than meets the ear: Sound in interactive environments," *IEEE Computer Graphics and Applications*, 40(4), pp. 62–70, 2020.

[30] Grimshaw-Aagaard, M., "The Importance of Sound to the Formation of Presence," in W. Gibbons and M. Grimshaw-Aagaard, editors, *The Oxford Handbook of Video Game Music and Sound*, Oxford University Press, 2024, online edn., Oxford Academic, 20 June 2024.

[31] Serafin, G. and Serafin, S., "Sound design to enhance presence in photorealistic virtual reality," Georgia Institute of Technology, Sydney, Australia, 2004.

[32] Kern, A. C. and Ellermeier, W., "Audio in VR: Effects of a soundscape and movement-triggered step sounds on presence," *Frontiers in Robotics and AI*, 7, p. 20, 2020.

[33] Sinclair, J.-L., *Principles of game audio and sound design : sound design and audio implementation for interactive and immersive media*, Routledge, New York, NY, 1st edition edition, 2020.

[34] Gospodarek, M., Genovese, A., Dembeck, D., Brenner, C., Roginska, A., and Perlin, K., "Sound design and reproduction techniques for co-located narrative VR experiences," in A. Andreopoulou and B. Boren, editors, *Audio Engineering Society Convention 147*, AES, 2019.

[35] Bridgett, R., *Leading with sound: Proactive sound practices in video game development*, Focal Press, London, 2021.

[36] Bridgett, R., "The big four, Understanding the four food-groups of audio (and how they interrelate)," in *Leading with Sound, Proactive Sound Practices in Video Game Development*, chapter 1, Focal Press, London, 1 edition, 2021.

[37] Bridgett, R., "What drives the mix of your game?" in *Leading with Sound, Proactive Sound Practices in Video Game Development*, chapter 2, Focal Press, London, 1 edition, 2021.

[38] Bridgett, R., "Taking a psychological approach to sound categorisation," in *Leading with Sound, Proactive Sound Practices in Video Game Development*, chapter 3, Focal Press, London, 1 edition, 2021.

[39] Kellaway, S.-A., "Virtually Replacing Reality: Sound Design and Implementation for Large Scale Room Scale VR Experiences," in A. Mayo and L. Gedemer, editors, *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality*, AES, Los Angeles, CA, USA, 2016.

[40] Walther-Hansen, M. and Grimshaw-Aagaard, M., "Don't extend! reduce! the sound approach to reality," in K. Groß-Vogt and R. Höldrich, editors, *Proceedings of the 15th International Audio Mostly Conference*, pp. 8–15, 2020.

[41] Samuels, C. T., "Encountering Directors: Interview with Robert Bresson," `https://web.archive.org/web/20021130085634/http://members.bellatlantic.net`

/~vze25jh7/, 1970, archived web version accessed 2026.

[42] Bosman, I. d. V., Tordiffe, R., Huang, T., Buruk, O., Jørgensen, K., and Hamari, J., "Soundtrek: A virtual reality audio perspective-switching game," in O. O. Buruk, G. Freeman, and J. Hamari, editors, *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play*, pp. 352–358, 2024.

[43] Garner, T. A., *Echoes of Other Worlds: Sound in Virtual Reality: Past, Present and Future*, Springer, Cham, Switzerland, 2017.

[44] Deacon, T. and Barthet, M., "Spatial Design Considerations for Interactive Audio in Virtual Reality," in M. Geronazzo and S. Serafin, editors, *Sonic Interactions in Virtual Environments*, pp. 181–217, Springer International Publishing, 2023.

[45] Jain, D., Junuzovic, S., Ofek, E., Sinclair, M., Porter, J., Yoon, C., Machanavajhala, S., and Morris, M. R., "A Taxonomy of Sounds in Virtual Reality," 2021.

[46] Hillman, N., *Sound for Moving Pictures: The Four Sound Areas*, Routledge, London, 2021.