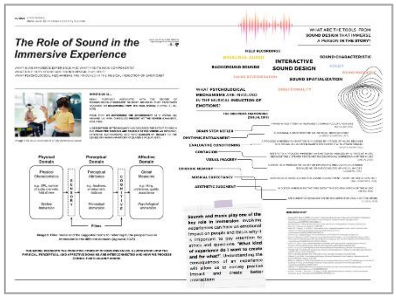
AVANT, Vol. XIV, No. 3, ISSN 2082-6710
https://doi.org/10.26913/ava3202406

The Role of Sound in the Immersive Experience
Poster article

Volha Saroka
Maria Curie-Sklodowska University, Lublin
soroka.olga.mih@gmail.com



Received 2023-07-05; accepted 2023-09-14; published Online First 2024-05-21.

Abstract

To create an immersive experience, it is not important to have a plethora of stimuli for different sensors, but to have their coherence, or in other words, multi-sensory integration. This coherence becomes the necessary basis for immersing oneself in a story, a game, a movie, or any other experience. Sound plays a key role in creating such an experience, both by accompanying the action and emphasizing the narrative flow, and as an independent actor that takes a central role and directs attention. The aim of the paper and poster is to explore the means by which sound can influence immersion in computer-generated virtual environments and the psychological mechanisms involved in the musical induction of emotions. The main question to be answered is how these techniques and models can be used to create a sense of immersion. The first part of the article examines immersiveness itself. The second part is devoted to the practical aspects of sound design for immersive technologies, describing the tools and techniques used to create realistic and compelling soundscapes. The third part touches briefly on the psychological mechanisms behind the induction of musical emotions. Taking a holistic view of these components and using them to reveal each other is how sound can create a sense of presence and place.

Keywords: immersion, experience, virtual reality, sound design, cognitive approach

1. Concepts of Immersion

The concept of immersion has emerged as a fascinating area of study. Immersion refers to a mental state of deep absorption and a sense of total involvement, often characterized by a sense of presence, time distortion, and heightened sensory perception that makes the individual more attuned to their surroundings. Immersion can manifest itself in various contexts and situations, such as media consumption, artistic creation, games, movies, sports, learning, social interactions, and more. What these contexts have in common is a deep engagement with the moment. For the purposes of this article, I will focus only on computer-generated virtual environments, where many parameters can be manipulated and controlled to design and simulate a particular experience.

The main goal of this article is to explore the role that sound plays in immersion and how it contributes to such experiences. Before we delve into answering these questions, let's provide some definitions of immersion and highlight the components of the experience that allow us to use the term.

According to Turner et al. (2016), immersion is "positively associated with the degree of technologically mediated sensory richness that facilitates isolation or decoupling from the real world." In this definition, the focus is on the extent to which technological enhancements provide an enhanced sensory experience that allows individuals to disconnect from their immediate physical environment and fully immerse themselves in an alternate, digitally mediated reality.

Another definition provided by Chalmers (2022, p. 148) is that "immersive means that we experience the environment as a world around us, with ourselves at the center. Here, the emphasis is on the comprehensive perception of the environment as an all-encompassing world in which the individual occupies a central and active position, fostering a profound sense of presence and engagement within the immersive space.

Both perspectives have cognitive connotations that emphasize the role of attention and perception as the primary gateways to the immersive experience, and both definitions underscore the idea that the individual becomes fully absorbed and present within an alternate reality. Another view leads to the notion of immersion as an objective property of the system or technology that facilitates the experience (Agrawal, 2022). Slater and Wilbur (1997) proposed that "immersion is a description of a technology and describes the extent to which computer displays are capable of providing a comprehensive, extensive, surrounding, and vivid illusion of reality to the senses of a human participant. They also suggest that a sense of presence may be closely related to immersion, representing a particular state of consciousness associated with a sense of being in a place (Slater & Wilbur, 1997).

These are just a few perspectives on immersion as a particular mental state, but they help us to see what is important. Taken together, they emphasize the complexity of the phenomenon and suggest that several factors contribute to the subjective feeling of being present and immersed in a virtual environment, including the objective characteristics of the technologies used. This leads us to the next question: what technical tools can we use to achieve this effect in practice? Let's return to the experience constructed in a computer-generated virtual environment and use sound to address these questions.

Virtual reality technology, with its use of input and output devices, software, and technical capabilities to create an audio-visual experience, becomes a powerful tool that allows us to manipulate the human senses and potentially create a subjectively immersive experience. Video sequences influence visual perception, while audio sequences influence sonic perception. In addition, storytelling can create context, and interactions increase the sense of participation. When an experience is presented, we perceive it as a whole, so it's not just about the sum of these elements. What matters is creating a sense of completeness and coherence in the experience for immersion.

The subjectively perceived continuity and coherence allows us to immerse ourselves in a virtual environment and follow the presented sequence with each element of the audiovisual storytelling. If the role of sound in creating immersion is crucial, what is the mechanism behind it? The filter model provided by Agrawal (2022) offers a possible answer. In the following sections of this article, I will provide specific examples of how this model can be applied to create immersive experiences through sound.

The model (Figure 2 on the poster) presents the cognitive concept of immersiveness, illustrating how the physical, perceptual, and affective domains are interrelated and how we process stimuli, particularly sound.

> The model starts with the physical domain, which contains a physical stimulus (e.g. a music signal played by a loudspeaker). The stimulus is characterized by the physical measurements of audio frequency content, spatial audio channels, etc. After passing through the sensory filter, the stimulus is perceived when it is transformed by the sensory system (e.g., the auditory system) into neural energy. The result is an auditory event consisting of attributes of sound (e.g., loudness, envelopment). Elicitation of the attributes and their strength depend on the characteristics of the physical stimulus and the sensory system. The auditory event can be evaluated by perceptual measurements in the perceptual domain. Finally, to form an overall impression of the auditory event, the perception passes through the cognitive filter, which takes into account emotional state, expertise, expectations, mood, context, etc. The cognitive factors and the individual attributes from the perceptual domain contribute to the overall impression, which requires an integrative state of mind. These affective or hedonic measures include the evaluation of quality, degree of liking, annoyance, acceptance, etc. (Agrawal, 2022).

The model guides us step-by-step through how environmental sound travels through various domains and filters to become a meaningful musical event that affects us. Thus, if we want to create an immersive experience, we can observe how an individual reacts and responds to specific sound stimuli, identify which filters the stimuli pass through, and explore why certain filters are not engaged.

2. Sound Design Tools to Create Immersive Experiences

The creation of immersive experiences relies heavily on the skilled use of sound design tools to transport people into compelling artificial environments. These tools offer a range of capabilities, from manipulating audio elements to creating spatial audio effects. Sound design tools allow creators to shape the soundscape, which includes audio elements such as music, sound effects, and spatial audio. These tools, in turn, are used to enrich narrative dimensions, evoke emotional responses, and enhance the palpability of virtual presence.

Sound design is mostly a creative task with high technical requirements that have improved and evolved over time. From the era of silent cinema to video games, sound design has adapted its technical and creative strategies to the transformations of the media in order to convey a compelling narrative (Salselas, 2020). With this in mind, we will outline some key tools for constructing an immersive encounter.

- Binaural audio. It is critical to creating a sense of spatialization and depth in immersive experiences. It involves capturing or creating audio that mimics the way the human ear perceives sound in three dimensions. Binaural recording and playback techniques, such as the use of special microphones and headphones, can enhance the realism of the sound experience.
- Field recording refers to a tool used to capture audio outside of a studio environment. These recordings include a variety of sources, both natural and man-made, including sounds such as water flowing, wind, birdsong, crowds, traffic, etc. Field recordings create the effect of presence and authenticity, transporting the listener to a specific place and time. By playing back such audio recordings, the listener can experience the feeling of being in the place where they were made.
- Foley refers to the recreation of sound effects from the real world. These sound effects are used in movies, TV shows, video games, and music productions. Examples of foley sound effects include footsteps, the sound of clothes moving, doors opening, glass breaking, and more. This practice adds authenticity and a sense of presence to immersive content.

- Sound spatialization. This is a way to enhance immersion by creating a realistic auditory environment where sounds can be perceived as coming from specific directions and distances. Sound spatialization adds fidelity to the location of sound sources and thus to the of sound sources and thus to the experience of space (Begault, 2000; Wenzel, 1988).
- Sound reverberation also contributes to a more realistic acoustic experience and appropriation of virtual space. As known in psychoacoustics, reverberation and early reflections are responsible for the sensation of auditory "spaciousness" and are one of the main cues responsible for the externalization of acoustic images (Begault, 1992, 2000).
- Directionality. Sounds often have a specific direction in which they are louder/unfiltered. Few sounds actually radiate equally in all directions. Think of your voice, for example, or a bullhorn. Using these controls can greatly enhance the "reality" of your soundscape (Stevens & Raybould, 2011, p. 103).
- Background sounds produced by objects in the environment and by the user interacting with the objects also play an important role in the sense of being part of the environment. This acoustic information is perceived on a primitive cognitive level, without conscious awareness of it, but it is crucial for the experience of the acoustic space (Pellegrini, 2011).
- Interactive sound design. This involves dynamically changing or adapting the sound based on the person's actions or inputs. This could include triggering specific sounds, adjusting volume levels, or changing the mix based on the user's position or behavior.

Certainly, the range of tools discussed is not exhaustive, and the selection of a particular tool will depend on the unique characteristics of the intended experience. For example, consider the scenario of immersing an individual in a virtual forest environment. In this context, the use of authentic field recordings from an actual forest can be instrumental. To further enhance immersion, incorporating sound spatialization techniques can simulate the sensation of walking through the virtual forest. Another example is the use of virtual reality as a therapeutic platform to alleviate conditions such as PTSD. This is where the use of Foley technology becomes a valuable asset. Foley technology excels at reproducing the authentic sounds inherent in different locations, contributing to a heightened sense of realism and immersion.

The use of software and sound design tools allows us to modify the stimulus and adjust its strength, volume, tone and other parameters to find those that are appropriate for the particular experience. In this way, the technical side helps to regulate the physical and perceptual domains, which in turn are responsible for system and perceptual immersion. Now let's discuss how we can approach the affective domain and what might contribute to psychological immersion.

3. Psychological Mechanisms Underlying the Musical Induction of Emotions

The study of the psychological dimensions inherent in immersive experiences explores the interplay between these technologies and our perceptual, cognitive, and emotional states. A critical point in this exploration is to understand how sound and music can induce emotions in immersive environments.

Mechanisms of emotion induction are viewed as information processing devices at different levels of the brain that use different types of information to guide future behavior (Juslin, 2011). It is further hypothesized that sound perception has survival value in part due to its ability to activate these mechanisms (Juslin, 2013). Based on this understanding, and in order to answer the question: how does music induce emotions, let's take a closer look at the BRECVEMA framework (Juslin, 2013; Juslin & Västfjäll, 2008).

> The BRECVEMA framework takes an evolutionary perspective. An evolutionary perspective on human sound perception suggests that the survival of our ancient ancestors depended on their ability to detect patterns in sounds, derive meaning from them, and adjust their behavior accordingly. From this assumption, it is theorized that there are multiple emotion mechanisms implemented by a number of distinct "brain networks" that have developed gradually and in a particular order during evolution, from simple reflexes to complex judgments. Each mechanism responds in its own unique way to specific configurations of information in the music, the listener, and the situation, collectively referred to as the "musical event" (Juslin, 2015).

The framework features eight mechanisms:

- Brain stem reflex: a hard-wired attentional response to simple acoustic features such as extreme or increasing loudness or speed. Brain stem reflexes to music are based on the early stages of auditory processing (Sokolov, 1963).
- Rhythmic entrainment: a gradual adjustment of an internal body rhythm (e.g., heart rate) to an external rhythm in the music (Harrer & Harrer, 1977).
- Evaluative conditioning: refers to a process in which an emotion is elicited by a piece of music simply because that stimulus has been repeatedly paired with other positive or negative stimuli (Blair & Shimp, 1992).
- Contagion: refers to a process whereby an emotion is induced by a piece of music because the listener perceives the emotional expres-

sion of the music and then "imitates" that expression internally, resulting in the induction of the same emotion either through peripheral feedback from muscles or through more direct activation of the relevant emotional representations in the brain (Juslin, 2000).

- Visual imagery: refers to a process whereby an emotion is induced in a listener because he or she conjures up visual images (e.g., of a beautiful landscape) while listening to music (Osborne, 1980).
- Episodic memory: a conscious recollection of a particular event from the listener's past that is triggered by the music (Baumgartner, 1992).
- Musical expectancy: refers to a process whereby an emotion is induced in a listener because a particular feature of the music violates, delays, or confirms the listener's expectations about the continuation of the music; an emotion is induced when expectations about the continuation of the music are violated (Meyer, 1956).
- Aesthetic judgment: a subjective evaluation of the aesthetic value of the music based on an individual set of weighted criteria (Juslin, 2013).

All of the mechanisms in the BRECVEM framework (and aesthetic judgments, discussed later) have in common that they involve psychophysical relationships between "external" features of the environment (i.e., the music and the context) and "internal" features of the perceiver. A particular piece of music might contain certain types of information, but whether that information activates a mechanism also depends on the characteristics of the listener and the context (Juslin, 2013).

In the context of immersive experiences, the BRECVEMA framework provides a comprehensive perspective on how sound and music can contribute to eliciting emotions in virtual environments. As we can see, following this framework, it is important to consider both the audience and the context when creating a virtual experience. All of this increases the chances of influencing the affective domain and creating musical events that have an impact on psychological immersion.

4. Conclusion

Now that we have explored the concept of immersion, considered the tools within the sound design arsenal for constructing immersive environments, and delved into the psychological mechanisms that underlie the induction of emotions through music, a central question remains: how can sound create an immersive experience? The poster and article present key elements: by thoughtfully orchestrating sound design tools within an understanding of the psychological mechanisms behind musical emotions, and by paying attention to audience, context, storytelling, and interactions, it is possible to transport individuals into captivating environments and create a sense of immersion. The role and power of sound is critical, and it can completely direct a person's attention, provide feedback about the environment, emphasize context, or evoke specific emotions. The choice of how to use it depends on the specific virtual environment and goals.

References

Agrawal, S.R. & Bech, S. (2023). Immersion in Audiovisual Experiences. In: Geronazzo, M., Serafin, S. (eds.). Sonic Interactions in Virtual Environments. Human–Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-031-04021-4_1

Baumgartner, H. (1992). Remembrance of things past: Music, autobiographical memory, and emotion. *Advances in Consumer Research*, 19, 613–620.

Begault, D. R. (1992). Perceptual effects of synthetic reverberation on Three-Dimensional audio Systems. *Journal of the Audio Engineering Society*, 40(11), 895–904. https://www.aes.org/e-lib/browse.cfm?elib=7027

Begault, D. R. (1995). 3-D sound for virtual reality and multimedia. *Computer Music Journal*, 19(4), 99. https://doi.org/10.2307/3680997

Blair, M. E., & Shimp, T. A. (1992). Consequences of an unpleasant experience with music: A second-order negative conditioning perspective. *Journal of Advertising*, 21, 35–43. https://doi.org/10.1080/00913367.1992.10673358

Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. Penguin UK.

Geluso, P. (2018). *Immersive sound: The art and science of binaural and multichannel audio*. New York, NY; Abingdon, Oxon: Routledge, an imprint of the Taylor& Francis Group, 2018.

Harrer, G., & Harrer, H. (1977). Music, emotion, and autonomic function. In M. Critchley & R. A. Henson (Eds.), *Music and the brain: Studies in the neurology of music* (pp. 202–216). London, England: William Heinemann. https://doi.org/10.1016/B978-0-433-06703-0.50019-X

Juslin P. N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Physics of life reviews*, 10(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. (2011). *Music and emotion: seven questions, seven answers*. In Oxford University Press eBooks (pp. 113–136). https://doi.org/10.1093/acprof:osobl/9780199581566.003.0007

Juslin, P. N. (2000). Vocal expression and musical expression: Parallels and contrasts. In A. Kappas (Ed.), *Proceedings of the Sixteenth Conference of the International Society for Research on Emotions* (pp. 281–284). Quebec City, QC: ISRE Publications.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559–621. https://doi.org/10.1017/S0140525X08005293

Juslin, P., Barradas, G., & Eerola, T. (2015). From Sound to Significance: Exploring the mechanisms underlying emotional reactions to music. *American Journal of Psychology*, 128(3), 281–304. https://doi.org/10.5406/amerjpsyc.128.3.0281

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: Chicago University Press.

Osborne, J. W. (1980). The mapping of thoughts, emotions, sensations, and images as responses to music. *Journal of Mental Imagery*, 5, 133–136.

Pellegrini, R.S. (2001). Quality assessment of auditory virtual environments. *Proceedings of the 7th international conference on auditory display. Laboratory of acoustics and audio signal processing and the telecommunications software and multimedia laboratory* (pp. 161–168).

Salselas, I., Penha, R. & Bernardes, G. *Sound design inducing attention in the context of audiovisual immersive environments*. Pers Ubiquit Comput 25, 737–748 (2021). https://doi.org/10.1007/s00779-020-01386-3

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6), 603-616. https://doi.org/10.1162/pres.1997.6.6.603

Sokolov, E. N. (1963). Higher nervous functions: the orienting reflex. *Annual Review of Physiology*, 25, 545–580. https://doi.org/10.1146/annurev.ph.25.030163.002553

Stevens, R., Raybould, D. (2011). *The Game Audio Tutorial. A Practical Guide to Creating and Implementing Sound and Music for Interactive Games*. New York: Routledge Roginska, A., In

Turner, S., Huang, C., Burrows, L., & Turner, P. (2016). Make-Believing virtual realities. In *Human-computer interaction series*. https://doi.org/10.1007/978-3-319-29553-4_3

Wenzel, E. M., Wightman, F. L., Kistler, D. J., & Foster, S. H. (1988). Acoustic origins of individual differences in sound localization behavior. *Journal of the Acoustical Society of America*, 84(S1), S79. https://doi.org/10.1121/1.2026486