A Pilot Study of the Relationship Between 3D Background Music and the Players' Perceptual Levels of Immersion in Game and Virtual Environments

Yaojia Chen[1], Wenzheng Cui[1]

[1]New York University

Correspondence should be addressed to Wenzheng Cui (wc2599@nyu.edu)

ABSTRACT

This study investigates the impact of 3D background music on player immersion in video games, focusing on the role of 3D audio in increasing the level of time loss and the level of player engagement. Building on the premise that ambient sounds and background music significantly influence gaming immersion [2,5,6], the research aligns with the increasing demand for advanced audio technologies in various applications [1]. Within the scope of game audio applications, 3D audio and channel-based audio are collaboratively employed to craft immersive virtual environments [11], yet the background music continues to be presented in stereo format. This observation prompts the primary inquiry and rationale behind this paper: does the incorporation of 3D background music enhance player immersion? A preliminary experiment has been conducted to analyse participants' game experiences differentiated by stereo and 3D background music formats in order to evaluate the feasibility of the future study. The outcomes of this pilot study reflect on the nuance made by 3D background music in video games on player immersion.

1. INTRODUCTION

The role of audio in games has evolved drastically in the last few decades. The transformation from creating simplistic sounds to integrating more information and enhancing the immersive experience has made game audio a crucial topic to discuss. The introduction of 3D sounds has arguably contributed to game immersion improvement [7]. To investigate the impact of 3D mixing format on player immersion level, it is essential to examine the role of 3D audio in contemporary video games. The implementation of 3D audio is crucial for crafting ambient sounds, environmental elements, and specific in-game events, providing spatial information and enhancing the player's sense of immersion [2]. The aspects of immersion being investigated in this study are the level of time loss while playing video games and the level of player engagement. Background music (BGM) has also significantly contributed to increasing the immersion level of video games as it has been recognised as a tool for immersion enhancement, especially when the background music is highly related to the game's contents [5,6].

The relevance of this investigation is underscored by the increasing demand for efficient algorithms and methodologies in applications like gaming, 3D video and audio reproduction, immersive teleconferencing,

and virtual and augmented reality [1]. Furthermore, the study aligns with previous research highlighting the significant role of background music in the immersive quality of video games [3]. The potential of the Dolby Atmos format to enhance players' immersive experience represents a novel implementation of affective audio technologies in video games, contributing to discourses in this area of research [4].

2. PLAYER IMMERSION

Immersion is a vague terminology that has not yet been clearly defined or standardised. Researchers have proposed several immersion models and discussed factors that affect the immersive experience. The similar terminologies that are related to immersion are sometimes used interchangeably which leads to an ambiguous definition in the standardisation of immersion in academia. Although different models of immersion have been proposed [8], this study uses specific narrowed-down factors to examine the impact of 3D background music on players' immersive experience in a video game. The level of time loss represents to what degree players lose the sense of time while playing video games. Researchers have found that time loss is a common phenomenon that happens to the majority of video game players [9, 10]. The level of player engagement is the other key factor that is used in this pilot study to examine the immersive experience.

3. 3D AUDIO

3D audio refers to audio technologies that simulate a three-dimensional auditory environment. Examples of 3D audio technologies are Ambisonics and Object-based audio, which are sometimes used in combination with 2D sounds (stereo and multi-channel audio) in video games to create an immersive environment [11].

3.1 Localisation Cues & HRTF

According to the duplex theory (Lord Rayleigh, 1907), two localisation cues, Inter-aural time difference and Inter-aural level difference are used for sound source localisation. Inter-aural time difference (ITD) refers to the time difference it takes for sound to reach both ears while Inter-aural level difference (ILD) measures the difference in the level of sound reaching both ears. It was proposed that ILDs, resulting from the shadowing effect of the head, are key in determining the localization of high-frequency sounds (approximately 2000 Hz in individuals with larger heads). Conversely, ITDs were believed to be

more crucial for low-frequency sounds, owing to phase ambiguities that arise at frequencies above about 1000-1500 Hz [13].

3DOF (three degrees of freedom) head trackers measure the orientation of the head (pitch, yaw, roll). 6DOF tracker measures the position of the listener (x, y, z), in addition to the orientation of the head [16].

Filters developed from ear-specific traits are Finite Impulse Responses (FIR) and are referred to as Head-Related Impulse Responses (HRIRs) in the time domain, or Head-Related Transfer Functions (HRTFs) in the frequency domain. Using these HRTF-based filters to process any sound can give it spatial qualities. Additionally, if this filtering is done in real-time, it can replicate the effects of both the sound sources' movement and the moment of the listener's head [13].

3.2 3D Audio Technologies

Ambisonic technology is a full-sphere, scene-based audio technology that captures and recreates the sound field using a combination of signals from a set of microphones arranged in a specific pattern, representing sound equally in all directions [13,17].

Object-based audio is a technology where sound sources are treated as individual sound objects in a three-dimensional space [17]. Sound objects have been defined as combinations of audio elements (audio waveforms) and related parameters (metadata) that represent the artistic purpose by specifying how these audio elements are transformed into loudspeaker outputs. In comparison to a pre-mixed audio delivery for a specific loudspeaker configuration, object-based audio provides more adaptability to various playback systems and the possibility to deliver a more immersive audio experience [13].

3.3 3D Audio Implementation in Video Games

In the context of game audio applications. 3D audio and channel-based audio are used in conjunction with each other to create compelling virtual environments. Since Ambisonic technology allows users to efficiently capture and render a full-sphere environment, it is well suited for non-primary audio sources such as the ambiences in games. Object-based audio, in most cases, is suitable for the sounds that need to be localised by the player, an example would be the enemy's weapon sound, which needs to

be localised in 3D by the player to provide crucial information in the game. Stereo audio files are often used for non-directional sounds such as in-game announcements, critical dialogue, UI (User Interface) sounds and music [11].

4. BACKGROUND MUSIC

Music in video games plays an important role in contributing to the players' immersion [12]. Prior Scholar work by Scott D. Lipscomb and Sean M. Zehnder has been conducted in the past that investigates the impact of musical soundtracks on the players' experience of playing video games. In Lipscomb and Zehnder's research, participants were divided into three groups: playing a game with music, playing a game without music, and only listening to music. Afterwards, the findings showed that while not all scales showed significant changes, some did exhibit statistically significant differences when a musical score was provided [15].

Stereophonic sound has long been the standard format for popular music, however, the 2019 launch of Dolby Atoms Music on Tidal HiFi and Amazon Prime Music HD marked a significant step in integrating 3D music into streaming services, reflecting a broader trend of popular music's venture into immersive and interactive audio technologies [14].

2. EXPERIMENT

This preliminary experiment is to evaluate the feasibility of the study on the influence of 3D background music in video games by exploring the impact of different auditory stimuli on participant experiences in a video game setting. The pilot study incorporates three distinct stimuli, each featuring both 3D-mixed and Stereo-mixed Background Music

	Question	Question Type
1	To what degree did you lose a sense of time or did you attend to the game?	Scale (1-10)
2	To what degree were you focused on the game?	Scale (1-10)
3	To what degree did you forget the plan of the day or the current concern?	Scale (1-10)
4	To what degree did you believe to be in the virtual environment in the game?	Scale (1-10)
5	How many hours do you play video games every week?	Short Answer

Table 1. Questionnaire

(BGM), to which all participants are exposed. All BGMs have the same integrated loudness (-18 dB LKFS), sample rate (44.1 kHz), and bit depth (16-bit). The preliminary experiment has been conducted to quantify the variations of players' immersion levels with different background music mixes (3D and Stereo). Twelve individuals (age 18 - age 35) are recruited to participate in all 3 stimuli of the study and most participants have some sort of gaming experience. After each stimulus, the participants are asked to complete a questionnaire.

The questionnaire is composed of 5 questions regarding the level of time loss, the level of engagement, as well as the personal gaming experiences. For all the questions with scales, participants have been informed that all questions refer to the experience of the game with the second demo of the stimulus (3D-mixed background music) in comparison to the first demo of the stimulus (stereo-mixed version). For these questions, "1" means the first demo has a way higher degree while "10" means the second demo has a way higher degree to the factor being examined. To optimise the accuracy of the duration of each stimulus, the background music and the game demo are applied on separate devices. The outputs of the devices are connected to an audio recorder and the output of the recorder is provided to participants. All the channels maintain the same level (40 dB Gain) to avoid the influence of the loudness difference.

Stimulus 1	Q1	Q2	Q3	Q4
	8	9	7	6
	7	6	8	8
	7	8	5	6
	8	8	7	5
	7	7	6	8
	4	5	5	5
	6	7	6	7
	8	8	9	9
	9	8	7	9
	7	8	7	6
	7	8	6	4
	8	5	3	5
Average	7.1667	7.2500	6.3333	6.5000
Standard Deviation	1.2673	1.2881	1.5570	1.6787

Table 2. Analysis of Stimulus 1.

Stimulus 2	Q1	Q2	Q3	Q4
	6	9	6	8
	8	8	8	6
	7	8	7	6
	5	4	5	5
	7	6	7	4
	5	4	4	5
	4	4	5	5
	7	7	7	9
	7	5	8	5
	8	9	8	6
	8	8	8	9
	4	5	3	6
Average	6.3333	6.4167	6.3333	6.1667
Standard Deviation	1.4974	1.9752	1.7232	1.6422

Table 3. Analysis of Stimulus 2.

Stimulus 3	Q1	Q2	Q3	Q4
	8	9	7	8
	8	7	7	7
	7	8	8	9
	6	7	6	4
	8	8	8	8
	4	4	5	5
	6	7	5	7
	3	4	4	4
	6	6	6	6
	3	3	8	9
	9	9	9	8
	6	6	4	7
Average	6.1667	6.5000	6.4167	6.8333
Standard Deviation	1.9924	1.9771	1.6765	1.7495

Table 4. Analysis of Stimulus 3.

5. RESULTS

The overall results of the collected data, after rounding to the fourth decimal place, indicate that players generally tend to have a more immersive experience while playing the video game with 3D-mixed background music than with stereo-mixed background music. All the responses gathered from the questionnaire, as detailed in Table 2, Table 3, and Table 4, have an average rating above 6 on a scale from 1 to 10. This indicates that, among all participants, a majority found the second demo (3D-mixed) of the stimuli to be more immersive than the first demo.
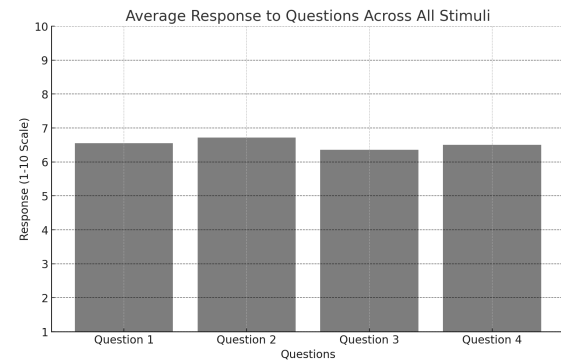

Fig. 1: Average Response to Questions Across All Stimuli

6. DISCUSSION

It is notable that the summation of the responses of all stimuli to each question has an average above 6 on a scale of 1 to 10 across all questions, as shown in Figure 1. While this pilot study indicates a mild tendency that 3D-mixed background music in video games could potentially increase player immersion, the collected data is not sufficient enough to support such a statement. This uncertainty arises from several contributing factors including the limited amount of participants involved in the study, individual preferences for video game types and music genres, and the order in which the demos (3 different stimuli) were presented to the participants. As such, more factors should be considered for future study in order to output more accurate results.

References

1. Belloch, J. A., Ramos, G., Badia, J. M., & Cobos, M. (2020). An efficient implementation of parallel parametric HRTF models for binaural sound synthesis in mobile multimedia. *IEEE Access, 8*, 49562-49573.
2. Broderick, J., Duggan, J., & Redfern, S. (2018, August). The importance of spatial audio in modern games and virtual environments. In *2018 IEEE Games, Entertainment, Media Conference (GEM)* (pp. 1-9). IEEE.
3. Fu, X., & Zhang, J. (2017). The influence of strategy video game and its background music

on cognitive control. *EC Psychology and Psychiatry*, *2*(1), 15-25.

4. McIntosh, T. H., Weinel, J., & Cunningham, S. (2022, September). Lundheim: exploring affective audio techniques in an action-adventure video game. In *Proceedings of the 17th International Audio Mostly Conference* (pp. 155-158).

5. Nelson, C., & Wünsche, B. C. (2007, January). Game/music interaction: an aural interface for immersive interactive environments. In *Proceedings of the eight Australasian conference on User interface-Volume 64* (pp. 23-26).

6. Zhang, J., & Fu, X. (2015). The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy*, *5*(4), 1.

7. Conway, F., & Paterson, N. (2010, October). Spatial audio and reverberation in an augmented reality game sound design. In *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*. Audio Engineering Society.

8. Lee, H. (2020). A conceptual model of immersive experience in extended reality.

9. Sanders, T., & Cairns, P. (2010). Time perception, immersion and music in videogames. *Proceedings of HCI 2010 24*, 160-167.

10. Wood, R. T., Griffiths, M. D., & Parke, A. (2007). Experiences of time loss among videogame players: An empirical study. *Cyberpsychology & behavior*, *10*(1), 38-44.

11. Sinclair, J. L. (2020). *Principles of game audio and sound design: sound design and audio implementation for interactive and immersive media*. CRC Press.

12. Munday, R. (2007). Music in video games. *Music, sound and multimedia: From the live to the virtua*l, 51-67.

13. Roginska, A., & Geluso, P. (Eds.). (2017). *Immersive sound: the art and science of binaural and multi-channel audio*. Taylor & Francis.

14. Bresler, Z. (2021). Immersed in Pop: 3D Music, Subject Positioning, and Compositional Design in The Weeknd's "Blinding Lights" in Dolby Atmos. *Journal of Popular Music Studies, 33*(3), 84-103.

15. Lipscomb, S. D., & Zehnder, S. M. (2004). Immersion in the virtual environment: The effect of a musical score on the video gaming experience. *Journal of Physiological Anthropology and Applied Human Science, 23*(6), 337-343.

16. Quackenbush, S. R., & Herre, J. (2021). MPEG standards for compressed representation of immersive audio. *Proceedings of the IEEE, 109*(9), 1578-1589.

17. Sun, X. (2021). Immersive audio, capture, transport, and rendering: a review. *APSIPA Transactions on Signal and Information Processing, 10*, e13.