Sonic Scope: New Approaches to Audiovisual Culture •

Fashioning the Immersive Fallacy at Five Nights at Freddy's: A New Approach to Music, Sound, and Their Relationship to the Immersive Process in Moving Image Media

Hannah Capstick

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

Fashioning the Immersive Fallacy at *Five Nights at Freddy's*: A New Approach to Music, Sound, and Their Relationship to the Immersive Process in Moving Image Media

[Hannah Capstick](#), University of Oxford

*Music is a key factor in creating the immersive environment of a film or a game. There is much scholarship exploring this fact, and as a result there are many different terms and emphases in use. To attempt to move closer to an understanding of the specific ways in which music and sound contribute to the immersive experience, we must probe into the terms that we use, our current theories and the mechanics thereof, and explore differing methodologies across a range of genres and formats, to bring us closer to a more full understanding of these immersive processes, and why we seek them out and enjoy them. An analysis of *Five Nights at Freddy's* applies pressure to the existing scholarship and forces us to reconsider how we define the concepts necessary to reach an understanding of immersion, and encourages us to seek new methods through which we can apply these to music in video games. I conclude with a suggestion of a new idea – that of the “Global Music Box” : a new approach to the way in which we think of sounds in moving-image media, considering music and sound as equally important and mutually supportive, going as far as to break down the boundary between the two. The concept of the “Global Music Box” recognises that a simple transplantation of our understanding of the relationship between music and sound onto video games (and other forms of moving image media) is problematic, as it does not consider that all sounds in media are artificial, whereas in “real life” we have a distinction between the natural and the artificial. In turn, this has repercussions for our understandings of related concepts, including the diegesis, interactivity, and more broadly definitions of sound and music themselves.*

Introduction

Through an analysis of the *Five Nights at Freddy's* franchise, I will critically approach a selection of the key theories and methodologies in the study of music in video games and demonstrate the value of my concept of a Global Music Box. In contrast to current approaches, which often pick apart the soundworld of a game into microscopic fragments, the Global Music Box considers the sum product of all sounds on a macroscopic level and considers the relationship between all its constituent parts. This is not to invalidate these approaches - indeed, their values certainly speak for themselves - but to argue that there is value in stepping back and seeing the product in its entirety. The GMB is another tool in our analytical toolbelt, so to speak, which can be used alongside existing and future methods of studying music and the moving image. Indeed, the broader perspective afforded us by the GMB can often provide more nuance to the detailed discussions we approach through in-depth analysis of specific features of a game.

First, I will provide an overview and analysis of the *Five Nights at Freddy's* franchise, noting both interesting events and instances that challenge our current understandings of music's role in moving image media. Following a brief review of prominent literature and ideas surrounding sound and music in film and games, I will then discuss these sources in relation to my analysis of *Five Nights at Freddy's* to demonstrate how this franchise applies pressure to our current understandings of music's role in moving image media, exploring how the nascent concept of the GMB can begin to shed further light on the relationships between music, sound, immersion, and the diegesis, as well as other related concepts.

Why *Five Nights at Freddy's*?

Before delving into my analysis, it is important to provide an introduction to the games of the *Five Nights at Freddy's* franchise (hereafter referred to as '*FNaF*'), and to note what elements of the franchise specifically attracted my attention. The concept of the series is simple. You play as an unnamed security guard, working the nightshift at Freddy's Pizzeria. The singing animatronics that entertain the children during the day have not been given a proper night mode, and, as a result, attempt to enter the security office and kill you on shift. Each entry in the *FNaF* series largely follows this formula and, while each sequel introduces new aspects and nuance to the story, such variation is not the focus here. There are eight main games in the franchise, all existing within the same

universe with an overarching narrative that can be traced between the games; however, the games can be played without needing to closely follow the main story. Since the first game's release in 2014, the series has gained a cult following, in no small part due to "Let's Plays" and analysis videos, such as those by YouTube channels Markiplier and The Game Theorists, whose most popular videos on the game have amassed 92 million and 27 million to date, respectively.¹ The majority of the games are played in a stationary point-and-click format (again, there are exceptions to this, but is not the point of this discussion) while the player must rely upon a combination of visual and auditory clues to survive for five nights.

The interesting role of music and diegetic sound in the series is what initially drew me to study these games, given their importance in guaranteeing the player's survival. Particularly in the later games in the series, such as the eighth instalment of the franchise, *Ultimate Custom Night*, the player must take in an incredible amount of information from a combination of visual and audio cues to survive. Indeed, there are many audio-only cues making the act of (active) listening necessary for survival and thus an integral part of the horror experience. In addition to the importance of music in terms of the gameplay itself, the ambient background sound of the first four games is also of interest. As is often the case in video games, the background music is designed to establish a mood, in this instance one of anxiety, without necessarily being consciously heard. As a result, there is nothing ear-catching about the music – it is simply a low, metallic hum that ebbs in pitch as the game is played. However, placed onto this canvas are seemingly random aural occurrences, such as distant laughter, loud footsteps, or a pipe organ fading in and out. This is to name but a few of the interesting sonic features of the series and I will discuss these instances in more detail below, but they have profound implications for the player's experience and are a prominent reason as to why these games are so scary and immersive for the player.

This immersion is perhaps one of the main reasons behind the popularity of *FNaF*. As such, it is worth quickly discussing why a study on immersion in video games would choose to focus on the horror genre, rather than other highly immersive genres, such as the Role-Playing Game. To create a more affective (i.e., scarier) experience, the player must truly believe they are part of the fictional world, and thus are genuinely under threat. In achieving such immersion, a network of normative methodologies are used, and in turn subverted, thereby creating a game that balances the player's expectations and desire for novelty, and an unpredictable, thus scarier, experience. As we will see, the horror genre often relies on the subversion of expectation and common tropes to create uncertainty in the player, which is key to creating a scary experience. While there are

many common tropes in the horror genre which are frequently employed with great success, this subversion extends to the ‘double-bluff,’ wherein horror tropes are subverted to further the anxiety and uncertainty created within the horror experience.

Current Academic Understandings

The current literature discussing the role of music and sound in enhancing player immersion should be critically discussed before becoming fully delving into an analysis of *Five Nights at Freddy's*. Music is key to the process of immersion in moving image media such as film and video games. Through music, an individual can lose track of real life and be transported into the game world. There has been much scholarship and literature exploring the term “immersion,” and as such, many uses and definitions have been ascribed to it. While there are generally agreed upon broad strokes that contribute towards a definition of immersion (which will be discussed below), there are also alternative usages and understandings that diverge from the common understanding. To reach an understanding of immersion, and music’s role in the process of immersion, we must also address the concepts (and their related concepts) of the diegesis and interactivity. Ultimately, it will be demonstrated that we are broadly in agreement as to the processes of immersion; though the terminology may vary, the concepts remain the same.

Claudia Gorbman discusses the role of music in film, and the techniques by which music contributes to an immersive experience in the introduction to her book, *Unheard Melodies: Narrative Film Music*.² Here, she states that music signifies according to musical cues, and cultural and cinematic codes. These signifiers are subsequently exploited by filmmakers, who use these associations to create an impactful and immersive experience. However, these cues and codes must first be recognised by the listener to work. Gorbman states that the “core musical lexicon has tended to remain conservatively rooted in Romantic tonality, since its purpose is quick and effective signification to a mass audience.”³ However, she goes on to relate this to the concept of “Muzak,” and notes that film music is not designed to be listened to, rather it relies on familiar language, bathed in Romantic affect, to reduce critical distance and subdue listeners.⁴ However, the horror genre and video game music requires more nuance from Gorbman’s discussion.

Firstly, the difference between *hearing* and *listening* should be noted. As Kristine Jørgensen puts it, “[t]here is a perceptual difference between these in the sense that *hearing* is an unintentional activity while *listening* is intentional and focused towards specific sounds.”⁵ To expand on this, we are not consciously listening to the music as we watch a film or play a game, but the emotional affect it elicits is still registered as we unconsciously hear it. It primes for certain emotions, as is discussed above, which reduces critical distance between the viewer/player and the fictional world, and thus contributes to the immersive process. I would argue that, in the case of video game music, there should be additional nuance surrounding music that subdues the audience. Indeed, the soundtrack (or the background-level sound) may have a subduing effect; however, the players exist in active participation in a game world, which they do not do while watching a film. Where does this come from? My Global Music Box theory comes into play here. In discussing immersion in a game, we can see the contrasting roles of sounds that passively immerse the player (those sounds that are heard but not listened to) and sounds that actively immerse the player (such as graphical user interface sounds, or instructional sounds; sounds that increase the sense that the player is interacting with the virtual world). This concept sits across boundaries of the diegesis, sounds, and music, which makes their consideration through the GMB useful. Here we see that, regardless of their specific definitions and purposes, all sounds are serving the process of immersing the player in the game.

Finally, Gorbman argues that music can simultaneously be part of and comment upon on the narrative of the film.⁶ I would argue that this supports my theory of a Global Music Box. The suggestion that the music and sound of a game can be both diegetic *and* extradiegetic spaces - perhaps existing across many of these spaces at the same time - presents a strong argument for in-depth analyses with broader, universal considerations of the roles of sound, as well as discussions of the implications of a sound existing within multiple spaces simultaneously or the function of a sound moving across or between these spaces throughout a game. Examples that demonstrate the significance of this will be discussed below.

To borrow from Robynn Stilwell, the player character is the “bridging mechanism between the audience [or player] and the diegesis as we enter into his or her subjectivity” in a place “beyond empathy.”⁷ This pushes against our understanding of the diegetic and non-diegetic poles, which, as Stilwell claims, exists only “‘behind’ the screen.”⁸ Yet, as she goes on to state, “music may osmose through this boundary.”⁹ This is evidenced in the way that music can act in relation to the character on the screen to give an indication as to how the external player should act, thereby, to

reference Stilwell' s terminology, bridging the fantastical gap between the real and virtual world and thus narrowing the critical distance between the player and the virtual world, allowing for immersion within the fiction diegesis.¹⁰

The fact that the boundary between diegesis and non-diegesis is often crossed by video game music does not invalidate it. Stilwell argues that this fluidity “calls attention to the act of crossing and therefore reinforces difference.”¹¹ This “fantastical gap” (or liminal space) between the two binary points of diegetic and non-diegetic represents a gap in our understanding – a place of instability and ambiguity.¹² A related term suggested by Kristine Jørgensen is that of the “transdiegetic space” ; the conceptual space that exists when in-game communications questions the boundaries of the virtual world.¹³ There are many possible subdivisions within this, including in-game sounds that communicate information to the external player (such as the sound of the player character' s thoughts or internal monologue), and extradiegetic accompaniment (such as the soundtrack). Ultimately, however, the combination of the diegesis, the player interface, and the extradiegesis (what Jørgensen defines as a communicating aesthetic feature with no direct reference to an actual source in the medium) leads us to this transdiegetic space.¹⁴ Evidently, the concept and understandings of the diegesis are incredibly diverse and nuanced, and we could further follow Jørgensen into increasingly specific terminology. However, this would be to digress from the point of this essay. It is more productive to instead note that music can traverse across these boundaries, and in doing so, bring the player with them.

As I alluded to above, the term “immersion” is often avoided or re-termed, and examples of this are discussed in detail in Jørgensen' s thesis. For example, she discusses Marie-Laure Ryan' s notions of “possible worlds,” wherein there exist multiple fictional worlds that exist only as products of human mental activity, what Jørgensen refers to as “satellites around the [real] world.”¹⁵ The process of “recentering” (a counterpart to immersion), is when the player relocates their focus to one of these satellite worlds.¹⁶ Similarly, Jørgensen makes reference to Mihaly Csikszentmihalyi' s discussions of “flow,” wherein a player becomes “engaged in an activity for its own sake,” during the actions of which nothing but the game seems to matter.¹⁷ Glassner describes four “levels” of immersion: curiosity, a casual desire to know; sympathy, where the player begins to see the world through the player-character, and identify with the character; empathy, where the player emotionally bonds with the character; and, ultimately, transportation, where the boundary between the real and the virtual is lost.¹⁸ Indeed, Tim Summers has claimed that the journey through these levels is pleasurable, and even desired by players.¹⁹ Transportation, in this discussion, sounds

incredibly familiar to the movement across the transdiegetic space that we have discussed previously.

While the above discussions point to the nuance and intricacies surrounding the notion of immersion, I will ultimately define it as when a game monopolises the senses and engages the player physiologically to the extent where the player begins to lose track of the real world, and begin to experience a transportation to the fictional world presented. This occurs when players involve themselves on an emotional level with the game, detaching themselves from the real world and submerging themselves in the new world. This is in contrast to Paul Toprac and Ahmed Abdel-Meguid's idea of "presence," which they define as where the player simply perceives a world that exists beyond the specific remit of their sensory organs.²⁰ Yet, while the idea that a reduced critical distance between the player and the game, is agreed to be "immersion" (or whichever similar term is used by the respective authors), there is an additional problematic concept - the "immersive fallacy."²¹ This is the understanding that a player can become so immersed in a game, they forget that they are playing a game, and lose the relationship between reality and the fictional world. My response is simple: do we care? If a game is so immersive that the player genuinely forgets they are playing a game, is that not indicative of an excellent game (if a fully immersive experience is what is desired of the specific game)?

The control and manipulation of a player's emotional state while playing a game is a key part of what stimulates the immersive process, and it is evident that music's significance in moving image media is in no small part related to its ability to stimulate certain feelings in a listener. Therefore, to understand this process, we must understand the differences between emotions and moods, through which we can begin to theorise as to what processes allow music to have an emotional effect on the listener. As Patrik Juslin discusses in his article "Emotional Responses to Music," emotion is a brief, intense feeling, generally in response to a change in an organism's environment, both the physical changes in the external environment and the internal physiological responses an organism experiences as a result of environmental changes.²² For instance, an emotion, such as fear, can be experienced in response to both an external environmental change (such as a loud noise or seeing a predator) as well as in response to their corresponding physiological processes (such as the increased heart rate or involuntary starts experienced when startled). In contrast, as Bernard Perron has stated, moods are longer lasting, less intense states which encourage specific emotions.²³ Moods encourage us to focus on the stimuli that evokes the assumed emotion, thus generating those emotions more quickly and more effectively. A mood is therefore a presupposition of an emotion.

Where then, does music and sound fit into the process of immersing a player in a fictional world? On the surface, it appears to be a paradox: constant music is not a ‘natural’ experience, so surely the player will be distracted from the immersive experience by a soundtrack or other form of ambient music.²⁴ Yet, as Jørgensen’s 2007 study found, playing games without sound can be shocking, awkward, or even difficult, with some participants likening the lack of sound to being in the dark, becoming blind, or even losing a leg.²⁵ Music, therefore, must play a significant role in the mediation between the real and the virtual in interactive media. As is seen in qualitative studies, such as that by Toprac and Abdel-Meguid, music can easily stimulate both moods and emotions (specifically in the horror genre: anxiety and fear, respectively), and these can be manipulated by game designers to enhance the specific experiences they aim to create for the players.²⁶ To briefly summarise Toprac and Abdel-Meguid’s study: ambient, low volume sound is used to generate anxiety, whereas sudden, high volume sounds generate fear.²⁷

Sound is particularly important for the horror genre: unlike the eyes, the ears unintentionally capture all nearby sounds. When structuring his theories of horror sound design, Perron makes reference to Friedrich Nietzsche, who notes that hearing is an inherent primal fear which developed at night and in the darkness as an additional sense to distinguish danger when the eyes cannot.²⁸ Through this, Perron goes on to state that “the intent of both horror and sound design...is to render an unfamiliar landscape filled with anxiety, fear and dread. Uncertainty is key to unlocking the unconscious and accessing primal terrors.”²⁹ Robynn Stilwell also discusses this, noting how such uncertainty makes us fearfully “look around for the visual grounding” of the threat, which is not only a scary feeling in itself, but also primes us for the inevitable scary experience of seeing the threat for the first time.³⁰ As Stilwell suggests, the only thing worse than knowing you are being hunted is not knowing where the predator is.

In contrast to the above discussion regarding music’s ability to contribute to immersion, music must also be considered in relation to the visual processes it supplements. Indeed, there is scope in my Global Music Box concept for extension into the visual, and perhaps further into the physical actions and intentions of the player to create a fully rounded understanding of the video game experience. The audio of a game works with the visual and the narrative components to form a complete, “compelling product.”³¹ Perron terms this “affective congruence,” and defines it as the “making of affective meaning in both music and visuals.”³² Drawing upon Michel Chion’s concept of “added value,” which he describes as “the expressive and/or informative value with which a sound enriches a given image,” Perron states that music within games adds:³³

Expressive and informative value in which a sound enriches a given image so as to create the definite impression...that this information of expression “naturally” comes from what is seen...added value is what gives the (eminently incorrect) impression that sound is unnecessary, that sound duplicates a meaning until in reality it brings about, either all on its own or by discrepancies between it and the image.³⁴

Much is betrayed when we realise that, when designing sounds for games, we do not necessarily want to use the most realistic sounds. Moreover, we aim to use the most *credible* sounds, for, as Jørgensen asserts, perceptual fidelity is more desirable than auditory fidelity.³⁵ This betrays a focus on the audience: that playing into the player’s expectations to draw them into the game world is a key desire of audio designers for video games. Perron’s discussions of “affective congruence” and “added value” explains the reasons for this. In both, the visual combines with the audio to create a more “natural” (realistic?) effect. In this, what something *should* sound like is irrelevant. Instead what is relevant is what the player *believes* something should sound like.

I have provided a broad overview of just some of the scholarship on immersion and sound in video games. This is, of course, not a complete or detailed study, which there simply isn’t time to undertake here; however, I have demonstrated the broadly unified strokes of the literature, where terms diverge, and where I feel there is space for more nuance. Having done so, we can now apply these understandings to an analysis of the *Five Nights at Freddy’s* franchise.

The Literature in the Light of *Five Nights at Freddy’s*

I will now address an example from *Five Nights at Freddy’s* that supports the discussions presented above, both in presenting and challenging the current literature, and stating the case for my Global Music Box.

The background soundscape of the first *Five Nights at Freddy’s* (*FNaF*) game demonstrates a simple, yet highly effective, method of establishing a mood of anxiety to prime the player for the emotional response of the horror genre’s characteristic jump scares. The game’s soundscape fits into Toprac and Abdel-Meguid’s proposed anxiety-generating sounds – that of ambient, low volume sound.³⁶ There is no melodic or harmonic characteristic to the *FNaF* soundtrack; it instead it consists of a low and quiet

metallic rolling sound, giving the suggestion of a large open space beyond the small security office the player is confined to. As we have noted above, a primary way in which sound generates fear and anxiety is by inferring that there is a threat outside of our visual perception, and so the implication that there is an extended world outside of the player's control through this simple sound is an excellent method of instilling anxiety. That this is confirmed visually by the player through the camera system mechanic only adds to this effect.

Interestingly, this contrasts against what Claudia Gorbman argues in *Unheard Melodies*.³⁷ Gorbman states that music signifies according to musical cues and cultural and cinematic codes, and that these associations are used in creating the most impactful immersive experience.³⁸ To be recognised, Romantic tonality is often used as it is designed for "quick and effective signification."³⁹ This clearly isn't the case in *FNaF*, and yet it quickly and effectively signifies and establishes an appropriate mood for the game. Clearly, Gorbman's discussion does not consider the role of atonality and other extended techniques, or indeed the roles of sound (such as the metallic sound of *FNaF*) in film and game scoring to generate affects. Regardless, Gorbman later notes that we are not consciously listening to film music as we watch a film, but the unconscious affect it provides is still heard.⁴⁰ It primes for certain emotions, as is discussed above, which reduces critical distance between the player and the fictional world and thus contributes to the immersive process. This is regardless of whether the music is Romantically-coded, or if it steps outside of this tradition. Ultimately, what matters is whether the audience understands the implications of the cues used, rather than by what method these cues are generated.

As I argued earlier, it is necessary for a nuanced understanding of the ways that music, in both video games and films, can subdue its audience. However, given the active participation necessary for a video game player, there are more interactive instances of music and sound in game soundtracks that does not exist in the same way in film soundtracks. As such, game audio more often crosses the boundaries of the diegesis, as seemingly diegetic aural occurrences happen following player interaction. As such, we can see that all sounds in the game are contributing to the player's immersion.

Let us consider this argument further in the context of *FNaF*. As previously discussed, the primary background sound is the metallic rolling sound. This effectively establishes a mood of anxiety and could be argued to act in subduing the player, as it forces them into an anxious state, primed for the sudden, high volume jump scares that generate fear. But,

this does not fully express the experience of playing a horror game, especially not *FNaF*. There is more going on here.

This can be explained through the Global Music Box, but first we should note the rest of the sounds in the game. In addition to the metallic sound, the background is interspersed randomly with additional sounds, including footsteps, a distant pipe organ, and deep, echoing laughter. These have no gameplay significance, other than scaring the player, and not allowing the player to settle into a fixed soundscape. These are double bluffs: the player has been primed by the soundscape for a fear response and so to hear these loud, fear-inducing sounds with no consequence or corresponding visual stimulus intensifies the anxiety response. Moreover, they cannot even trust that a loud, sudden noise is simply a jump scare; potentially it means something else is happening outside the visual range of the player, or perhaps the player has missed an important visual cue necessary to their survival. Thus, we have a series of sounds that may well derive from within the diegesis but, in reality, act upon the player extradiegetically. Indeed, this develops theories such as Perron's "affective congruence": meaning does not have to be made in both music and their visuals, instead the two can be separated to generate a specific effect.⁴¹

In addition to these sounds, we must also consider the sounds communicating with, and generated by, the player. In *FNaF*, these sounds are simple – the buzzing of the door lights, the slamming of the doors, and the buzzing of interactions with the camera. It would be simple to write these off as only confirmatory sounds that the user utilises to confirm they have interacted successfully with the user interface, or that they are intended to generate a sense of verisimilitude by sounding as we would expect an object in the real world to, but through the lens of the Global Music Box they take on additional significance. Surely, we do not distinguish between that which we consciously listen to and that which we unconsciously hear? By that logic, therefore, these interactive sounds also contribute to the overall soundscape of a game. As I mentioned earlier, an inconsistent soundworld contributes to the anxiety being established in the game. I would also argue that the perception of both the interactive and the affective background sound being united transdiegetically by music contributes to immersion. The affect that unconsciously draws the player in and subdues them into the confines of the fictional world, is then made "real" by the player being able to interact with and partially generate the fictional world. This would explain the gap previously noted in Gorbman's theory of film music, wherein the music seeks to subdue the listener, implying a passive reception not typical of (survival horror) video games. The interactivity and active participation of a video game does not contradict Gorbman's statement that the sound of

a film or game is a passive process. Rather, it enhances it by the nature of the active participation contributing to the generation of the pacifying sound. In turn, this contributes to the theory of the immersive fallacy – the idea that a player may lose track of the fact they are playing a game. Indeed, they may, as they are contributing to the generation of a new world on their terms, thereby restructuring their reality as they see fit.

As has been demonstrated here, the Global Music Box is a beneficial analytical method, especially when used alongside existing detailed theories. Through the above analysis of *Five Nights at Freddy's*, I have shown that there is value in my theory as a methodology that approaches all the sounds in a game as a whole and evaluates the relationships between them and how they contribute to the global immersive experience. Of course, the existing methodologies and theories that would look in detail at specific aspects of the game are still useful, and would certainly support my theory, but as I have stated, I am offering a new, additional tool to add to our analytical toolbox, and have demonstrated the value of this approach in the light of the first *Five Nights at Freddy's* game.

Reappraising Our Understanding: Where Do We Go Now?

My concept of the Global Music Box, as the name implies, seeks to treat all sounds in a game as part of the same system with the same goal. It steps outside distinctions between sound, music, interaction, and diegesis, and seeks to consider the individual characteristics of each sound, and how it relates to the network of other sounds and features that work towards creating a specific emotion and, in turn, an immersive experience. This allows us to treat each case in a unique manner, allowing for the nuance of each game and its soundworld to be approached on its own terms. I have been clear in stating that this is not to replace or diminish the analytical approaches and theories approached above. Moreover, it is another tool that can be used to approach the challenging topic of immersion in games, and the role music and sound plays in enhancing the process. There is precedent for this concept in the scholarly literature. Stilwell notes that “although fine distinctions [between sounds’ diegetic status] may be fascinating to explore, they also risk recapitulating the stratifying or branching taxonomic approach.” ⁴² Similarly, yet more broadly, Arnt Maasø suggests that we should study all

filmic elements, not just its single features in isolation – all of a film’ s (or a game’ s) features contribute to the meaning making process.⁴³

Stillwell also refers to Noël Carroll’ s description of Theory and theorising in discussing the concept of the fantastical gap. Carroll warns against a single unitary Theory that “presumes to explain everything.”⁴⁴ Instead, he argues for theorising, which breaks the large undertaking of Theory into piecemeal chunks that are (in Stillwell’ s terms) “easily digestible but nourishing.”⁴⁵ This would indeed be the primary argument against my concept of the Global Music Box. However, to continue Carroll’ s metaphor, while understanding the constituent parts of the dish is important, it is irrelevant should the complete dish be unpalatable. As I am at pains to stress, my theory is not to invalidate previous approaches to games, as they certainly are beneficial and important in understanding how games work. Yet my concept exists more as a reminder that the broader picture, and the relationships between these small-scale understandings, is a beneficial process, reminding us of the broader reasons for the smaller studies and the complete games we love to play.

Ultimately, I have proved the benefits of my Global Music Box theory, the ways in which it complements existing theories, and how it pushes beyond what is already understood. As is demonstrated in my analysis of the first *Five Nights at Freddy’ s* game, this theory has usefully explained the underlying processes leading to immersion, both specifically in terms of the game, as well as contributing more broadly to the discourse around sound and music in games and their role in immersion. This is, of course, a nascent theory, and I am certain that there is more research that can be done to expand and develop my ideas in order to investigate the broader applications of the theory. However, for now, I have laid the groundwork for a theory that considers the relationships between the constituent sounds of video games, and how together this can allow us to generate a new or more detailed understanding of the immersive processes taking part as a player engages in a game such as *Five Nights at Freddy’ s*.

Bibliography

Carroll, Noël. "Prospects for Film Theory: A Personal Assessment." In *Post-Theory: Reconstructing Film Studies*, edited by David Bordwell and Noël Carroll, 37-68. Madison: University of Wisconsin Press, 1996.

Castelvecchi, Stefano. "On 'Diegesis' and 'Diegetic' : Words and Concepts." *Journal of the American Musicological Society* 73, no. 1 (April 2020): 149-171.

Chion, Michel. *Audio-Vision: Sound on Screen*. Translated by Claudia Gorbman. New York: Columbia University Press, 1994.

Collins, Karen. *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design*. Cambridge: MIT Press, 2008.

Csikszentmihalyi, Mihaly. *Finding Flow: The Psychology on Engagement with Everyday Life*. New York: Basic Books, 1998.

Farkas, Tomas. "How Interactive? Is Music in Video Games Just Film Music in Disguise?" *Marketing Identity* 3, no 1/1 (2015): 438-499.

Glassner, Andrew. *Interactive Storytelling: Techniques for 21st Century Fiction*. Boca Raton, FL: CRC Press, 2004.

Gorbman, Claudia. *Unheard Melodies: Narrative Film Music*. Bloomington: Indiana University Press, 1987.

Jørgensen, Kristine. "'What are Those Grunts and Growls Over There?' Computer Game Audio and Player Action." PhD diss., Copenhagen University, 2007.

Juslin, Patrik N. "Emotional Responses to Music." In *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Micheal Thaut, 131-140. Oxford: Oxford University Press, 2008.

Maasø, Arnt. "Lyden av Levende Bilder." IMK report no. 14 from the Department of Media and Communication, University of Oslo, 1994.

Perron, Bernard. *The World of Scary Video Games: A Study in Videoludic Horror*. New York: Bloomsbury, 2018.

Roberts, Rebecca. "Fear of the Unknown: Music and Sound Design in Psychological Horror Games." In *Music in Video Games: Studying Play*, edited by Kevin J. Donnelly, William Gibbons and Neil William Lerner, 138-158. New York: Routledge, 2014.

Ryan, Marie-Laure. *Narrative as Virtual Reality. Immersion and Interactivity in Literature and Electronic Media*. Baltimore: Johns Hopkins University Press, 2001.

Stilwell, Robynn J. “The Fantastical Gap Between Diegetic and Nondiegetic.” In *Beyond the Soundtrack: Representing Music in Cinema*, edited by Daniel Goldmark, Lawrence Kramer and Richard D. Leppert, 184-202. Berkeley: University of California Press, 2007.

Summers, Tim. *Understanding Video Game Music*. Cambridge: Cambridge University Press, 2018.

Toprac, Paul, and Ahmed Abdel-Meguid. “Causing Fear, Suspense, and Anxiety using Sound Design in Computer Games.” In *Game Sound Technology and Player Interaction: Concepts and Developments*, edited by Mark Grimshaw, 176-191. Hershey: Information Technology Reference, 2010.

van Elferen, Isabella. “Analysing game Musical Immersion: The ALI Model.” In *Ludomusicology: Approaches to video game music*, edited by Michiel Kamp, Tim Summers, Mark Sweeney, 32-52. Sheffield: Equinox, 2016.

Media Cited

Five Nights at Freddy's. Video game. Texas: Scott Cawthon, 2014.

Five Nights at Freddy's 2. Video game. Texas: Scott Cawthon, 2014.

Five Nights at Freddy's 3. Video game. Texas: Scott Cawthon, 2015.

Five Nights at Freddy's 4. Video game. Texas: Scott Cawthon, 2015.

Five Nights at Freddy's: Sister Location. Video game. Texas: Scott Cawthon, 2016.

Five Nights at Freddy's: Help Wanted. Video game. Texas; California: Scott Cawthon; Lionsgate Games, 2016.

Markiplier. “WARNING: SCARIEST GAME IN YEARS | Five Nights at Freddy's – Part 1.” YouTube video. 17:43. August 12, 2014. Accessed July 21, 2021. <https://www.youtube.com/watch?v=iOztnsBPrAA&t=556s>

The Game Theorists. “Game Theory: Five Nights at Freddy’ s SCARIEST Monster is You!” YouTube video. 16:36. October 23, 2014. Accessed July 21, 2021.
https://www.youtube.com/watch?v=th_LYe97ZVc.

Ultimate Customer Night. Video game. Texas: Scott Cawthon, 2018.

Biography

Hannah Capstick is a third-year undergraduate student reading music at University College, Oxford, where she is President of the College Music Society. She is events director and editor for the Broad Street Humanities Review, a journal dedicated to publishing the works of undergraduate academics, where she is establishing a conference for aspiring academics to present their work. Her current research explores the impact of interactivity on the music, diegesis and immersion, and how these change our understanding of sound and music in video games and film. She is also a keen performer, studying the flute at the Royal Academy of Music, managing and playing in the Oxford University Philharmonic Orchestra, as well as giving regular solo recitals and singing in the college Chapel Choir.

Footnotes

1.

Markiplier, “WARNING: SCARIEST GAME IN YEARS | Five Nights at Freddy’ s – Part 1,” August 12, 2014, accessed July 21, 2021, <https://www.youtube.com/watch?v=iOztnsBPrAA&t=556s>;

The Game Theorists, “Game Theory: Five Nights at Freddy’ s SCARIEST Monster is You!,” October 23, 2014, accessed July 21, 2021, https://www.youtube.com/watch?v=th_LYe97ZVc. ↵

2. Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (Bloomington: Indiana University Press, 1987), 1-7. [↵](#)
3. *Ibid.*, 4. [↵](#)
4. *Ibid.*, 5. [↵](#)
5. Kristine Jørgensen, “ ‘What are Those Grunts and Growls Over There?’ Computer Game Audio and Player Action” (PhD diss., Copenhagen University, 2007), 58. [↵](#)
6. Gorbman, *Unheard Melodies*, 82-89. [↵](#)
7. Robynn J. Stilwell, “The Fantastical Gap Between Diegetic and Nondiegetic,” in *Beyond the Soundtrack: Representing Music in Cinema*, ed. Daniel Goldmark, Lawrence Kramer and Richard D Leppert (Berkeley: University of California Press, 2007), 196. [↵](#)
8. *Ibid.* [↵](#)
9. *Ibid.* [↵](#)
10. *Ibid.*, 184-202. [↵](#)
11. *Ibid.*, 184. [↵](#)
12. *Ibid.*, 186-187. [↵](#)
13. Jørgensen, “Grunts and Growls,” 74. [↵](#)
14. Jørgensen, “Grunts and Growls,” 75;
Ibid., 84. [↵](#)
15. *Ibid.*, 47; Marie-Laure Ryan, *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media* (Johns Hopkins University Press: Baltimore, 2001). [↵](#)
16. Ryan, *Narrative as Virtual Reality*, 103. [↵](#)
17. Mihaly Csikszentmihalyi, *Finding Flow: The Psychology on Engagement with Everyday Life* (New York: Basic Books, 1998), 29-30. [↵](#)

18. Andrew Glassner, *Interactive Storytelling: Techniques for 21st Century Fiction* (Boca Raton, FL: CRC Press, 2004), 53-92. [↵](#)
19. Tim Summers, "Texturing and the Aesthetics of Immersion," in *Understanding Video Game Music* (Cambridge: Cambridge University Press, 2018), 58. [↵](#)
20. Paul Toprac and Ahmed Abdel-Meguid, "Causing Fear, Suspense, and Anxiety using Sound Design in Computer Games," in *Game Sound Technology and Player Interaction: Concepts and developments*, ed. Mark Grimshaw (Hershey: Information Technology Reference, 2010), 176-191. [↵](#)
21. Karen Collins, "Gameplay, Genre, and the Functions of Game Audio," In *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design* (Cambridge: MIT Press, 2008), 133. [↵](#)
22. Patrik N. Juslin, "Emotional Responses to Music," in *The Oxford Handbook of Music Psychology*, ed. Susan Hallam, Ian Cross, and Micheal Thaut (Oxford: Oxford University Press, 2008), 131-135. [↵](#)
23. Bernard Perron, *The World of Scary Video Games: A Study in Videoludic Horror* (New York: Bloomsbury, 2018), 304. [↵](#)
24. Tomas Farkas, "How Interactive? Is music in video games just film music in disguise?" *Marketing Identity* 3, no 1/1 (2015): 447. [↵](#)
25. Jørgensen, "Grunts and Growls," 132. [↵](#)
26. Toprac and Abdel-Meguid, "Causing Fear," 176-191. [↵](#)
27. *Ibid.* [↵](#)
28. Perron, *The World of Scary Video Games*, 288. [↵](#)
29. *Ibid.*, 287. [↵](#)
30. Stilwell, "The Fantastical Gap," 188. [↵](#)
31. Rebecca Roberts, "Fear of the Unknown: Music and Sound Design in Psychological Horror Games," in *Music in Video Games: Studying Play*, ed. Kevin J. Donnelly, William Gibbons and Neil William Lerner (New York: Routledge, 2014), 138. [↵](#)
32. Perron, *The World of Scary Video Games*, 300. [↵](#)

33. Michel Chion, *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), 221. [↵](#)
34. Perron, *The World of Scary Video Games*, 293-294. [↵](#)
35. Jørgensen, “Grunts and Growls,” 48-49. [↵](#)
36. Toprac, “Causing Fear,” 176-191. [↵](#)
37. Gorbman, *Unheard Melodies*, 1-7. [↵](#)
38. *Ibid.*, 2-3. [↵](#)
39. *Ibid.*, 4. [↵](#)
40. *Ibid.*, 1-7. [↵](#)
41. Perron, *The World of Scary Video Games*, 300. [↵](#)
42. Stilwell, “The Fantastical Gap,” 196. [↵](#)
43. Arnt Maasø, “Lyden av Levende Bilder,” IMK report no. 14 from the Department of Media and Communication, University of Oslo, 1994. [↵](#)
44. Noël Carroll, “Prospects for Film Theory: A Personal Assessment,” in *Post-Theory: Reconstructing Film Studies*, eds. David Bordwell and Noël Carroll (Madison: University of Wisconsin Press, 1996), 40. [↵](#)
45. Stilwell, “The Fantastical Gap,” 185. [↵](#)