CrossMark

Inside the Loop: The Audio Functionality of Inside

Mathew Arnold[1]

Abstract

The manner in which soundscapes evolve and change during gameplay can have many implications regarding player experience. *INSIDE* (Playdead in INSIDE. Released on Microsoft Windows, Playstation 4, Xbox One, Nintendo Switch and iOS, 2016) features a gameplay section in which rhythmic audio cues loop continuously both during gameplay and after player death. This paper uses this aspect of the soundtrack as a case study, examining the effects of looping sound effects and abstract musical cues on player immersion, ludic functionality, and episodic engagement. The concept of spectromorphology proposed by Smalley (Organised Sound 2(2):107–126, 1997) is used to analyse the way in which musical cues can retain ludic functionality and promote immersion in the absence of diegetic sound design. The "musical suture" (Kamp, in: Ludomusicology: approaches to video game music, Equinox, Sheffield, 2016) created by continuously looping audio during death and respawn is also examined with regards to immersing the player within an evolving soundscape.

Keywords Spectromorphology · Immersion · Sound design · Ludomusicology · Musical suture

1 Introduction

INSIDE is a two-dimensional puzzle-platform game that features a gameplay sequence involving distant explosions and shockwaves. During this sequence, a regularly repeating explosion in the background creates a rhythmic shockwave that moves from background to foreground, and the player must avoid being caught out in the open when the shockwave arrives. Failure to find cover results in a swift and

---

✉ Mathew Arnold
   mat.arnold@me.com

1   Leeds Beckett University, Leeds, England, UK

Springer

graphic death, as would be expected, but the manner in which *INSIDE* respawns the player is where it differs from many other games. Instead of curtailing the soundtrack and reverting the game back to a previous state, *INSIDE* allows the repeating sound of the shockwave to continue and respawns the player at an appropriate time within the cycle of explosions. This audio-led system creates the opportunity for unique ludic functionality, and also serves as a method of deepening immersion, creating narrative progression and delineating the game's episodic structure.

2 Ludic Functionality

Collins (2008) posits that interactive games must ensure that the player hears sounds that provide feedback based on their actions, instruct them as to their objectives, and orientate them within the world of the game. These ludic functions of sound are unique to the medium of video games, and *INSIDE* provides a valuable case-study in combining multiple functions using the single core mechanic of persistently looping audio. The core challenge of this section of the game is to avoid the shockwave, which is a design decision that has implications regarding the level geometry. The player character's maximum movement speed must be taken into account in order to ensure that cover is physically positioned in such a way that it is close enough to be reached within one shockwave cycle, but far enough from other areas of cover to provide a suitable challenge. This combination of carefully placed cover and a constantly recurring deadly threat leaves a relatively small margin for player error. In order to avoid player frustration, the developer's expectations of the player must be clearly communicated. The audio provides valuable instruction to the player regarding the manner in which they should approach this section of gameplay. The sound of the shockwave is loud, has a large amount of low frequency content, and also causes the ambient sound of the level to duck when it occurs. Juslin and Vastfjall (2008) found that sounds with large amounts of low frequency content and noise-like spectra resulted in increased nervous system activity in the listener, creating feelings of unease and sensory arousal due to brain stem activation. This brain stem response communicates clearly to the player that the shockwave is a threat, prompting the player to find cover before the shockwave arrives. Whilst behind cover, this feeling of safety is reinforced by removing some of the noise-like components via low-pass filtering of the explosion sounds.

The shockwave is preceded by a rising sound similar to that of a falling explosive shell and a muffled explosion. The delay between the explosion and the shockwave also conveys some spatial information about the game environment, as it creates the impression that the explosion must be occurring very far away in order for the shockwave to take so long to arrive. This orientates the player within the physical space of the game-world, giving a sense of depth to the environment along the spatial axis that the player is unable to explore due to the two-dimensional nature of the gameplay. It should also be noted that this series of sounds would normally be physically impossible, as in reality the sound of the explosion

and the shockwave would be one and the same phenomenon. Despite its unrealistic nature, the sequence of sounds conforms to player expectations via what Chion (1994, p. 109) refers to as *rendering*. As he explains:

> We must distinguish between the notions of *rendering* and *reproduction*. The film spectator recognizes sounds to be truthful, effective, and fitting not so much if they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation.

The player most likely expects some form of auditory notification from the initial explosion to appear synchronously with the visual flash that appears in the background, as it contextualises the following shockwave by alluding to its initial source. Furthermore, it provides an additional rhythmic point of reference to the player, allowing them more easily to interpret the timing information encoded within the sound design.

This sequence of sounds allows the player to time their movements between aras of cover and to avoid a fail state (death). As the cycle of warning sound, explosion and shockwave repeats, it establishes a rhythm in the player's mind. The player may then rely on this internal sense of rhythm to overcome the puzzles in the section rather than relying on a more literal interpretation of the sound cues to avoid danger. If the player was simply waiting for the sound of an incoming explosive shell to denote imminent danger before moving to safety, some of the section's puzzles would be much more difficult.

One in particular involves using a heavy door as movable cover and, due to its inertia, the player must pre-empt the warning cue with their movement input in order to be successful. These kinds of puzzles are solvable only by paying attention to the rhythmic nature of the audio, so a respawn system that breaks the rhythmic continuity could potentially lead to player frustration or confusion. In keeping the rhythmic continuity intact throughout death and respawning, *INSIDE* not only avoids player frustration but also continually conditions the player to listen to the audio cues and respond accordingly.

This appears to have been one of Playdead's key motivators in designing this section of the game. Sound designer Martin Stig Andersen (2016) states that a similar section in their previous release *Limbo* was compromised due to the fact that the audio would not loop continuously, making the puzzle harder to solve. The section of gameplay involved gravity becoming periodically inverted after an audible cue. *INSIDE* was subsequently designed from the ground up to allow for continuously looping audio. The rhythmic nature of the gameplay allows for parallels to be drawn with rhythm action games such as *Guitar Hero* (RedOctane 2005) and *Rock Band* (MTV Games 2007). In *INSIDE,* however, it is not a player's reflexes and hand eye co-ordination skills that are challenged through the use of rhythmic gameplay, but rather their ability to plan ahead by judging speed, distance, and positioning.

3 Challenge-Based and Sensory Immersion

Ermi and Mäyrä (2005, pp. 7–8) delineate immersion into three main categories; sensory immersion as created by the graphics and audio of the game, challenge-based immersion created by the player applying their skills to overcome the game's challenges, and imaginative immersion created by the player empathising with their player character or becoming absorbed in the story of the game. Death and respawn mechanics present a unique challenge in terms of preserving sensory and challenge-based immersion.

Upon dying the player has triggered a fail state and no longer has physical control of their character. This interrupts the challenge-based immersion until the player respawns and regains agency. Furthermore, the game will usually have to reset itself back to a previous checkpoint. This most likely means fading the audio out and fading the screen to black, breaking sensory immersion through lack of audiovisual feedback. Collins (2008, p. 134) states that "Audio plays a significant role in the immersive quality of a game. Any kind of interruption in gameplay—from drops in frame rate playback or sluggish interface reactions—distracts the player and detracts from the immersion and from audio's playback, particularly interruptions in music such as hard cut transitions between cues". If the audio were simply to stop at the point of death and start again when the player respawned, it would create exactly the kind of discontinuity that Collins mentions here, which may have a detrimental effect on player immersion. By continuing to loop the shockwave sound effects even after the player dies, *INSIDE* retains sensory immersion throughout, keeping the player immersed until they are respawned and challenge-based immersion can resume. Once the player respawns sound plays an active role in promoting challenge-based immersion, providing both a constant reminder of the challenge that the player faces, and the rhythmic information required to overcome it.

4 Imaginative Immersion

The uninterrupted nature of the audio has implications regarding the physical state of the world in which the game takes place. Often in video games, dying and respawning have the effect of rewinding time to a prior checkpoint, thus allowing the player to reattempt the section of gameplay anew. This is often reinforced sonically by the interruption of the character's diegesis at the point of death, with music and sound effects also returning to a previous point in time. The sound and music in this section of *INSIDE* eschew this by continuing to move temporally forwards in the event of the player character's death. One implication of this is that the physical state of the world may persist throughout the death state. This persistent world state hypothesis can be fairly easily disproven, however, as all physical objects revert to their previous positions upon respawn. This creates a dissonance between the player's visual and auditory perceptions of the game world, potentially causing them to question both the nature of the world and their relationship to it. Is the player character an avatar onto which the player should project themselves, or is the player an

outside observer repeatedly guiding a hapless character to his death until the desired outcome is achieved? The themes of control and agency are prevalent within the game, with a secret ending showing the player character unplugging a large cable before immediately slumping motionless to the floor, as though the player's control of the character has finally been severed. The continuously looping sound and music subtly reinforce the notion that not only are the player and their in-game avatar separate entities, but also that the character may in fact see the player as a malevolent controlling force from which they wish to be freed. By prompting these questions to the player, the continuously looping audio serves as a means of deepening imaginative immersion. Therefore, it can be seen that this one simple mechanic can have an effect on sensory, challenge-based, and imaginative immersion simultaneously.

4.1 Mood Induction

There is a puzzle part-way through the section where the player is required to synchronise a piece of cover that moves in a circular pattern with the rhythm of the shockwave. If the player is successful then the cover will move across a ladder at just the right time to protect the player as they ascend it. By its very nature the puzzle can be a difficult one to solve, and the game provides a clear signifier to the player when they achieve the solution. This is achieved by removing the shockwave sound design entirely and replacing it with a synthesised musical cue that follows the same rhythmic pattern. First and foremost this notifies the player that they have solved the puzzle, but it also encourages a deeper form of immersion. Grau (2003, p. 13) states that "…immersion is mentally absorbing and a process, a change, a passage from one mental state to another. It is characterized by a diminishing critical distance to what is shown and an increasing emotional involvement in what is happening." During this puzzle *INSIDE* promotes imaginative and sensory immersion by switching from sound effects to emotionally engaging music very suddenly. This exploits the heightened sense of peril as the player ascends the ladder, using it to more effectively induce a different mood. As Collins (2008, p. 133) states:

> Mood induction and physiological responses are typically experienced most obviously when the player's character is at significant risk of peril… In this way, sound works to control or manipulate the player's emotions, guiding responses to the game.

As the musical cue enters, the player is halfway up a ladder, exposed until the last moment when the cover swings into place to protect them. Due to the stressful nature of this predicament the mood induction can be highly effective, arriving at the precise moment the player experiences the relief of solving the puzzle and the endorphin rush of a reward for competency. The synthesised sounds are analogous to the original sound effects in terms of their duration, envelope, and the way in which their spectral content varies with time. Smalley (1997, p. 1) refers to these time-dependent qualities of sound using the term *spectromorphology,* defining the term as "the interaction between sound spectra (*spectro-*) and the ways they change and are shaped through time (*-morphology*)." The similar

spectromorphology between the sound and music ensures that the player remains explicitly aware that the musical cue is now acting as a surrogate for the preceding sound design. It also ensures that the exact same ludic information is communicated to the player regarding timing, whilst allowing for the encoding of additional meaning via harmony and timbre. To use Chion's (1994) terminology, the player shifts from a causal listening state wherein they are concerned with the perceived source of the sound (a hazardous explosion) into a semantic listening state, wherein they are focussed on the encoded meaning of the sound itself (the timing information). Smalley (1997) describes sound in terms of extrinsic and intrinsic qualities, where the extrinsic qualities of a sound are linked to the process that created it, and the intrinsic qualities are the spectromorphological qualities of the sound itself. The player shifts from focusing on the extrinsic qualities of the sound to focusing on the intrinsic qualities of the music. This can be thought of as the logical extreme of Chion's concept of *rendering* as discussed previously. The sound has become utterly divorced from anything relating to its initial cause, and arguably expresses a deeper sense of the feelings associated with the player's journey because of this.

As the player progresses through the section, higher-pitched reverberant synth layers gradually fade into the mix, drawing the player deeper into the abstract soundscape and even further from the literal auditory events of the scene. This creates what Gorbman (1987, p. 68) refers to as *spectacle,* where non-diegetic music "lends an epic quality to the diegetic events. It evokes a larger than life dimension which, rather than involving us *in* the narrative, places us in contemplation *of* it." This reinforces the notion that the player is an observer, a controlling force sending a hapless child to his death multiple times until their goals have been achieved. This further demonstrates the manner in which imaginative immersion can be promoted via the additional emotional content that can be encoded into music.

5 Suturing and Emphasis

Breaking or maintaining sonic continuity upon entering the fail state can be used as a form of instruction in and of itself. Kamp (2016, p. 83) states that:

> …the way death is represented affects how the player experiences the flow of gameplay. Musical suturing can de-emphasize the harshness of obstacles and fail states, just as its absence can highlight them.

Here the term "musical suturing" refers to continuous music that can be used to cover up sudden visual changes. *INSIDE* uses musical suturing in this way to de-emphasise the fail state. As the player has an effectively infinite number of lives and the only penalty for death is being reset to the last checkpoint, there is no reason to emphasise player deaths at the expense of immersion, as death is fairly inconsequential. By contrast, Kamp (2016) uses *Super Mario Bros* (Nintendo 1985) as an example where the death state is important to the player due to a limited number of lives.

Each death is a step closer to the ultimate fail state of the game over screen, which results in the player losing all progress and being forced to start back at the very first level. The lack of musical suturing in *Super Mario Bros* emphasises the consequences of death, discouraging the player from dying too often. In *INSIDE,* each death is treated merely as a stepping stone towards deducing the correct solution to a particular puzzle, and the suturing present during the death state serves to incentivise the player to engage with this iterative process of progression through the game.

5.1 Episodic Engagement

Whilst musical suture is used to retain immersion within this section of gameplay, its absence is also used to signify the transition to the next section. The game establishes a clearly defined set of rules and expectations for the player via the use of sound during this section, so therefore it must also be clearly communicated to the player when those rules no longer apply. At the end of the section, the shockwave equivalent musical layers cease and the score gradually fades out to signify that the section has come to a close. This section of the game requires that the player be highly engaged and focussed, and the absence of musical suture at its close signals to the player that they can revert to a lower state of engagement. Salmond (2016) argues that interest curves are an important factor in keeping players engaged for longer periods of time. An interest curve essentially plots the level of intensity of an audiovisual experience over time, and a well-designed interest curve will ensure that players are neither exhausted by prolonged sections of intense engagement, nor bored by extended periods of relative inactivity. The loading screen is an intrinsic feature of many games that contributes towards the shape of the interest curve. Although loading screens can be immersion-breaking, they can also provide some of the necessary troughs in the interest curve. They also serve to delineate episodes, providing that the game is designed in such a way that they occur at appropriate junctures within the narrative. *INSIDE* utilises a streaming system that dynamically loads the assets for the next gameplay section as the game progresses, and as such avoids any loading screens whatsoever during gameplay. This combines with the respawn mechanic noted earlier to create a perpetual state of engagement in the player, which if not properly managed could lead to fatigue. The music and sound design in this section serve as a method of episodic delineation, providing a clear beginning, middle and end to the episode within a continuous gameplay experience. The diegetic explosion effects at the beginning of the section provide narrative exposition for the episode, communicating the rules governing the section. Once established, the switch to music provides a sense of narrative progression at the section's mid-point, and when the score is reduced to the higher pitched synth layers at the end of the section it provides a sense of narrative closure. Thompson et al. (1994) found that film scores that end on a strong beat tend evoke a stronger feeling of closure in the audience. *INSIDE* takes advantage of this effect, as the musical elements representing the shockwave are removed abruptly after the final shockwave arrives. This ensures that the section always ends on a strong beat, as the remaining synth layers contain little to no rhythmic information. This serves to notify the

player that the danger has passed, and that they can proceed normally from here on. Once the player enters an elevator at the very end of the section, the music fades out entirely, before the shockwave sounds return suddenly, causing the elevator to crash into the water below. This final subversion of expectations catches the player unaware and serves to bring them back to the game's more literal reality very suddenly. This unexpected tonal shift creates feelings of stress in the player, setting them up for the next mood induction which comes in the form of a claustrophobic underwater gameplay section.

6 Future Implications

INSIDE demonstrates that the use of spectromorphologically similar sound design and music can create new opportunities for mood induction and imaginative immersion. These techniques could be expanded and applied to a number of different gameplay contexts in future. For example, the balance of sound and spectromorphologically similar music could be altered based on a continuous variable, rather than simply switching fully once a virtual threshold has been crossed by the player. Perhaps the most obvious application of this would be to control the balance based on the player's remaining health, with the music becoming more predominant as the player approaches death. Emotional content encoded into the music could be used to great effect to induce a mood of dread or anxiety in the player in this way, simultaneously serving to notify and deepen immersion.

The disadvantage with using a technique such as this would result from the inherently non-musical timing of events that occur within the game. For example, a pseudo-random sound such as the approaching footsteps of an enemy squad would not translate easily into a musically viable instrumental layer. Care would have to be taken apply this technique only to musically viable sounds. Alternatively, in-game events would have to be constricted to occur at musically appropriate intervals of time. Stevens and Raybould (2015) propose a possible solution to this conflict, suggesting that each game event could be assigned a window of opportunity in which to occur. They suggest that within this window, the game could select the closest musical subdivision and snap the event to occur in synchronicity. *INSIDE* circumvents this pitfall by applying the technique only to one set of sound effects with a fixed rhythmic structure, set within a level designed specifically for the purpose of utilising this effect. For the technique to be more widely adopted across multiple genres it would require that games be designed from the ground up to trigger events based on musical information.

7 Conclusion

In conclusion, *INSIDE* utilises audio in this gameplay section to imbue one fairly simple mechanic with a number of different functions. By giving priority to the looping audio and ensuring that all other gameplay elements conform

appropriately, *INSIDE* is able to effectively communicate the game mechanics whilst retaining and deepening player immersion, even through repeated player deaths. The usage of spectromorphologically similar sound and music allows for effective mood induction, narrative progression and the deepening of sensory, challenge-based and imaginative immersion. The specific advantage that this technique possesses is that it allows for all of this whilst still performing the same ludic functions as traditional diegetic sound design. By altering the soundscape in this manner as the player progresses, episodic compartmentalisation can be achieved as part of a continuous gameplay experience, eschewing the need for more traditional methods of delineation such as cut scenes or chapter title cards.

References

Andersen, M. (2016) Inside: A game that listens. [Online Video] Available at https://www.youtube.com/watch?v=Dnd74MQMQ-E. (Accessed 4 Feb 2018).

Chion, M. (1994). *Audio vision: Sound on screen*. New York: Columbia University Press.

Collins, K. (2008). *Game sound: An introduction to the history, theory and practice of video game music and sound design*. Cambridge (MA): The MIT Press.

Ermi, L. & Mäyrä, F. (2005) Fundamental components of the gameplay experience: Analysing immersion. In *DiGRA '05—Proceedings of the 2005 DiGRA international conference: Changing views: Worlds in play*.

Gorbman, C. (1987). *Unheard melodies: Narrative film music*. Bloomington (IN): Indiana University Press.

Grau, O. (2003). *Virtual art: From illusion to immersion*. Cambridge (MA): The MIT Press.

Juslin, P., & Vastfjall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioural and Brain Sciences, 31,* 559–621.

Kamp, M. (2016). Sutures and peritexts: Music beyond gameplay and diegesis. In M. Kamp, T. Summers, & M. Sweeney (Eds.), *Ludomusicology: Approaches to video game music.* Sheffield: Equinox.

MTV Games (2007): Rock Band. Released on Playstation 2, Xbox 360, Playstation 3 and Nintendo Wii.

Nintendo (1985): Super Mario Bros. Released on the Nintendo Entertainment System.

Playdead (2016): INSIDE. Released on Microsoft Windows, Playstation 4, Xbox One, Nintendo Switch and iOS.

RedOctane (2005): Guitar Hero. Released on Playstation 2, Xbox 360.

Salmond, M. (2016). *Video game design: Principles and practices from the ground up*. London: Bloomsbury Press.

Smalley, D. (1997). Spectromorphology: Explaining sound-shapes. *Organised Sound, 2*(2), 107–126.

Stevens, R. & Raybould, D. (2015): Designing a game for music. In K. Collins, B. Kapralos, & H. Tessler (Eds.), *The Oxford handbook of interactive audio.* Oxford: Oxford University Press.

Thompson, W. F., et al. (1994). Effects of underscoring on the perception of musical closure. *Psychomusicology, 13,* 9–27.