

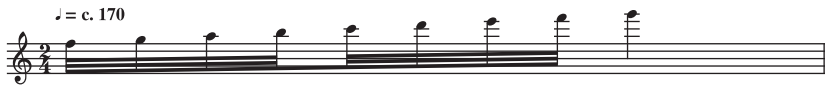
16 | Audio and the Experience of Gaming: A Cognitive-Emotional Approach to Video Game Sound

DANA PLANK

It was June of 1990; I was four years old, waiting with my mom in our car, which was parked on the searing hot asphalt of a mall parking lot. My brother, who was nine, was inside with my father picking out his birthday present. When they finally returned, my brother was carrying a huge grey, black and red box with the words ‘Nintendo Entertainment System’ printed on the side. Without this day, impatient and blazing hot in my memory, I might never have known Mario and Link and Kid Icarus and Mega Man, and my life would have been much poorer for it. We brought home two games that day: the promotional 3-in-1 game that came with the system (*Super Mario Bros./ Duck Hunt/Track Meet*; 1985), and *The Legend of Zelda* (1986). It is almost impossible to imagine the rich, diverse game world of *Zelda*’s Hyrule without its characteristic sounds. How would the player experience the same level of satisfaction in restoring their health by picking up a heart container or lining their coffers with currency without that full, round plucking sound as they apprehend the heart, or the tinny cha-ching of picking up a gemlike rupee (see Example 16.1)?

Without a variety of sounds, video games would lose a great deal of their vitality and power. Without the occasional punctuation of ostensibly diegetic sound effects, the player would not feel as deeply engrossed in a game and could grow irritated with its musical tracks, which are looped throughout the duration of a particular zone. Without sound effects to flesh out the world of a game, the aural dimension could fall flat, lacking novel stimuli to keep the player motivated and invested in the outcome of play. Aural stimuli elicit emotional, psychological and physiological responses from players, whether that response is ultimately meant to support narrative, influence gameplay decisions, foster player agency or facilitate incorporation into the game body. This discussion draws on empirical literature from music cognition and psychology, indicating the potent effects of sound. Video games add one more dimension to this conversation that will be discussed later in the chapter: interactivity. The soundscape is a site of incorporation, a dynamic

Example 16.1 *The Legend of Zelda*, small treasure sound effect



and vital bridge between the bodies of player and avatar. In this chapter, I argue that sound is one of the most important modalities through which the game incorporates its players into a complexly embodied hybrid of the material and virtual; sound largely determines the experience of play.

Game Sound, Arousal and Communication

Music and game sound serve as a source of intense and immediate communication with the player by engaging our physiological and psychological systems of arousal and attention. Our bodies do not remain fully alert at all times; a state of perpetual readiness to respond to the environment would be particularly taxing on physical resources, a costly expenditure of energy.¹ Therefore, the body undergoes many changes in arousal levels throughout the day, both on the order of minutes or hours (called tonic arousal, as in the diurnal cycle of sleep and wakefulness) and on the order of seconds or minutes (called phasic arousal, as triggered by various stimuli in the environment). Phasic arousal relates to the appearance of salient stimuli in the auditory environment. For example, the sound of a slamming door will drastically raise phasic arousal for a brief moment, until the cortex assesses the sound and determines that it is not an impending threat. There are many studies that explore the connection of music and sound to phasic arousal, focusing on elements such as dynamics and tempo.²

¹ David Huron, 'Arousal', *An Ear for Music* (n.d.), <https://web.archive.org/web/20100620200546/http://csml.som.ohio-state.edu/Music838/course.notes/ear04.html>. Huron states that music engages four of our 'readiness' mechanisms: arousal, attention, habituation and expectation. He also demonstrates how costly constant arousal would be: 'if a person were in a constant state of high arousal, they would need to consume 6 to 10 times their normal caloric intake.'

² Francesca R. Dillman Carpentier and Robert F. Potter, 'Effects of Music on Physiological Arousal: Explorations into Tempo and Genre', *Media Psychology* 10, no. 3 (2007): 339–63; Gabriela Husain, William Forde Thompson and E. Glenn Schellenberg, 'Effects of Musical Tempo and Mode on Arousal, Mood, and Spatial Abilities', *Music Perception* 20, no. 2 (2002): 151–71; Alf Gabrielsson, 'Emotions in Strong Experiences with Music', in *Music and Emotion: Theory and Research*, ed. Patrik N. Juslin and John A. Sloboda (New York: Oxford University Press; 2001), 431–49; Carol L. Krumhansl, 'An Exploratory Study of Musical Emotions and Psychophysiology', *Canadian Journal of Experimental Psychology* 51 (1997): 336–52; Isabelle Peretz, 'Listen to the Brain: A Biological Perspective on Musical Emotions', in *Music and*

A new sound in a game soundscape is a stimulus that increases phasic arousal, causing the temporary activation of the sympathetic nervous system (SNS). SNS activation leads to a number of physiological changes that prime the body for a fight-or-flight response until the sense of danger is assessed by the cortex as being either hazardous or innocuous. The physiological responses to SNS activation include: (a) the release of epinephrine (also known as adrenaline) from the adrenal gland, acting in the body to raise heart rate and increase muscle tension to prepare for potential flight; (b) the release of norepinephrine (or noradrenaline) in an area of the brain called the *locus coeruleus*, leading to raised sensitivity of the sensory systems, making the player more alert and mentally focused; (c) the release of acetylcholine in the nervous system, a neurotransmitter that increases muscle response; (d) increased blood pressure; (e) increased perspiration; and (f) increased oxygen consumption.³ The stimulus startling the player will be relayed to the thalamus and then move into the amygdala, causing an automatic, quick fear response.⁴ From there, the signal will travel to the sensory cortex in the brain, which evaluates the signal and then sends either an inhibitory or reinforcing signal back to the amygdala. In the video game, the ‘danger’ is virtual instead of proximate, and so these startle responses are most likely to evoke an eventual inhibitory signal from the sensory cortex. However, activation of the SNS by sound in the video game will lead to short bursts of energy – physiological changes that lead to a spike in the player’s overall alertness, priming them for action.

There are several acoustical features that affect this kind of sympathetic arousal, such as: an increase in the loudness and speed of the stimulus; physically proximal sounds (aka, close sounds); approaching sounds, unexpected or surprising sounds; highly emotional sounds, and sounds that

Emotion: Theory and Research, ed. Patrik N. Juslin and John A. Sloboda (New York: Oxford University Press, 2001), 105–134; Louis A. Schmidt and Laurel J. Trainor, ‘Frontal Brain Electrical Activity (EEG) Distinguishes Valence and Intensity of Musical Emotions’, *Cognition and Emotion* 15 (2001): 487–500; John A. Sloboda and Patrik N. Juslin, ‘Psychological Perspectives on Music and Emotion’, in *Music and Emotion: Theory and Research*, ed. Patrik N. Juslin and John A. Sloboda (New York: Oxford University Press, 2001), 71–104.

³ Epinephrine/adrenaline, norepinephrine/noradrenaline and acetylcholine are all neurotransmitters – organic chemical substances in the body that transfer impulses between nerves, muscle fibres and other structures. In other words, neurotransmitters are compounds that foster communication between the brain and various parts of the body.

⁴ The thalamus relays sensory and motor signals to other parts of the brain and has the vital function of regulating consciousness, sleep and alertness. The amygdala is composed of two almond-shaped structures deep in the middle of the brain; the amygdala handles memory processing, decision-making and emotional responses (such as fear, anxiety and aggression), and is therefore vital in recognizing and responding to a startling signal.

Example 16.2 *The Legend of Zelda*, secret

have a learned association with danger or opportunity or that are personally addressed to the player (e.g., using the player's name in a line of dialogue). Video game sound effects often serve many of these functions at once: they tend to be louder than, or otherwise distinguished from, the background musical texture in order to stick out; they are often surprising; they tend to invoke the player's learned association with the meaning of the effect (danger, reward, opportunity, discovery of a secret, etc.; see Example 16.2); and all are directly addressed to the player as a communicative device conveying information about the gameplay state.⁵

A sound effect is meant to evoke an orienting response – in other words, it commands attention from the player.⁶ If an orienting response is triggered, the player will experience additional physiological changes: pupil dilation, a bradycardic heart response (where the heart rate goes down, then up and then back to the base line), cephalic vasodilation (the blood vessels

⁵ Sound effects are usually short clips lasting a couple of seconds or less that serve a signalling function to the player and are typically triggered by direct operator (player) actions on the console controller. Sound effects – whether they are samples or recordings of real-world sounds, short musical cues, or abstract inventions for fantasy or science-fiction sounds – can be utilized for several purposes in the game context: confirming the spatiality of the playing field (making the space sound realistic or conveying its size through reverberation), aiding in identification with the game character and serving as multimodal sensory feedback (by connecting button presses and other player actions to the game character's responses) and modifying player behaviour. Definitions of sound effects can also include Foley, the reproduction of human-generated sounds such as footsteps (or atmospheric sounds such as wind and rain). Foley sounds provide subtlety and depth to the gameworld, and are typically mixed behind the musical cues as a backdrop to the more salient aural elements. Most early games use Foley rather sparingly, due to the lack of channel space to accommodate fully independent Foley, sound effects and musical tracks. For example, on the Nintendo Entertainment System, sound effects often impede on one of the five audio channels, briefly interrupting the melody or harmony in order to signal the player. Sound effects based on player actions are privileged in these early soundscapes for their communicative function – conveying critical information and serving as multimodal feedback to the player. For more on these distinctions, see Axel Stockburger, 'The Game Environment from an Auditive Perspective', in *Level Up: Proceedings from DiGRA* (2003), 1–5 at 5, accessed 9 April 2020, www.audiogames.net/pics/upload/gameenvironment.htm; Kristine Jørgensen, 'Audio and Gameplay: An Analysis of PvP Battlegrounds in World of Warcraft', *Game Studies: The International Journal of Computer Game Research* 8, no. 2 (2008), accessed 29 October 2020, <http://gamestudies.org/0802/articles/jorgensen>.

⁶ Alvin Bernstein, 'The Orienting Response and Direction of Stimulus Change', *Psychonomic Science* 12, no. 4 (1968): 127–8.

in the head become dilated), peripheral vasoconstriction (blood vessels in the extremities constrict) and increased skin conductance.⁷ All of these physiological changes due to SNS activation or orienting response prime the player for action by temporarily raising their phasic arousal via short, deliberate signals. However, it is not ideal to remain at a high level of phasic arousal; not only is this a costly energy expenditure, it can cause fatigue and stress. In the context of a game, a player will be less likely to continue play if they experience an incessant barrage of stimuli activating the SNS.

Just as sound can increase phasic arousal, it can also reduce it, especially when it is low-energy; for example, slow, predictable or soft. Effective game sound design will seek a balance; maintaining or lowering phasic arousal (e.g., with background music and silence) to maintain player concentration while raising it to draw attention to particular elements and prime the player to respond (e.g., with sound effects). Repetitive background sounds may serve an additional function, beyond merely preventing player stress from overstimulation of the SNS. According to the Yerkes–Dodson law, for simple tasks, higher phasic arousal leads to better performance. For complex tasks, lower phasic arousal leads to better performance.⁸ Therefore, the more complicated the video game task (in terms of motor skills and/or critical thinking), the less active, arousing and attention-getting the soundscape should be – if the goal of the game developers is to enhance player performance. Effective sounds work alongside other elements of the soundscape (such as background music for a particular area) to pique the player’s phasic arousal to a level optimal for game performance at any given moment.

Game composers may opt to intentionally manipulate the player, however, and use short musical cues as sound effects that will lead to a higher arousal level in a simpler area, in order to increase the difficulty of the level (and increase player satisfaction upon successful completion of the level). Koji Kondo famously used music this way in *Super Mario Bros.* (1985), doubling the tempo of the background-level track as a signal that the player was running out of time; at the 100-second mark a short ascending chromatic signal would play, startling the player and potentially raising their heart rate and stress levels (see Example 16.3). The question remains

⁷ John W. Rohrbaugh, ‘The Orienting Reflex: Performance and Central Nervous System Manifestations’, in *Varieties of Attention*, ed. Raja Parasuraman and D. R. Davies (Orlando, FL: Academic Press, 1984), 325–48; Evgeny N. Sokolov, ‘The Neural Model of the Stimulus and the Orienting Reflex’, *Problems in Psychology* 4 (1960): 61–72; Evgeny N. Sokolov, *Perception and the Conditioned Reflex* (New York: Macmillan, 1963).

⁸ Robert M. Yerkes and John Dillingham Dodson, ‘The Relation of Strength of Stimulus to Rapidity of Habit-Formation’, *Journal of Comparative Neurology and Psychology* 18 (1908): 459–82.

Example 16.3 *Super Mario Bros.*, ‘Hurry!’

$\text{♩} = 200$

Pulse 1

Pulse 2

Triangle

whether this sound signal inherently spikes a player’s arousal through its acoustic features, or the physiological response arises from familiarity with the sound’s meaning from previous play attempts or from similarly structured musical cues in other games.

Game Sound and Attention

Attention relates directly to arousal; an orienting response occurs when a sound commands a player’s attention. Sounds that lower phasic arousal can help a player to concentrate or remain invested in a task. However, attention is not synonymous with arousal. In the real world, we learn how to filter sounds in a noisy environment to determine which ones are the most important. This process of focusing on certain sounds is known as the cocktail party effect (or selective attention), first described by E. C. Cherry in 1953.⁹ In games, however, the sound design must perform some of this work for us, bringing elements to the forefront to highlight their importance and convey information to the player.¹⁰ If we remove the need for the player to selectively attend to game sound, there are greater opportunities to manipulate the two other forms of attention: exogenous (also known as passive attention; when an event in the environment commands awareness, as in a sudden sound effect that startles the player), and endogenous (when the player wilfully focuses on a stimulus or thought, as when the player is concentrating intently on their task in the game).

⁹ E. C. Cherry, ‘Some Experiments on the Recognition of Speech, with One and Two Ears’, *Journal of the Acoustical Society of America* 25 (1953): 975–9.

¹⁰ There are often options in the main menu screen to control for the volume of the dialogue, sound effects and music separately. An experienced player can thus customize the soundscape to avoid audio fatigue or to become more sensitive and responsive to auditory cues, but a novice might want to rely on the default mix for optimal performance.

Music, in addition to potentially lowering phasic arousal, can shift players into a mode of endogenous attention, which could be one reason for the pervasive use of wall-to-wall background music in early video games.¹¹ The musical loops facilitate endogenous attention, so that the introduction of communicative sound effects could act as exogenous attentional cues and raise phasic arousal as needed. A study by MacLean et al. from 2009 investigated whether a sudden-onset stimulus could improve sensitivity during a task that required participants to sustain attention; they found that exogenous attention can enhance perceptual sensitivity, so it would appear that the combination of music (facilitating endogenous attention) with effects (facilitating exogenous attention) can help players to focus and stay involved in their task.¹² This manipulation of attention has a powerful effect on neural patterns during play.¹³ After game sounds engage attentional processes, they can then do other work, such as communicating information about the game state or eliciting emotion.

The most important function of sound effects is to provide information to the player; therefore, game sounds have several of the same features as signals in the animal kingdom. They are obvious (standing out from the background track in terms of tempo, register, timbre or harmonic implication); frequently multi-modal (e.g., the sound is often joined to a visual component; and, in later games, haptic feedback through ‘rumble packs’ on controllers); employ specific sounds for a singular purpose (sound effects rarely represent multiple meanings); and are meant to influence the behaviour of the player (informing a player of impending danger will make them more cautious, whereas a sound indicating a secret in the room helps the player to recognize an opportunity). Signals are meant to influence or change the behaviour of the observer, just as game sounds may serve to alter a player’s strategic choices during play.

Background musical loops typically play over and over again while the player navigates a specific area. Depending on the number of times the

¹¹ Thomas Schäfer and Jörg Fachner, ‘Listening to Music Reduces Eye Movements’, *Attention, Perception & Psychophysics* 77, no. 2 (2015): 551–9.

¹² Katherina A. MacLean, Stephen R. Aichele, David A. Bridwell, George R. Mangun, Ewa Wojciulik and Clifford D. Saron, ‘Interactions between Endogenous and Exogenous Attention During Vigilance’, *Attention, Perception, & Psychophysics* 71, no. 5 (2009): 1042–58; when the stimulus was unpredictable, sensitivity went up, but performance did not. When the sudden-onset stimulus was more predictable (as in a video game, when a certain action prompts a sound and the player knows to expect this response), the participant did not suffer the decrement in performance.

¹³ Annabel J. Cohen, ‘Music as a Source of Emotion in Film’, in *Handbook of Music and Emotion: Theory, Research, Applications*, ed. Patrik N. Juslin and John A. Sloboda (New York: Oxford University Press, 2010), 879–908 at 894.

track repeats, this incessant looping will likely lead to habituation, a decreased sensitivity to a musical stimulus due to its repeated presentation.¹⁴ A level loop that repeats over and over will eventually be less salient in the soundscape; the player will eventually forget the music is even present. However, the music will continue to exert influence on the player, maintaining arousal levels even if the music falls away from conscious awareness. One can regain responsiveness to the initial stimulus if a novel, dishabituating stimulus is presented (engaging exogenous attentional mechanisms and raising phasic arousal); in the case of the video game, this is usually in the form of a sound effect. Thus, depending on the specific conditions of gameplay, game sound can serve two main functions with regards to attention and arousal. First, sound can create a muzak-like maintenance of optimal player arousal without attracting attention.¹⁵ The second function is to directly manipulate the player's attention and arousal levels through dishabituating, obvious musical changes in dynamics, tempo or texture. Combining music's regulatory capabilities with the signalling mechanisms of sound effects, the game composer can have a tremendous amount of control over the player's affective experience of the video game.

Emotion and Measurement

Affect is a broad, general term involving cognitive evaluation of objects, and comprising preference (liking or disliking), mood (a general state), aesthetic evaluation and emotion.¹⁶ Emotion is thus a specific type of affective phenomenon. Many definitions emphasize that emotions require an eliciting object or stimulus, differentiating them from moods.¹⁷ Music often serves as the eliciting object, a potentially important event that engenders a response.¹⁸ Changes in emotion can be captured in terms of

¹⁴ David Huron, 'A Psychological Approach to Musical Form: The Habituation-Fluency Theory of Repetition', *Current Musicology* 96 (2013): 7–35 at 9.

¹⁵ David Huron, 'The Ramp Archetype and the Maintenance of Passive Auditory Attention', *Music Perception* 10, no. 1 (1992): 83–92.

¹⁶ Patrik N. Juslin, 'Emotional Reactions to Music', in *The Oxford Handbook of Music Psychology*, 2nd ed., ed. Susan Hallam, Ian Cross and Michael Thaut (Oxford: Oxford University Press, 2009), 197–214 at 198.

¹⁷ Nicholas Cook, *Analysing Musical Multimedia* (Oxford: Clarendon, 1997), 23; Cohen, 'Music as a Source of Emotion in Film', 880.

¹⁸ Ian Cross and Caroline Tolbert, 'Music and Meaning', in *The Oxford Handbook of Music Psychology*, 2nd ed., ed. Susan Hallam, Ian Cross and Michael Thaut (Oxford: Oxford University Press, 2009), 33–46.

shifts in arousal levels (as described in the preceding section) or valence (positive or negative attributions).¹⁹

Empirical studies have shown consistently that music has a measurable effect on the bodies of listeners.²⁰ Researchers can monitor psychophysiological changes induced by musical stimuli in order to access internal, invisible or pre-conscious responses to music that may belie self-reported arousal levels or perceived emotional changes; these measures can include heart rate, systolic and diastolic blood pressure, blood volume, respiration rates, skin conductance, muscular tension, temperature, gastric motility, pupillary action or startle reflexes.²¹ Additionally, researchers have used mismatch negativity (MMN) from electroencephalogram (EEG) data to understand neuronal responses to the event-related potentials (ERP) of particular sounds; these neural processes result in increased oxygenation of the blood that changes the local magnetic properties of certain tissues, allowing researchers to capture these changes through the use of functional magnetic resonance imaging (fMRI).²² In other words, a person's body can register responses to the eliciting stimulus at a pre-conscious level, even if they report that they are not experiencing an emotion. However, some bodily responses might be idiosyncratic and influenced by the subjective experiences of the player.²³ Though emotions have a biological basis, there are also a range of socio-cultural influences on the expression of particular emotions, and psychological mechanisms that serve as mediators between external events and the emotional response they elicit.²⁴

¹⁹ James A. Russell, 'A Circumplex Model of Affect', *Journal of Personality and Social Psychology* 39 (1980): 1161–78.

²⁰ Ivan Nyklíček, Julian F. Thayer and Lorenz J. P. Van Doornen, 'Cardiorespiratory Differentiation of Musically-Induced Emotions', *Journal of Psychophysiology* 11, no. 4 (1997): 304–21; this study investigated autonomic differentiation of emotions – whether cardiorespiratory measures could differentiate between discrete emotions. They found successful differentiation based on two components: respiratory (relating to arousal levels) and chronotropic effects (changes in heart rate).

²¹ David A. Hodges, 'Bodily Responses to Music', in *The Oxford Handbook of Music Psychology*, 2nd ed., ed. Susan Hallam, Ian Cross and Michael Thaut (Oxford: Oxford University Press, 2009), 183–96.

²² Laurel J. Trainor and Robert J. Zatorre, 'The Neurobiology of Musical Expectations from Perception to Emotion', in *The Oxford Handbook of Music Psychology*, 2nd ed., ed. Susan Hallam, Ian Cross and Michael Thaut (Oxford: Oxford University Press, 2009), 285–306.

²³ Hodges, 'Bodily Responses to Music.'

²⁴ It is worth noting that there is a small but developing literature specifically on the psychology of game music; see for example Mark Grimshaw, Siu-Lan Tan and Scott D. Lipscomb, 'Playing with Sound: The Role of Music and Sound Effects in Gaming', in *The Psychology of Music in Multimedia*, ed. Siu-Lan Tan, Annabel J. Cohen, Scott D. Lipscomb and Roger A. Kendall (Oxford: Oxford University Press, 2013), 289–314; Inger Ekman, 'A Cognitive Approach to the Emotional Function of Game Sound', in *The Oxford Handbook of Interactive Audio*, ed.

The BRECVEMA Framework

One thing is clear from the existing bodies of research: sound induces emotional responses with direct physiological implications that are objectively measurable. However, researchers are still exploring exactly *what* emotions music can evoke, and the mechanisms behind *how* sound or music influences listeners. One influential model for understanding this process is the 'BRECVEMA framework', developed by Juslin and Västfjäll.²⁵ The BRECVEMA framework comprises eight mechanisms by which music can evoke an emotion:

- *Brain Stem Reflex*: an automatic, unlearned response to a musical stimulus that increases arousal (e.g., when a player is startled).
- *Rhythmic Entrainment*: a process where a listener's internal rhythms (such as breathing or heart rate) synchronize with that of the musical rhythm.
- *Evaluative Conditioning*: when a musical structure (such as a melody) has become linked to a particular emotional experience through repeated exposure to the two together.
- *Contagion*: a process where the brain responds to features of the musical stimulus as if 'to mimic the moving expression internally'.²⁶
- *Visual Imagery*: wherein the listener creates internal imagery to fit the features of the musical stimulus.
- *Episodic Memory*: the induction of emotion due to the association of the music with a particular personal experience.
- *Musical Expectancy*: emotion induced due to the music either failing to conform to the listener's expectations about its progression, delaying an anticipated resolution or confirming an internal musical prediction.
- *Aesthetic Judgement*: arising from a listener's appraisal of the aesthetic value of the musical stimulus.

Video game sounds can involve several domains from this framework. For example, a tense sound might cause an anxious response in the player, demonstrating contagion; a sound of injury to the avatar might cause a player to recoil as if they have been hit (which could draw on entrainment, evaluative conditioning and visual imagery). Players might activate

Karen Collins, Bill Kapralos and Holly Tessler (New York: Oxford University Press, 2009), 196–214.

²⁵ Juslin, 'Emotional Reactions to Music', 206–7; Patrik N. Juslin and Daniel Västfjäll, 'Emotional Responses to Music: The Need to Consider Underlying Mechanisms', *Behavioral and Brain Sciences* 3, no. 5 (2008): 555–75.

²⁶ Juslin, 'Emotional Reactions to Music', 204.

one or more of these mechanisms when listening; this may account for individual differences in emotional responses to particular sound stimuli. While studies of game sound in isolation can demonstrate clear effects on the bodies of players, sound in context is appraised continuously and combined with other stimuli.²⁷ Mark Grimshaw suggests a connection between aural and visual modalities while highlighting the primacy of sound for evaluating situations in games.²⁸ Inger Ekman suggests that although the visual mode is privileged in games, this is precisely what grants sound its power.²⁹

Emotion and Game Sound

Although I have been reviewing some of the literature from *music* psychology and cognition, we have seen that many of the same processes apply to communicative *sounds* in the game. The functional boundaries between categories of sounds are largely perceptual, based on the clarity of the auditory image evoked by each stimulus. And yet, there is slippage, and the boundaries are not finite or absolute: Walter Murch has written that even in film, most sound effects are like ‘sound-centaurs’; simultaneously communicative and musical.³⁰ As William Gibbons asked of the iconic descending tetrachord figure in *Space Invaders* (Taito, 1978), ‘Is this tetrachord the sound of the aliens’ inexorable march, is it a musical underscore, or is it both?’³¹ As a result of this inherent ambiguity, the BRECVEMA framework serves as a useful starting point for understanding some of the potential sources of emotion in game sound. Emotion is a clear site of bodily investment, implicating a player’s physiological responses to stimuli, phenomenological experiences and subjective processes of making and articulating meaning. As we have also seen, emotion relates to perception, attention and memory – other emphases in the broader field of cognitive psychology.

²⁷ Ekman, ‘Cognitive Approach’, 197.

²⁸ Mark Grimshaw, ‘Sound and Player Immersion in Digital Games’, in *The Oxford Handbook of Sound Studies*, ed. Trevor Pinch and Karin Bijsterveld (New York: Oxford University Press, 2009), 347–66 at 347–8; Cook, *Analysing Musical Multimedia*, 22.

²⁹ Ekman, ‘Cognitive Approach’, 200; Karen Collins, *Playing With Sound: A Theory of Interacting with Sound and Music in Games* (Cambridge, MA: The MIT Press, 2013), 27.

³⁰ Walter Murch, ‘Dense Clarity – Clear Density’, *Transom Review* 5, no. 1 (2005): 7–23 at 9.

³¹ William Gibbons, ‘The Sounds in the Machine: Hirokazu Tanaka’s Cybernetic Soundscape for *Metroid*’, in *The Palgrave Handbook of Sound Design and Music in Screen Media*, ed. Liz Greene and Danijela Kulezic-Wilson (London: Palgrave, 2016), 347–59 at 348.

If emotion comprises responses to potentially important events in the environment, then sound effects in a video game serve as potent stimuli. These sounds represent salient events that are cognitively appraised by the player using any number of mechanisms from the BRECVEMA framework, leading to changes in their emotional and physiological state and behaviour during gameplay. The sounds also serve as important sites of incorporation, linking the game body of the avatar to that of the player.

Aki Järvinen describes five different categories of emotional response in video games: prospect-based, fortunes-of-others, attribution, attraction and well-being.³² Prospect-based emotions are associated with events and causal sequence and involve expectations (eliciting emotions such as hope, fear, satisfaction, relief, shock, surprise and suspense). Fortunes-of-others emotions are displays of player goodwill and are most often triggered in response to events in massively multiplayer online role-playing games (MMORPGS), where the player feels happy or sorry for another player. Attribution emotions are reactions geared towards agents (other human players, a figure in the game or the game itself): Järvinen states that the intensity of these emotions ‘is related to how the behavior deviates from expected behavior’; thus, a player may experience resentment of an enemy, or frustration at the game for its perceived difficulty.³³ Attraction emotions are object-based, including liking or disliking elements of the game settings, graphics, soundtrack or level design. These emotions can change based on familiarity and are invoked musically by the player’s aesthetic appraisal of the music and sound effects. Finally, well-being emotions relate to desirable or undesirable events in gameplay, including delight, pleasant surprise at winning or achieving a goal, distress or dissatisfaction at game loss. Well-being emotions are often triggered (or at least bolstered) musically, through short fanfares representing minor victories like obtaining items (see Example 16.4), or music representing death (Example 16.5). The intensity of the elicited emotion is proportional to the extent that the event is desirable or undesirable, expected or unexpected in the game context. Well-being emotions frequently relate to gameplay as a whole (victory or failure), as opposed to the more proximal goal-oriented category of prospect emotions. Sound is an important elicitor of ludic emotion (if not *the* elicitor) in four out of Järvinen’s five categories.

³² Aki Järvinen, ‘Understanding Video Games as Emotional Experiences’, in *The Video Game Theory Reader*, ed. Bernard Perron and Mark J. P. Wolf (New York: Routledge, 2008), 85–108.

³³ Järvinen, ‘Emotional Experiences’, 91.

Example 16.4 *Metroid*, 'Item Found'

♩ = 110

Pulse 1

Pulse 2

Triangle

Example 16.5 *The Legend of Zelda*, 'Death'

♩ = 90

Interactivity, Immersion, Identification, Incorporation and Embodiment

In discussing the invented space between the real world and the bare code, game scholars speak of interactivity, immersion, transportation, presence, involvement, engagement, incorporation and embodiment, often conflating the terms or using them as approximate synonyms. What these terms have in common is that they evoke a sense of motion towards or into the game. Interactivity is sometimes broken down into two related domains depending on which end of the process the researcher wants to explore: the experience of the user (sometimes described as spatial presence) and the affordances of the system that allow for this experience (immersion).³⁴ The player does not enter the console, but instead a world between; represented and actively, imaginatively constructed. Discussions of terms such as interactivity have tended to privilege either the player or the system, rather than the *process* of incorporation through play.³⁵ Sound helps to create both a site of interactive potential and the process, incorporating the body of the player into the avatar. It is to these modes of traversing and inhabiting the game that I now turn, in order to bring my arguments about sound and gamer experience to a close.

³⁴ Werner Wirth, Tilo Hartmann, Saskia Böcking, et al., 'A Process Model of the Formation of Spatial Presence Experiences', *Media Psychology* 9 (2007): 493–525.

³⁵ Lev Manovich, *The Language of New Media* (Cambridge, MA: The MIT Press, 2001), 56.

Example 16.6 *Super Mario Bros.*, jump and stomp enemy sound effects

Definitions of immersion use a somewhat literal metaphor – a feeling of being surrounded by the game or submerged as if into a liquid.³⁶ However, Gordon Calleja's definition of incorporation involves a process of obtaining fluency through the avatar, 'the subjective experience of inhabiting a virtual environment facilitated by the potential to act meaningfully within it while being present to others'.³⁷ Conscious attention becomes internalized knowledge after a certain amount of play.³⁸ Calleja's work emphasizes process; a player becomes more fluid in each domain over time.³⁹ Incorporation is a more cybernetic connection between player and game; instead of merely surrounding the player, the feedback mechanisms of the game code 'make the game world present to the player while simultaneously placing a representation of the player within it through the avatar'.⁴⁰ Incorporation invokes both presence and the process; it suggests both a *site* in which the bodies of the player and avatar intertwine, and the *stages of becoming*, involving simultaneous disembodiment and embodiment.

As Mark Grimshaw suggests, immersion resulting from game sound is 'based primarily on contextual realism rather than object realism, verisimilitude of action rather than authenticity of sample'.⁴¹ Sounds related to jumping and landing in platformer games help the player to feel the weight and presence of their actions and give them information about when to move next. The speed of play in *Mega Man* (1987) is faster than in *Super Mario Bros.*; the sound effect is tied to landing rather than springing off the momentum of stomping an enemy (see Example 16.6), but it still has an upward contour. Rather than suggesting the downward motion of landing,

³⁶ Winifred Phillips, *A Composer's Guide to Game Music* (Cambridge: The MIT Press, 2014), 38; she describes immersion as 'sinking completely within' a game.

³⁷ Gordon Calleja, 'Digital Games Involvement: A Conceptual Model', *Games and Culture* 2, no. 3 (2007): 236–60 at 254.

³⁸ Gordon Calleja, 'Digital Games and Escapism', *Games and Culture* 5, no. 4 (2010): 254–7.

³⁹ Calleja, 'Digital Game Involvement', 254.

⁴⁰ *Ibid.*

⁴¹ Grimshaw, 'Sound and Player Immersion', 362; see also Mark Grimshaw, 'Sound', in *The Routledge Companion to Video Game Studies*, ed. Mark J. P. Wolf and Bernard Perron (New York: Routledge, 2014), 117–24 at 119: 'What typically characterizes these sound effects is that they conform to a realism of action; *do a sound, hear a sound* (a play on the film sound design mantra of *see a sound, hear a sound*) . . . Through synchronization and realism of action, the sound *becomes* the sound of the depicted event.'

Example 16.7 *Mega Man*, Mega Man landing

this effect gives the player a precise indication of the instant when they can jump again (see Example 16.7).⁴² Jump sounds are immersive because of their unrealism (rather than in spite of it), because of how they engage with the player's cognitive processes and embodied image schema.

James Paul Gee theorizes video games according to studies of situated cognition, arguing that embodied thinking is characteristic of most video games.⁴³ Gee describes the avatar as a 'surrogate', and describes the process of play in this way: 'we players are both imposed on by the character we play (i.e., we must take on the character's goals) and impose ourselves on that character (i.e., we make the character take on our goals)'.⁴⁴ Waggoner takes up the notion of a projective identity as a liminal space where games do their most interesting and important work by influencing and inflecting the bodies of players and game characters.⁴⁵ In his ethnographic work on MMORPG players and their avatars, Waggoner found that players tended to distance themselves from their avatars when speaking about them, claiming that the avatar was a distinct entity or a tool with which to explore the game. Yet, those same players tended to unconsciously shift between first- and third-person language when talking about the avatar, slippage between the real and the virtual that suggests a lack of clear boundaries – in other words, experienced players tended to speak from this space of projected identity.

The avatar's body is unusual in that it becomes something both inhabited and invisible; a site of both ergodic effort and erasure.⁴⁶ This has led

⁴² The speed of this effect evokes a sensation of one foot landing just slightly before the other.

⁴³ James Paul Gee, 'Video Games and Embodiment', *Games and Culture* 3, nos. 3–4 (2008): 253–63 at 258; I was similarly inspired to explore this topic after a realization that not only are most games embodied, but they use health and ability as a core game mechanic.

⁴⁴ Gee, 'Video Games and Embodiment', 260; see also James Paul Gee, *Why Video Games are Good for Your Soul* (Altona, Victoria: Common Ground, 2005), 56; James Paul Gee, *What Video Games Have to Teach Us About Learning and Literacy* (New York: Palgrave Macmillan, 2007).

⁴⁵ Zach Waggoner, *My Avatar, My Self: Identity in Video Role-Playing Games* (Jefferson: McFarland, 2009), 15.

⁴⁶ Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature* (Baltimore, MD: The Johns Hopkins University Press, 1997), 1; Aarseth gives several examples of ergodic (or non-trivial) effort in traversing a text, such as eye movement and the turning of pages. But, on page 94 he gives a very different (and less-cited) definition of particular use to game studies: 'a situation in which a chain of events (a path, a sequence of actions, etc.) has been produced by the nontrivial efforts of one or more individuals or mechanisms'. Both imply qualities of the text itself that are

some theorists to treat games as a ‘simultaneous experience of disembodied perception and yet an embodied relation to technology’, a notion I find compelling in its complexity.⁴⁷ Despite the myriad contested models, what is clear is that gameplay creates a unique relationship to embodied experience, collapsing boundaries between the real and virtual and suggesting that a person can exist in multiple modes simultaneously through identifying *with* and *as* a digital avatar. The player body remains intact, allowing the sensation and perception of the gameworld that is vital to begin the process of incorporation.⁴⁸ The controller serves as a mediator and even as a kind of musical instrument or conductor’s baton through which the player summons sound, improvises and co-constitutes the soundscape with the game. Through the elicitation of sound and movement in the game, the controller allows for the player’s body to become technologically mediated and more powerfully incorporated. But the controller does not extend the body into the screen – our embodied sensations and perceptions of the gameworld do that. Sound is the modality through which the gameworld begins to extend out from the screen and immerse us; sound powerfully engages our cognitive and physiological mechanisms to incorporate us into our avatars. The controller summons sound so that we may absorb it and, in turn, become absorbed.

Despite the numerous technological shifts in game audio in the past thirty years, my response to the sounds and musical signals is just as powerful as it was at the age of four. It is *still* impossible to imagine Hyrule without its characteristic sounds, though the shape and timbre of these cues in 2017’s *The Legend of Zelda: Breath of the Wild* have a slightly different flavour from those of the original games in the franchise, with sparse, pianistic motives cleverly playing against the expansiveness of the open world map. I still experience a sense of achievement and fulfilment finding one of the hundreds of Korok seeds hidden throughout the land (Example 16.8), the

activated by the audience or player. The word (particularly its association with non-trivial effort) appears often in work on cybertexts, electronic literature, interactive fiction and game studies, as in the work of Noah Wardrip-Fruin.

⁴⁷ Martti Lahti, ‘As We Become Machines: Corporealized Pleasures in Video Games’, in *The Video Game Theory Reader*, ed. Mark J. P. Wolf and Bernard Perron (London: Routledge, 2003), 157–70 at 168.

⁴⁸ Timothy Crick, ‘The Game Body: Toward a Phenomenology of Contemporary Video Gaming’, *Games and Culture* 6, no. 3 (2011): 259–69 at 266: ‘If there were no reasons to assume that the virtual body in a game world is the same as our body in the actual world, then the US Marines’ licensing of the classic FPC video game Doom (Id Software, 1993) in the mid-1990s may suggest otherwise. Licensed and rebuilt as ‘Marine Doom’, the use of video games as military training tools is particularly instructive in establishing a phenomenological link between embodied perception in the virtual and real world.’

Example 16.8 *The Legend of Zelda: Breath of the Wild*, Korok seed/small collectible item sound effect



Example 16.9 *The Legend of Zelda: Breath of the Wild*, heart container/stamina vessel sound effect



Example 16.10 *The Legend of Zelda: Breath of the Wild*, spotted by a Guardian cue, bars 1–2



flush of pride and triumph from exchanging spirit orbs for heart containers or stamina vessels (Example 16.9) and a rush of panic from the erratic tingling figure that indicates that I have been spotted by a Guardian and have mere seconds to avoid its searing laser attack (Example 16.10).

Game sound is one of the most important elicitors of ludic emotion. Sound is uniquely invasive among the senses used to consume most media; while the player can close their eyes or turn away from the screen, sound will continue to play, emitting acoustical vibrations, frequencies that travel deep inside the ear and are transmitted as electrical signals to the auditory processing centres of the brain. Simply muting the sound would be detrimental to game performance, as most important information about the game state is communicated, or at least reinforced, through the audio track in the form of sound effects.⁴⁹

⁴⁹ Kristine Jørgensen, 'Left in the Dark: Playing Computer Games with the Sound Turned Off', in *From Pac-Man to Pop Music: Interactive Audio in Games and New Media.*, ed. Karen Collins (Aldershot: Ashgate, 2008), 163–76 at 167; Jørgensen's work suggests that turning off the sound results in decreased performance in the game and increased anxiety for players.

Thus, the game composers and sound designers hold the player enthralled, immersing them in affect, manipulating their emotions, their physiological arousal levels, their exogenous and endogenous attention and their orienting responses. The player cannot escape the immense affective power of the soundscape of the game. Empirical work in game studies and the psychology of music has a lot of work to do in order to fully understand the mechanics behind these processes, but an appreciation for the intensity of the auditory domain in determining the player's affective experience will help direct future investigation. Through all of these mechanisms and processes, game sound critically involves the body of the player into the game by way of the soundscape; I argue that the soundscape is vital to the process of incorporation, joining the material body of the player to those in the game.