



University of
Salford
MANCHESTER

Psychophysiological approaches to sound and music in games

Williams, DAH

<http://dx.doi.org/10.1017/9781108670289.019>

Title	Psychophysiological approaches to sound and music in games
Authors	Williams, DAH
Publication title	The Cambridge Companion to Video Game Music
Publisher	Cambridge University Press
Type	Book Section
USIR URL	This version is available at: http://usir.salford.ac.uk/id/eprint/62440/
Published Date	2021

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: library-research@salford.ac.uk.

Psychophysiological approaches to sound and music in games

Duncan Williams, University of Salford

d.a.h.williams@salford.ac.uk

Abstract

Psychological research investigating sound and music has increasingly been adapted to the evaluation of soundtracking for games, and is now being considered in the development and design stages of some titles. This chapter summarizes the main findings of this body of knowledge, including approaches to memory, expectation and cliché. It also describes the application of emotional measurement techniques, and player responses using biophysiological metering and analysis. The distinction between different types of psychophysiological responses to sound and music are explored, with the advantages and limitations of such techniques considered. The distinction between musically-induced and perceived emotional response is of particular relevance to future game design, and the use of biophysiological metering presents a unique opportunity to create fast, continuous control signals for gaming and game sound to maximise player experiences. The world of game soundtracking also presents a unique opportunity to explore sound and music evaluation in ways which traditional musicology might not offer (for example, deliberately antagonistic music, non-linear sound design and so on). The chapter concludes with directions for future research based on the paradigm of biofeedback and bio-controlled game audio.

Introduction

As readers of this book will doubtlessly be very familiar, creating effective soundtracking (which readers should take to include music, dialogue, and/or sound effects), presents particular challenges that are distinct from creating music for its own sake, or music for static sound-to-picture formats (films, TV). This is because gameplay typically affords players agency over the timing of events in the game. One of the great benefits of effective soundtracking is the ability to enhance emotional engagement and immersion with the gameplay narrative for the player – accentuating danger, success, challenges, and so on. Different types of game can

employ established musical grammars which are often borrowed from linear musical domains. Psychophysiological assessment can provide clues as to player states and how previous experience with sound and music in linear domains (traditional non-interactive music and media) might influence a player in a game context (learned associations with major or minor keys, for example) – and moreover suggests that these states are not necessarily distinct (for example, a player might be both afraid and excited – consider, for example, riding a rollercoaster, which can be both scary and invigorating, with the balance depending on the individual). In other words, there is not necessarily an emotional state which is universally ‘unwanted’ in gameplay, but there may be states which are ‘inappropriate’ or emotionally incongruous. These are difficult phenomena to evaluate through traditional measurement techniques, especially in the midst of gameplay, yet they remain of vital importance for the game sound designer or composer. Biophysiological measurement, which includes central and peripheral nervous system measurement (i.e., heart and brain activity, as well as sweat and muscle myography) can be used both to detect emotional states, and to use physical responses to control functional music, for example, in the design of new procedural audio techniques for gaming (AlZoubi, Koprinska, & Calvo, 2008). Music can have a huge impact on our day-to-day lives, and is increasingly being correlated with mental health and wellbeing (MacDonald, Kreutz, & Mitchell, 2012). In such contexts, music has been shown to reduce stress, improve athletic performance, aid mindfulness, and increase concentration. With such functional qualities in mind, there is significant potential for music in game soundtracking, which might benefit from biophysiologicaly-informed computer aided music or sound design. This chapter gives an overview of this type of technology and considers how biosensors might be married with traditional game audio strategies (e.g., procedural audio, or algorithmically generated music), but the general methodology for design and application might be equally suited to a wide variety of artistic applications. Biophysiological measurement hardware is becoming increasingly affordable and accessible, giving rise to music specific applications in the emerging field of brain-computer *music* interfacing (BCMI) (Chew & Caspary, 2011; Daly et al., 2016).

There is a distinction to be made between active and passive physiological control over the audio, or more generally the game environment. Active control means that the user must be able to exert clear agency over the resulting actions, for example,

by imagining a player action and seeing an avatar follow this input with a direct correlation in the resulting sound world. Passive control would include detection of control signals which are not directly controllable by the end user (for example, heart rate, or galvanic skin response for emotional state estimation). This has been proposed as a control signal in horror games such as 2011's *Nevermind* (Giżycka & Nalepa, 2018), in which the gameplay is responsive to electrodermal activity – world designs which elicit higher sweat responses from the player are chosen over worlds with less player response (Velloso, Löhnert, & Gellersen, 2015) - and becomes particularly interesting in game contexts when the possibility of adapting the music to biofeedback is introduced (Blumenstein, Bar-Eli, & Tenenbaum, 1995; Labbé, Schmidt, Babin, & Pharr, 2007; Teitelbaum, 1976). Here, the potential to harness unconscious processes (passive control) suggests that augmented music and sound generation, for example, individually adaptive, responsive, or context-dependent remixing, may be a possibility. Such technology could be married together with the significant advances in music information retrieval (MIR), non-linear music creation (Berndt, 2009), and context-adaptive music selection in the future. For example, it could be used to create systems for unsupervised music selection based on individual player preference, performance, narrative, or directly from biophysiological activity. As with any such system, there are ethical concerns, not just in the storage of individual data, but also in the potential to cause harm to the player, for example by being too intense or scary in this context.

For these types of system to be truly successful, the mapping between neurophysiological cues and audio parameters must be intuitive for the audience (players) and for the composer/sound designers involved. Several such systems have been presented in research (Berndt, 2011; Eladhari, Nieuwdorp, & Fridenfolk, 2006; D. Williams et al., 2015) amongst others, but commercial take up is limited. To understand why this might be, we must first consider how biosensors might be used in a game music context, and how emotions are generally measured in psychological practice, as well as how game music might otherwise adapt to game contexts which require non-linear, emotionally congruous soundtracking in order to maximise immersion. Therefore, this chapter will begin with an overview of emotion and measurement strategies, including duration, scale, and dimensionality, before considering experimental techniques using biosensors, including a discussion of the

use of biosensors in overcoming the difficulty between differentiation of perceived and induced emotional states. This section will include a brief overview of potential biosensors for use in game audio contexts. We will then review examples of non-linearity and adaptive audio in game soundtracking, before considering some possible examples from the future through the lens of a single series of releases. The *Doom* franchise is a useful example because it has a notable stylistic fingerprint, yet also a commitment to continual adaptation in each release, most recently through using voice as a highly effective emotional communication tool.

We know a lot about the power of cutscenes to engage player emotional responses, (Hooper, 2018). Anecdotally, games testers who worked at XBox UK on Quality Assurance for *Lost Odyssey* reported tears across the workspace at a cut-scene depicting the death of character Lirum, for example¹. Casually there is a tendency to treat cut-scenes as entirely discrete objects - in other words as though they were pieces of traditional, linear media (for example television, film etc). However, this somewhat does a disservice to the cutscene in a video game, which will often involve characters that the player has inhabited. Again this gives us a chance to explore the use of biosensors in an evaluation context. For example do players feel heightened emotional response to game cutscenes, in comparison to traditional linear storytelling formats? We might look for for increased heart rate variability, skin conductance (in other words, more sweat), particular patterns of brain activity (particularly frontal asymmetry (Allen & Kline, 2004) or beta frequencies), and so on. In the interactive soundtracking context we consider throughout this chapter, these types of question essentially draw on the distinction between perceived and induced emotions, and which is typically so individual depending on the listeners preferences and current emotional state. The player might ask themselves: "*I understand this section of play is supposed to make me feel sad*" As opposed to "*I feel sad right now at this point in the game*".

Doom and predecessor *Wolfenstein 3D* are the games often credited with the creation of the entire genre of first-person shooters. It would be a sweeping statement to suggest that there was a particular target audience, but the makers Id Games definitely fit a demographic in terms of their own musical taste: they are heavy metal fans. The imagery behind *DOOM* and its marketing owe much to the

¹ <https://www.youtube.com/watch?v=Qemzis8bDVo>

thrash metal bands of the 1980s (Slayer, Exodus, etc). It's no surprise that music in the original game, which used a single piece of music for each level, played homage to that genre (Robert Prince's soundtrack borrowing from Pantera, Alice in Chains, Judas Priest, Slayer, Metallica, and host of other big name metal bands).

More recent versions of *DOOM* have enjoyed a significant upgrade in terms of soundtrack. Mick Gordon, the composer behind most of *DOOM*'s most recent releases (including the VR version), has talked about some of the ideas that inspired his work on the game in public lectures². Whilst interested readers should make the most of the ability to listen to Gordon describe his work first hand, he certainly acknowledges the metal stylings of his soundtracking work, whilst ironically stating that the 'new' soundtrack wasn't supposed to sound "too heavy metal". One particularly intriguing timbre is Gordon's convolved synthesiser, which uses a process of dynamic convolution (Farina & Armelloni, 2005; Gallo & Adviser-Selesnick, 2010) to modulate a synthesiser until it sounds somewhat akin to a Djent (Marrington, 2017) guitar tone. The soundtrack to *DOOM* now makes use of a few levels of pre-recorded music, zoned to intensity according to the gameplay: in particular, low level ambience in scenes without enemy characters, a general action state, and a markedly intense metal-style action sequence which occurs whenever there are significant threats (end of scene bosses, large numbers of enemies in the scene etc). If we imagine that the target audience for the game probably have a passing likeness for metal, this makes perfect sense and requires no personalisation in so far as musical taste or preference, unlike the biofeedback driven examples mentioned earlier.

For more information on the literature behind this see the section on distinguishing felt emotion in this chapter.

[Emotion: what is it, how is it measured, and why should we care?](#)

Psychological evaluations of music as a holistic activity (composition, performance, and listening) exist, as do evaluations of the specific contributions of individual musical features to overall perception. These generally use self-report mechanisms (asking the composer, performer, or listener what they think or feel in response to stimulus sets which often feature multiple varying factors). The domain of

² <https://www.youtube.com/watch?v=U4FNBMZsqrY>

psychoacoustics tends to consider the contribution of individual, quantifiable acoustic features to perception (for example, correlation between amplitude and perceived loudness, spectro-temporal properties and timbral quality, or acoustic frequency and pitch perception). More recently this work has begun to crossover with consideration of the contribution of such factors to emotional communication in a performance and emotional state in a listener (the player in game contexts). For work which considers such acoustic cues as emotional correlates in the context of speech and dialogue, see, for example, (Juslin & Laukka, 2004) — note, though, that as above, these acoustic cues do not necessarily correspond to single musical features.

These domains often borrow from perceptual evaluation in other senses (e.g., food science for taste perception), and most pertinently in the case of considering video game soundtracking, have begun to acknowledge the influence of sound and music when respondents react to multimodal contexts.

At this stage it is useful to define common terminology and the concepts that shape psychophysiological assessments of perceptual responses to sound and music. We do not aim to be exhaustive, rather we would like to provide the reader who might not be familiar with these concepts, with an overview of the psychological phenomenon that might inform the development of music systems in future gaming contexts. Interested readers can find more exhaustive reviews on the link between music and emotion in (Klaus R. Scherer, 2004) and a special issue of *Musicae Scientiae* (Lamont & Eerola, 2011). Much of this literature describes models of affect, including definitions of emotions, emotional components, and emotional correlates that may be modulated by musical features.

There are generally three categories of emotional response to music:

1. *Emotion*: A shorter response; usually to an identifiable stimulus, giving rise to a direct perception, and possibly an action (e.g., fight/flight)
2. *Affect/Subjective feeling*: A longer evoked experiential episode.
3. *Mood*: The longest category, and more diffuse. Perhaps indiscriminate to particular eliciting events, and as such less pertinent to gameplay, though it may influence cognitive state and ability (e.g., negative or depressed moods).

Emotion and action are the most relevant to gameplay scenarios, where player state might be influenced by soundtracking, and as such this chapter will henceforth refer to emotion unilaterally.

Methods for measuring listeners emotional responses to musical stimuli have included self-reporting and biophysiological measurement, and in many cases combinations thereof.

Models of emotions in response to music

A very common model for emotions is the 2-Dimensional or circumplex model of affect (Russell, 1980), wherein positivity of emotion (valence) is placed on the horizontal axis, and intensity of emotion (arousal) on the vertical axis of the cartesian space. This space has been proposed to represent the blend of interacting neurophysiological systems dedicated to the processing of hedonicity (pleasure–displeasure) and arousal (quiet–activated) (Russell, 2003; Russell & Barrett, 1999). In similar ways, the Geneva Emotion Music Scale (GEMS) describes nine dimensions that represent the semantic space of musically evoked emotions (M. Zentner, Grandjean, & Scherer, 2008), but unlike the circumplex model, no assumption is made as to the neural circuitry underlying these semantic dimensions (Klaus R. Scherer, 2004). Other common approaches include single bipolar dimensions (such as *happy/sad*), or multiple bipolar dimensions (*happy/sad*, *boredom/surprise*, *pleasant/frightening* etc.). Finally, some models utilise free-choice responses to musical stimuli to determine emotional correlations. Perhaps unsurprisingly, many of these approaches share commonality in the specific emotional descriptors used. Game music systems that are meant to produce emotionally loaded music could potentially be mapped to these dimensional theories of emotion drawn from psychology. Finally, much of the literature on emotion and music uses descriptive qualifiers for the emotion expressed by the composer or experienced by the listener in terms of basic emotions, like “sad”, “happy” (Ekman, 2007; Izard, 2007). We feel important to note here that, even though listeners may experience these emotions in isolation, the affective response to music in a gameplay context is more likely to reflect collections and blends of emotions, including of course gameplay narrative.

Distinguishing felt emotion

The distinction between perceived and induced or experienced emotions in response to musical stimuli is very important, and has been well documented (see for example (Västfjäll, 2001; J. K. Vuoskoski & Eerola, 2011), though the precise terminology used can vary widely. For example induced (“*I felt this*”, “*I experienced this*”), or perceived (“*this was intended to communicate*”, “*I observed*” and so on).

Perceived emotions are more commonly noted than induced emotions, “Generally speaking, emotions were less frequently felt in response to music than they were perceived as expressive properties of the music.” (M. Zentner et al., 2008) p502fo³. This represents a challenge which becomes more important when considering game soundtracking that aims to induce emotionally congruent states in the player, rather than to simply communicate them. In other words, rather than communicating an explicit goal, for example “*play sad music == sad player*”, the goal is rather to measure the player’s own bioresponse and adapt the music to their actual arousal as an emotional indicator as interpreted by the system.

Immersion as a quality criterion in gameplay

Immersion is a popular criterion in the design and evaluation of gaming. Immersion is often used interchangeably in this context with ‘flow’, (Weibel & Wissmath, 2011), yet empirical research on the impact of music on immersion in the context of videogames is still in its infancy. A player focussed definition (as opposed to the perceptual definition of Slater, (Slater, 2003) wherein simulated environments combine with perception to create coherence) has been suggested as a sense of the loss of time perception (Sanders & Cairns, 2010), but this was contested in later work (Nordin et al., 2013) as anecdotal. Some studies have considered players perceived, self-reported state in an ecological context (Tobin, Bisson, & Grondin, 2010), i.e. whilst actually playing – suggesting that players underestimate how long they actually play for if asked retrospectively. Later research has suggested that attention might be better correlated to immersion (Nordin et al., 2013).

³ For more information on method and epistemology in regards to perceived and induced emotional responses to music, the interested reader may wish to refer to (Gabrielsson, 2001; K. R. Scherer, Zentner, & Schacht, 2002; M. R. Zentner, Meylan, & Scherer, 2000).

Whichever definition the reader prefers, the cognitive processes involved are likely to be narrative dependent, in other words the music needs to be emotionally congruent with gameplay events (Jørgensen, 2008). This has implications for soundtrack timing, and for emotional matching of soundtrack elements with gameplay narrative (D. Williams et al., 2015).

Despite the increasing acknowledgment of sound or music on immersion in multimodal scenarios (which would include video game soundtracking, sound in film, interactive media, VR and so on), there is a problem in the traditional psychophysical evaluation of such a process: simply put, the experimenter will break immersion if they ask players to actively react to a change in the stimulus (e.g., stop playing when music changes), clearly this makes self-report mechanisms problematic when evaluating game music in an authentic manner.

Biosensors in psychophysiological evaluation

Whilst biosensors are now becoming more established, with major brands like FitBit and Apple Watch giving consumers access to biophysiological data, harnessing this technology for game music, or even more general interaction with music (for example, using sensors to select music playlists autonomously) is less common and something of an emerging field. The challenges can be broadly surmised as 1) extracting meaningful control information for improving accuracy and signal clarity (*utility*), and 2) design of game audio soundtracks which respond in a way which that is narratively congruent with both the narrative and the player state as measured in the extracted control signal (*congruence*). In the case of 1), Often research has a tendency to focus on issues of technical implementation, beyond the selection and definition of signal collection metrics outlined above, for example, it can be challenging to remove data artefacts caused by blinking, eye movement, head movement, and other muscle movement related errors (Bhardwaj, Jadhav, Adapa, Acharyya, & Naik, 2015; Jadhav et al., 2014). Related research challenges include increased speed of classification of player emotional state (as measured by biofeedback), for example, by machine learning techniques (Liu, Dale, & Belliveau, 2002), or accuracy of the interface (Lay-Ekuakille et al., 2013). Generally these are challenges related to the type of data processing employed, either statistical data reduction or noise reduction and artefact removal techniques, which can be borrowed from other measurement paradigms, such as functional magnetic

resonance imaging (fMRI) (Feis et al., 2015). Hybrid systems, utilizing more than one method of measurement include joint sensor studies correlating affective induction by means of music with neurophysiological cues (Nicolaou et al., 2017), measurement of subconscious emotional state as adapted to musical control (Ramirez & Vamvakousis, 2012) and systems combining galvanic skin responses (GSR) for emotional assessment of audio and music (Daly, Malik, et al., 2015). One such sensor which has yet to find mainstream use in game soundtracking but is making steady advances in the research community is the Electroencephalogram (EEG) a device for measuring electrical activity across the brain via electrodes placed across the scalp (Niedermeyer & Silva, 2005). Placement varies but there is a standard 10/20 arrangement for electrode placement (Allen & Kline, 2004) – the number of electrodes, successful placement, and quality of signal are all barriers to mainstream use in gaming at the time of writing, and potentially insurmountable in VR contexts due to the need to place equipment around the scalp (which might interfere with VR headsets, and be adversely affected by the magnetic field induced by headphones or speakers).

There is a marked difference between systems for controlling music in a game directly by means of these biosensors, and systems that translate the biophysiological data itself into audio (*sonification* or *musification*) (Baier, Hermann, & Stephani, 2007; Hinterberger & Baier, 2005). The distinction between *sonification* and *musification* is the complexity and intent of the mapping, and has been discussed in (D. Williams, 2016). Several metrics for extracting control data from EEG for music are becoming common. The P300 ERP (Event Related Potential, or 'oddball' paradigm) has been used to allow active control over note selection for real-time sound synthesis (M. Grierson, 2008; Mick Grierson & Kiefer, 2011). Such methods are not dissimilar to systems that use biophysiological data for spelling words, which are now increasingly common in the BCI world, e.g., (Krusienski et al., 2006; Neuper, Müller-Putz, Scherer, & Pfurtscheller, 2006), though the system is adapted to musical notes rather than text input. Stimulus-responsive input measures, for example, SSVEP responses (a measurable response to visual stimulation at a

given frequency)⁴ (Middendorf, McMillan, Calhoun, Jones, & others, 2000), have been adapted to real-time score selection (Miranda & Castet, 2014). SSVEP allows users to focus their gaze on a visual stimulus oscillating at a given rate. As well as active selection in a manner which is analogous to the ERP described above, SSVEP also allows for second level of control by mapping the duration of the gaze with non-linear features. This has clear possibilities in game soundtracking, wherein non-linearity is crucial, as it might allow for a degree of continuous control, for example, after selecting a specific musical motif or instrument set, the duration of a user's gaze can be used to adjust the ratio for the selected passages accordingly. A similar effect could be achieved using eye-tracking in a hybrid system, using duration of gaze as a secondary non-linear control. Active control by means of brain patterns (Mu frequency rhythm) and the player's imagination of motor movement are also becoming popular as control signals for various applications, including avatar movement in virtual reality, operation of spelling devices, and neuroprostheses (McFarland, Miner, Vaughan, & Wolpaw, 2000; Neuper et al., 2006). The challenge, then, is in devising and evaluating mappings which are most suited to game-specific control (non-linear and narratively congruent mappings).

A hypothetical example

Non-linearity is a relatively straightforward concept for game soundtracking – music with a static duration may not fit synchronously with player actions. Emotional congruence is also quickly explained: Let us imagine a MMORPG such as *World of Warcraft*. The player might be in a part of the narrative which requires generically happy sounding music – no immediate threat, for example. The tune which is selected by the audio engine might be stylistically related to ambient, relaxing music, but in this particular case, the player has a phobia of open spaces, wind chimes and drone notes. This soundscape might then create the opposite of the intended emotional state in the player. In related work, there is a growing body of evidence (Jonna K. Vuoskoski & Eerola, 2012; Jonna K. Vuoskoski, Thompson, McIlwain, & Eerola, 2012) which suggests that when sad music is played to listeners who are in a similar emotional state, the net effect can actually be that the listeners' emotional

⁴ SSVEP is a measurable response to visual stimulation at a given frequency seen in the visual cortex. For a detailed explanation the reader is referred to (Middendorf, McMillan, Calhoun, Jones, & others, 2000; Müller-Putz, Scherer, Brauneis, & Pfurtscheller, 2005), and to (D. A. H. Williams & Miranda, 2018) for a review of use in various music related cases.

response is positive, due to an emotional mirroring effect which releases some neurosympathetic responses. There is some suggestion that music has the power to feel as though it is a sympathetic listener, and to make people in negative emotional states feel listened to – essentially, mirroring. So, simply giving generically sad music to the player at an opposing point in the narrative (e.g., low player health, death of an NPC or similar point in the narrative) might also be inappropriate.

One solution might be to apply procedural audio techniques to an individual player's own selection of music (as in the *Grand Theft Auto* radio tuner options), or in an ideal world, to use biophysiological measures of emotion to manipulate a soundtrack according to biofeedback response to best maximize the intended, induced emotional response in an individual player on a case-by-case basis.

Figure 1 shows a suggested signal flow illustrating how existing biosensor research might be implemented in this fashion, in for example a non-linear, biofeedback informed game audio system, based on metrics including heart-rate, electro-dermal activity, and so on, mentioned above.

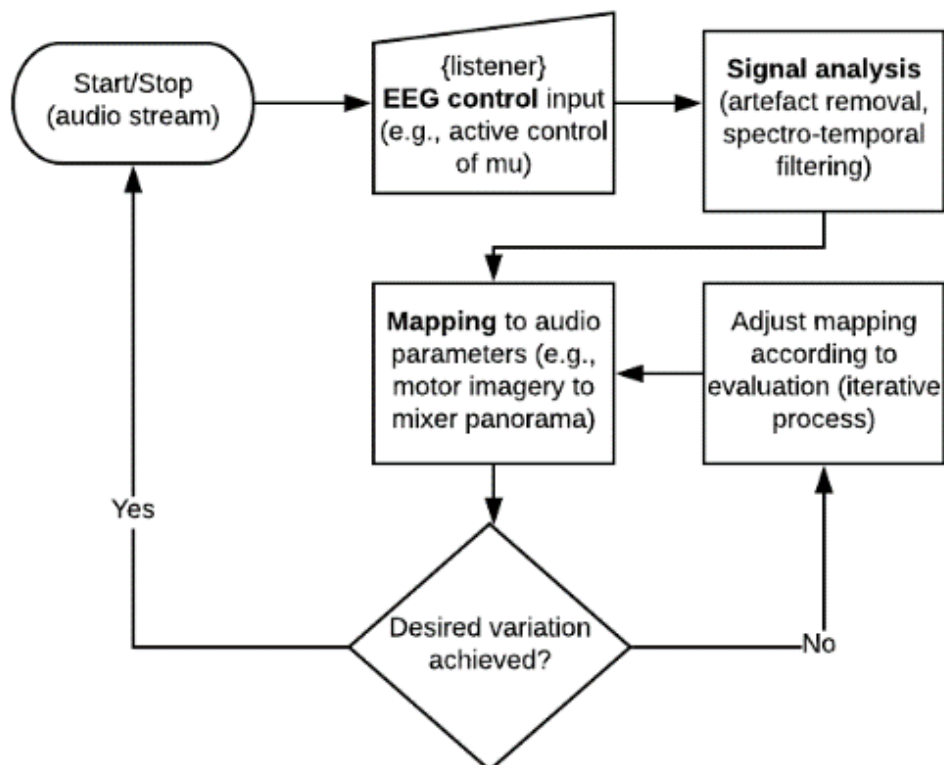


Figure 1. Overview of signal flow and iterative evaluation process. Real-time input is analyzed and filtered, including artefact removal to produce simple control signals. Control signals are then mapped to mixer parameters; for example combining different instrumentation or musical motifs. In the case of a horror game, is the user 'scared' enough? If not, adjust the music. Is the user 'bored', if so, adjust the music, and so on. Any number of statistical data reduction techniques might be used in the signal analysis block - For details of previous studies using principal component analysis for music measurement when analyzing EEG, the interested reader is referred to previous work in (Daly et al., 2014, 2016; Daly, Williams, et al., 2015; Nicolaou et al., 2017).

If we consider one of the best-known examples of interactive, narrative congruous music, the iMUSE system by LucasArts (popularised in *Monkey Island 2*, amongst other games), musical material adapts to the player character's location and the game action (Strank, 2013). For example musical transformations include vertical sequencing (adding or removing instrumentation), and horizontal (time based) resequencing (such as looping sections, or jumping to different sections of a theme). Adding biosensor feedback would allow such changes to respond to a further control signal: that of the player emotional state.

Concluding remarks

Music has been shown to induce physical responses on a conscious, and unconscious level (Grewe, Nagel, Kopiez, & Altenmüller, 2005, 2007). Such measurements can be used as indicators of affective states. There are established methods for evaluating emotional responses in traditional psychology and cognitive sciences. These can, and have, been adapted to the evaluation of emotional responses to music. One such popular model is the 2-dimensional (or circumplex) model of affect. However, this type of model is not without its problems. For example, let us consider a state which is very negative and also very active (low valence and high arousal). How would you describe the player in this condition, afraid? Or angry? Both are very active, negative states, but both are quite different types of emotional response. Hence many other dimensional or descriptive approaches to emotion measurement have been proposed by music psychologists.

This is a difficult question to answer based on the current state of research – we might at this stage simply have to say that existing models do not yet currently give us the full range we need when thinking about emotional response in active gameplay. Again, biosensors potentially offer a solution to overcoming these challenges in evaluating player emotional states (Reuderink, Mühl, & Poel, 2013). However, player immersion remains well-understood by the gaming community and as such is a desirable, measurable attribute (Lipscomb & Zehnder, 2004). Emotional congruence between soundtracking and gameplay is also a measurable attribute, providing the players evaluating this attribute have a shared understanding of exactly what it is they are being asked to evaluate. Considering the affective potential of music in video games is a useful way of understanding a player's experience of emotion in the gameplay.

There are ethical concerns when collecting a player's biophysiological responses beyond the immediate data protection requirements. Should the game report if it notes unusual biophysiological responses? The hardware is likely to be well below medical grade, and false reporting could create psychological trauma. What if in the case of a game like *Nevermind*, the adaptations using biofeedback were too effective – again this could result in psychological trauma. We proposed earlier in this chapter borrowing the principle of the Hippocratic oath – firstly do no harm – when considering these scenarios, and moreover it could be argued that playing a game gives some degree of informed consent in the case of extremely immersive experiences. Within the context of gameplay, the idea of 'harm' becomes somewhat nebulous, but there is certainly room to consider levels of consent for the player in particularly intense games. Returning once more to *DOOM*'s creators, Id games, who offered a number of intensity levels to the player at the beginning of gameplay in *Wolfenstein 3D* (from “*Can I Play Daddy?*” to “*I am Death Incarnate!*”). Nevertheless these are not solutions and are offered to the reader here for consideration.

In order for videogame soundtracking to move towards enhancing player immersion, we might suggest that simple modulation and procedural audio techniques are not yet sophisticated enough in terms of musical features (reviewing the literature correlating emotional response to music), synchronously with player narrative. A fully

realised system might need to incorporate elements of surprise, for example. Biomarkers indicating lack of attention or focus could be used as control signals for such audio cues, depending on the gameplay – ERP in brain activity would be one readily available biophysiological cue. As such, there is a serious argument to be made for the use of active biophysiological sensor mapping to assist access in terms of inclusion for gaming: players who might otherwise be unable to enjoy gaming could potentially take part using this technology, for example in the field of EEG games (Coyle, Garcia, Satti, & McGinnity, 2011; Thomas, Vinod, & Guan, 2013), which remains primarily the domain of research activity at the time of writing, though Facebook have very recently announced the acquisition of the MYO brand of “mind reading wrist wearables”. Significant advances have already been made in the field of brain-computer music interfacing – might bio-responsive control signals be able to take these advances to the next level in terms of inclusive – by which we mean accessible to players with a wide range of access issues, as in the case of the EEG games above - game soundtracking design? The use of musical stimuli to mediate or entrain a player's emotional state (i.e., through biofeedback) also remains a fertile area for research activity (Daly, Williams, et al., 2015; Hinterberger & Baier, 2005). Neurofeedback is becoming increasingly common in the design of brain-computer music interfacing for specific purposes such as therapeutic applications, and this emerging field might be readily adaptable to gaming scenarios. This would potentially offer game soundtrack augmentation to listeners who would benefit from context specific or adaptive audio, including non-linear immersive audio in gaming or virtual reality by harnessing the power of music to facilitate emotional contagion, communication, and perhaps most importantly, *interaction* with other players in the context of multiplayer games (e.g., MMORPG games). At the time of writing these solutions sound somewhat far-fetched, but we have already seen a vast increase in consumer level wearable biosensor technology, and decreasing computational cost associated with technology like facial recognition of emotional state – so these science fiction ideas may well become commercially viable in the world of game audio soon.

References

- Allen, J. J. B., & Kline, J. P. (2004). Frontal EEG asymmetry, emotion, and psychopathology: The first, and the next 25 years. *Biological Psychology*, *67*(1–2), 1–5.
<https://doi.org/10.1016/j.biopsycho.2004.03.001>
- AlZoubi, O., Koprinska, I., & Calvo, R. A. (2008). Classification of brain-computer interface data. *Proceedings of the 7th Australasian Data Mining Conference-Volume 87*, 123–131. Retrieved from <http://dl.acm.org/citation.cfm?id=2449309>
- Baier, G., Hermann, T., & Stephani, U. (2007). Multi-channel sonification of human EEG. *Proceedings of the 13th International Conference on Auditory Display*.
- Berndt, A. (2009). *Musical Nonlinearity in Interactive Narrative Environments*. Ann Arbor, MI: MPublishing, University of Michigan Library.
- Berndt, A. (2011). Diegetic Music: New Interactive Experiences. *Game Sound Technology and Player Interaction: Concepts and Developments*, 60–76.
- Bhardwaj, S., Jadhav, P., Adapa, B., Acharyya, A., & Naik, G. R. (2015). Online and automated reliable system design to remove blink and muscle artefact in EEG. *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 6784–6787. IEEE.
- Blumenstein, B., Bar-Eli, M., & Tenenbaum, G. (1995). The augmenting role of biofeedback: Effects of autogenic, imagery and music training on physiological indices and athletic performance. *Journal of Sports Sciences*, *13*(4), 343–354.
- Chew, Y. C. (Denise), & Caspary, E. (2011). MusEEGk: A brain computer musical interface. *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems*, 1417–1422. <https://doi.org/10.1145/1979742.1979784>
- Coyle, D., Garcia, J., Satti, A. R., & McGinnity, T. M. (2011). EEG-based continuous control of a game using a 3 channel motor imagery BCI: BCI game. *2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, 1–7. IEEE.

- Daly, I., Malik, A., Hwang, F., Roesch, E., Weaver, J., Kirke, A., ... Nasuto, S. J. (2014). Neural correlates of emotional responses to music: An EEG study. *Neuroscience Letters*, *573*, 52–57.
- Daly, I., Malik, A., Weaver, J., Hwang, F., Nasuto, S. J., Williams, D., ... Miranda, E. (2015, September). *Towards human-computer music interaction: Evaluation of an affectively-driven music generator via galvanic skin response measures*. 87–92.
<https://doi.org/10.1109/CEEC.2015.7332705>
- Daly, I., Williams, D., Hallowell, J., Hwang, F., Kirke, A., Malik, A., ... Nasuto, S. J. (2015). Music-induced emotions can be predicted from a combination of brain activity and acoustic features. *Brain and Cognition*, *101*, 1–11.
- Daly, I., Williams, D., Kirke, A., Weaver, J., Malik, A., Hwang, F., ... Nasuto, S. J. (2016). Affective brain–computer music interfacing. *Journal of Neural Engineering*, *13*(4), 46022–46035.
- Ekman, P. (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. Holt Paperbacks.
- Eladhari, M., Nieuwdorp, R., & Fridenfalk, M. (2006). The soundtrack of your mind: Mind music-adaptive audio for game characters. *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, 54. Retrieved from <http://dl.acm.org/citation.cfm?id=1178887>
- Farina, A., & Armelloni, E. (2005). Emulation of Not-Linear, Time-Variant Devices by the Convolution Technique. *Congresso AES Italia, Como*.
- Feis, R. A., Smith, S. M., Filippini, N., Douaud, G., Dopper, E. G., Heise, V., ... others. (2015). ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Frontiers in Neuroscience*, *9*, 395.
- Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Spec Issue, 2001-2002*, 123–147.
- Gallo, M. N., & Adviser-Selesnick, I. (2010). *Vacuum tube emulation, distortion curves, and zero-latency fast convolution methods for digital modeling of tube amplifiers*.

- Giżycka, B., & Nalepa, G. J. (2018). Emotion in models meets emotion in design: Building true affective games. *2018 IEEE Games, Entertainment, Media Conference (GEM)*, 1–5. IEEE.
- Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2005). How Does Music Arouse “Chills”? *Annals of the New York Academy of Sciences*, *1060*(1), 446–449.
- Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion*, *7*(4), 774–788. <https://doi.org/10.1037/1528-3542.7.4.774>
- Grierson, M. (2008). Composing with brainwaves: Minimal trial P300b recognition as an indication of subjective preference for the control of a musical instrument. *Proceedings of International Cryogenic Materials Conference (ICMC'08)*. Retrieved from <http://classes.berklee.edu/mbierylo/ICMC08/defevent/papers/cr1450.pdf>
- Grierson, Mick, & Kiefer, C. (2011). *Better brain interfacing for the masses*. 1681. <https://doi.org/10.1145/1979742.1979828>
- Hinterberger, T., & Baier, G. (2005). Poser: Parametric orchestral sonification of eeg in real-time for the self-regulation of brain states. *IEEE Trans. Multimedia*, *12*, 70.
- Hooper, G. (2018). Sounding the Story: Music in Videogame Cutscenes. In *Emotion in Video Game Soundtracking* (pp. 115–141). Springer.
- Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, *2*(3), 260–280.
- Jadhav, P., Shanamugan, D., Chourasia, A., Ghole, A., Acharyya, A., & Naik, G. (2014). Automated detection and correction of eye blink and muscular artefacts in EEG signal for analysis of Autism Spectrum Disorder. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 1881–1884. IEEE.
- Jørgensen, K. (2008). *Left in the dark: Playing computer games with the sound turned off*. Ashgate.

- Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, 33(3), 217–238.
- Krusiński, D. J., Sellers, E. W., Cabestaing, F., Bayouduh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, 3(4), 299–305. <https://doi.org/10.1088/1741-2560/3/4/007>
- Labbé, E., Schmidt, N., Babin, J., & Pharr, M. (2007). Coping with stress: The effectiveness of different types of music. *Applied Psychophysiology and Biofeedback*, 32(3), 163–168.
- Lamont, A., & Eerola, T. (2011). Music and emotion: Themes and development. *Musicae Scientiae*, 15(2), 139–145. <https://doi.org/10.1177/1029864911403366>
- Lay-Ekuakille, A., Vergallo, P., Caratelli, D., Conversano, F., Casciaro, S., & Trabacca, A. (2013). Multispectrum approach in quantitative EEG: Accuracy and physical effort. *IEEE Sensors Journal*, 13(9), 3331–3340.
- Lipscomb, S. D., & Zehnder, S. M. (2004). Immersion in the virtual environment: The effect of a musical score on the video gaming experience. *Journal of Physiological Anthropology and Applied Human Science*, 23(6), 337–343.
- Liu, A. K., Dale, A. M., & Belliveau, J. W. (2002). Monte Carlo simulation studies of EEG and MEG localization accuracy. *Human Brain Mapping*, 16(1), 47–62.
- MacDonald, R., Kreutz, G., & Mitchell, L. (2012). What is music, health, and wellbeing and why is it important. *Music, Health and Wellbeing*, 3–11.
- Marrington, M. (2017). From DJ to djent-step: Technology and the re-coding of metal music since the 1980s. *Metal Music Studies*, 3(2), 251–268.
- McFarland, D. J., Miner, L. A., Vaughan, T. M., & Wolpaw, J. R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography*, 12(3), 177–186.

- Middendorff, M., McMillan, G., Calhoun, G., Jones, K. S., & others. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8(2), 211–214.
- Miranda, E. R., & Castet, J. (Eds.). (2014). *Guide to brain-computer music interfacing*. London: Springer.
- Müller-Putz, G. R., Scherer, R., Brauneis, C., & Pfurtscheller, G. (2005). Steady-state visual evoked potential (SSVEP)-based communication: Impact of harmonic frequency components. *Journal of Neural Engineering*, 2(4), 123.
- Neuper, C., Müller-Putz, G. R., Scherer, R., & Pfurtscheller, G. (2006). Motor imagery and EEG-based control of spelling devices and neuroprostheses. *Progress in Brain Research*, 159, 393–409.
- Nicolaou, N., Malik, A., Daly, I., Weaver, J., Hwang, F., Kirke, A., ... Nasuto, S. (2017). Directed motor-auditory EEG connectivity is modulated by music tempo. *Frontiers in Human Neuroscience*.
<https://doi.org/10.3389/fnhum.2017.00502>
- Niedermeyer, E., & Silva, F. H. L. da. (2005). *Electroencephalography: Basic principles, clinical applications, and related fields*. Philadelphia (Mass.): Lippincott Williams and Wilkins.
- Nordin, A. I., Ali, J., Animashaun, A., Asch, J., Adams, J., & Cairns, P. (2013). Attention, time perception and immersion in games. *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1089–1094. ACM.
- Ramirez, R., & Vamvakousis, Z. (2012). Detecting Emotion from EEG Signals Using the Emotive Epoc Device. In F. Zanzotto, S. Tsumoto, N. Taatgen, & Y. Yao (Eds.), *Brain Informatics* (pp. 175–184). Retrieved from http://dx.doi.org/10.1007/978-3-642-35139-6_17
- Reuderink, B., Mühl, C., & Poel, M. (2013). Valence, arousal and dominance in the EEG during game play. *International Journal of Autonomous and Adaptive Communications Systems*, 6(1), 45–62.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161.

- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145.
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805.
- Sanders, T., & Cairns, P. (2010). Time perception, immersion and music in videogames. *Proceedings of the 24th BCS Interaction Specialist Group Conference*, 160–167. British Computer Society.
- Scherer, K. R., Zentner, M. R., & Schacht, A. (2002). Emotional states generated by music: An exploratory study of music experts. *Musicae Scientiae*.
- Scherer, Klaus R. (2004). Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them? *Journal of New Music Research*, 33(3), 239–251. <https://doi.org/10.1080/0929821042000317822>
- Slater, M. (2003). A note on presence terminology. *Presence Connect*, 3(3), 1–5.
- Strank, W. (2013). The Legacy of iMuse: Interactive Video Game Music in the 1990s. *Music and Game*, 81–91.
- Teitelbaum, R. (1976). In tune: Some early experiments in biofeedback music (1966-1974). *Biofeedback and the Arts, Results of Early Experiments*. Vancouver: Aesthetic Research Center of Canada Publications.
- Thomas, K. P., Vinod, A. P., & Guan, C. (2013). Design of an online EEG based neurofeedback game for enhancing attention and memory. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 433–436. IEEE.
- Tobin, S., Bisson, N., & Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: A study involving gamers. *PLoS One*, 5(2), e9271.
- Västfjäll, D. (2001). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae, Spec Issue, 2001-2002*, 173–211.

- Velloso, E., Löhnert, T., & Gellersen, H. (2015). *The body language of fear: Fearful nonverbal signals in survival-horror games*.
- Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae, 15*(2), 159–173. <https://doi.org/10.1177/1029864911403367>
- Vuoskoski, Jonna K., & Eerola, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts, 6*(3), 204.
- Vuoskoski, Jonna K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception, 29*(3), 311–317.
- Weibel, D., & Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology, 2011*, 6.
- Williams, D. (2016). Utility Versus Creativity in Biomedical Musification. *Journal of Creative Music Systems, 1*(1). Retrieved from <http://jcms.org.uk/issues/Vol1Issue1/4/4.html>
- Williams, D. A. H., & Miranda, E. R. (2018). BCI for Music Making: Then, Now, and Next. In C. S. Nam, A. Nijholt, & F. Lotte (Eds.), *Brain-Computer Interfaces Handbook* (pp. 191–205). CRC PRESS-TAYLOR & FRANCIS GROUP.
- Williams, D., Kirke, A., Miranda, E. R., Daly, I., Hallowell, J., Hwang, F., ... Eaton, J. (2015). Dynamic game soundtrack generation in response to a continuously varying emotional trajectory. *Proceedings of the 56th Audio Engineering Society Conference*. Presented at the 56th Audio Engineering Society Conference - Audio for Games, Queen Mary, University of London.
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*(4), 494–521. <https://doi.org/10.1037/1528-3542.8.4.494>

Zentner, M. R., Meylan, S., & Scherer, K. R. (2000). Exploring musical emotions across five genres of music. *Sixth International Conference of the Society for Music Perception and Cognition (ICMPC)*, 5–10.