TRANSMISSIONS: THE JOURNAL OF FILM AND MEDIA STUDIES 2017, VOL.2, NO. 2, PP. 83-101.

Katherine Mancey
University of Liverpool

Warping Diegesis: The Evolving Role of the Soundtrack in Virtual Reality Gaming

Abstract

Whilst the study of music in audio-visual media is a well-established academic field, the fast and constant evolution of gaming technology creates the need for a simultaneous evolution in the analytical methodologies of video game music, and results in a need to draw from a wide range of disciplines, from film music studies to cognitive science. The rising popularity of virtual reality gaming, with the mass market introduction of VR headsets such as the Oculus Gear VR and Sony's PSVR opens up a new field of discussion for the evolving role of non-diegetic music in gaming, and its subsequent effect on the player experience. My article addresses the shifting role of music from 2D screens to the in-world experience of virtual reality, and how this alters our perception of diegesis. This is supported by a discussion of varying neuropsychological explanations for the relationship between music and emotion, its effect on various theories of immersion/presence, and their application to this new visual immediacy. Through an analysis of a selection of films and video games, including the most recent *Batman: Arkham VR*, I will be questioning whether the same need for a non-diegetic underscore exists, or if this new visual experience calls for more musical fluidity across the diegesis where soundtrack and in-game audio flow seamlessly. Now that the physical distance between the player and the graphics has been drastically reduced, I question whether traditional non-diegetic video game scores will break the fourth wall rather than enhancing the player's immersive experience. By examining where we draw lines within diegesis, this paper acknowledges the need for a re-evaluation of the relationship between soundtrack and in-game audio, as gaming technology moves away from the traditional static screen experience and into an edgeless game world, highlighting the significance of analysing all game sound as a guide to emotional response.

Key words: music, virtual reality, video games, soundtrack, Batman, film

The soundtrack has been a vital part of audio-visual media since its introduction in early film. However, before discussing the evolution of the soundtrack from silent film to vir-

tual reality gaming, it is important to first contextualise virtual reality within both the gaming world and the wider technology sector. Whilst VR video games and 360º visual experiences have only recently emerged on the mass market, the design principle of Darlene Wheatstone's 1838 stereoscope is still used today for the Google Cardboard and other lower budget VR head-mounted displays for mobile phones, using two different 2D images side by side to build a 3D image. One of the first 'VR' gaming experiences for the mass market was Nintendo's table top console, Virtual Boy. Released in 1995, it used the same stereoscopic principles to create a 3D display but without free head movement. This proved unpopular with global sales of only 770,000 units[1]. Since then, developments in technology, such as an increase in graphical fidelity, have led to greater sophistication in virtual reality technology and in 2016 three premium VR headsets were released, the HTC Vive, the Oculus Rift, and the PSVR. Of the three, the PSVR has been the most commercially successful, perhaps due to its lower price point and easier set-up, requiring a PS4 to use rather than a high specification gaming PC, selling roughly one million units in its first six months.[2]

The HTC Vive and the Oculus Rift have applications outside of the gaming world, with their virtual reality technology starting to be introduced into a range of industries. For example, VR headsets allow architects to better 'feel' the spaces they are designing[3]. There has been a rise in 360º video production too, with popular entertainment outlets such as Buzzfeed sharing numerous 360º videos on their YouTube channel. To be experienced, their videos use stereoscopic principles which require lower-end VR headsets such as Google Cardboard, or they can be viewed without a headset by scrolling around the video, bringing this new cinematic experience to a much wider audience, and beginning to normalise the virtual reality experience in popular culture. Although these videos are not interactive, it shows a rise in the popularity of this visual engulfment, with their 'How Animals See the World' video being viewed over 4 million times[4]. In contrast, the PSVR is only used for gaming as it works in conjunction with the PS4. Due to its popularity and sole gaming function, I will be focusing on the PSVR and games made for the PSVR specifically whilst discussing the shifting role of music from film to 2D video games, through to virtual reality.

The soundtrack is well established as an essential part of film, television, and video games, serving many functions both practically and aesthetically. Before the advent of synchronised dialogue in 1927, films were completely silent but the theatres themselves were far from silent. The machinery involved in projecting films was loud, and audiences treated

[1] Edwards, B. Unraveling The Enigma Of Nintendo's Virtual Boy, 20 Years Later. *Fast Company.* <https://www.fastcompany.com/3050016/unraveling-the-enigma-of-nintendos-virtual-boy-20-years-later> accessed 13th May 2017.
[2] Good, O. S. PlayStation VR sales figures revealed in exec's interview. *Polygon.* <https://www.polygon.com/2017/2/26/14742880/playstation-vr-sales> accessed 17th May 2017.
[3] Fruchter R., Lia, A. *Developing a collaborative workflow from BIM to virtual reality.* (Stanford, CA: Stanford University) (2015).
[4] How Animals See the World, *Buzzfeed.* <https://www.youtube.com/watch?v=HqmBa8FPMx8> accessed 28th August 2017.

going to the movies as a social occasion. This gave music a functional purpose as it could drown out the real world, drawing attention to the screen. Music was seen as part of the performance of a film, with composers Hans Erdmann and Giuseppe Becce questioning in 1927 whether films needed music at all[5], as it was believed that non-diegetic music during dialogue scenes could annoy viewers who were trying to concentrate on the dialogue. But this opinion was not universal; when film makers managed to synchronise dialogue with image, the fidelity of the sound playback was poor. Actors had to stand awkwardly still and close to hidden microphones which lead to unnatural line delivery. This resulted in actors on screen appearing ghostly, as suggested by Adorno and Eisler, 'For the talking picture, too, is mute. The characters in it are not speaking people but speaking effigies'[6]; without music films became 'emaciated, bloodless, lacking in emotional appeal'[7]. It was music's job to bring emotion to scenes and to breathe life back into the ghostly figures, overcoming the 'questions of belief, immediacy and illusion'[8] from the audience. Musical ensembles differed between theatres, with some using full orchestras, but most movie theatres employed solo theatre organists or pianists to accompany the film. After reading the director's 'emotion' annotation for the scene, they would use a mixture of improvisation, original scores, and compiled scores from pre-existing music to express this emotion to their audience; as noted by Stuart Oderman, 'The tremolo increased suspense...A diminished chord called for resolution'[9], creating the foundations of a soundtrack. This use of soundtrack to convey human emotion has carried through to video games. As Collins suggests, sound in video games is a 'simulacrum of the real'[10], whilst this also applies to the reactive sound effects in games, the principles behind the use of sound are still the same, to stop the on-screen action from feeling alien and detached.

We can draw parallels between the use of musical cues in early film and in video games. Just as audiences had to read slides, in early video games the players had to read text to continue the narrative, reinforced by music (see Fig 1.1). Without the capacity to play dialogue, music had to bring the text to life. Whilst still very obviously non-diegetic, this music was a key communicator of diegetic mood. Square Co.'s *Final Fantasy I* (1987) employs this technique. After reaching the Light Warrior Memorial Bridge, a prologue screen opens accompanied by the Final Fantasy overture (see Example 1.1)

[5] Discussed by Reay in *Music in Film: Soundtracks and Synergy* (New York: Wallflower Press) (2014).
[6] Adorno, T. Eisler, H. *Composing for the Films,* (New York: Oxford University Press) (1947) p76.
[7] Bakshy, A. With Benefit of Music. *The Nation,* 27 May (1931) p359.
[8] Lastra, J. Sound Technology and the American Cinema (New York: Columbia University Press) (2000) p64.
[9] Oderman, S. "The sound of silents", *Films in Review* 47:3 (1996) p38.
[10] Collins, K. *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games.* (Cambridge: MIT University Press) (2013) p135.

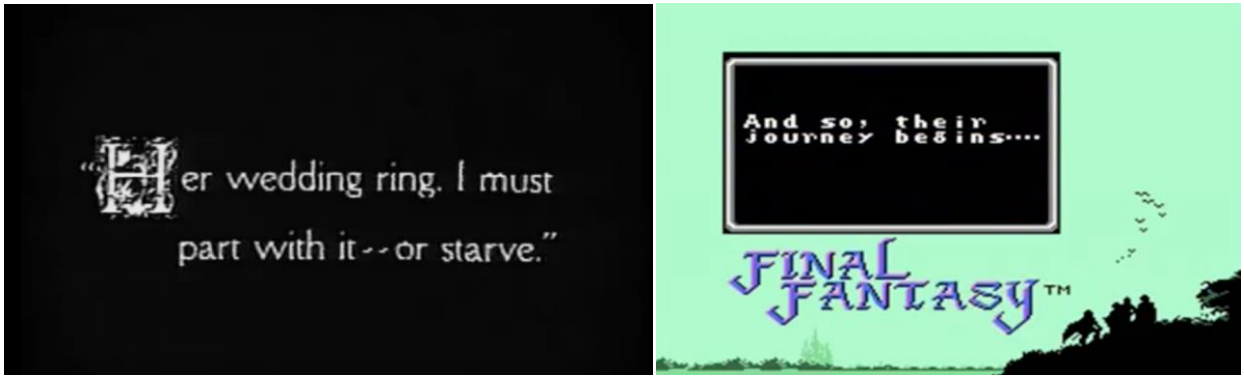


Figure 1.1 (left) On-screen text from Charlie Chaplin's *The Pawnshop* (1916) and (right) text from the prologue of Square Co.'s *Final Fantasy I* (1987).

Example 1.1: <https://www.youtube.com/watch?v=_nWX1M6xTA0>

The music during the prologue has to set the scene for adventure, achieving this through a variety of compositional techniques. If we analyse the opening melodic statement (See Fig 1.2 and 0'0" to 0'13" from Example 1.1) we can see how its structure mirrors the beginning of a journey. It starts with the tonic note in the bass and the fifth in both melodic lines, followed by a minor seventh leap. The seventh pushes us to the expansion of an upper octave in the safety of the tonic key opposed to a more nostalgic return to the original octave[11]; therefore, its use in the opening of the prologue readies the player for expansive adventure as the on-screen text reads, 'And so, their journey begins…'. The use of the seventh is repeated in sequence; transposed up a tone to G minor it accompanies the text, 'What awaits the Four, they do not know'. By using rising sequential repetition

[11] Maman, F. *The Role of Music in the Twenty-first Century.* (California: Tama-Dõ) (1997).

and a minor chord, tension builds which creates an air of mystery to mirror the text, finishing with a C major 7 chord which propels the music forwards as the ear wants it to resolve, again creating an expectation of exploration and a sense of adventure.

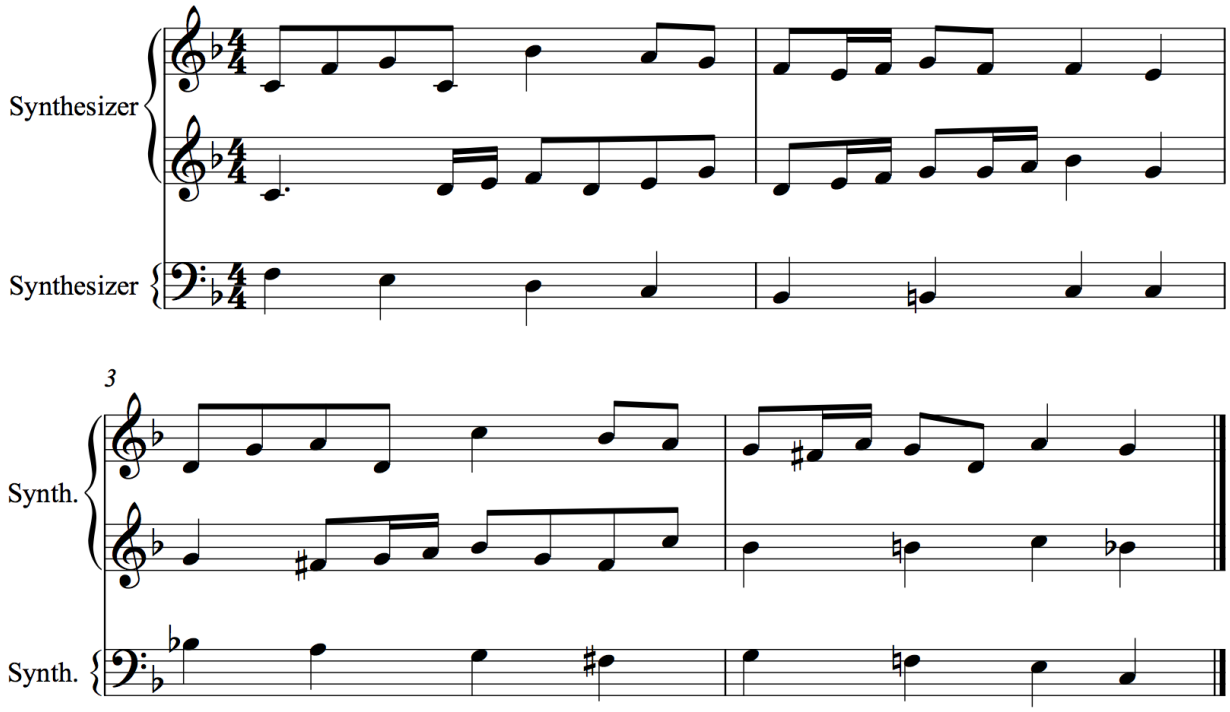


Figure 1.2 Author transcription of the opening four bars from the prologue of Square Co.'s *Final Fantasy I* (1987)

The melody, harmonic structure and overall musical phrasing is reminiscent of baroque music, using a continuo bass and melodic counterpoint alongside binary form and faster-paced chord changes. However, *Finally Fantasy I* is not set in 17th Century Western Europe, therefore the decision to use this baroque style is not reflective of the diegetic world. When *Final Fantasy I* was released for the 8-bit Nintendo Entertainment System, only three audio channels were available, and during gameplay one channel had to be left open for sound effects. As Collins suggests, 'technological constraint has shaped aesthetic decisions'[12]. The use of this baroque style fit nicely within the three-channel system, it allowed for melodic lines to be effective whilst sounding in similar timbres akin to the contrapuntal melodic lines in Bach's piano preludes, building and resealing tension to add depth to the pixilated characters. It also gave a sense of grandeur to the otherwise thin textures, making the resolve to the major tonic feel triumphant.

Final Fantasy I has spawned a series of sequels since its 1987 release. The most recent, *Final Fantasy XV* (2016), has a vastly greater capacity for audio with huge development in both

[12] Collins, K. "In the loop: Creativity and constraint in 8-bit video game audio", *Twentieth-century music* 4:2 (2008) p209.

the technology within the game engine and consoles (such as memory and processing capacity), and the technology used to record the music itself (such as microphone fidelity and digital processing) since the 8-bit three channel system, resulting in a high-fidelity and far more musically intricate soundtrack. It uses the same prologue theme from *Final Fantasy I*, but this time it is heard at the end of the game and in much richer scoring (see Example 1.2).

Example 1.2: <https://www.youtube.com/watch?time_continue=2&v=EojQj4UI58A>

As Summers suggests, 'The total time, on average, that a player spends with any one game is significantly longer than the running time of most mainstream films'[13], and the player's exposure to this prologue motif is significantly longer if they have played through all fifteen games in the series. This familiarity and long-term association with the motif adds greater emotional value to the piece above the sociocultural meaning accompanying its musical style, reinterpreting the prologue theme from the opening game play of *Final Fantasy I* as a finale theme to bring the whole series full circle. Yoko Shimomura's re-orchestration of the prelude theme takes the music from a string quartet, to a piano-led quintet, through to a fanfare using brass, percussion, and a choir. This mirrors the journey the player has taken throughout the game (and throughout the whole series,), moving away from the baroque scoring techniques used in *Final Fantasy I* and using more contemporary scoring techniques with no technological restriction. For example, the opening string quartet uses homophonic harmonic movement with a leading violin melody on top, previously impossible via earlier 8-bit technology (See Fig. 1.3). However, both Example 1.1 and 1.2 are taken from cut scenes. These passive moments from otherwise interactive experiences allow for a linear score which lends itself to filmic musical tropes: scoring can be thicker, and the certainty of linear playback allows for greater melodic and harmonic development. Whilst this interpretation of the theme is much grander and has been humanised through the use of acoustic timbres and performance techniques such as varying vibrato, the music

[13] Summers, T. "Wagner and Video Game Music" in: *Music in Video Games: Studying Play* eds. Donnelly, K. J., Gibbons, W., Lerner, W. (UK: Routledge) (2014) p206.

is still firmly non-diegetic and there is a clear distinction between the sounds and timbres of the game world and those of the soundtrack.

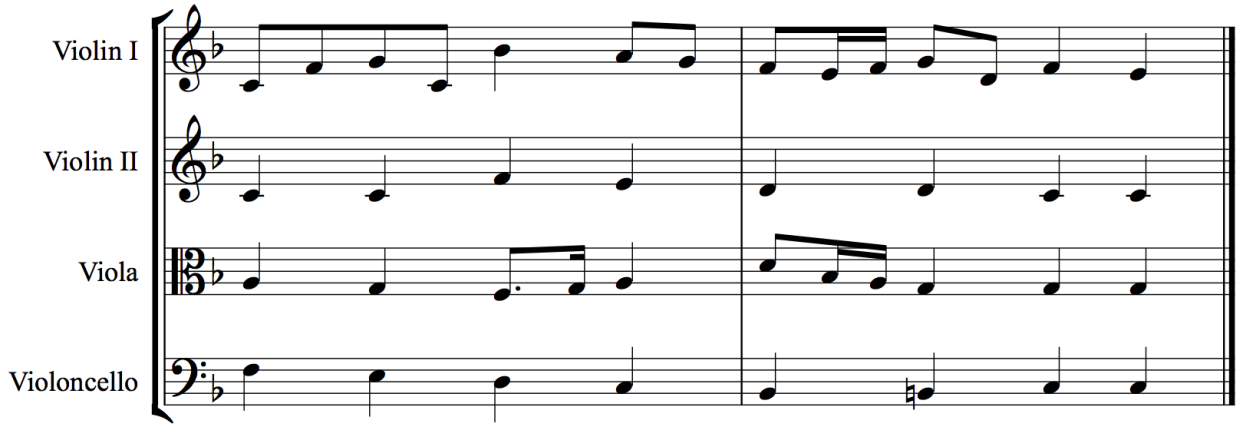


Figure 1.3 Author transcription of the opening string quartet scoring from the prologue theme in Square Exin's *Final Fantasy XV* (2016)

Numerous studies have been conducted analysing the effect that music has on emotion; Schäfer et al. found in their study that the faster the rhythms, the more stressful and dangerous a situation the listener experiences[14]; this correlates with the use of fast-paced music in chase scenes. Chia-Hung Yeh et al. found that participants in their study could easily and accurately identify the emotions behind a piece of popular music even through different cover versions of the piece using varying lyrics in different languages[15], thus showing how powerful melody can be. Lipscomb and Kendall analysed participants' ability to determine which music best fit a scene from Leonard Nimoy's *Star Trek IV: The Voyage Home* from five musical examples. They found that participants could identify the 'correct' music for each scene with a high degree of accuracy, and this level of accuracy increased when scenes included human characters as the music worked to mirror human gestures[16].

This reaction to musical cues works in reverse too, as found in Bullerjahn and Güldering's study in which participants watched intentionally ambiguous film clips accompanied by a range of music, and found that musical genre had a great impact on the participants film outcome predictions, such as 'thriller' music leading them to imagine a more adventurous and violent film ending than the 'melodrama' score which lead to visions of happier, family-centric plot conclusions[17]. Imagining an accompanying image is a commonly reported

[14] Schäfer, T., Huron, D., Shanahan, D., Sedlmeier, P. "The sounds of safety: stress and danger in music perception", *Frontiers in Psychology* 6:9 (2015).
[15] Yeh, CH., Tseng, WY., Chen, CY. et al. "Popular music representation: chorus detection & emotions recognition", *Multimedia Tools and Applications* 73:3 (2014).
[16] Lipscomb, S. D., Kendall, R. A. "Perceptual judgment of the relationship between musical and visual components in film", *Psychomusicology* 13:1 (1994).
[17] Bullerjahn, C., Güldenring, M. "An empirical investigation of effects of film music using qualitative content analysis", *Psychomusicology* 13:1 (1994)

reaction to music[18] and whilst each image in the mind of the listener may differ from the next, Osborne found that listeners often associated relatively similar images to specific music[19], which shows how music can be used as a powerful tool in video games to underpin gameplay and create a vaster picture of a game world than can be shown on the screen. Musical foreshadowing is a powerful tool in video games too; as Whalen suggests, 'the paradigmatic atmosphere music also acts as melodic foreshadowing' in relation to the interactive ocarina feature in Nintendo's *The Legend of Zelda: Ocarina of Time,* in which 'the melodies they learn have an eerie familiarity' as they are heard throughout the game [20]. Therefore, as discovered in each of these studies, music has a profound impact on the players' emotional reading of a scene, with musical cues both aiding the emotional understanding of current gameplay and also creating narrative expectations for gameplay sequences to come.

There are varying neuropsychological explanations for these emotional responses to music. Juslin and Västfjäll suggest that the emotional response is controlled by brainstem reflexes, and that emotions are induced by music because of their fundamental acoustical characteristics, signalling a potentially important and urgent event[21]. For example, a large shift in dynamic from pianissimo to fortissimo is alarming to the brain and can result in an emotion of fear or shock, manipulating an evolutionary response to sound. Berlyne suggests this emotional-arousal property of music, regulated by the brainstem, explains why there is a preference for certain music at certain times[22], such as soft and harmonically predictable music being relaxing. Another explanation for the relationship between music and emotion is evaluative conditioning, which suggests that a formerly neutral stimulus such as music is transformed into a conditioning stimulus by its repeating pairing with an already existing emotionally charged stimulus[23], such as the repeated pairing of kiss scenes in film with lyrical string melodies to create an association between lush string scoring and romance. Another theory, emotional contagion, relates musical expression and gesture with an internal mimicking of that gesture by the listener, which in turn reproduces the emotions felt by the performer. This phenomenon has been seen in facial expressions in which musical expressions of fear have been recreated in the tensing of certain facial muscles by the listener[24]. Davies's theory of emotional episodic memory works with theories of the personal semiotic values of music, as it suggests that specific emotions can be induced by the relationship between an emotional memory and the sounds of that time, also

[18] Tan, S. L., Kelley, M. E. "Graphic representations of short musical compositions", *Psychology of Music* 32:2 (2004).
[19] Osborne, J. W. "A phenomenological investigation of the musical representation of extra-musical ideas", *Journal of Phenomenol Psychology* 20:1 (1989).
[20] Whelan, Z. N. *Play Along: Video Game Music as Metaphor and Metonymy* (Florida: University of Florida) (2004) p43.
[21] Juslin, P. N., Västfjäll, D. "Emotional responses to music: the need to consider underlying mechanisms", *Behavioural Brain Science* 31:1 (2008).
[22] Berlyne, D. E. *Aesthetics and Psychobiology.* (New York: Appleton-Century-Croft) (1971).
[23] Field, A. P., Moore, A. C. "Dissociating the effects of attention and contingency on awareness on evaluating conditioning effects in the visual paradigm", *Cognition & Emotion.* 19:1 (2005).
[24] Lundqvist, L. "Facial expressions are contagious" *Journal of Psychophysiol.* 9:1 (1995).

called the 'Honey, they are playing our tune' phenomenon[25], and can also be used to explain the emotional currency of nostalgia. All of these theories attempt to describe the way music can trigger emotion within the listener, and whilst exact emotion responses are unique to each person, they all suggest there is some commonality in emotional reactions to certain audio stimuli, whether it be an evolutionary response or a culturally embedded and subconsciously learned response to certain musical stimuli.

This use of music to aid emotional depth is a significant contributing factor to a player's level of immersion, so that the player can experience the sensation of feeling spatially located in the game world environment[26]. There are a range of theories as to what creates this feeling of immersion and how it can be sustained, but there is a general consensus amongst scholars that immersion is a multi-faceted process. Wirth et al. suggest that this immersion process happens on two levels, influenced by media factors, process components, user actions and user factors. Primarily, the media factors and user factors work together to create both controlled attention, such as the player actively choosing to play a video game, and automatic attention, such as game music drawing the player's eye to the screen. Once the player is engaging with the game, they suggest that immersion moves to a secondary level where the player experiences a 'suspension of disbelief' as they adapt to the laws and physics of the game world[27]. Brown and Cairns suggested a three-tier system: The first is engagement, which includes learning the controls and game mechanics when a conscious effort has to be made to play the game. The second is engrossment, when the controls have become second nature the player is susceptible to emotional influence from the game as their focus moves from learning to experiencing. The third is immersion, when the player is totally emotionally invested in the game at that moment and their own emotions are influenced by the characters and narrative in the game world[28].

Sweester and Wyeth focused on quantifying what creates immersion, developing their GameFlow model, using a table of 'GameFlow criteria for player enjoyment in games' in an attempt to explain the process of becoming immersed in video games[29]. From this table they can award games an individual GameFlow score, but their calculation does not take into account the role of music on any of their criteria, which is arguably a prominent

[25] Davies, S. "Philosophical perspectives on music's expressiveness", in *Music and Emotion: Theory and Research*, eds. Juslin, P. N., Sloboda, J. A. (Oxford: OUP) (2001).
[26] Groner, R., Weibel, D., Wissmath, B. "Dubbing or Subtitling? Effects on Spatial Presence, Transportation, Flow, and Enjoyment", *Journal of Media Psychology* 21:3 (2009).
[27] Bocking, S., Hartmann, T., Klimmt, C., Schramm, H., Vorderer, P., Wirth, W. "Spatial Presence Theory: State of the Art and Challenges Ahead", in *Immersed in Media: Telepresence Theory, Measurement & Technology,* eds. Biocca, F., Freeman, J., IJsselsteijn, W., Lombard, M., Schaevitz, R. J. (New York: Springer) (2015).
[28] Brown, E., Cairns, P. "A Grounded Investigations of Game Immersion". *CHI'04 Extended Abstracts on Human Factors and Computing Systems* (Vienna, April 2004).
[29] Sweester, P., Wyeth, P. "GameFlow: A Model for Evaluating Player Enjoyment in Games", *ACM Computers in Entertainment* 3:3 (2005).

influencer on player immersion, as suggested by Lipscomb and Zehnder's study[30]. Using game play sequences from Electronic Arts' *Lord of the Rings: The Two Towers,* they measured player enjoyment with and without music using a verbal scale and found a positive correlation between music and enjoyment. Nacke, Grimshaw and Lindly's study furthers this notion, analysing the impact of both music and sound effects on player enjoyment in a first-person shooter game, and found through their gamer experience questionnaire that the most enjoyable play occurred when both music and sound effects were present, and that playing without music and sound effects resulted in a tenser and less enjoyable gameplay experience[31]. The fact that these two studies focused on games from distinctly different genres but resulted in similar findings highlights the universal importance of music for player enjoyment, as it created a better environment for player immersion to occur.

The Batman franchise serves well as a case study to discuss the varying use of music to evoke emotion from early film through to virtual reality gaming, with a long lineage of films, video games and now a virtual reality video game for PSVR to analyse. Batman is a long-standing figure in popular culture, with the first Batman story, 'The Case of The Chemical Syndicate' published in Detective Comics issue #27 in May of 1939[32]. He first appeared in films during the 1940s and has since been a staple character of the blockbuster, rebranded in 2005 with Christopher Nolan's *Batman Begins,* taking him further away from the underpants-over-tights comic book hero and into the more sophisticated 'dark' hero that Batman is best recognised as in current popular culture. This new image for Batman proved popular, with the two sequels *The Dark Knight* (2008) and *The Dark Knight Rises* (2012), also directed by Christopher Nolan, earning over $1 billion worldwide.[33]

Throughout his appearances in film, television and video games, Batman has existed at the extremes of musical gesture. Lambert Hillyer's 1943 *The Batman,* released during Hollywood's 'golden age', uses a full orchestral score synonymous with that era of film, with culturally embedded musical tropes such as tremolo strings and rising melodic lines to create suspense (See Example 2.1). In 1966, Batman was brought to television with a live-action series based on the comic books, its theme tune is in stark contrast to the music of Lambert Hillyer's film, with obvious influence from popular music of the 1960s in both musical style and instrumentation, using a jovial, catchy melody (See Example 2.2). In further contrast, the use of music in Christopher Nolan's *Batman Begins* moves away from the heavy orchestral scoring of the 1940s and the up-tempo theme of the 1960s, and takes a more reduced, almost acousmatic approach to scoring Batman's scenes (See Example

[30] Lipscomb, S. D., Zehnder, S. M. "Immersion in the Virtual Environment: The Effect of a Musical Score on the Video Gaming Experience", *Journal of Physiological Anthropology and Applied Human Science* 23:1 (2004).
[31] Nacke, L. E., Grimshaw, M. N., Lindly, C. A. "More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game", *Interacting with Computers* 22:5 (2010).
[32] Daniels, L. *Batman - The Complete History: The Life and Times of the Dark Knight.* (San Francisco: Chronicle Books) (2004).
[33] Subers, R. Around-the-World Roundup: 'Dark Knight Rises' Joins Billionaire Club. *Box Office Mojo.* < http://www.boxofficemojo.com/news/?id=3514&p=.htm> accessed 18th May 2017.

2.3). The first Batman video game, Ocean Software's *Batman,* was released in 1986 for the ZX Spectrum, Amstrad CPC, and other 8-bit personal home computers. Even with the limitations of the three-channel audio system, the game opens with a chip-tune version of the Batman theme from the 1966 television series (See Example 2.4). This ease of adaptation to fit changes in contemporary popular culture aesthetics throughout the past century shows the robust nature of Batman and his sustained popularity as a character, from early film to television through to video games.

Example 2.1: <https://www.youtube.com/watch?v=6zVmIit0PXU>

Example 2.2: <https://www.youtube.com/watch?v=FDWSHNEbkII>

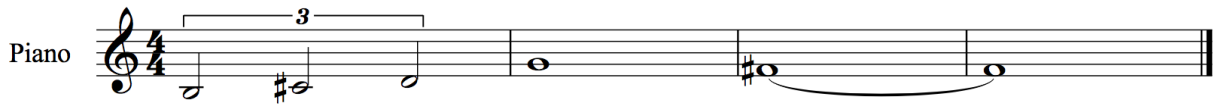
Example 2.3: <https://www.youtube.com/watch?v=r8tT_2QMwxI>

Example 2.4: <https://www.youtube.com/watch?v=I5_cHljImI0>

Just as there is a long and varied history of Batman films, Batman has been a staple character of video games across gaming platforms for decades, appearing in over sixty titles, of which he is the main character in over thirty. The portrayal of Batman in these games has largely stayed in line with his portrayal in film and television. Due to technological limitations of the 8-bit systems, the earliest Batman video games used soundtracks typical of 8-bit games during gameplay but still took musical cues from other Batman media where possible (See Example 2.4). In 1990, SunSoft developed four video games with the title *Batman*; each game was tailored to its platform (Nintendo Entertainment System, Game Boy, Sega Genesis, PC) resulting in a difference in gameplay across platforms, but it is best associated with the release for the Nintendo Entertainment System, selling over a million copies[34]. Whilst it was released just after Burton's film *Batman* (1989)*,* the game's narrative does not bear much resemblance to the film, and Naoki Kodaka's soundtrack is independent of Elfman's film score. However, they do share some melodic characteristics. For example, in the first level (See Example 2.5) the overall melodic contour is similar to that of Elfman's scoring for the film's opening scene (See Example 2.6), with both melodies rising up from the tonic to the sixth and falling to a sustained fifth (See fig 2.1).

Example 2.5: <https://www.youtube.com/watch?v=HktXo2AgWBk>

[34] Strangman, R. *Memoirs of a Virtual Caveman.* (North Carolina: Lulu Press) (2014).



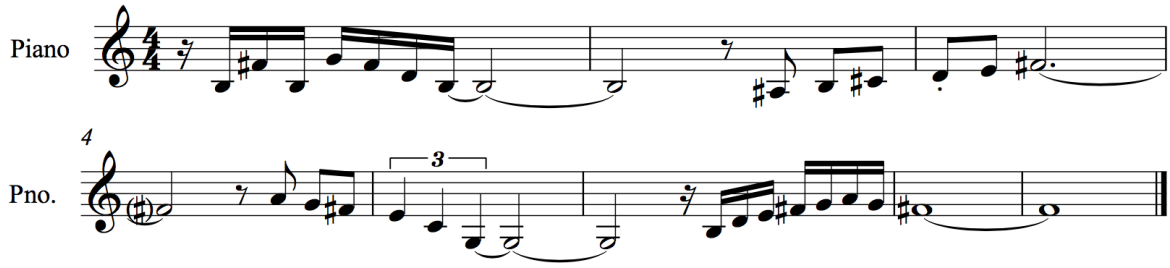


Figure 2.1 Author transcription of (above) the 'Batman' theme from the opening scene of Burton's *Batman* (1989) and (below) the melody from the first level of Sunsoft's *Batman* (1990)

Example 2.6: <https://www.youtube.com/watch?v=vXGXllCMOQg>

As the possibilities of game audio increased, the music of Batman video games became more stylistically reminiscent of film scores as it was possible to use a full orchestral linear score during interactive cinematic cut-scenes and more sophisticated modular cues that are triggered during gameplay. Tsunami Music's score for Eurocom's Batman Begins (2005) is far more closely linked to Nolan's film Batman Begins (2005) than Sunsoft's Batman (1990) was to Burton's Batman (1989) in both character representation (See fig 2.3) and soundtrack, due to improvements in both graphic and audio fidelity. This considerable increase in game audio fidelity caused a shift in video game soundtracks as composers could broaden their use of timbres and move away from the synthesized sounds of 8-bit and 16-bit consoles.

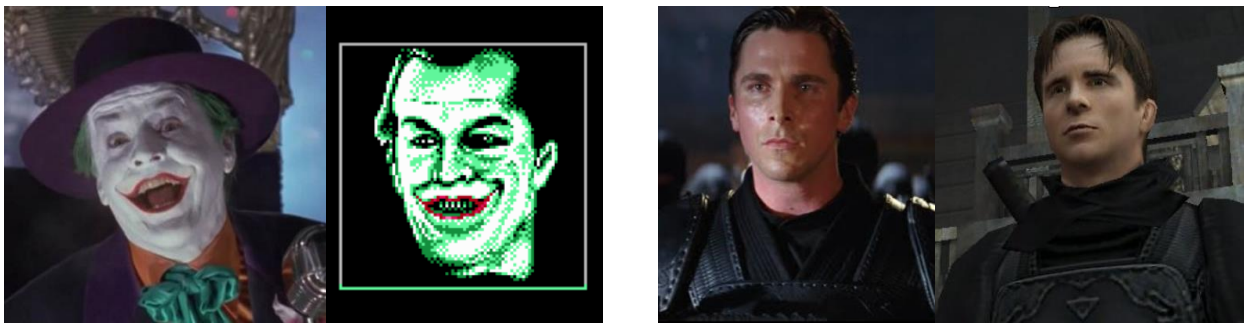


Figure 2.3 (left) Jack Nicholson as the joker in Burton's *Batman* (1989) and Sunsoft's *Batman* (1990) and (right) Christian Bale as Batman in Nolan's *Batman Begins* (2005) and Eurocom's *Batman Begins* (2005)

Whilst the music in Batman video games does reference its use in film, especially during interactive cinematics, film uses a fixed linear narrative which creates a passive audience experience, while video games use a more flexible interactive narrative and active player

engagement which calls for a difference in compositional approach during gameplay. Ryan suggests that interactivity appears on two levels, one constituted by the medium and one intrinsic to the work itself. She suggests these can exist as internal, where the user plays the role of an individual in the narrative, or as external, where the user has a god-like role. This creates two main types of user experience: the first is exploratory, in which the user has no creative power in the story world, such as in films; the second is ontological, in which user actions create objects and cause effects that bring lasting change, as in video games. These can be combined to create four types of interactivity: external-explanatory, external-ontological, internal-ontological, and internal-exploratory[35]. Dogramain and Liptay discuss the relationship between 'solid' and 'liquid' realities, suggesting that the anticipation of liquid spaces breaks down the barrier between the viewer and media[36]. Although they discuss this in relation to visual art, this anticipation of liquid space works within Ryan's descriptions of interactivity as the premise of a liquid narrative environment invites player immersion, as seen in the external-ontological and internal-ontological narrative structures. This ontological narrative results in a need for dynamic, reactive scoring, based on modular cues. As discussed, music has a prominent effect on emotion, but the player's ability to direct the narrative leads to the condensing of musical tropes to short looping cues which can seamlessly fade in and out of the non-diegetic score during gameplay. This use of adaptive scoring further cements the internal-ontological status of the player within the game world, as they are leading the soundtrack through their narrative choices.

Though there is a need for different approaches to music with linear versus modular scoring, film and 2D video games share a need for music to draw focus to the screen and to help transport the audience/player into that 2D world. However, virtual reality gaming brings a new visual experience to the player, totally separate from that of 2D video games, film and even 3D film. It creates for the first time an edgeless game world, where the player experiences total visual engulfment. With the headset on, there is no room to look away from the screen, giving the game world an immediate sense of vastness. There is no separation between the player and the action, with characters appearing to be standing directly in front of their face. This alters music's function as it no longer needs to attract the player's attention to the screen, and this change in function can be seen in comparing the opening of Rocksteady's *Batman: Arkham Asylum* (2009) with their *Batman: Arkham VR* (2016). The opening of *Batman: Arkham Asylum* (See Example 3.1), uses orchestral timbres such as horn swells, rising sequential string melodies and a choral pad. In stark contrast to this, the opening of *Batman: Arkham VR* (See Example 3.2) uses a more processed string sound, synthesised pad swells and a melody heard once in octaves on horn and strings before launching straight into thick atmospheric sounds. There is no melodic development, more synthesised timbres, thinner scoring, and a much shorter title sequence. It still depicts the darkness of Gotham City and the image of Batman as the dark hero, but in a musically simplified way in comparison with the 2D game. The 'Suit Up'

[35] Ryan, M. Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity in Literature and Electronic Media (Baltimore: Johns Hopkins University Press) (2015).
[36] Dogramain, B., Liptay, F. "Immersion in the Visual Arts and Media", in *Immersion in the Visual Arts and Media* eds. Dogramain, B., Liptay, F. (Leiden: Brill) (2015).

scene in *Batman: Arkham VR* (See Example 3.3) does use a slightly thicker score for this interactive cinematic, but it is still far sparser than the music of the 2D game. We hear the theme but again in octaves with no countermelody or obvious harmonic reinforcement, underpinned by a synthesised pulse and low horn and string pad which swells when the player is not engaging with the game mechanics. This use of more prominent processed timbres and less prominent orchestral timbres both reinforces the visuals of the technology surrounding the player and also compliments the new 'high-tech' experience of virtual reality.

Example 3.1: <https://www.youtube.com/watch?v=DHXR4QIQPsA>

Example 3.2: <https://www.youtube.com/watch?v=RTiTnY9FPgk>

Example 3.3: <https://www.youtube.com/watch?v=_bedb1UBtEc>

Another explanation for this use of reduced scoring could be the change in the physical delivery of the sound to the player. More than ever before players are encouraged to use surround sound headphones with the PSVR; a pair is included with the headset and there is a headphone jack on the headset control with its own volume control, making it easy to adjust the sound. This change in delivery plays with the notion of diegesis as the sound feels closer to the player and there is greater accuracy in the reception of stereo imaging. A busy soundtrack using orchestral timbres runs the risk of breaking the fourth wall for players as it becomes more obviously non-diegetic whilst they are totally visually situated within this game world. Using these electronic timbres that are more congruent with sounds of the game world creates a less obviously non-diegetic soundtrack which can aid player immersion and comfort without making players hyper aware that they are playing a video game. This soundtrack is reduced further during gameplay, when moving sound creates a 360º audio picture to accompany the image and to further reinforce the player's feeling of being surrounded by the game and at the centre of the image (See Example 3.4). Just as the player is totally encompassed by visuals, they are totally cloaked in sound. Techniques applied to orchestral scores to create tension are instead applied to game world sounds to create a sense of unease without it being obviously musical. There is a constant pitter-patter of rain, setting a pulse, then rhythmic fragments are created by the blowing wind which moves between left and right and occupies the higher-end frequencies usually inhabited by violins and upper woodwind in an orchestral setting. This is underpinned by a consistent sub-bass rumble, then dissonances used to create tension come from the moving sounds of trams which pan from side to side and slide in pitch from high to low or low to high depending on their starting position in relation to the player, reminiscent of the doppler effect. Where these sounds intersect, dissonances are created, as if these tram sounds were an evolving instrumental pad; the dissonances do not resolve, they simply fade away as the trams pass, building suspense as the ear waits for a solid resolution.

Example 3.4: <https://www.youtube.com/watch?v=sTsjbn-6e30>

This use of music alters our perception of diegesis as it uses a mixture of diegetic and non-diegetic timbres that work together to create a soundtrack, existing somewhere in between the two sound worlds. It also changes the relationship between music and sound effects: rather than existing as two independent devices, they exist as one multifunctional device. Barnabé suggests that the notion of space in video games can be broken down into three concepts. The first, cosmos, is the game world that exists regardless of play participation, game narrative, or game objective. The second, diegesis, is the cosmos which is oriented by the story as its construction is necessary to allow for gameplay. The third, ludiegesis, is the diegesis as ruled by the player's actions, the space that exists as a result of the player's decisions[37]. If we apply this framework to virtual reality, the collaboration between in-game and external sound sources facilitates the smooth transitions between these virtual spaces as the player moves from the diegetic, such as an interactive cinematic scene, to the ludiegesis during gameplay. For example, Batman transitions from the ludiegesis, where he has the ability to capture the Penguin, to the diegesis, where he listens to the Penguin's dialogue, back to the ludiegesis, without an obvious break between the fixed scene and the gameplay as there is no change in the sound world (See Example 3.4). Virtual reality easily accommodated this transition as the characters appear to be directly in front of the player, so a full screen close-up of their face can appear in both gameplay and the cinematic without needing an obvious change in camera angle. The seemingly randomised interactions between the various sound sources that creates the soundtrack also means there is no sense of interruption to the underlying score as the player transitions between these game states, allowing the player to feel as if they are always actively engaged in play even when a non-player character is delivering dialogue to progress the narrative.

From Wheatstone's stereoscopic invention to modern day virtual reality technology, there has been growing interest in experiencing total visual engulfment. Although it has many cross-industry applications, the largest mass market for virtual reality is currently in video games, and an increasing number of game developers are creating virtual reality-only games and adding virtual reality downloadable content to their other titles. Throughout history, sound has accompanied moving image and has adapted to developments in visual and in audio technology. By analysing the way soundtracks have been used from early film through to virtual reality, we can see its evolution from the thick romantic scoring of the 'golden age', to the baroque stylings of early video games, to the use of modern film music technique. However, what has been shared throughout all these iterations of soundtrack is the use of a clear distinction between the diegetic world and the non-diegetic music, whether this is the live accompaniment during silent films or the full orchestral scores used in Rocksteady's *Batman: Arkham Asylum* (2009). Virtual reality brings with it another change in music scoring technique, sound design and sound delivery to the player. It shares

[37] Barnabé, F. 'Storytelling in Videogames: From 'Narraive' Towards 'Fictional Universe', presented at: *Poetics of the Algorithm - Narrative, the Digital, and 'Unidentified' Media.* Liège, 18th June 2016. <http://orbi.ulg.ac.be/handle/2268/198777> accessed 5th January 2017.

some fundamental characteristic with 2D video games, such as anchoring the player in the game world, but there is a contrast between the way these are executed. Now, the player's perception of diegesis is warped as they are wholly visually engulfed in the game world, and the use of greater fluidity between music and game-world sound reflects this. We are still at the beginning of mass market virtual reality gaming, and whilst it is popular, it could take a few generations of virtual reality technology and a decrease in system prices before more players become accustomed to the new visual experience. However, through combining theories of immersive play, music's representation of emotion, and its influence on feelings, with a discussion of the development of soundtrack over the past century, it is clear that music and sound are still vital to the gaming experience. The new game world of virtual reality may be edgeless, but our field of vision is still 114º, needing sound to add emotional context and depth to build a 360º experience.

References

Adorno, T. Eisler, H. *Composing for the Films,* (New York: Oxford University Press) (1947).

Bakshy, A. With Benefit of Music. *The Nation,* 27 May (1931).

Barnabé, F. 'Storytelling in Videogames: From 'Narraive' Towards 'Fictional Universe', presented at: *Poetics of the Algorithm - Narrative, the Digital, and 'Unidentified' Media.* Liège, 18th June 2016. <http://orbi.ulg.ac.be/handle/2268/198777> accessed 5th January 2017.

Berlyne, D. E. *Aesthetics and Psychobiology.* (New York: Appleton-Century-Croft) (1971).

Bocking, S., Hartmann, T., Klimmt, C., Schramm, H., Vorderer, P., Wirth, W. "Spatial Presence Theory: State of the Art and Challenges Ahead", in *Immersed in Media: Telepresence Theory, Measurement & Technology,* eds. Biocca, F., Freeman, J., IJsselsteijn, W., Lombard, M., Schaevitz, R. J. (New York: Springer) (2015).

Brown, E., Cairns, P. "A Grounded Investigations of Game Immersion". *CHI'04 Extended Abstracts on Human Factors and Computing Systems* (Vienna, April 2004).

Bullerjahn, C., Güldenring, M. "An empirical investigation of effects of film music using qualitative content analysis", *Psychomusicology* 13:1 (1994).

Collins, K. "In the loop: Creativity and constraint in 8-bit video game audio", *Twentieth-century music* 4:2 (2008).

Collins, K. *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games.* (Cambridge: MIT University Press) (2013).

Curtis, M. E., Bharucha, J. J. The Minor Third Communicated Sadness in Speech, Mirroring Its Use in Music, *Emotion* 10:3 (2010)

Daniels, L. *Batman - The Complete History: The Life and Times of the Dark Knight.* (San Francisco: Chronicle Books) (2004).

Davies, S. "Philosophical perspectives on music's expressiveness", in *Music and Emotion: Theory and Research*, eds. Juslin, P. N., Sloboda, J. A. (Oxford: OUP) (2001).

Dogramain, B., Liptay, F. "Immersion in the Visual Arts and Media", in *Immersion in the Visual Arts and Media* eds. Dogramain, B., Liptay, F. (Leiden: Brill) (2015).

Edwards, B. Unraveling The Enigma Of Nintendo's Virtual Boy, 20 Years Later. *Fast Company.* <https://www.fastcompany.com/3050016/unraveling-the-enigma-of-nintendos-virtual-boy-20-years-later> accessed 13th May 2017.

Field, A. P., Moore, A. C. "Dissociating the effects of attention and contingency on awareness on evaluating conditioning effects in the visual paradigm", *Cognition & Emotion*. 19:1 (2005).

Fruchter R., Lia, A. *Developing a collaborative workflow from BIM to virtual reality*. (Stanford, CA: Stanford University) (2015).

Good, O. S. PlayStation VR sales figures revealed in exec's interview. *Polygon.* <https://www.polygon.com/2017/2/26/14742880/playstation-vr-sales> accessed 17th May 2017.

Groner, R., Weibel, D., Wissmath, B. "Dubbing or Subtitling? Effects on Spatial Presence, Transportation, Flow, and Enjoyment", *Journal of Media Psychology* 21:3 (2009).

Halfyard, J. K. "Cue the Big Theme? The Sound of the Superhero", in *The Oxford Handbook of New Audiovisual Aesthetics*, eds. Richardson, J., Gorbman, C., Vernallis, C. (2013).

Juslin, P. N., Västfjäll, D. "Emotional responses to music: the need to consider underlying mechanisms", *Behavioural Brain Science* 31:1 (2008).

Lastra, J. Sound Technology and the American Cinema (New York: Columbia University Press) (2000).

Lipscomb, S. D., Kendall, R. A. "Perceptual judgment of the relationship between musical and visual components in film", *Psychomusicology* 13:1 (1994).

Lipscomb, S. D., Zehnder, S. M. "Immersion in the Virtual Environment: The Effect of a Musical Score on the Video Gaming Experience", *Journal of Physiological Anthropology and Applied Human Science* 23:1 (2004).

Lundqvist, L. "Facial expressions are contagious" *Journal of Psychophysiol.* 9:1 (1995).

Maman, F. *The Role of Music in the Twenty-first Century.* (California: Tama-Dõ) (1997).

Nacke, L. E., Grimshaw, M. N., Lindly, C. A. "More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game", *Interacting with Computers* 22:5 (2010).

Oderman, S. "The sound of silents", *Films in Review* 47:3 (1996).

Osborne, J. W. "A phenomenological investigation of the musical representation of extramusical ideas", *Journal of Phenomenol Psychology* 20:1 (1989).

Reay, P. *Music in Film: Soundtracks and Synergy* (New York: Wallflower Press) (2014).

Ryan, M. *Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity in Literature and Electronic Media* (Baltimore: Johns Hopkins University Press) (2015).

Schäfer, T., Huron, D., Shanahan, D., Sedlmeier, P. "The sounds of safety: stress and danger in music perception", *Frontiers in Psychology* 6:9 (2015).

Strangman, R. *Memoirs of a Virtual Caveman.* (North Carolina: Lulu Press) (2014).

Subers, R. Around-the-World Roundup: 'Dark Knight Rises' Joins Billionaire Club. *Box Office Mojo.* < http://www.boxofficemojo.com/news/?id=3514&p=.htm> accessed 18th May 2017.

Summers, T. "Wagner and Video Game Music" in: *Music in Video Games: Studying Play* eds. Donnelly, K. J., Gibbons, W., Lerner, W. (UK: Routledge) (2014).

Sweester, P., Wyeth, P. "GameFlow: A Model for Evaluating Player Enjoyment in Games", *ACM Computers in Entertainment* 3:3 (2005).

Tan, S. L., Kelley, M. E. "Graphic representations of short musical compositions", *Psychology of Music* 32:2 (2004).

Whelan, Z. N. *Play Along: Video Game Music as Metaphor and Metonymy* (Florida: University of Florida) (2004).

Wingstedt, J., Brändström, S., Berg, J. "Narrative Music, Visual and Meaning in Film", *Visual Communication.* 9:2 (2010).

Yeh, CH., Tseng, WY., Chen, CY. et al. "Popular music representation: chorus detection & emotions recognition", *Multimedia Tools and Applications* 73:3 (2014)

Films

Batman (1989, Tim Burton)

Batman and Robin (1949, Spencer Gordon Bennet)

Batman Begins (2005, Christopher Nolan)

Star Trek IV: The Voyage Home (1986, Leonard Nimoy)

The Batman (1943, Lambert Hillyer)

The Dark Knight (2008, Christopher Nolan)

The Dark Knight Rises (2012, Christopher Nolan)

The Pawnshop (1916, Charlie Chaplin)

Video Games

Batman (1986, Ocean Software)

Batman (1990, SunSoft)

Batman: Arkham Asylum (2009, Rocksteady)

Batman: Arkham VR (2016, Rocksteady)

Batman Begins (2005, Eurocom)

Final Fantasy I (1987, Square Co.)

Final Fantasy XV (2016, Square Enix)

Lord of the Rings: The Two Towers (2002, Electronic Arts)

The Legend of Zelda: Ocarina of Time (1998, Nintendo)