Investigating Sound Design in Film:
A Commentary on Kock and Louven

SIU-LAN TAN[1]
Kalamazoo College

ABSTRACT: Kock and Louven (in this issue) examined the effects of music, sound effects, full sound design (music and sound effects) and no sound on self-reported measures of immersion and suspense in real time, as viewers watched very brief original films. This commentary discusses the method, analysis, and implications of their findings within the broader context of the state of the art of film music research, and future directions for investigations in this area.

IN an innovative exploratory study, Kock and Louven (this issue) examined the effects of film sound on viewers' perceived immersion and suspense while watching original films of under 2 minutes in length. Two types of films were included in their study: an animatic film (computer-generated, with dialogue and sound track added later) and a live-action film (shot on a movie set with simultaneous dialogue and sound recording). A total of 240 media production and technology students served as participants, and indicated their responses on an iPad touchscreen, using software that was co-developed by one of the authors. Especially notable was their inclusion of several audio conditions: no audio, music only, sound effects only, or full sound design (which consisted of music and sound effects). These were the key features of Kock and Louven's study that most caught my attention and serve as focal points for the present discussion. This commentary will consider the potential contributions of this study for the empirical study of film music, within the broader context of the state of the art of film music research, and future directions for investigations in this area.

BEYOND THE MUSIC TRACK

Studies employing simple audiovisual stimuli (such as flashing lights and clicks or bell strikes) have been conducted since the early days of psychological research. Indeed, research on how people respond to simultaneously-presented auditory and visual stimuli was conducted in the first psychology laboratory established in 1879 by Wilhelm Wundt, often recognized as the father of experimental psychology. However, investigations of the role of music and other elements of the soundtrack on viewers' perception of film have accumulated relatively slowly. One notable landmark was an issue of *Psychomusicology* journal published in 1994, the first special issue devoted to empirical studies of film music, guest edited by Annabel Cohen. Motivating this collection of articles was the central question: 'What does film music contribute to the perception of film?' (Cohen, 1994, p. 5). Almost two more decades passed before the publication of the first book devoted to experimental studies on the role of music within the larger scope of media, *The Psychology of Music in Multimedia* (2013), for which I served as one of the editors (along with Cohen, Lipscomb, and Kendall). Despite the broad interest that this topic stirs up, and the growing number of tools (such as customized computer software, eye-tracking, various biometric measures, brain-imaging, MEG) that have been applied to this area of study, there is still much to discover about our responses to the auditory elements of the film experience.

One of the limitations of the existing body of research on film music is that most studies have treated music as an isolated track. Disentangling the elements was a necessary step in the early investigations in this area, as researchers faced the challenges of working with complex and dynamic stimuli (i.e., music and moving images). However, the focus on the isolated music track in the majority of studies in this area has inadvertently led to the neglect of other audio channels, namely dialogue and sound

effects – and lack of insight into how music, dialogue, and sound effects intertwine to create the psychological effects that they do. Speech works intrinsically with music in many film scenes, such as in the pivotal scene in which King George VI declares war on Germany in an international radio broadcast in *The King's Speech*, against the backdrop of an excerpt from the second movement of Beethoven's Symphony No. 7 [2]. Further, the line between music and sound effects is often hard to draw, as the orchestra often doubles sound effects such as cymbal crashes for shattering glass. A special case can be found in the film *Gravity*, in which Steven Price's score gives the impression of something akin to sound effects to enhance the impact of shuttles breaking apart and massive explosions, as there is no sound in the vacuum of space. There is an increasing need for researchers to recognize that films employ many (audio) channels, including music, speech, and sound effects (e.g., as described by film theorist Christian Metz [1974], echoed in the widely-cited Congruence-Associationist Model [Cohen, 2013]). Few film music studies have examined the separate effects of music and sound effects on viewers' experience of motion picture, and in that regard, Kock and Louven are headed in the right direction.

Among the most useful contributions of Kock and Louven's exploratory study (to be confirmed by further research) is the finding that sometimes the music track alone does not have the most powerful impact on such fundamental audience responses as perceived suspense or immersion. It was particularly striking to see how the mix of sound effects and music seem to have a synergistic effect that makes them more effective than either one alone. This makes sense, as composers often have to carefully coordinate the music with other elements of the soundtrack – such as creating 'beds' for important moments of dialogue to be clearly heard, steering clear of certain frequencies that may mask voices or Foley (such as approaching footsteps), and blending effectively with sound effects. During post-production, sound editors and music editors also meticulously balance the interplay between the different elements of dialogue, music, and sound effects, as elucidated in some of the technical books on film audio referenced in this article. A focus on the isolated music track was an important starting point for the early research in this area. However, if film music research is to 'come of age', we must broaden our scope of study to all aspects of audio, and seek to understand how the different sound elements work together to bring about the psychological effects they have on film audiences. Along these lines, it would have been helpful to read more detailed descriptions of what Kock and Louven found to be the "efficient audio mix" or "well-balanced and congruent audio mix of music and sound effects", which they only referred to in passing.

ELUSIVE ELEMENTS OF THE VIEWERS' EXPERIENCE

The method of presenting only one version of the video to each group of participants, as used in Kock and Louven's study, is standard procedure in the research in this field. This 'between-subjects' design [3] is the standard method of choice for film music researchers, not only because it allows researchers to capture fresh responses during the first viewing of a film clip, but also because it helps conceal the fact that the soundtrack has been manipulated so that the viewer does not pay undue attention to the music track. [4]

What is more innovative is the use of the *emoTouch* application on an iPad to record participants' responses in (near) real time, and its use of the touchscreen to indicate responses directly on the screen using the x and y dimensions for immersion and suspense, respectively. It would have been good to specify where the response marker (cursor) was placed at the start of each trial, as the starting position could influence the participants' responses or the measures obtained. This reviewer would also have been interested in a rationale for why immersion was always on the x-axis and suspense always on the y-axis; as we may have a tendency to treat vertical and horizontal space differently, would it have been better to counterbalance these, or was there a clear rationale for the fixed x/y designation? For instance, in Russell's (1980) circumplex model on which this scheme is based, the placement of valence (negative to positive) and arousal (low to high) on the x- and y-axes respectively makes intuitive sense, and they are conjoined as opposed to independent judgments. Specific strengths of the procedure should also be noted: The authors devised an inventive means of collecting responses in a way that does not require the viewer to take their eyes off the screen; they employed a touchscreen in a way that is very familiar to participants in their mid-twenties; and they provided a practice session to familiarize themselves with the set-up.

"Immersion" and "suspense" are both rather complex subjective aspects of experience, that have been defined in different ways in various studies (see Lehne & Koelsch, 2015; Visch, E. Tan, & Molenaar, 2010) [5]. In this vein, it would have been good for the researchers of the present study to include some relevant literature, and clearly operationalized definitions, as well as to include any definitions of "immersion" and "suspension" given to the participants (and a translation if presented in German). There

are few empirical studies on cinematic immersion, as most immersion studies have focused on interactive kinds of media (such as immersion in the "world" of a video game or virtual reality environment). Immersion has often been tied to "transportation" (see Green & Brock, 2000), which seems to come from being so immersed in the world of a story that one loses the sense of awareness of the present real-world surroundings. As such, some researchers and scholars question whether elements of experience such as "immersion" can be accessed or monitored in any conscious or deliberate manner while in the flow of experiencing it. If immersion involves a loss of awareness of the real physical environment, would asking subjects to gauge their own level of immersion interfere with one's experience of it? For this reason, some previous researchers have addressed immersion and related aspects of the viewer's experience in indirect ways.

For instance, Cohen and Siau (2008) found that when a film sequence was played with music that was congruent with the action in the scene, viewers' response time was slower when pressing a button to indicate when they saw "X" marks that were embedded in the far corners of the screen, compared to viewers who watched the same action sequence without music. They concluded that a congruent music track led to greater absorption in the film, leading to slower reaction time to notice something irrelevant appearing in the periphery of one's vision. Bezdek and colleagues also monitored viewers' visual attention to the periphery of highly suspenseful film clips (such as the crop duster chase sequence in Hitchcock's *North by Northwest*), by using an fMRI machine to record brain activity in regions of the brain associated with visual attention to peripheral areas of the screen, where they had added a flashing checkerboard border (Bezdek et al., 2015). They found that peaks in suspense elicited greater activation in central visual processing regions and stimulus-driven attention networks, while suppressing activation in peripheral visual processing regions [6]. In our own study of sound and video game play, my colleagues and I used time estimation (stopping a player once to ask how much time had passed) as an indirect measure of immersion or absorption (though we did not find any significant differences for playing with music and sound effects, sound effects only, an unrelated music track, or absence of sound; see Tan, Baxa, & Spackman, 2010). Perhaps Kock and Louven could have added an indirect measure to assess immersion, as it is so elusive.

FUTURE DIRECTIONS IN FILM AUDIO RESEARCH

Another distinctive component of Kock and Louven's study was the use of two original short films created by graduate media students. Although a few previous studies have used original material (e.g., Bullerjahn & Güldenring, 1994; Shevy, 2007; Thompson, Russo, & Sinclair, 1994), by and large, most researchers have opted for pre-existing film or television clips. Perhaps the most ambitious of early studies employing original materials was undertaken by Bullerjahn and Güldenring (1994), who used an original film and commissioned 5 professional film and television composers to score it in different styles. The study showed that different scores led to significantly different interpretations of the same film's narrative, including perceptions of character's intentions, back story, and perceived relationships between characters.

With original films, researchers run the risk of participants being distracted by any amateurish elements in the quality of the film or acting, for this generation of participants who are accustomed to high production values (see Cohen, MacMillan, & Drew, 2006, on a study focusing on music and sound effects in a scene from *Witness* [7]). This is of course particularly important in a study focusing on suspense and immersion, as both may be disturbed if viewers are distracted by low production values or unconvincing actors. On the other hand, the clear advantage of using original films for such a study is that the researchers have more control over the content of the materials and are assured that participants have not been previously exposed to the films, thus working with more pristine stimuli. Additionally, it was great to see two different types of film (animatic, live-action) being employed, and to read how they led to some nuanced differences in viewers' responses. If permission can be obtained from the graduate student film-makers, it would be helpful if links to the short films (and their four audio conditions) could be included with the article, to engage readers or for use for classroom demonstrations, and most importantly to give researchers and scholars greater context to understand the findings of this study, and for replication purposes.

Reflecting on their findings, Kock and Louven conjecture that a more fitting sound design of the *Catacombes* film, being a live-action film set in underground tunnels, may be "non-diegetic sound effects like sound layers of reverb and drones" (rather than thematic music); or in the *Goldenberg* short, it could be diegetic fighting and action sounds that are not shown on screen. It is indeed possible that these appropriate

matches may enhance the impact of the sound effects and lead to greater viewer immersion or suspense. For instance, Boltz (2017) recently showed that the tempo of diegetic sounds (such as the pace of the sounds of footsteps, putts of a motorcycle, or rotations of a helicopter propeller) influences viewers' perception of the speed of the visual information. One might conjecture that perceiving more rapid movement (either of the viewer's impression of their own movement, or of elements within the scene) may heighten degree of suspense and immersion in a chase sequence, for example.

Given the present authors' ability to access able film makers, it might be interesting to design a future study in which another layer is added to the audio conditions: the presence of diegetic music. A scene could be set up in which diegetic or non-diegetic music might be equally fitting - for instance, music either treated to sound like it is playing over a jukebox in a bar within a scene (diegetic), or as dramatic scoring (non-diegetic). Very few studies have manipulated diegetic and non-diegetic music to study the effects on various aspects of the viewers' experience of a film. In our own study, for instance, my colleagues and I found that diegetic music ratcheted up the perceived tension and conflict between the characters in an action sequence, compared to a louder version of the same music that was mixed to suggest a non-diegetic dramatic score (Tan, Spackman, & Wakefield, 2017). This may be because presenting a serene ballad as if emanating from loudspeakers inside a shopping mall may heighten the viewers' impression of tension of a scene and characters as it contrasts with the mood of the scene in a seemingly natural or coincidental way. On the other hand, we found that the same gentle music mixed at a louder level to suggest non-diegetic music may be soothing as it might be read as a commentary on the scene.

Another interesting follow-up study would be one that is similar to Kock and Louven's present procedure but compares the effects of a film scene with sound effects alone vs. *diegetic* music and sound effects. This may allow for a cleaner comparison of the effects of absence or presence of music, as both of these kinds of sound are presented as if existing within the world of the characters.

Interest in investigating viewers' real-time responses to film and film music is growing, in both industry and academic spheres. To cite a recent example: In 2016, the bioanalytics technology company Lightwave teamed up with 20th Century Fox to take biometric measures (such as heart rate, skin temperature, and body movements) while test audiences were watching *The Revenant* [8]. Further, academic researchers are conducting studies in controlled environments that simulate aspects of the theater experience more closely than in earlier studies. For instance, researchers at the Dublin City University established a CDV*Plex* cinema lab, equipped with a large-screen digital projector and 5.1 surround sound speaker system and Smart Chairs that record body movements and changes in posture. The Dublin researchers also monitor viewers' responses as they watch entire movies, as opposed to isolated excerpts, and then segment the film into smaller parts for their analysis. For instance, Rothwell et al.'s (2006) study employed 37 full-length films over 10 weeks, yielding 500+ hours of biometric data. Their studies focus on the role of music in particular, and their ongoing research strongly suggests that emotional responses to particular parts of the film are influences by the score, as manifested by the pattern of responses shown in the biometric measures (Rothwell et al., 2006; Smeaton & Rothwell, 2009).

As the empirical research on film music comes into a new age, it must take on an increasingly interdisciplinary flavor, ideally involving collaborative ventures that draw on the specialized expertise of many people representing various academic subjects and factions of industry, to truly gain insight into how we engage with motion pictures and their rich sound design. The focus of investigations of the soundtrack must also encompass more than the isolated music track. It is important to remember that music does not act alone or play a solo role in film audio, but is part of an 'ensemble cast' in the interplay of dialogue, music, and sound effects.

ACKNOWLEDGMENTS

NOTES

[1] Correspondence can be addressed to: Dr. Siu-Lan Tan, Department of Psychology, Kalamazoo College, 1200 Academy Street, Kalamazoo MI 49006, USA. Siu-Lan.Tan@kzoo.edu

[2] See Bashwiner (2013) in the references for a particularly insightful analysis of this scene from *The King's Speech*.

[3] Elsewhere, I have provided a brief overview of methods employed in this area of research for readers in the humanities (see Tan, 2017, in references).

[4] As opposed to a 'within-subjects design' (in which participants are exposed to multiple conditions, so that each would see all versions of the altered film clip).

[5] Ed S. Tan is no relation to the author of this commentary.

[6] It should be noted that in a variation of the study with a music manipulation, it was found that during points of high suspense, a congruent music track produced a similar pattern of activation in stimulus-driven attention networks while producing a broader increase in activity in central and peripheral visual processing regions (see Bezdek, Wenzel, & Schumacher, 2017).

[7] See Cohen, MacMillan, and Drew (2006) for discussion on the possible limitations of this study, and the challenges of isolating and manipulating sound effects.

[8] This information was obtained from Lightwave's Press Release posted on January 12 2016, which can be found at http://www.marketwired.com/press-release/lightwave-partners-with-20th-century-fox-measure-audiences-physical-emotional-responses-2087379.htm

REFERENCES

Bashwiner, D. (2013). Musical analysis for multimedia A perspective from music theory. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.). *The psychology of music in multimedia* (pp. 89-117). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608157.003.0005

Bezdek, M. A., Gerrig, R. J., Wenzel, W. G., Shin, J., Revill, K. P., & Schumacher, E. H. (2015). Neural evidence that suspense narrows attentional focus. *Neuroscience*, *303*, 338-345. https://doi.org/10.1016/j.neuroscience.2015.06.055

Bezdek, M. A., Wenzel, W. G., & Schumacher, E. H. (2017). The effect of visual and musical suspense on brain activation and memory during naturalistic viewing. *Biological Psychology*, *129*, 73-81. https://doi.org/10.1016/j.biopsycho.2017.07.020

Boltz, M. G. (2017). Auditory driving in cinematic art. *Music Perception, 35*, 77-93. https://doi.org/10.1525/mp.2017.35.1.77

Bullerjahn, C., & Güldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, *13*, 99-118. https://doi.org/10.1037/h0094100

Cohen, A. J. (1994). Introduction to the special volume on psychology of film music. *Psychomusicology, 13*, 2-8. https://doi.org/10.1037/h0094104

Cohen, A. J. (2013). Congruence-Association Model of music and multimedia: Origin and evolution. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.). *The psychology of music in multimedia (pp. 17-47).* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608157.003.0002

Cohen, A. J., & Siau, Y.-M. (2008). The narrative role of music in multimedia presentations: The Congruence-Association Model (CAM) of music and multimedia. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10) Sapporo, Japan* (pp. 77-82). Adelaide, Australia: Causal Productions.

Cohen, A. J., MacMillan, K. A., & Drew, R. (2006). The role of music, sound effects and speech on absorption in a film: The congruence-associationist model of media cognition. *Canadian Acoustics, 34*, 40-41.

Green, M. C., & Brock, T. C. (2000) The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*, 701-721. https://doi.org/10.1037/0022-3514.79.5.701

Kock, M. & Louven, L. (2018). The power of sound design in a moving picture: An empirical study with *emoTouch* for iPad. *Empirical Musicology Review, 13(3-4),* 132-148. https://doi.org/10.18061/emr.v13i3-4.6572

Lehne, M., & Koelsch, S. (2015). Toward a general psychological model of tension and suspense. Frontiers in Psychology, 6, 79. https://doi.org/10.3389/fpsyg.2015.00079

Metz, C. (1974). F*ilm language: A semiotics of the cinema.* New York, NY: Oxford University Press.

Rothwell, S., Lehane, B., Chan, C. H., Smeaton, A. F., O'Connor, N. E., Jones, G. J. F., & Diamond, D. (2006). The CDV*Plex* Biometric Cinema: Sensing physiological responses to emotional stimuli in film. *Adjunct Proceedings of Pervasive Computing, 207*, 103-106.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178. https://doi.org/10.1037/h0077714

Shevy, M. (2007). The mood of rock music affects evaluation of video elements differing in valence and dominance. *Psychomusicology, 19* (2), 57- 78. https://doi.org/10.1037/h0094034

Smeaton, A. F., & Rothwell, S. (2009). Biometric responses to music-rich segments in films: The CDV*Plex. IEEE Computer Society*, pp. 162-168, 2009. https://doi.org/10.1109/CBMI.2009.21

Tan, S.-L. (2017). From intuition to evidence: The experimental psychology of film music. In M. Mera, R. A. Sadoff, & B. Winters (Eds.), *The Routledge Companion to Screen Music and Sound* (pp. 517-530). New York: Routledge. https://doi.org/10.4324/9781315681047-41

Tan, S.-L., Baxa, J. P., & Spackman, M. P. (2010). Effects of built-in audio versus unrelated background music on performance in an adventure role-playing game. *International Journal of Gaming and Computer-Mediated Simulations, 2*, 1-23. https://doi.org/10.4018/jgcms.2010070101

Tan, S.–L., Cohen, A. J., Lipscomb, S. D., & Kendall, R. A. (2013). *The psychology of music in multimedia.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608157.001.0001

Tan, S.-L., Spackman, M. P., & Wakefield, E. M. (2017). Effects of diegetic and non-diegetic music on viewers' interpretations of a film scene. *Music Perception, 34*, 605-623. https://doi.org/10.1525/mp.2017.34.5.605

Thompson, W. F., Russo, F. A., & Sinclair, D. (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology*, *13*, 99-118. https://doi.org/10.1037/h0094103

Visch, V. T., Tan, E. S., & Molenaar, D. (2010). The emotional and cognitive effect of immersion in film viewing. *Cognition and Emotion, 24*, 1439-1445. https://doi.org/10.1080/02699930903498186