



eCOMMONS

Loyola University Chicago
Loyola eCommons

Psychology: Faculty Publications and Other Works

Faculty Publications

6-2017

The Effects of Diegetic and Nondiegetic Music on Viewers' Interpretations of a Film Scene

Elizabeth M. Wakefield
Loyola University Chicago, ewakefield1@luc.edu

Siu-Lan Tan
Kalamazoo College

Matthew P. Spackman
Brigham Young University

Follow this and additional works at: https://ecommons.luc.edu/psychology_facpubs



Part of the [Musicology Commons](#), and the [Psychology Commons](#)

Recommended Citation

Wakefield, Elizabeth M.; Tan, Siu-Lan; and Spackman, Matthew P. The Effects of Diegetic and Nondiegetic Music on Viewers' Interpretations of a Film Scene. *Music Perception: An Interdisciplinary Journal*, 34, 5: 605-623, 2017. Retrieved from Loyola eCommons, Psychology: Faculty Publications and Other Works, <http://dx.doi.org/10.1525/mp.2017.34.5.605>

This Article is brought to you for free and open access by the Faculty Publications at Loyola eCommons. It has been accepted for inclusion in Psychology: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
© The Regents of the University of California 2017

THE EFFECTS OF DIEGETIC AND NONDIEGETIC MUSIC ON VIEWERS' INTERPRETATIONS OF A FILM SCENE

SIU-LAN TAN
Kalamazoo College

MATTHEW P. SPACKMAN
Brigham Young University

ELIZABETH M. WAKEFIELD
Kalamazoo College

PREVIOUS STUDIES HAVE SHOWN THAT PAIRING a film excerpt with different musical soundtracks can change the audience's interpretation of the scene. This study examined the effects of mixing the same piece of music at different levels of loudness in a film soundtrack to suggest *diegetic music* ("source music," presented as if arising from within the fictional world of the film characters) or to suggest *nondiegetic music* (a "dramatic score" accompanying the scene but not originating from within the fictional world). Adjusting the level of loudness significantly altered viewers' perceptions of many elements that are fundamental to the storyline, including inferences about the relationship, intentions, and emotions of the film characters, their romantic interest toward each other, and the overall perceived tension of the scene. Surprisingly, varying the loudness (and resulting timbre) of the same piece of music produced greater differences in viewers' interpretations of the film scene and characters than switching to a different music track. This finding is of theoretical and practical interest as changes in loudness and timbre are among the primary post-production modifications sound editors make to differentiate "source music" from "dramatic score" in motion pictures, and the effects on viewers have rarely been empirically investigated.

Received: October 23, 2010, accepted January 26, 2017.

Key words: film music, diegetic, interpretation, loudness, emotion

IN ACADEMIC DISCUSSIONS OF FILM, THE TERM *diegesis* has come to refer to "all that belongs, by inference, to the narrated story, to the world

supposed or proposed by the film's fiction" (Souriau, as cited by Gorbman, 1987, p. 21). Film music is often described with respect to its relation to this fictional universe. *Diegetic music* is "produced within the implied world of the film" (Kassabian, 2001, p. 42). The sound of a mother's voice as she is shown singing to her infant is clearly diegetic, and the faint sounds of a music box might be assumed to be coming from inside the nursery, even if the source is not shown on screen.

In contrast, *nondiegetic music* refers to music that accompanies a scene but is external to the fictional world, such as dissonant orchestral chords sounding as a figure appears in a deserted alley or throbbing music punctuating a car chase. Nondiegetic music is not part of the sensory world that film characters are assumed to be able to see, touch, smell, feel, or hear (Brown, 1994). The terms *source music* (diegetic) and *dramatic score* (nondiegetic) are more commonly used in industry parlance, and are similar in a practical sense.

Early sound films often solely employed diegetic music, limiting the use of nondiegetic music to the opening and closing credits, out of concern that the inclusion of nondiegetic music in the rest of the film would puzzle the audience as its presence could not be explained by the images (Cooke, 2001). But by the early 1930s, Max Steiner and other Hollywood composers began to depart from this convention by producing dramatic scores (Larsen, 2005; Schrader et al., 2017), which quickly gained popularity. Musical soundtracks consisting of mainly nondiegetic music soon became the mainstay of narrative films.

In practice, film music often explores ambiguities between the diegetic and nondiegetic. In a scene in Mel Brooks' (1977) *High Anxiety*, two men discuss the possibility that someone they know may have been murdered. One of the men seems surprised to hear dramatic orchestral music punctuating their conversation like a nondiegetic dramatic score—until it is revealed that the music is coming from a passing bus, where an entire symphonic orchestra is busy rehearsing! In another memorable example, in Robbe-Grillet's (1968) *The Man Who Lies*, a fugitive cowers at the sound of a snare drum in the nondiegetic musical score—as if it reminds him of gunfire—though the nondiegetic drumming should

only be heard by the film audience. Audiences seem to make inferences about the implied sources of music and react with surprise or amusement when expectations are thwarted.

Some film theorists have challenged the binary assumption of the diegetic/nondiegetic distinction. Kasabian (2001) argues that the distinction between diegetic and nondiegetic music “obscures music’s role in producing the diegesis itself” (p. 42), and assumes a film apart from music “that constructs its narratively implied world silently” (p. 42). Others advocate keeping the distinction but acknowledging the many instances that may occur in the “fantastical gap” or zone between diegetic and nondiegetic (Stilwell, 2007; see also Neumeier, 2009, and Winters, 2010). Despite its complexities, the “conceptual distinction regarding music’s relation to narrative space is one of the cornerstones of film music theory” (Smith, 2009, p. 2). Notwithstanding the rich theoretical discourse on the diegetic/nondiegetic distinction and its importance to film music theory, the effects of diegetic versus nondiegetic presentation of music on the viewer’s experience of a film scene is understudied. We address this question with our present work.

Effects of (Nondiegetic) Musical Soundtracks on Interpretations of Film Scenes

Previous studies have shown that pairing a film excerpt with different musical soundtracks can alter viewers’ perceptions and interpretations of a film scene (e.g., Boltz, 2001; Bullerjahn & Güldenring, 1994; Marshall & Cohen, 1988; Shevy, 2007; Tan, 2017a; Vitouch, 2001) and how they relate to film characters (Hoeckner, Wyatt, Decety, & Nusbaum, 2011). For instance, Boltz (2001) showed three ambiguous film scenes and asked participants to judge the relationship between the two main characters. Participants tended to give more positive interpretations when the images were accompanied by “positive” music (i.e., pieces in major mode, consistent tonality scheme, predictable rhythm), more negative interpretations when accompanied by “negative” music (i.e., minor key, atonal, irregular rhythm), and film excerpts viewed in the no-music condition were interpreted as relatively neutral. For instance, negative music was associated with significantly more inferences about one character intending to harm another character and more negative adjectives describing the main character’s personality.

The music accompanying a scene may also shape viewers’ expectations about what events may unfold next. Vitouch (2001) played the opening scene to Billy

Wilder’s 1945 film *The Lost Weekend* with either Miklós Rózsa’s (original) score for the film, or with Barber’s *Adagio for Strings* Opus 11. The excerpt showed a man through the open window of an apartment building, pausing while packing a suitcase. When asked to describe what would happen next, participants wrote more positive or ambivalent plot continuations when the scene was accompanied by Rózsa’s romantic film score, and more “negatively tinted continuations” when paired with Barber’s “sad” or “melancholic” pieces (p. 77). A content analysis revealed that music even colored some participants’ impressions of the weather and setting of the scene, as reflected by phrases such as “a nice warm autumn day” among those who viewed the scene with Rózsa’s score (p. 76), versus “all gray and desperate” among those watching with Barber’s music (p. 79).

In another study, Tan, Spackman, and Bezdek (2007) paired film excerpts that featured a solitary film character with music that conveyed “happiness,” “sadness,” “fear,” or “anger.” The music either ended before a close-up of a character’s face (with neutral or subdued emotional expression) was shown, or started after the character had left the scene. Participants interpreted the film character’s emotions in ways that were generally congruent with the emotion of the music, even though the music was mainly played before or after the character was shown. Vitouch’s (2001) and Tan et al.’s (2007) studies show that music may not only influence viewers’ interpretation of concurrent images, but can play a role in how viewers form expectations about subsequent scenes and reframe viewers’ understanding of scenes they have previously viewed (see also Boltz, Schulkind, & Kantra, 1991).

Music not only influences immediate responses to a film scene but also memory for the visual content of the film (e.g., Boltz, 2001, 2004; Boltz et al., 1991). In the aforementioned Boltz (2001) study, for instance, a second group of participants were asked to return one week after viewing film scenes paired with “positive” or “negative” music (as described earlier). When presented with a list and asked to identify which items had been present in the film clips, participants’ memory for onscreen items followed a general pattern of mood congruency. Specifically, they correctly remembered more positive items (such as flowers and candles) when scenes had been accompanied by “positive” music, and more negative items (such as skulls and tombs) when the scenes were accompanied by “negative” music. Further, positive items were more likely to be falsely recognized when the scene had been shown with “positive” music, and more negative items were incorrectly recognized when the scene had been shown with “negative” music.

Of course, the effects of the music track are not always simple or predictable, particularly when scenarios depicted on the screen are more determinate (e.g., Cohen, 1993) or when audiovisual combinations are more complex or incongruent (e.g., Pavlović & Marković, 2011). For instance, Cohen (1993) found that the music track only influenced viewers' perceptions of a man chasing after a woman in an interaction that could be interpreted as either playful flirtation or a more aggressive pursuit, but had no effect on a less ambiguous scenario of a fight between two men. In another study, music altered the degree of perceived intensity of an emotion, but not the valence of the emotion (Boli-var, Cohen, & Fentress, 1994, Experiment 3).

Over the past three decades, a body of research has steadily accumulated on the influence of the music track on meaning, memory, and engagement in the viewer's experience (Cohen, 2013, 2014; Tan, 2017a, 2017b). However, since its emergence in the early 1980s, research in this area has focused almost solely on *nondiegetic* music or music tracks that served as background musical accompaniment to film scenes. Most studies have compared the effects of pairing the same film scene with music tracks of contrasting moods or emotions (e.g., Boltz, 2001; Tan et al., 2007), contrasting styles or genres of music (e.g., Bullerjahn & Güldenring, 1994; Hoeckner et al., 2011), or congruence or incongruence with the emotional tone and action of a film scene (e.g., Bezdek, Wenzel, & Schumacher, 2014; Boltz, 2004; Cohen & Siau, 2008). Typically, film excerpts that do not inherently include any music in the original scene are used, making it possible to add different music tracks to accompany the film clip. Thus, film music studies have very rarely employed film scenes involving diegetic music (Tan, Cohen, Lipscomb, & Kendall, 2013, p. 392).

As far as we are aware, only one published empirical study has addressed questions relevant to diegetic versus nondiegetic presentations of music in film (Fujiyama, Ema, and Iwamiya, 2013). Fujiyama et al. played an excerpt from Akira Kurosawa's 1949 film *Stray Dog* in which the mood of the music and image were incongruent with each other. The scene showed two men fighting in a garden while soft, classical piano music is heard in the distance. In the original clip, a woman practicing the piano is shown through the open window of a nearby house. In the altered version, the part of the film revealing the pianist was deleted so that the source of the music was not shown.

Participants gave poorer overall ratings for the audiovisual combination (on a scale from good to bad) when the source of the mood-incongruent music was not

shown on screen (making it more ambiguous whether the music is diegetic or nondiegetic) than when the source of the mood-incongruent music was shown (to clearly convey that the music is diegetic). However, the focus of this study was on the impact of incongruence between the music and moving image on judgments of the "goodness" of the music-film pairing in the film scene, as opposed to any effects on the audience's impressions of the characters and scene (as is the focus of the present study).

Diegetic Music: Research Challenges

One reason that the effects of diegetic versus nondiegetic presentations of music have rarely been examined in empirical studies may stem from the particular challenges involved in working with diegetic music in scenes from major motion pictures (with the production values that participants are accustomed to seeing). For instance, there is the logistical challenge of modifying a film excerpt to produce both diegetic and nondiegetic music versions of the same film scene, when researchers rarely have access to the separate layers of audio (i.e., different components of the music track, dialogue, and sound effects) in order to remix them. Further, both presentations of the music must be plausible accompaniments to the scene, so that participants do not become aware of the alterations to the film clip. A piece of music may no longer "work" with a scene if converted from diegetic to nondiegetic, or the reverse.

In the aforementioned study, Fujiyama et al. (2013) cleverly addressed these issues by leaving the audio unchanged, and editing the video to conceal the source of the music in one version of the scene. However, the faint and distant quality of the music in the selected scene from Kurosawa's *Stray Dog* strongly suggests that the piano tones supposedly emanate from within the environment of the characters. In the present study, we used another method to create a version to suggest a nondiegetic musical counterpart to a scene that originally included diegetic music (as described further in the Materials section).

Another challenge stems from the illusory nature of the distinction between diegetic and nondiegetic music. Diegetic music is very rarely recorded during the filming of the scene, as the presence of live music would interfere with voice pickup and make the footage impossible to edit (Rose, 2008, p. 77). Like nondiegetic music, diegetic music is usually added to the film during post-production. Even music that appears to be coming from performers within a scene, such as a band, is typically recorded separately and played back during

filming for musicians to mime to it. Diegetic music is then set to the appropriate loudness to match the implied source within the scene, and usually modified with sound filters, equalizers, and reverberation during post-production to give the impression that it is coming from within the environment of the film characters (Cross, 2013; Rose, 2013; Sonnenschein, 2001, Chapters 1, 5, 7). In sound designer Jay Rose's words, "source [i.e., diegetic] music always needs equalization and reverb—often with drastic settings—to make it match the ambience of a scene" (2008, p. 327). Thus the distinction between diegetic and nondiegetic music is usually based on a *suggested or implied*, rather than actual, source.

As diegetic and nondiegetic music are usually both achieved through sound and mixing techniques during post-production, this can pose a methodological challenge for researchers. There is often no clear way to separate the idea of diegetic and nondiegetic music (or "source music" versus "dramatic score") from the physical manipulations to achieve them, as these are indistinguishable from a film production standpoint. Thus, as described in our Materials section, we manipulated the level of loudness of the music to *suggest* nondiegetic versus diegetic music within a particular scene in a film, in a manner similar to the way in which it is done during post-production of motion pictures. This point is further addressed in our discussion. The current study therefore represents an early step toward exploring the psychological impact of diegetic and nondiegetic presentation of music with a film scene, an important but slippery facet of film music.

Aims of the Present Study

The present study examines the effects of mixing the level (i.e., loudness) of music in a film soundtrack, to suggest nondiegetic versus diegetic presentation of the same piece of music. We compared participants' interpretations of three versions of a film excerpt from Steven Spielberg's film *Minority Report*. The original version was an action scene accompanied by diegetic music (i.e., with music that had been processed to sound faint and distant, as if it were playing over the public sound system inside a shopping mall in which the scene took place). We created a second version by mixing the same instrumental piece used in the original film at a louder level to suggest a nondiegetic dramatic score. As the original music track (a gentle ballad) was incongruent with the suspenseful mood of the scene, we also created a third version by pairing the scene with another piece of music that was tense and dissonant (i.e., "chase music") to create a mood-congruent nondiegetic version.

Our primary goal was to examine the effects of the same piece of music presented *diegetically* or *nondiegetically* (as signaled by mainly varying the loudness of the music) on the audience's interpretation of the scene. Except in rare cases, in which the source of the music is intended to be ambiguous with respect to the diegesis (see Kassabian, 2001, p. 47), some differences in intensity and/or timbral qualities are involved when mixing diegetic versus nondiegetic music in post-production, as described earlier. Therefore, the question of whether altering the loudness of the same musical track influences the audience's interpretations of a scene is informative for both practitioners and theoreticians.

A secondary aim was to examine the effects of nondiegetic music that matches the mood and onscreen events of the scene (*mood-congruent*), and nondiegetic music that is contrasted with the mood and onscreen action (*mood-incongruent*). As reviewed earlier, previous studies showed that viewers tend to interpret and remember scenes in ways that are congruent with the emotion conveyed by the music. Thus, we expected the nondiegetic music to influence participants' interpretations of onscreen events in ways that are generally congruent with the character of the music paired with the film excerpt, especially with respect to mood. As the only other published study of which we are aware employing diegetic music (Fujiyama et al., 2013) did not focus on effects on interpretation of the scene, we did not formulate specific hypotheses as to the possible effects of the diegetic versus nondiegetic presentation of the music.

Method

MATERIALS

Film excerpt. The film excerpt used in the study was a sequence taken from *Minority Report* (Goldman, Shuett, & Spielberg, 2002), as shown in Figure 1. The scene opens as the two main characters, John and Agatha, enter a shopping mall (frame 1). As the pair makes their way hurriedly through the mall (frame 2), armed police can be seen pursuing them, racing up an escalator to the floor above them (frame 3). Agatha instructs John to stop and stand still (frame 4). A man selling balloons crosses in front of them (frame 5) and as the police reach the platform and peer down over the balcony, John and Agatha are completely hidden from their view behind the large bunch of balloons (frame 6).

The sequence appears at 1:35:33 to 1:36:57 of the 2002 DVD version of the film and is 84 seconds long. The two characters are played by Tom Cruise and Samantha Morton, and Morton has almost all of the speaking lines



FIGURE 1. Six frames from the *Minority Report* film sequence viewed by participants. © 2002 Dream Works and Twentieth Century Fox.

(as transcribed in the Appendix). The scene is accompanied by an instrumental version of Henry Mancini's ballad "Moon River," played faintly and distantly, as if coming through loudspeakers inside the mall.

This film excerpt was selected as it met several technical and methodological criteria. First, it employs diegetic music that is played quietly enough to be masked by other music tracks, while preserving speech and sound effects at a sufficiently audible level. Second, the diegetic music in the original mall scene is available for purchase as a single recording of "Moon River" (Mancini & Mercer, 1961) enabling us to create a nondiegetic version of the soundtrack by laying the same piece of music over the original soundtrack at a louder level. Third, the situation and relationship between the characters in the scene are somewhat open to interpretation by those who have not seen the film.

Musical soundtracks. The music for the two altered versions was mixed by a professional video editor with 20 years of experience, using Final Cut Pro Version 7 video editing software. The dialogue and sound effects from the original soundtrack were preserved by pulling down the volume of the music for all dialogue, as is common practice for film scores.

The three versions of the film excerpt were as follows:

- *Diegetic "Moon River" (the original version):* The original musical soundtrack that accompanied the scene in the Spielberg film was an orchestral rendition of the song "Moon River," played softly as if transmitted over distant speakers inside an expansive indoor space. To achieve the illusion that the song is playing over the public sound system inside the shopping mall, the music was mixed at a low (i.e., soft) level relative to the dialogue and sound

effects. In addition, post-production processes such as applying equalizers and filters were likely to have been used to give a low-fidelity sound (Cross, 2013; Rose, 2008) though the music in the scene was soft, so these effects were heard at a low level. The main cue that signals that the music is supposed to be coming from inside the mall is the low level of the music. As the slow-paced and gentle ballad unfolds within a suspenseful scene, the music is *diegetic* and *mood-incongruent*.

- *Nondiegetic "Moon River" version:* To create a nondiegetic version of the same film scene, we asked an experienced professional video editor to add the same audio recording of the "Moon River" music that was used in the original film to the scene, mixing it louder than in the original scene to suggest a dramatic score accompanying the scene.¹ Although we did not directly manipulate the timbre of the music ourselves, the music also sounded richer, as the same instrumental piece in the original film had been processed to sound like it was playing distantly over loudspeakers inside the mall, as previously described. The main parameter we varied to suggest a nondiegetic dramatic score was increased loudness of the music (see Procedure for sound level meter readings when played for participants).

Although the original music was masked by the new music, the dialogue from the original soundtrack was preserved and did not differ significantly in clarity (i.e., intelligibility) from the

¹ Diegetic music can also be loud and clear, so the manipulation involved to create a nondiegetic version differs with the particulars of the scene.

original Spielberg version (as rated by 120 participants in our second pilot test). As the music was mixed to resemble a dramatic score that was external to the environment of the film characters, and is a slow and gentle romantic ballad accompanying a suspenseful scene, the music track is *nondiegetic* and *mood-incongruent*.

- *Nondiegetic Chase Music version*: To create this version, we used music from John Williams' score for Steven Spielberg's *Empire of the Sun* that originally accompanied a chase scene (at 0:39:49 in the 2001 DVD version of the film). The music excerpt was taken from Track 10, 2:03-2:30 and 2:42-3:42² of the CD recording of the original motion picture soundtrack for *Empire of the Sun* (Shapiro & Spielberg, 1987/2001), and added to the excerpt to mask the original music in the scene. Music from this film was selected as it was created by the same director-composer team that collaborated on *Minority Report*. Again, speech and sound effects from the original soundtrack were preserved. The "chase music" is relatively fast with a variable tempo, dissonant chords, and no clear melody line. Thus, this music is *nondiegetic* and *mood-congruent*.

We did not create a *diegetic*, mood-congruent soundtrack as "chase music" playing over the sound system inside the mall would not have been convincing and this would have posed a significant problem for the study. Also, it was not technically possible to create a version with speech and sound effects alone (without any music) as we did not have access to separate audio and music tracks for the film, and therefore a no-music condition was not included.

Six DVDs were created to address all possible orderings of the three versions of the film clips. All film and music excerpts were imported, edited, and rendered using Final Cut Pro Version 7 software. DVD-quality video was used to avoid reductions in picture quality. Two pilot studies³ were conducted as manipulation checks of our audiovisual materials.

PILOT 1: MUSIC TRACKS AND FILM EXCERPT

The music tracks and film excerpt were pilot tested with 39 undergraduates (aged 22-23 years) enrolled in their

² Twelve seconds of music was cut so that the end of the film scene would coincide with the end of a phrase, as in the original music accompanying the scene in the *Minority Report* film.

³ The second pilot test was conducted after the main procedure was completed to address concerns about whether adding music masked the dialogue, and is reported in the method section for clear presentation.

TABLE 1. Mean Ratings for Perception of Mood of Music for Pilot 1

Music rating criteria	Moon River-nondiegetic (incongruent)		Chase Music-nondiegetic (congruent)	
	Mean	SD	Mean	SD
Fit for romantic scenes	4.56	0.71	1.33	0.49
Grandiose	3.00	1.50	3.39	1.20
Cheerful	2.94	1.35	1.06	0.24
Fit for comedy	2.94	1.31	1.33	0.59
Fit for action scenes	1.22	0.55	4.61	0.50
Tense	1.11	0.32	4.67	4.61

5th year of an Audiovisual Communication program at a large metropolitan university. The main goals of the pilot were to ensure that (1) the mood evoked by the music and film scene were congruent in the Nondiegetic Chase music version and incongruent in the Nondiegetic "Moon River" version; and (2) that both Chase music and "Moon River" tracks were plausible as nondiegetic film scores. We selected advanced Audiovisual Communication students for this pilot, as they were generally familiar with the terms "diegetic" and "nondiegetic," had completed film appreciation and film production courses, and had experience selecting and/or composing music tracks for films or videos.

One group of students ($n = 18$) was shown a silent version of the *Minority Report* film excerpt (unaccompanied by music, speech, or sound effects) and the nondiegetic "Moon River" and Chase music alone (unaccompanied by the film). The other group ($n = 21$) watched the film with the two soundtracks we created. Students responded to a series of questions using Likert-Type scales of agreement ranging from 1 (*I disagree completely*) to 5 (*I agree completely*).

Students who listened to the music⁴ alone (without the film) gave the highest ratings for "romantic," "grandiose," "cheerful" and "fit for comedy" to describe the "Moon River" music, whereas the highest ranked descriptors for the Chase music (audio only) were "tense," "fit for action," and "grandiose" (see Table 1 for means). The same group was also asked to identify the genre of the film (after watching it without sound), and gave the three highest ratings for "science fiction," "action," and "thriller" (see Table 2 for means). Taken together, the responses confirmed that the film scene was perceived to be a suspenseful action scene, therefore being congruent with the Chase music and incongruent

⁴ Only the two musical recordings were compared. The music in the original film excerpt was not included in this comparison as the speech and sound effects could not be separated from the music.

TABLE 2. Mean Ratings for Perception of Genre of Film Excerpt for Pilot 1

Movie Genre	Mean	SD
Science fiction	4.51	0.94
Action	4.41	0.68
Drama	3.39	1.08
Thriller	3.51	1.28
Reality	2.37	1.08
Romantic	1.79	1.11
Horror	1.59	0.88
Historical	1.13	0.52
Comedy	1.08	0.42

with the “Moon River” ballad as reflected by the participants’ responses to the music.

Finally, the other group of students (who viewed the film together with the musical soundtracks) gave the highest rating for “creates tension” for the Chase music ($M = 4.86$, $SD = 0.48$) and the lowest rating for “creates tension” for the Nondiegetic “Moon River” version ($M = 1.76$, $SD = 1.34$). The group exposed to the music and film separately also rated both pieces of music as more suitable for nondiegetic than diegetic presentation with the film excerpt, giving a higher rating for the Chase music ($M = 3.94$, $SD = 1.43$) than the “Moon River” music ($M = 2.33$, $SD = 1.85$). These responses confirm that the two music tracks were perceived to have a different effect on the accompanying film scene, and that both music tracks were judged to be plausible pairings with the scene.

PILOT 2: ORIGINAL DIEGETIC AND NONDIEGETIC “MOON RIVER” VERSIONS

As the comparison between the original film excerpt and Nondiegetic “Moon River” version was of primary interest, we also conducted a pilot test to check (1) accuracy in identifying diegetic versus nondiegetic music; (2) clarity of speech of the original Diegetic “Moon River” compared to the altered nondiegetic version; and (3) possible effects of directing participants’ attention to the story versus the sound (as our main procedure did not reveal that the focus was on the music track and directed attention away from the music).

For this pilot test, it was important to test participants similar to those in our main sample, so we recruited 120 undergraduates at one of the two schools from which we drew our sample for the main study. A total of 164 participants completed the procedure, but only the 120 who reported they had no previous familiarity with the *Minority Report* film were included in analyses. Each participant was randomly assigned to one of four

conditions. Two groups were instructed to focus on the story line, and two groups were instructed to focus on the sound in the film (but were not explicitly told to pay attention to the music). Within both the story and sound conditions, one of the two groups was shown the Diegetic “Moon River” clip (original Spielberg version) and the other was shown the Nondiegetic “Moon River” clip. After viewing the film excerpt, participants completed a short questionnaire.

As a check on the music source manipulation of the main study, it was important to determine whether participants attributed the source of the “Moon River” track to the mall in the diegetic version and to the sound track in the nondiegetic version. Overall, music in the Diegetic clip was correctly identified as being diegetic 72.0% of the time and music in the Nondiegetic clip was identified as being nondiegetic 73.0% of the time, suggesting that the music source manipulation was largely noted by participants. There was a significant interaction between music source attribution and experimental condition (attention to story vs. sound), $\chi^2(3) = 28.67$, $p < .01$. Participants’ attributions were more accurate across conditions when they were told to focus on the sound (79% correct identification in the sound condition vs. 64% correct in the story condition), and they were most accurate when instructed to focus on the sound in the diegetic condition (83.3% accuracy vs. 75.7% accurate identification in the nondiegetic condition). It is important to note that responses are based on participants’ recollection of the film clip, as even in the sound condition, participants were not told beforehand that they would be asked questions about the source of the music.

Given that we manipulated the soundtrack in the original film by adding the music used in the scene at a louder level for the Nondiegetic condition, it was also important to ensure that the speech in the original film remained audible. Participants rated the dialogue from the clips they viewed on two scales, one with anchors of “Hard to hear” and “Audible and clear” and the other with anchors of “Could not understand any of the speech” and “Every word was distinct and clear.” In both cases, participants were instructed to place an “X” on a line with an anchor statement on each end of it. Although participants rated dialogue on a continuous line, for the purposes of analysis, anchor values of 0 and 100 were assigned, 0 indicating that the dialogue was “Hard to hear” or that the participant “Could not understand any of the speech.”

Two, two-way analyses of variance were conducted to evaluate participants’ responses. Responses to the two scales served as dependent variables and music source

(diegetic and nondiegetic) and focus (on the scene or the sound) were independent variables. There were no significant main or interaction effects from these analyses. Whether participants were told to concentrate on the sound or story and whether the music was diegetic or nondiegetic did not affect perceptions of the actors' speech on either the "Hard to hear" vs. "Audible and clear" scale: Sound ($M = 64.85$, $SD = 27.76$) vs. Story ($M = 59.46$, $SD = 27.4$); Diegetic ($M = 68.18$, $SD = 27.03$) vs. Nondiegetic ($M = 56.92$, $SD = 27.24$), or the "Could not understand any of the speech" vs. "Every word was distinct and clear" scale: Sound ($M = 73.04$, $SD = 26.03$) vs. Story ($M = 71.96$, $SD = 25.34$); Diegetic ($M = 78.48$, $SD = 22.40$) vs. Nondiegetic ($M = 67.27$, $SD = 27.26$). The finding that ratings of audibility and clarity of speech did not differ between the two versions was of particular importance, indicating that the louder music in our Nondiegetic "Moon River" version did not mask the dialogue to such a degree as to significantly diminish the intelligibility of the speech, compared to the original film excerpt.

THE MAIN STUDY

Participants. A total of 245 participants (110 male, 134 female, and one who did not identify sex) were recruited from introductory psychology courses at two American institutions, a large, private university in the west and a Midwestern liberal arts college. The mean age was 20.25 years ($SD = 2.19$) for males and 19.92 years ($SD = 4.20$) for females. Sixty-two percent of participants reported having taken music lessons. Students indicating music training had studied an average of 2.09 instruments ($SD = 1.06$), including voice, for an overall average of 10.56 years of training ($SD = 6.36$).⁵

Procedure. The 245 participants in the main study were run in six groups of approximately 40, with each group viewing one of the six possible orderings of the three film clips.⁶ (Although the focus of the study was on the first part of the procedure in which each group answered questions on only one version of the film clip, the other two versions were revealed in the second part of the procedure to examine questions of secondary interest). Participants were seated in a large windowless room with tiered seating so that each had a clear view of a large screen with dimensions of 115 x 115 cm. The film images were projected onto the screen via an NEC Model MT1056 projector, and the sound was presented through speakers on either side of the screen. The

loudness of the original Spielberg film excerpt ranged from 57-70 dB, and both altered versions with nondiegetic music tracks ranged from 60-77 dB (intentionally set to the same range of loudness by our video editor during mixing), as measured on a GenRad 1565-B sound-level meter in the occupied room with participants present. These ranges include dialogue and sound effects, which were preserved for the altered tracks.

Participants were told that the aim of the study was "to examine how people understand story lines of film scenes." During recruitment, they were only told that they would take part in "a film study." The topic of film music was also not mentioned during the introduction to the procedure. By directing attention to the film and not the soundtrack, we hoped to create a situation that would resemble film viewing in real world conditions, in which the audience's attention is rarely focused on the music.

In the first part of the procedure, only one of the three film excerpts (lasting 84 seconds) was shown. After viewing the excerpt, participants responded to a series of questions about the scene they had just watched. These items addressed participants' interpretations of the story line, the perceived relationship between the main characters and their intentions toward each other, and the emotions of the male and female characters. Questions about the music accompanying the film were then given on a separate questionnaire after the first questionnaire had been collected, as we did not want to direct attention to the music at the outset.

In the second part of the procedure, the nature of the manipulated tracks was revealed and the two other versions of the film were shown, with the order of presentation counterbalanced across experimental sessions. They were then asked which of the three versions they thought was from the original film. Finally, participants completed a post-questionnaire that assessed level of film literacy and music training, and familiarity with the film and music used in the procedure.

Results

PRELIMINARY ANALYSES

An initial analysis was conducted to determine whether participants could identify the implied source (i.e., diegetic or nondiegetic) of the music in the first film excerpt they viewed. We assessed perceptions of the first film only, as we were most interested in participants' responses before the manipulation of the musical soundtrack was revealed. Participants were asked to check one of the following responses to indicate the supposed source of the music:

⁵No effects for music training were found in any of the analyses.

⁶No effects for order of presentation were found.

TABLE 3. Percent of Participants Indicating the Source of the Film Music was Diegetic or Nondiegetic

Familiarity with <i>Minority Report</i>	Film Excerpt			
	Source attribution	Original Diegetic	Moon River Nondiegetic	Chase Music Nondiegetic
No	Diegetic	41.7	21.9	3.7
	Nondiegetic	58.3	78.1	96.3
Yes	Diegetic	51.7	19.4	5.9
	Nondiegetic	48.3	80.6	94.1

_____ From inside the mall. (Even though the music may have been added in later, it is supposed to sound like it is coming from inside the mall).

_____ From the soundtrack to the film. (The characters in the film could not hear it; it is from the soundtrack accompanying the film).

_____ Other. (Please explain)

The terms “diegetic” and “nondiegetic” were not used, and clear descriptions suitable for a general audience were provided in parentheses as shown above.

Table 3 lists the percent of participants indicating whether the music associated with the first film excerpt they viewed was diegetic or nondiegetic. Chi-square tests for those unfamiliar and for those familiar with *Minority Report* indicated significant differences in the categorization of the implied sources of the three film clips, unfamiliar: $\chi^2(2, 95) = 12.26, p < .01$ familiar: $\chi^2(2, 116) = 27.80, p < .001$. As expected, the nondiegetic congruent music (the Chase Music) was most accurately identified by participants as being nondiegetic; very few participants (less than 4% and 6%) did not correctly identify the Chase Music as an external dramatic score. Interestingly, a fair number of participants (21.9% of those who were not familiar with *Minority Report* and 19.4% of those who were) indicated that they thought the nondiegetic, incongruent music (the “Moon River” dramatic score version) was diegetic. It is important to keep in mind that participants’ attention was not drawn to the music track and they were not told that they would be asked questions about the music, so the participants had to draw on their memory of the film music during an implicit listening task. Although about one-fifth of the participants viewing the music with the Nondiegetic “Moon River” recalled it as originating from inside the world of the characters, other results show that participants responded to the film excerpt paired with the Nondiegetic “Moon River” soundtrack quite differently from the original Spielberg film excerpt with the Diegetic “Moon River” soundtrack.

We also assessed the degree to which participants who were “accurate” in their attributions of the supposed sources of the music in the film excerpts perceived the excerpts differently from those who were “inaccurate.” For each of the analyses conducted, there were no differences between the perceptions of those who were and were not accurate in their attributions of the implied source of the music. In other words, as discussed further in our findings, our diegetic/nondiegetic manipulation had significant effects upon viewers’ perceptions, regardless of whether or not they correctly attributed the supposed source of the music.

PERCEPTIONS OF THE FILM

As familiarity with the film *Minority Report* might affect perceptions of the film excerpts, we assessed participants’ familiarity with the film. A total of 111 participants indicated no familiarity with the film and 130 indicated some familiarity. (“Familiarity” was indicated by the ability to identify the title of the film or to give a brief synopsis of a few sentences). Four participants who did not reveal whether or not they were familiar with the film were dropped from all analyses in which familiarity was included as a variable.

For the purposes of this study, we were particularly interested in the responses of the 111 participants who were unfamiliar with *Minority Report* prior to participating in the study. The first analysis was an investigation of this group’s perceptions of the film scene. A one-way MANOVA was conducted, with ratings on film perception rating scales as dependent variables and film music type as the independent variable (See Figure 2 for anchors associated with the scales). A significant main effect was found for film music type, Wilks Lambda (20, 198) = .57, $p < .001$. Tests of between-subjects effects revealed that each of the univariate tests for all 10 items was significant at the $p < .001$ level. Figure 2 shows the standardized means for each of the film perception scales for each of the three music conditions. Because participants’ ratings were made by placing an X on the

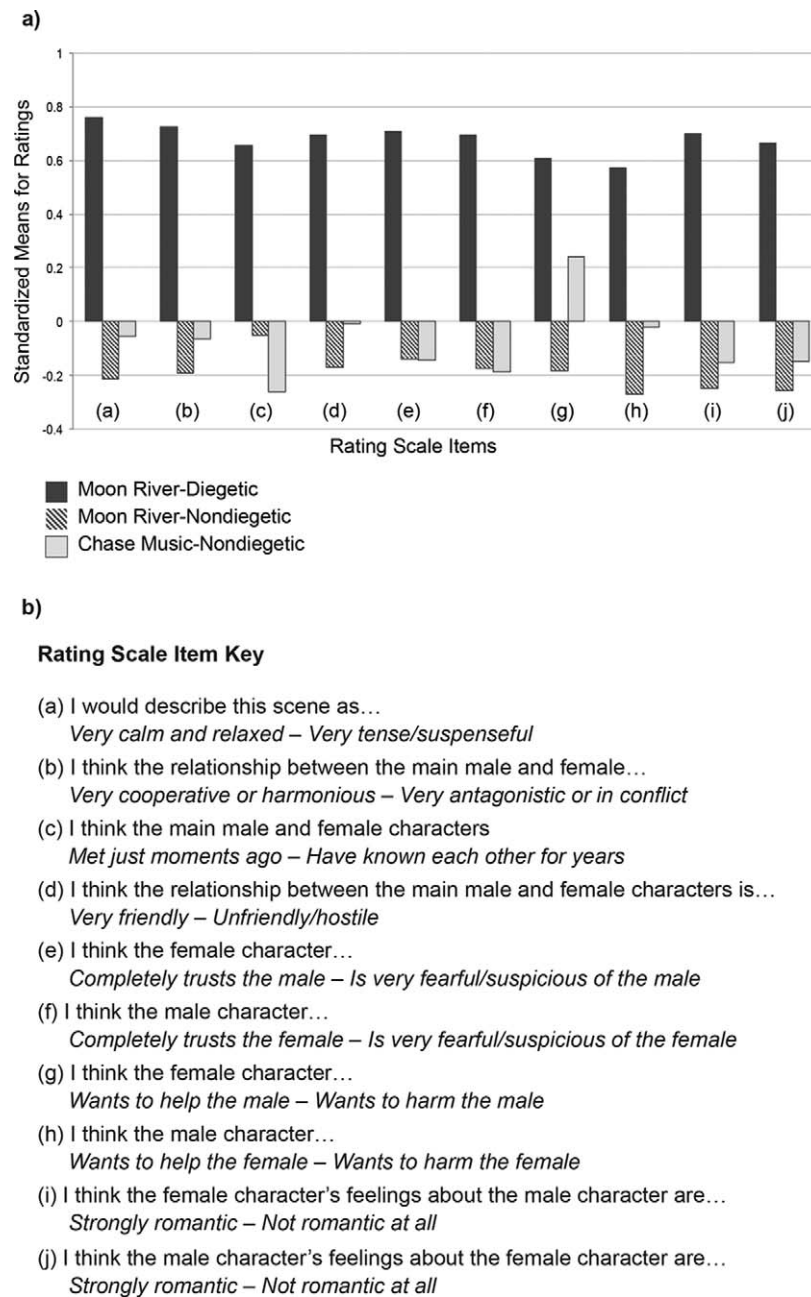
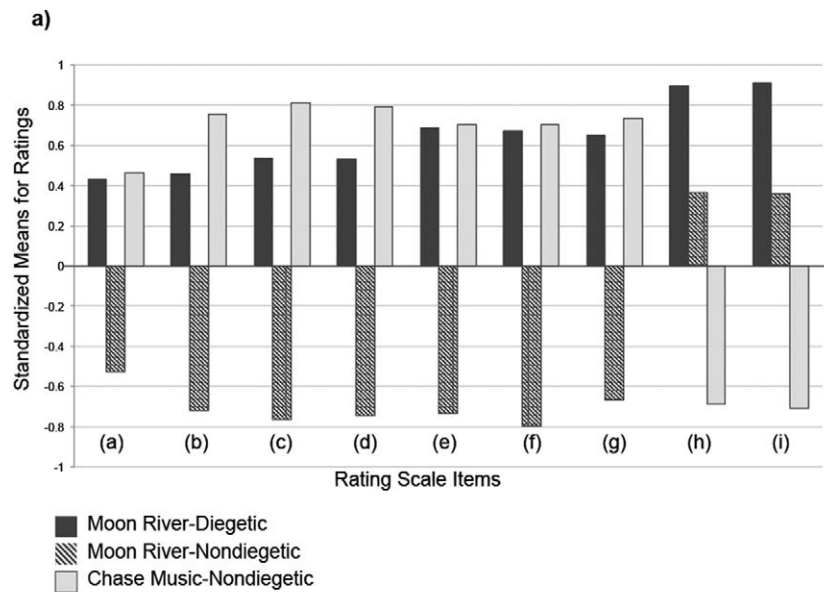


FIGURE 2. a) Means of the standardized ratings on film perception rating scale items for each film music type are displayed on the y-axis. (b) Key for the film perception rating scale items: Prompts and anchors for each item are given, with the first anchor listed corresponding to a more negative rating.

rating line for each item, these ratings were standardized to avoid differences in participants' anchoring of their ratings.

In Figure 2, it may be seen that ratings of the content of the film scene (including such items as mood of scene, relationship between the two main characters, attitude of the main characters toward each other,

and intentions of the main characters) differed significantly for the three film music conditions for participants who were unfamiliar with the film. Specifically, comparison of the means of the standardized ratings on the film perception scales showed that the original Diegetic "Moon River" condition was rated higher *on all of the items*, indicating more intense or antagonistic



b)

Rating Scale Item Key

- (a) Loud – Soft
- (b) Slow – Fast
- (c) Calm/Relaxed – Tense/Exciting
- (d) Gentle/Sweet – Rough/Boisterous
- (e) Reassuring – Unsettling
- (f) Tender/Romantic – Cold/Aloof
- (g) Pleasant – Unpleasant
- (h) Matched mood/emotion shown on screen – Did not match mood/ emotion shown on screen
- (i) Matched actions/events happening on screen – Did not match actions/events happening on screen

FIGURE 3. (a) Means of the standardized ratings on music perception rating scale items for each film music type are displayed on the y-axis. (b) Key for the music perception rating scale items: Anchors for each item are given, with the first anchor listed corresponding to a more negative rating.

interpretations of the characters and scene. Interestingly, the Nondiegetic “Moon River” and the Nondiegetic Chase Music conditions were rated more similarly on the film perception items and both differed from the Diegetic “Moon River” condition. In other words, the differences across the music condition on the film perception scales is largely a function of the *diegetic/nondiegetic* dimension and not the mood *congruence/incongruence* dimension, as one might expect.

PERCEPTIONS OF THE MUSIC

The second analysis focused on participants’ perceptions of the music used in the film clip, and once again included only the 111 participants who reported being unfamiliar with the film. A one-way MANOVA was

again conducted, with ratings on the music perception rating scales as dependent variables and film music type as the independent variable (see the anchors of Figure 3 for these rating scales). A significant main effect was found for film music type, Wilks Lambda (18, 180) = .29, $p < .001$. Tests of between-subjects effects revealed that each of the univariate tests for all 9 items (See Figure 3) were significant at the $p < .001$ level. The means of the standardized ratings on the music perception scales are shown in Figure 3.

As shown in Figure 3, the first seven questions captured general characteristics and mood of the music, whereas the last two questions focused on how well the music matched the emotion and actions on the screen. Interestingly, with the exception of the last two

questions, the original Diegetic “Moon River” and Non-diegetic “Moon River” conditions differed greatly from one another, despite the fact that the same recording of the song was used. These findings suggest that very different moods and characteristics were perceived in the two versions of the same song, depending on the perceived source of the music in relation to the diegesis (as cued by differences in the quality of the sound such as loudness and timbre). It may also be seen that the Nondiegetic “Moon River” and the Nondiegetic Chase Music ratings were nearly mirror images of one another. That is, the items on which one had high means, the other had low means, and vice versa.

PERCEPTIONS OF CHARACTERS' EMOTIONS

In addition to ratings of the film scene and musical pieces, the 245 participants (both those familiar and unfamiliar with *Minority Report*) indicated their perceptions of the male and female characters' emotional states on two different scales. In each case, the anchors were: *fear, depression, anger, boredom, excitement, sadness, anxiety, happiness, distress, peace, contentment, and love (romantic)*. Ratings on each of these scales were analyzed in one of two different MANOVAs (one each for male and female characters' emotions). In each analysis, film music type and familiarity with *Minority Report* served as independent variables and the respective rating scale as dependent variables. As our primary interest was in the effect of the three different musical pieces on perceptions of the film excerpts, main effects of the familiarity factor on perceptions of characters' emotions were not of primary interest. Therefore, only effects of the film music type and film music type by familiarity interaction on perceptions of characters' emotions will be reported.

For the male character, there was a significant main effect of film music type on participants' ratings of his emotional state, Wilks Lambda (2, 236) = .79, $p < .001$. Follow-up univariate analyses indicate significant differences on the fear, $F(2, 234) = 3.88$, $p = .02$, and excitement scales, $F(2, 234) = 5.21$, $p < .01$, and a marginal effect for the love scale, $F(2, 234) = 2.89$, $p = .06$. For the fear and excitement scales, the original Diegetic “Moon River” version (fear $M = 5.10$, excitement $M = 3.82$) and Chase Music means (fear $M = 5.01$, excitement $M = 3.66$) were higher than the Nondiegetic “Moon River” mean (fear $M = 4.68$, excitement $M = 2.93$). For the love scale, the Nondiegetic “Moon River” mean ($M = .93$) was highest, followed by the original Diegetic “Moon River” version ($M = .78$), and then Chase Music means ($M = .48$). These findings suggest the original Diegetic “Moon River” music served to heighten perceptions of

the tension in the scene, just as the Chase Music did, whereas the Nondiegetic “Moon River” did not have this same effect, though it was perceived as adding a romantic element to the scene.

A significant interaction between film music type and familiarity was found for ratings of the female character's emotions, Wilks Lambda (2, 234) = .85, $p = .05$. Follow-up univariate tests on each of the dependent variables indicated significant interactions on the excitement, $F(2, 234) = 3.18$, $p = .04$, and contentment scales, $F(2, 234) = 4.31$, $p = .02$. Means for these interactions are plotted in Figures 4a and 4b. For the excitement scale, the means for the unfamiliar participants in the Original Diegetic “Moon River” version ($M = 2.90$) and Chase Music ($M = 2.82$) conditions were lower on this scale than the means for those familiar with *Minority Report* (Original Version $M = 3.55$, Chase Music $M = 2.96$). However, for the Nondiegetic “Moon River” condition, those unfamiliar with *Minority Report* rated the female as being more excited ($M = 3.13$) than those who were familiar ($M = 2.29$). For the contentment scale, the means for the unfamiliar participants in the Original Diegetic “Moon River” version ($M = .68$) and Chase Music ($M = .56$) conditions were higher on this scale than the means for those familiar with *Minority Report* (Original Version $M = .23$, Chase Music $M = .53$). However, for the Nondiegetic “Moon River” condition, those unfamiliar with *Minority Report* rated the female as being less contented ($M = .36$) than those who were familiar ($M = .90$). It may be seen then, that participants' knowledge, or lack of knowledge, of the plot of *Minority Report* influenced their perceptions of the female character's emotions differently for each music type.

No significant main effects were found for perceptions of the male or female character's emotions when analyses were run for only the participants who were not familiar with the film. Also, there were no gender effects for perception of either character's emotions when responses of male and female participants (who were familiar or unfamiliar with the film) were compared.

IDENTIFICATION OF THE ORIGINAL SOUNDTRACK

One final question of interest was whether participants could identify the music originally paired with the film excerpt. After they had viewed all three scenes (in counterbalanced order), participants were asked which musical soundtrack they thought was used in the original film. Table 4 lists the percent of participants identifying each of the musical pieces as the original music associated with the film scene, separated into those familiar and unfamiliar with the movie *Minority Report* and by

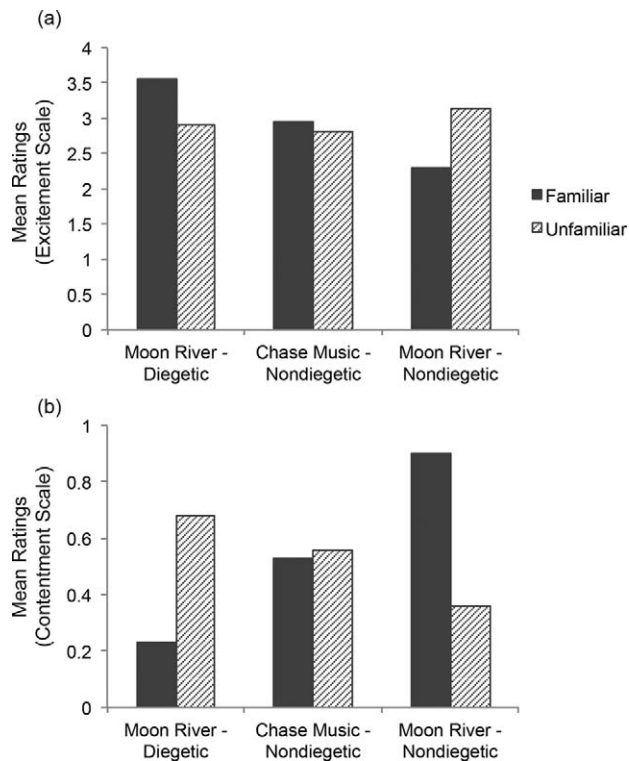


FIGURE 4. (a) Mean ratings on the excitement scale for each film music type for participants familiar and unfamiliar with *Minority Report*. (b) Mean ratings on the contentment scale for each film music type for participants familiar and unfamiliar with *Minority Report*.

which of the three film excerpts they saw first. A hierarchical loglinear analysis was conducted to assess the effects of familiarity with the movie and of the first film excerpt seen on identifications of which music was from the original movie. A significant effect was found for which excerpt was seen first, $\chi^2(4, 244) = 64.12, p < .001$. No other effects were significant. As may be seen in Table 4, when the Diegetic “Moon River” film excerpt (i.e., the original Spielberg version) was seen first, participants were more likely to correctly identify it as the original music. However, when either of the altered versions (the excerpts with either the congruent music or incongruent, nondiegetic music) was seen first, participants were most likely to identify the nondiegetic, congruent music (the Chase Music that was actually taken from *Empire of the Sun*) as the original music in this scene.

Discussion

Previous studies have shown that switching the musical soundtrack accompanying a film scene can lead to

dramatically different interpretations of the scene (e.g., Boltz, 2001; Bullerjahn & Gldenring, 1994; Marshall & Cohen, 1988; Tan et al., 2007; Vitouch, 2001). The present study examined the effects of mixing the same piece of music at a louder level in a film soundtrack to imply either *diegetic music* (presented as if emanating from within the fictional world of the film characters) or to suggest *nondiegetic music* (a dramatic score accompanying the scene but not originating from within the scene). As the diegetic music had been processed to sound like it was transmitted over a public sound system inside a shopping mall in the original soundtrack, timbre also varied somewhat, and because the diegetic music was soft this effect was heard at a low level. Thus, loudness was the main cue for diegetic versus nondiegetic music in this particular film scene.

We discovered that varying the loudness of the music track (to signal diegetic or nondiegetic music) produced significantly different readings of a scene from the film *Minority Report*, with respect to perceptions about the relationship, motivations, and emotions of the characters and degree of overall tension (as described in more detail later in this discussion). Further, it was found that varying the level of loudness (and some timbral differences due to the low fidelity inherent in the processed diegetic music) produced stronger effects than switching to a different musical soundtrack of contrasting mood and character, the method used in most previous studies.

Viewers were not always accurate (that is, with respect to artistic intention) in recalling whether the film excerpt had included diegetic or nondiegetic music in the two versions employing the “Moon River” song. In the main study, the implied source of the altered (nondiegetic) “Moon River” music was correctly identified by 80.6% of those familiar with the film and 78.1% of those not familiar with the film. Correct identification for the diegetic music in Spielberg’s version dropped to chance levels or below: only 51.7% and 41.7% of participants familiar and unfamiliar with the film correctly recalled the music as originating from within the world of the film characters in the original Hollywood version of the film.

However, it should be noted that participants identified the source based on their recollection of the film clip, as they were not told that they would be asked questions about the source of the music beforehand. Our second pilot study showed that directing participants’ attention to the sound in general (though not explicitly directed to the music) improved accuracy for the diegetic version, as 83.3% correctly recalled hearing the music as if playing inside the mall in the original

TABLE 4. Percent of Participants Indicating Each of the Musical Pieces was From the Original Movie

Familiarity with film	Musical piece identified	Film excerpt seen first		
		Original Version	Moon River Nondiegetic	Chase Music
No	Original Version	47.7	28.2	3.6
	Moon River Nondiegetic	4.5	7.7	0.0
	Chase Music	47.7	64.1	96.4
Yes	Original Version	62.5	29.8	5.8
	Moon River Nondiegetic	0.0	14.9	1.9
	Chase Music	37.5	55.3	92.3

Spielberg version, and 75.7% correctly recalled hearing the altered version as a dramatic score. In all reported analyses in the remainder of our findings, there were no differences found between the responses of those who were and were not accurate in their attributions of the supposed source of the music.

INTERPRETATIONS OF THE FILM SCENE

Whether based on their perception of music emanating from within the scene or from a dramatic score, or based on the relative loudness of the music, participants' interpretations of the scene and perceived qualities of the music were dramatically different for the three versions of the film excerpt. Specifically, participants who were unfamiliar with the film and viewed the original Spielberg film excerpt perceived the two main characters as being more antagonistic and hostile toward each other, more fearful and suspicious of each other, less romantically inclined, and having more harmful and negative intentions toward each other, compared to those who viewed the (altered) film excerpt versions paired with either the "Moon River" song or Chase Music as a dramatic score external to the world of the characters. The scene as a whole was also perceived to be more tense and suspenseful in the original diegetic version compared to the Nondiegetic "Moon River" song or Chase Music versions.

It is possible that the gentle ballad music sustains the level of tension of this suspenseful scene if presented as if it could be incidental music playing in the background, inside a shopping mall (diegetic). On the other hand, the same mood-incongruent music (which listeners in the first pilot study described as "romantic," "grandiose," and "cheerful") may be assumed to serve as a commentary to the scene when mixed to sound more like a nondiegetic dramatic score (see Kassabian, 2001, p. 59), thus softening the perceived tension and conflict in this sequence and imbuing the scene with romantic undertones.

Although it might seem surprising that fast music composed for an action scene led to lower ratings for perceived tension for the chase scene than the slow romantic "mall music," *ironic contrast* between music and images may increase tension. Whereas previous studies (e.g., Boltz, 2004; Boltz et al., 1991) have examined ironic contrast provided by nondiegetic musical soundtracks, the present study demonstrates that this intensifying effect may extend to the use of mood-incongruent diegetic music. As film theorist Gorbman (1987) has described: "By taking music meant as extra-narrative comment and rendering it diegetic... the narration motivates, naturalizes the music, makes its disparity with the filmed events acceptable" (p. 23). Indeed, the aforementioned research by Fujiyama et al. (2013) provides empirical evidence that audiences may be more accepting of incongruence between sound and image in instances when the music is clearly diegetic.

Those who had seen the film *Minority Report* before taking part in the study were not as easily influenced by the manipulation of the soundtrack. Nonetheless, the diegetic/nondiegetic manipulation of the "Moon River" music track influenced some perceptions of the film excerpt for the 245 familiar and unfamiliar participants as a whole. For instance, when the same "Moon River" music was presented louder and clearer to suggest a nondiegetic musical soundtrack, viewers perceived the male character to be less fearful, less excited, and somewhat more romantically interested in the female character. Thus, the musical soundtracks seemed to color perceptions of the scene somewhat, even among viewers who had seen the film and were familiar with the story line.

PERCEPTIONS OF THE MUSICAL TRACKS

The finding that the same instrumental piece was perceived quite differently in the two conditions was somewhat unexpected. The "Moon River" music in the original film excerpt was judged by participants

unfamiliar with the film to be significantly faster, softer, more tense and exciting, more rough and boisterous, more unsettling, colder and more aloof, and more unpleasant than the same piece of music presented nondiegetically. However, to achieve the illusion that a recording of music is playing over a public sound system in a crowded shopping mall, the music in the original film was mixed at a low level and almost certainly processed with artificial reverberation, equalizers, and other post-production techniques (Cross, 2013; Rose, 2013) that altered the timbre of the music. Specifically, adding reverberation delays the offset of sounds to create a sense of distance or spaciousness (Holman, 2005, p. 215), and equalizers boost certain elements of the sound spectrum while cutting others (Rose, 2013, p. 356), and these affect both the frequency content and time-based aspects of timbre.

Loudness and timbre—the primary cues signaling to participants that the music was diegetic or nondiegetic in this particular scene—have both been shown to influence listeners' perceptions of emotion in music (e.g., see Juslin & Timmers, 2010; Le Groux & Verschure, 2012). Specifically, many studies point to loudness as an important variable in both emotion expression and emotion induction in the context of music (e.g., Gabrielsson & Lindström, 2010; Juslin & Timmers, 2010). Whereas loudness expresses anger, softness is associated with sadness, tenderness, and fear (Juslin, 1997). Wide fluctuations in loudness express fear whereas consistent loudness tends to convey happiness and pleasantness (Scherer & Oshinsky, 1977). Listeners are also particularly sensitive to timbre (encompassing attack time, spectral energy distribution, and spectral flux) as a cue for emotion in music, and can detect subtle changes in emotional expression in the timbre of single notes (Eerola, Ferrer, & Alluri, 2012; Goydke, Altenmüller, Möller, & Münte, 2004).

Even fine differences in loudness and nuances in timbre may influence the perceived emotion of a piece of music. For instance, in a study using a synthetic system (SMuSe) to generate musical sequences in order to isolate and vary different parameters of music, the same musical excerpts were judged by listeners to be more arousing and negative and dominant when played louder, and damped sounds were perceived as less arousing and less dominant. The attack time and brightness parameters of timbre also influenced emotion ratings, with brighter sounds and shorter attack times perceived as more arousing (Le Groux & Verschure, 2012; see also Juslin & Sloboda, 2010, Chapters 14 and 17). Although the same instrumental song was used in the two versions employing “Moon River” in the present study, the

softer diegetic music and louder nondiegetic version differed greatly in expressive parameters that are strongly linked to emotion.

Our findings also speak to the film director, composer, or music editor's sharp instincts⁷ concerning the most effective integration of music into a scene. Based on the responses from our first pilot study, the Chase Music was more congruent or fitting with the emotion and events depicted in the chosen film scene. Indeed, the majority of participants (161 out of 245 or 65.7%) in our main study thought that the nondiegetic congruent soundtrack (which was actually “chase music” from the soundtrack to *Empire of the Sun*) was the original soundtrack used in the scene. However, the original diegetic presentation of the music—with “Moon River” sounding as if it were playing inside the mall—received the highest ratings of tension and suspense in the overall scene, most distrust and conflict between the characters, and most perceived fear and excitement in the male character. In each case, the greatest degree of tension, conflict, and intensity was yielded by Spielberg's original version of the scene. Naturally, the specific effects of “migrating” a piece of music from diegetic to nondiegetic depend on how the music and moving images and other elements interact within a particular scene; whereas gentle nondiegetic music may have softened this scene, it might have wound up the tension for another.

LIMITATIONS OF THE STUDY AND AVENUES FOR FUTURE RESEARCH

The present investigation is an opportunistic study arising from a film sequence that offered favorable conditions for exploring how viewers respond to diegetic and nondiegetic presentation of film music. It ventures into a fairly new area in the empirical literature on film music on which there has been little prior investigation. As such, the research reported here represents one of many steps needed to elucidate the psychological impact of diegetic and nondiegetic music.

An important limitation of the present study was the use of a single film excerpt. This was mainly due to the practical challenges of finding material meeting the methodological and technical criteria for our study, as described in detail in the Method section. In particular, we were limited by technical constraints, as the

⁷ Whether the choice of music was made by the director, music editor, or another member of Spielberg's team is not known by the authors. Composer John Williams stated that in keeping with standard practice, he had no hand in selecting most of the source music in *Minority Report* (Ostreich, 2002, p. 26). “Moon River” won the Academy Award for Best Original Song for the 1961 film *Breakfast at Tiffany's*. It has been used in numerous films including Oliver Stone's 1989 film *Born on the Fourth of July*, in which Tom Cruise danced to this music at a prom.

diegetic music in the original film scene had to be soft enough to be masked by different music tracks that were laid over it, while preserving speech and sound effects at a sufficiently audible and intelligible level. (Of course, diegetic music can also be very loud and clear, so the manipulation involved in creating a diegetic/nondiegetic versions of a film scene differs with the particulars of the scene).

More importantly, as the fainter and lower fidelity sound of the music served as cues that the music in the original Spielberg version was supposedly playing over distant loudspeakers in a large public space, loudness and timbre of the music were intertwined with the diegetic/nondiegetic manipulation in this study. Thus, as noted in the introduction, we cannot rule out the possibility that the participants' responses are due to the physical differences in the sound parameters of the two different renditions of "Moon River" music—rather than the more conceptual relationship of the perceived source of the music in relation to the diegesis (not clearly separable from a pragmatic film production standpoint).

Further, the diegetic or nondiegetic effect of the music is an *interpretation by the viewer* (Cross, 2013, p. 6) and the question was asked retrospectively, thus we did not see unanimous agreement among the participants as to the supposed source of the music—even for the original Hollywood version of the original scene in *Minority Report* produced by Spielberg's team. Future work should disentangle whether the stark differences in viewers' film interpretation in the current study were driven mainly by different levels of loudness versus viewers' subjective experience of music's source. Perhaps this could be done by collecting measures related to both factors, and determining which is more predictive of how the film is ultimately interpreted.

However, finding that variations in loudness and timbre in the "Moon River" musical track (as cues for diegetic versus nondiegetic music) produced dramatic differences in the viewers' impression and interpretation of the scene and characters, is in itself informative and interesting. As we have established, these are among the main post-production modifications that sound editors make to sound effects, dialogue, and music *to suggest* the source of music (e.g., live or if transmitted through devices such as loudspeakers), and *to suggest* whether the music is inside or outside of the story world (e.g., see Cross, 2013; Holman, 2010, Chapters 7 and 12; Sonnenschein, 2001, pp. 159-164). It would be interesting to repeat the present study with a film scene in which diegetic music is played loudly in the original scene, and to soften the diegetic music to see the effects

on viewers' ratings for tension of the scene and conflict between the characters.

The most direct route for future study would be through collaboration with professionals in the film industry with access to the separate audio elements of the soundtrack. Short of having access to these original resources, future studies could examine whether the audience's impressions of a film scene differ in instances in which the loudness and timbre of diegetic and nondiegetic music are as closely matched as possible, or in rare scenes where the source of the music is designed to be ambiguous (see Kassabian, 2001, p. 47). Moreover, assuming a diegetic/nondiegetic dichotomy overlooks the fluidity and variety of functions of music in relation to the fictional world of the film (see Hagen, 1971; Kassabian, 2001; Neumeyer, 2009; Stilwell, 2007; Winters, 2010). There is also "transdiegetic" music that may weave in and out of the diegesis; 'ambidiegetic' music in which diegetic music functions much like nondiegetic music, contributing to the dramatic development of plot, character, and other themes; and 'metadiegetic' music that seems to reflect a character's thoughts or imaginings (Gorbman, 1987, pp. 22-23; Holbrook, 2011). Thus the present study begins by exploring the broadest contrast and invites finer comparisons in future investigations.

Finally, this study branched out from the standard "switch-the-music-track" paradigm that emerged from the earliest published empirical studies on film music in the 1980s and is still valuable today (Tan, 2017b), to explore the effects of modifications to the musical score originally paired with the scene. Gorbman said, "change the score on the soundtrack, and the image-track can be transformed" (p. 30). Our findings suggests that varying the loudness of the same piece of music as cues to signal diegetic ("source music") versus nondiegetic music ("dramatic score") may significantly change the audience's perception of many elements of a scene, without having to replace the score.

Author Note

The authors are grateful to Lola Cuddy and William Forde Thompson for their valuable advice, and to our dedicated reviewers for detailed and insightful comments on earlier drafts. The authors also thank Danny Kim at Power Up Productions for preparing the audiovisual materials, Blas Payri at the Audiovisual Communication Department at the Universidad Politécnica de Valencia for an enlightening technical analysis of the film excerpt and for testing materials in his audiovisual classes, and acclaimed music editor Scott Stambler at

Chapman University's Dodge College of Film and Media Arts for informative conversations about film audio. Our thanks also to Dana Allswede, Clay Garnett, Rachel Jeffery, Jennifer McCutchen, and Emily Rohn (Kalamazoo College) and Justin Hale, Paul Hansen, Suzanne Lopes, Lara Schwicht, and Drew Smith (Brigham Young University) for assistance with data gathering and data entry, and to Matthew Bezdek for assistance with a figure. A portion of the preliminary findings was reported in the Conference Proceedings for

the 10th International Conference of Music Perception and Cognition in Sapporo Japan. At the time the final manuscript revision was submitted, E. Wakefield's institutional affiliation was the Department of Psychology at Loyola University Chicago.

Correspondence concerning this article should be addressed to Siu-Lan Tan, Psychology Department, Kalamazoo College, Kalamazoo Michigan 49006. E-mail: tan@kzoo.edu

References

- BEZDEK, M., WENZEL, W., & SCHUMACHER, E. (November, 2014). *The impact of music on neural activity during suspenseful film viewing: An fMRI study*. Poster session presented at the Annual Meeting of the Psychonomic Society, Long Beach, California.
- BOLIVAR, V. J., COHEN, A. J., & FENTRESS, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology, 13*, 28-59.
- BOLTZ, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception, 18*, 427-454.
- BOLTZ, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory and Cognition, 32*, 1194-1205.
- BOLTZ, M., SCHULKIND, M., & KANTRA, S. (1991). Effects of background music on the remembering of filmed events. *Memory and Cognition, 19*, 593-606.
- BROOKS, M. (DIRECTOR & PRODUCER). (1977). *High anxiety* [Motion picture]. United States: Twentieth Century Fox, Crossbow Productions.
- BROWN, R. (1994). *Overtones and undertones: Reading film music*. Berkeley, CA: University of California.
- BULLERJAHN, C., & GÜLDENRING, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology, 13*, 99-118.
- COHEN, A. J. (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review, 9*, 163-178.
- COHEN, A. J. (2013). Film music from the perspective of cognitive science. In D. Neumeier (Ed.), *The Oxford handbook of film music studies* (pp. 96-130). Oxford, UK: Oxford University Press.
- COHEN, A. J. (2014). Resolving the paradox of film music through a cognitive narrative approach to film comprehension. In J. C. Kaufman & D. K. Simonton (Eds.), *The social science of cinema* (pp. 57-83). New York: Oxford University Press.
- 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.
- COOKE, M. (2001). Film music. In *The new Grove dictionary of music and musicians* (Vol. 8, pp. 797-810). New York: Macmillan.
- CROSS, M. (2013). *Audio postproduction: Form film and television*. Boston, MA: Berklee Press.
- EEROLA, T., FERRER, R., & ALLURI, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Perception, 30*, 49-70.
- FUJIIYAMA, S., EMA, K., & IWAMIYA, S. (2013). Effect of technique of conflict between music and moving picture employed in a movie directed by Akira Kurosawa, *The Journal of the Acoustical Society of Japan, 69*, 387-396. (Published in Japanese⁸)
- GABRIELSSON, A., & LINDSTRÖM, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367-400). New York: Oxford University Press.
- GOLDMAN, G., & SHUSETT, R. (EXECUTIVE PRODUCERS), & SPIELBERG, S. (DIRECTOR). (2002). *Minority Report* [Motion picture]. United States: DreamWorks LLC and Twentieth Century Fox.
- GORBMAN, C. (1987). *Unheard melodies: Narrative film music*. Bloomington, IN: Indiana University Press.

⁸ A short English summary of this Japanese publication can be found in: Fujiyama, S., Ema, K., & Iwamiya, S. (2015). Effect of technique of conflict between music and moving picture employed in a movie directed by Akira Kurosawa, *Proceedings of the Kyushu Youngnam Joint Conference on Acoustics, Nagasaki, Japan, 202*, 1-4.

- GOYDKE, K. N., ALTENMÜLLER, E., MÖLLER, & J., MÜNTE, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research*, 21, 351-359.
- HAGEN, E. (1971). *Scoring for film*. New York: Criterion.
- HOECKNER, B., WYATT, E. W., DECETY, J., & NUSBAUM, H. (2011). Film music influences how viewers relate to movie characters. *Psychology of Aesthetics, Creativity, and the Arts*, 52, 146-153.
- HOLBROOK, M. (2011). *Music, movies, meanings, and markets*. New York: Taylor and Francis.
- HOLMAN, T. (2005). *Sound for digital video*. New York: Focal Press.
- HOLMAN, T. (2010). *Sound for film and television*. New York: Focal Press.
- JUSLIN, P. N. (1997). Emotion communicated in music performance: A functionalist perspective and some data. *Music Perception*, 14, 383-418.
- JUSLIN, P. N., & SLOBODA, J. A. (Eds.) (2010). *Handbook of music and emotion: Theory, research, applications*. New York: Oxford University Press.
- JUSLIN, P. N., & TIMMERS, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 453-489). New York: Oxford University Press.
- KASSABIAN, A. (2001). *Hearing film: Tracking identifications in contemporary Hollywood film music*. New York: Routledge.
- LARSEN, P. (2005). *Film music* (J. Irons, Trans.). London, England: Reaktion Books.
- LE GROUX, S., & VERSCHURE, P. F. M. J. (2012). Subjective emotional responses to musical structure, expression, and timbre parameters: A synthetic approach. In *Proceedings of the 9th International Symposium on Computer Music Modeling* (pp. 160-176). Queen Mary University of London, UK.
- MARSHALL, S. K., & COHEN, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception*, 6, 95-112.
- NEUMEYER, D. (2009). Diegetic/nondiegetic: A theoretical model. *Music and the Moving Image*, 2, 26-39.
- MANCINI, H., & MERCER, J. (1961). Moon River (Recorded by H. Mancini and his orchestra). On *Minority Report* [Motion picture DVD]. New York: RCA Records.
- OESTREICH, J. R. (2002, June 30). Schubertizing the movies. *The New York Times* (pp. 26-27).
- PAVLOVIĆ, I., & MARKOVIĆ, S. (2011). The effect of music background on the emotional appraisal of film sequences. *Psihologija*, 44, 71-91.
- ROBBE-GRILLET, A. (DIRECTOR), HALFON, S., & TOMASKOVIC, J. (Producers) (1968). *L'Homme qui ment* [The man who lies] [Motion picture]. France: Como Films.
- ROSE, J. (2008). *Producing great sound for film and video: Expert tips from preproduction to final mix* (3rd ed.). New York: Focal Press.
- ROSE, J. (2013). *Audio postproduction for film and video: After-the-shoot solutions, professional techniques, and cookbook recipes to make your project sound better* (2nd ed.). New York: Focal Press.
- SCHERER, K. R. & OSHINSKY, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, 1, 331-346.
- SCHRADER, M. (DIRECTOR), & KRAFT, R., WILLBANKS, J., THOMPSON, T., HOLMES, K., GOLD, N., CHAVARRIA, C. (PRODUCERS). (2017). *Score: A film music documentary* [Motion picture]. United States: Epicleff.
- SHAPIRO, R. (EXECUTIVE PRODUCER), & SPIELBERG, S. (DIRECTOR). (1987/2001). *Empire of the Sun* [Motion picture]. United States: Warner Brothers.
- SHEVY, M. (2007). The mood of rock music affects evaluation of video elements differing in valence and dominance. *Psychomusicology*, 19, 57-78.
- SMITH, J. (2009). Bridging the gap: Reconsidering the border between diegetic and nondiegetic music. *Music and the Moving Image*, 2, 1-25.
- SONNENSCHNEIN, D. (2001). *Sound design: The expressive power of music, voice, and sound effects in cinema*. Studio City, CA: Michael Wiese.
- STILWELL, R. (2007). The fantastical gap between diegetic and nondiegetic. In D. Goldmark, L. Kramer & R. Leppert (Eds.), *Beyond the soundtrack: Representing music in cinema* (pp. 187-202). Berkeley, CA: University of California Press.
- TAN, S.-L. (2017a). Scene and heard: How music shapes interpretations of film. In R. Ashley, & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 363-376). New York: Routledge.
- TAN, S.-L. (2017b). From intuition to evidence: The experimental psychology of film music. In M. Mera, R. Sadoff, & B. Winters (Eds.), *The Routledge companion to screen music and sound* (pp. 517-530). New York: Routledge.
- TAN, S.-L., COHEN, A. J., LIPSCOMB, S. D., & KENDALL, R. A. (2013). Future research directions for music and sound in multimedia. In S. L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall, (Eds.), *The psychology of music in multimedia* (pp. 391-405). Oxford, England: Oxford University Press.
- TAN, S.-L., SPACKMAN, M. P., & BEZDEK, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting music before or after a character is shown. *Music Perception*, 25, 135-152.
- VITOUCH, O. (2001). When your ear sets the stage: Musical context effects in film perception. *Psychology of Music*, 29, 70-83.
- WINTERS, B. (2010). The nondiegetic fallacy: Film, music, and narrative space. *Music and Letters*, 91, 224-244.

Appendix: Dialogue in Film Excerpt Used in Study

Agatha: Can you see the umbrella? Take it. Take it! A man in a blue suit. He drops his briefcase. You see a woman in a brown dress? She knows your face. Turn in here.

Police (off screen): Let's go!

Agatha: Wait.

John: No, we can't stop here.

Agatha: Can you see the balloon man? Wait. Wait.

John: What're we waiting for?

Agatha: Wait. Wait. Wait. Wait.

Source:

GOLDMAN, G., & SHUSETT, R. (EXECUTIVE PRODUCERS), & SPIELBERG, S. (DIRECTOR). (2002). *Minority Report* [Motion picture]. United States: DreamWorks LLC and Twentieth Century Fox.