

The role of music-induced emotions on recognition memory of filmed events

Psychology of Music

1–16

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/03057356211033344

journals.sagepub.com/home/pom**Ljubica Damjanovic**  and **Agata Kawalec**

Abstract

The present study investigated the effects of pairing a comedic movie trailer with emotive music on subsequent recognition memory of the events depicted in the trailer. In an independent groups design, the comedic trailer was paired with happy music (congruent condition) or sad music (incongruent condition). A no music condition served as the control condition. The results showed that participants in the incongruent condition displayed a recognition memory advantage for visual test items over participants in the congruent and control conditions. While changes in self-reported positive and negative affect did not correlate significantly with recognition memory, the perception of emotion-specific categories did. These findings help to establish an empirical basis of ironic contrast techniques and propose an affective component in the integration and representation of audiovisual action that is likely to emerge where a participant perceives or recognizes expressed emotions in music, without necessarily feeling an overall positive or negative affect.

Keywords

music-induced emotions, sadness, recognition memory, ironic contrast, perception

A long-held view by film directors and film theorists alike is that music can have a subtle but powerful influence on the cinematic experience (e.g., Carroll, 1988; Eisenstein, 1949). Despite relying predominantly on the visual medium to tell a story, the careful composition of a musical score's emotional properties can contribute just as much to the story's meaning, influencing our perception of film characters (e.g., Hansen & Hansen, 1988; Herget, 2021; Hoeckner et al., 2011; Marshall & Cohen, 1988; Steffens, 2020), the recall of details of a film scene (Boltz, 2001), and how we perceive friendly and aggressive behaviors (e.g., Bolivar et al., 1994; Bullerjahn & Gldenring, 1994; Nosal et al., 2016). It appears that even with durations as short as 15 s, the emotional impact of the musical excerpt is sufficient enough to influence an

School of Psychology, Liverpool John Moores University, Liverpool, UK

Corresponding author:

Ljubica Damjanovic, School of Psychology, Liverpool John Moores University, Tom Reilly Building, Byrom Street, Liverpool L3 3AF, UK.

Email: L.Damjanovic@ljmu.ac.uk

audience's interpretations of a character's emotions. Thus, cueing participants with "fear" music, primes participants to look for signs in the facial expressions of the characters that match the music's emotions (Tan et al., 2007).

This tight coupling between sight, sound, and other modalities enable us to develop robust representations between perception and action (e.g., Blake & Shiffrar, 2007; Damjanovic et al., 2018; Prinz, 1997). In some instances, however, one modality tends to dominate over another, altering the overall perception of the combined channels. For example, in the McGurk effect (McGurk & MacDonald, 1976), the auditory perception of the syllable (ba) is heard as (da) when it is combined with the visually presented lip movement of the syllable (ga). A similar effect has also been observed in audiovisual integration of emotional information, such that the identification of facial emotion is biased in the direction of the emotion conveyed in the voice (e.g., De Gelder & Vroomen, 2000; see also Ethofer et al., 2006; Pan et al., 2019; Vuoskoski et al., 2014; Weijkamp & Sadakata, 2017). These biases are not only restricted to the encoding of the incoming stimulus, but can affect the representations stored in long-term memory (LTM) as well. For instance, the "face advantage" documented in the person recognition memory literature shows that people are better at retrieving episodic and semantic memories when cued with a person's face than their voice, even though we rely heavily on both sources of identity information in our everyday interactions with others (e.g., Barsics & Brédart, 2011; Brédart et al., 2009; Damjanovic, 2011; Damjanovic & Hanley, 2007; Hanley & Damjanovic, 2009).

In many ways, the perception of film and music is comparable to that of faces and voices, which suggests activation of a common set of cognitive and neural mechanisms in audiovisual integration (e.g., Boltz, 2017; Holmes et al., 2009). In the context of the current work, a natural question to ask here is whether the emotional properties conveyed by music has any effect on the nature of the memory representation of a movie scene. One theoretical framework specifically developed to address the audiovisual integration of music and film information is the Congruence-Associationist Model (CAM: Cohen, 2000, 2013). The main sources of information within a film—namely, speech, music, visual action, and more recently, vibrotactile sensations—are processed at four levels (Boltz, 2004; Cohen, 2013; Tan, 2017). At the sensory level (A), the physical surface features of each domain are processed separately. At the next level (B), cross-modal analyses occur, in which the information is preattentively assessed for its degree of structural congruencies or redundancies. When the auditory information and visual information are structurally congruent, the visually congruent information becomes the focus of attention. From Level B, the audiovisual information is then sent to consciousness, short-term memory (STM), in which people construct a visual or working narrative and attempt to make sense of the film as a whole (Baddeley, 1986). This process (C) is facilitated by top-down processing from Level D in which LTM is accessed to generate inferences (e.g., how characters typically behave in particular contexts). Information in STM is attended if matched by information sent from LTM. Experiences in LTM include affective tone, represented by a network of emotion-specific concepts (e.g., happiness, sadness, anger, fear, etc.) typically referred to as *nodes*. When a particular node is activated, this activation spreads through its network of connections to evoke emotion-related memories and cognitions. Thus, during an emotional state, information that is associated with that emotion is more likely to come to mind and to influence the interpretation of ongoing stimulus events (e.g., Bower, 1981; Bower & Cohen, 1982).

In support of the CAM, Boltz et al. (1991) observed how films with a distinctly positive or negative overall affect were associated with better recall accuracy when accompanied by a musical score with similar affect than those with incongruent music. Here the mood-congruency effect created by a match between the affective tone of the film scene and the music enhanced the visual or working narrative for subsequent recall. However, these

mood-congruency effects seem to be dependent upon the relative placement of the music, with the strongest effects emerging when music accompanies rather than foreshadows a visual scene. In the foreshadowing condition, a mood-incongruence effect was found such that participants were better at scene recall when the musical score differed in affective meaning to the visual scene. Indeed, filmmakers often rely on a range of mood-incongruency techniques to create enduring and distinctive movie scenes. One such technique is the *ironic contrast* (e.g., Boltz, 2001, 2004; Boltz et al., 1991; Bordwell & Thompson, 1979; Giannetti, 1982). This involves combining emotionally negative scenes (e.g., sadness, fear, and anger) with emotionally positive music (e.g., happiness; Boltz et al., 1991), such that the incongruous background music leads to an emotional neutralization of the film scene and partly to a sarcastic effect (Rosenfeld & Steffens, 2019). The movies *Bowling for Columbine* and *A Clockwork Orange* provide examples of violent episodes that are accompanied by incongruent music.

Under laboratory movie viewing conditions, Boltz (2004) argues that the affective properties of the accompanying music attentionally highlights common patterns of audiovisual action, serving to establish a more coherent and integrated memory representation. For example, happy soundtracks typically display wide pitch variations, “bouncy” rhythms, and a relatively fast tempo (e.g., Hevner, 1936; Levi, 1982; Rigg, 1964; Scherer, 1979). It is proposed that when such a soundtrack accompanies a “happy” scene, attention is guided toward similar properties within the ongoing action, resulting in an integrated representation that can subsequently facilitate both comprehension and memory (e.g., see also Cohen, 2001). In scenes where the audiovisual action differs in affective content from the accompanying soundtrack, as in the case of ironic contrast, events across the two modalities are processed independently of each other. Under these circumstances, attention is misguided toward irrelevant information, such that the underlying goal of behavior and event relationships lacks coherence and meaning, resulting in representation which is much more fragmented in memory (Boltz, 2001; Boltz et al., 1991).

Music also has the capacity to both “express” or “represent” emotions that are perceived by the listener and “induce” emotions that are felt by the listener (e.g., Evans & Schubert, 2008; Gabrielsson, 2002; Juslin & Västfjäll, 2008; Kallinen & Ravaja, 2006; Kivy, 1990; Timmers, 2017). In terms of the cognitive processing of film and movie soundtracks, the selection of emotive music has generally been approached along a single bipolar scale such as “positive” and “negative” (e.g., Boltz, 2001, 2004; Boltz et al., 1991), yet whether such manipulations correspond in any way to changes in participants’ affective experiences or perceptions of emotion-specific categories has rarely been examined. As such, any significant relationships to emerge between the emotions felt and the specific emotions perceived in response to viewing a movie scene accompanied by music would point toward an affective component in the integration and representation of audiovisual action. The current work contributes to this important theoretical distinction between perceived and felt emotions by comparing changes in self-reported positive and negative affect and the perceived emotions in the music pieces on the events during a video clip.

Another important issue to address is whether exposing participants to music-based emotions mediates cognition in different ways depending on the stimuli and the testing format of the task at hand. For instance, both mood-congruency (e.g., Boltz, 2004; Boltz et al., 1991; Clark & Teasdale, 1985) and mood-incongruency effects (Vuoskoski & Eerola, 2012) have been observed in free recall tasks for words and film scenes, whereas ratings for facial expressions of emotions tend to elicit stronger mood-congruency effects. For example, participants exposed to depressing music perceived more rejection/sadness in ambiguous faces and less happiness in clear faces (e.g., Bouhuys et al., 1995). Personal relevance of the music used in the

experimental conditions also appears to play an important role in mood congruency effects with participants exposed to self-selected sad music giving lower happiness ratings to facial expressions than participants exposed to unfamiliar sad music and neutral music (Vuoskoski & Eerola, 2012).

The present research provides a preliminary investigation of these issues by considering the potential influence of background music on recognition memory for details from a movie trailer for the comedy *Table 19*. Movie trailers provide a unique film exhibition in which promotional discourse and narrative pleasure are combined to create expectations of what viewers will see in the film (Strobin et al., 2015). According to Finsterwalder et al. (2012) they are constructed with three specific elements: the first introduces the viewers to the characters and environment in which the film occurs; the second suggests some brief mode of tension or change in storyline; and the third presents a potential resolution of the storyline. The *Table 19* trailer predominantly contains visual sequences and behaviors that are typically associated with joyful emotions including instances of smiling faces, being bouncy and bubbly, giggling, laughing and physically hyper displays such as dancing, among others (e.g., Shaver et al., 1987). The structural correlates of different moods and emotions are remarkably invariant across domains such as music, walking gaits (e.g., Montepare et al., 1987), facial expressions (e.g., Damjanovic et al., 2010, 2010; Damjanovic & Santiago, 2016), and speech (e.g., Collingnon et al., 2008; Cosmides, 1983). Given that the accompanying music forms an essential part of the trailer viewing experience, by attracting the audience's attention, setting the stage, and creating the overall tone for the visual sequences contained within, we selected musical movie scores that have previously been validated by Eerola and Vuoskoski (2011) for their perceived emotionality to accompany the trailer for the comedy *Table 19*.

In one of our experimental conditions, we paired the trailer with happy music to create our music-congruent condition. In another, we paired the trailer with sad music to create our music-incongruent condition. For the third condition, which functioned as the control, no music or sound was presented at all. We devised a 4AFC recognition memory test consisting of 16 questions relating to the movie trailer. Half the AFC test items were presented with visual cues and half were presented only with verbal cues. In studies where the effects of different stimuli have been compared (e.g., visual vs. verbal), the nature of the task has varied substantially, as has the dependent variable of interest (e.g., Vuoskoski & Eerola, 2012). Thus, it is currently unclear to what extent music-based congruency/incongruency effects can be explained on the basis of stimulus type or task demands. Furthermore, while there have been some previous efforts made to establish music-congruency effects on recognition memory for visual items from a movie scene, these have largely been derived from test items presented exclusively in a verbal format (e.g., Boltz, 2001). By creating our recognition memory task in this unique way, we will be able to address this gap in knowledge by offering a new insight into whether music has a differential impact on recognition memory for visual and verbal test items while keeping the nature of the task constant.

According to the CAM, when there is a match in the affective component between the visual scenes and music a stronger STM visual narrative is created based on the predictions and experiences stored in LTM (Bower, 1981). Based on this assumption, we predicted that the music-congruent condition would yield better recognition memory performance than the control condition. This is because the "happy" music would serve as an effective cue to direct participants' attention to similar parts of the film, thus facilitating the encoding of a unified representation of the visual scene (e.g., Boltz, 2001, 2004). If the mismatch in emotionality between the sad musical piece and the comedic visual scene directs participants' attention toward irrelevant information, effectively disrupting the encoding of the scene in a unified way and

resulting in a weaker STM visual narrative, then we would expect the music-incongruent condition to result in poorer recognition performance from the music-congruent and control conditions (e.g., Boltz, 2004; Boltz et al., 1991).

Changes in positive and negative affect as measured by the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) will be used as a measure of music's capacity to induce an affective change in the participant (e.g., Juslin & Västfjäll, 2008), whereas self-reported ratings of perceived happiness and sadness will be used as a measure of each musical piece's ability to represent the intended emotion in the listener. These measures will be used to establish the relationship between affective experience and perceived emotion on the recognition memory test items. It is currently unclear whether such correlations can be found for recognition memory tasks associated with music-based emotions, although previous research with other cross-modal interactions and emotion-based cognitive tasks suggest that under certain conditions they may play an important role (e.g., Damjanovic et al., 2014, 2018, 2020). Thus, any significant correlations that emerge from these self-reported measures would make an important theoretical advance by highlighting how the representation of audiovisual action may be grounded by music's capacity to represent emotions and induce affective states in the listener.

Method

Ethics statement

The participants provided written consent to procedures approved by the ethics committee of the University's School of Psychology.

Participants

A total of 60 participants were recruited for the study, from which 55 contributed data to the analysis (see "Design and Analysis" section for exclusion details). The participants were students recruited from the University campus. Twenty participants were randomly assigned to one of the three between-subject conditions. All participants self-reported that they possessed normal to normal-to-corrected vision and normal hearing.

Apparatus and stimuli

Video stimulus. The movie trailer from the comedy *Table 19* (Levy et al., 2017) served as the video stimulus for the study. The clip lasted 2 min 24 s in duration and depicted scenes that conveyed superficially by gestures and facial expressions of the actors the emotion of happiness (e.g., happy facial expressions, laughter, smiling, dancing, and scenes from a wedding party). The scenes included close-ups and medium and long shots of one or more actors and contained dialog among its main characters. The soundtrack, including the background music and the spoken dialogue between characters was removed by recapturing the video with the audio feature disabled using video editing software Adobe Premier Pro (Adobe Inc., San Jose, California). The final version of the video was exported as a QuickTime movie file.

Music stimuli. The selected music examples were taken from the study by Eerola and Vuoskoski (2011). Among others, the authors evaluated a broad and structured pool of musical stimuli according to the basic emotions of anger, fear, happiness, sadness, and tenderness. From this pool, film music was used that, according to Eerola and Vuoskoski (2011), conveyed the

emotions happiness and sadness. To create the music-congruent condition, we selected from the happy emotion category, "The Beginning of the Partnership" (Warbeck, 1998, track 1) from the film *Shakespeare in Love* (Parfitt et al., 1998). The music was a fast-paced orchestral piece performed in the classical style. As reported by Eerola and Vuoskoski (2011), the extract was associated with a perceived mean happiness score of 7.17 on a response scale that ranged from 1 to 9. The musical piece lasted 2 min in duration and was looped to match the duration of the video clip. Adobe Premier Pro (Adobe Inc., San Jose, California) was used to combine the audio and video. The final version of the video was exported as a QuickTime movie file.

To create the music-incongruent condition, we selected from the sad emotion category, "Black Nights" (Yared, 1996, track 10) from the film *The English Patient* (Zaentz & Minghella, 1996). The music was a slow-paced orchestral piece performed in the classical style. As reported by Eerola and Vuoskoski (2011), the extract was associated with a perceived mean sadness score of 7.50 on a response scale that ranged from 1 to 9. The musical piece lasted 1 min 53 s in duration and was looped to match the duration of the video clip. Adobe Premier Pro (Adobe Inc., San Jose, California) was used to combine the audio and video. The final version of the video was exported as a QuickTime movie file. All QuickTime movie files were presented to participants via the QuickTime Movie Player program on a laptop PC with built-in speakers.

Measures

The PANAS. The PANAS (Watson et al., 1988) was used to measure positive affect (PA) and negative affect (NA) across all groups. The PANAS contains 20 one-word adjective items reflecting PA and NA. Participants rated the degree to which certain emotions were felt using a 5-point Likert-type scale, ranging from 1 (*very slightly or not at all*) to 5 (*extremely*) using a general time frame to assess trait affectivity. Across PA and NA scales, the PANAS has demonstrated adequate reliability, in terms of internal consistency and test–retest reliability (Watson et al., 1988). PA scores can range from 10 to 50, with higher scores representing higher levels of PA. NA scores can also range from 10 to 50, with lower scores representing lower levels of NA.

The 4AFC recognition memory task. A paper-based 16 item recognition memory task was devised for the study relating to details depicted in the movie trailer. Half the questions were accompanied by visual items captured as screenshots from the trailer (e.g., picture of a target wedding cake for the following test item: "which is the cake that the main character almost walks into?") and half without (e.g., "how many guests were already at the table when the character joined?"). Distracters were selected to closely resemble the correct answer. The location of the correct answer was randomly determined for each question.

Previous viewing experience questions. Participants in all conditions were required to answer the following question: "Have you ever seen the film *Table 19*, if yes how recently?" Participants in the music-congruent condition were additionally asked: "Have you ever seen the film *Shakespeare in Love*, if yes how recently?" and participants in the music-incongruent condition were additionally asked, "Have you ever seen the film *The English Patient*, if yes how recently?"

Emotion rating task. Participants in the music-congruent and music-incongruent condition were instructed to rate the perceived emotions in the music pieces using the five rating scales (happy, tender, sad, angry, and fearful) ranging from 1 (*not happy*) to 9 (*very happy*).

Procedure

Participants were tested individually. Participants first signed the informed consent and completed the PANAS scale (Watson et al., 1988). They were then instructed to watch one of the three movie files. As soon as the movie finished, participants were presented with a response booklet to complete the 4AFC task. They were provided with a time limit of 5 min. Participants did not receive any feedback on their recognition memory performance. Upon completion of the 4AFC task, participants provided written responses to the previous viewing experience questions. Participants in the experimental conditions also completed the emotion rating task. Finally, participants were provided with the PANAS to complete as a posttest measure of affective state. Each individual testing session took approximately 20 min to complete.

Design and analysis

Given that five participants had previously seen *Table 19*, their entire dataset was eliminated from the analysis (see also Boltz, 2001). Percentage correct scores on the 4AFC task were analyzed using a 3 (group: control, congruent, or incongruent) \times 2 (test item: visual or verbal) mixed-measures analysis of variance (ANOVA) with repeated measures on the last factor. A significant interaction was followed up with focused simple main effects analyses to examine the hypothesis of whether music has a differential impact on the processing of visual versus verbal test items. Significant simple main effects were followed up with planned comparisons *t*-tests for the following hypotheses-driven group comparisons: control versus congruent, congruent versus incongruent, and incongruent versus control. We analyzed the PA and NA scores separately using a 3 (group: control, congruent, or incongruent) \times 2 (time: pre- or post-test) mixed-measures ANOVA with repeated measures on the last factor. Significant interactions for the PANAS measures were followed up with simple main effects analyses to establish whether there were existing mood differences between groups for either PA or NA at pretest and whether there were any significant fluctuations in affect between the pre- and posttest for each group. An alpha level of $p < .05$ was set for these analyses.

We calculated the change in PA and NA scores separately (before and after the experiment) with positive values indicating an increase in affect and negative values indicating a decrease in affect and correlated these with percentage correct scores for the visual and verbal test items in the music-congruent and music-incongruent conditions. Correlations were also performed between happy and sad ratings and percentage correct scores in the music-congruent and music-incongruent conditions. These correlations were performed with a Bonferroni's adjustment to the alpha level, resulting in a new alpha level of $p < .006$. Two Mann-Whitney *U* tests were carried out to compare median happiness and sadness ratings between the experimental groups. Two one-sample Wilcoxon signed-rank tests were also performed on happiness ratings reported in the congruent group and the sadness ratings reported in the incongruent group with the happiness and sadness ratings reported for the selected musical excerpts in Eerola and Vuoskoski (2011). The alpha level for these analyses was set at $p < .05$.

Results

Memory effects

Each participant's score on the visual and verbal test items on 4AFC memory task was converted to a percentage (maximum score of 8 per test item).

Table 1. Performance on the 4AFC Task and Self-Report Measures as a Function of Musical Congruency.

| Measure | Group | | | | | |
|---------------------|--------------------------|-----------|----------------------------|-----------|------------------------------|-----------|
| | Control (<i>n</i> = 19) | | Congruent (<i>n</i> = 17) | | Incongruent (<i>n</i> = 19) | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Verbal score (%) | 63.82 | 21.20 | 62.50 | 17.68 | 66.45 | 14.47 |
| Visual score (%) | 49.34 | 28.10 | 51.47 | 17.61 | 68.42 | 14.05 |
| PA (pretest) | 29.11 | 9.40 | 29.35 | 6.13 | 29.21 | 5.05 |
| PA (posttest) | 23.63 | 7.71 | 36.29 | 7.49 | 30.84 | 5.15 |
| PA change | -5.47 | 8.73 | 6.94 | 6.33 | 1.63 | 6.97 |
| NA (pretest) | 17.00 | 5.39 | 15.41 | 4.37 | 18.58 | 4.05 |
| NA (posttest) | 18.32 | 6.55 | 12.41 | 2.27 | 18.42 | 3.79 |
| NA change | 1.32 | 6.42 | -3.00 | 4.26 | -0.16 | 4.10 |
| | Median | Min-max | Median | Min-max | Median | Min-max |
| Perceived happiness | - | - | 8 | 5-9 | 3 | 1-5 |
| Perceived sadness | - | - | 2 | 1-6 | 7 | 4-9 |

PA: positive affect; NA: negative affect.

The mean percentages of correct response for each item type as a function of group are displayed in Table 1. A 3(group: control, congruent, or incongruent) \times 2 (test item: visual or verbal) mixed measures ANOVA revealed a nonsignificant main effect of group (control: $M = 56.58$, $SD = 25.62$; congruent: $M = 56.99$, $SD = 18.26$; incongruent: $M = 67.43$, $SD = 14.10$), $F(2, 52) = 2.28$, $p = .112$, $\eta_p^2 = .08$, a significant main effect of test item such that recognition memory was better with verbal test items ($M = 64.32$; $SD = 17.75$) than visual ones ($M = 56.59$; $SD = 22.29$), $F(1, 52) = 10.92$, $p = .002$, $\eta_p^2 = .17$, and a significant group \times test item interaction, $F(2, 52) = 4.45$, $p = .016$, $\eta_p^2 = .15$. Simple main effects analyses revealed that while there were group differences in recognition memory for visual test items, $F(2, 104) = 5.39$, $p = .006$, $\eta_p^2 = .09$, there were no group differences for verbal test items, $F(2, 104) = 0.19$, $p = .825$, $\eta_p^2 = .00$. Independent sample *t*-tests showed that while the difference in means between the control and congruent groups was found to be nonsignificant, $t(34) = 0.27$, $p = .790$, $d = .09$, the incongruent group had a significantly higher score for visual test items than both the control $t(36) = 2.65$, $p = .012$, $d = .86$ and the congruent groups $t(34) = 3.21$, $p = .003$, $d = 1.07$; see Table 1.

Self-reported measures of PA

A 3 (group: control, congruent, or incongruent) \times 2 (time: pretest or posttest) mixed-measures ANOVA revealed a significant main effect of group, (control: $M = 26.37$, $SD = 8.92$; congruent: $M = 32.82$, $SD = 7.61$; incongruent: $M = 30.03$, $SD = 5.10$), $F(2, 52) = 5.45$, $p = .007$, $\eta_p^2 = .17$, and a nonsignificant main effect of time, (pretest: $M = 29.22$, $SD = 7.01$; posttest: $M = 30.04$, $SD = 8.50$), $F(1, 52) = 1.06$, $p = .309$, $\eta_p^2 = .02$. There was a significant group \times time interaction, $F(2, 52) = 12.78$, $p = .001$, $\eta_p^2 = .33$. Simple main effects revealed that there were no existing mood differences between the groups prior to watching the movie trailer, $F(2, 104) = 0.01$, $p = .994$, $\eta_p^2 = .00$. There were significant fluctuations in PA in both the control, $F(1, 52) = 10.26$, $p = .002$, $\eta_p^2 = .17$, and congruent groups, $F(1, 52) = 14.76$, $p = .001$, $\eta_p^2 = .22$,

such that self-reported levels of PA decreased significantly in the control group, but increased significantly in the congruent group. There were no significant changes found in PA between pre- versus posttest levels in the incongruent group, $F(1, 52) = 0.91, p = .344, \eta_p^2 = .02$; see Table 1.

Self-reported measures of NA

A 3 (group: control, congruent, or incongruent) \times 2 (time: pretest or posttest) mixed-measures ANOVA revealed a significant main effect of group, (control: $M = 17.66, SD = 5.95$; congruent: $M = 13.91, SD = 3.75$; incongruent: $M = 18.50, SD = 3.87$), $F(2, 52) = 7.23, p = .002, \eta_p^2 = .22$, a nonsignificant main effect of time, (pretest: $M = 17.05, SD = 4.74$; posttest: $M = 16.53, SD = 5.32$), $F(1, 52) = 0.81, p = .374, \eta_p^2 = .02$. There was a significant group \times time interaction, $F(2, 52) = 3.43, p = .004, \eta_p^2 = .12$. Simple main effects revealed that there were no existing mood differences between the groups prior to watching the movie trailer, $F(2, 104) = 2.10, p = .128, \eta_p^2 = .04$. There were significant fluctuations in NA in the congruent group, $F(1, 52) = 5.96, p = .002, \eta_p^2 = .10$, with a significant decrease in NA observed. There were no significant changes found in NA between pre- versus posttest levels in the incongruent, $F(1, 52) = 0.02, p = .893, \eta_p^2 = .00$, or the control groups, $F(1, 52) = 1.28, p = .263, \eta_p^2 = .02$; see Table 1.

Self-reported measures of perceived happiness and sadness

The happy musical excerpt received higher levels of perceived happiness than the sad musical excerpt, $U = 1.50, z = -5.12, p = .001, r = -.85$, while the sad musical excerpt received higher levels of perceived sadness than the happy musical excerpt, $U = 6.50, z = -4.97, p = .001, r = -.83$. A one-sample Wilcoxon signed-rank test indicated that the median level of perceived happiness obtained in the current study for the happy excerpt was comparable to the value ($M = 7.17$) obtained by Eerola and Vuoskoski (2011), $z = 0.64, p = .520, r = .16$. Furthermore, a one-sample Wilcoxon signed-rank test indicated that the median level of perceived sadness obtained in the current study for the sad excerpt was comparable to the value ($M = 7.50$) obtained by Eerola and Vuoskoski (2011), $z = -1.06, p = .290, r = -.24$. Thus, the selected musical excerpts elicited the perceived level of emotion associated with their category to similar levels that have been previously validated in the literature; see Table 1.

Individual differences in changes in affect and perceived happiness and sadness

To investigate whether changes in PA and NA and self-reported perceptions of happiness and sadness were associated in any way with recognition memory we conducted correlations between affect change scores and ratings of happiness and sadness with memory performance separately for visual and verbal test items. Table 2 displays these correlations for the musical congruent and incongruent conditions combined.

Levels of self-reported happiness and sadness correlated significantly with recognition memory performance, but only for visual test items. Higher levels of perceived happiness were associated with lower levels of recognition memory for filmed events, $r_s = -.48, p < .006$, while higher levels of perceived sadness were associated with higher levels of recognition, $r_s = .53, p < .006$. There were no statistically significant correlations for positive and negative measures of affect for either the verbal or visual test items. These results suggest that the perception of

Table 2. Correlations between Self-Report Measures and Performance on the 4AFC Task as a Function of Test Item Combined across the Musical Groups.

| Self-report measures | Memory performance (%) | |
|----------------------------------|------------------------|--------|
| | Verbal | Visual |
| PA change ^a | .09 | -.04 |
| NA change ^a | -.14 | -.01 |
| Perceived happiness ^b | -.20 | -.48* |
| Perceived sadness ^b | .31 | .53** |

PA: positive affect; NA: negative affect.

^aPearson's correlation.

^bSpearman's rho.

* $p < .004$. ** $p < .001$.

happiness and sadness, rather than general changes in mood, provide a closer connection to recognition memory performance for visual scenes accompanied by emotive music.

Discussion

This study aimed to investigate whether music-induced emotions generated a differential impact on visual versus verbal recognition memory items for filmed events by pairing a comedic movie trailer with music that was either congruent (i.e., happy music) or incongruent (i.e., sad music) with the visual content. Although the theoretical framework offered by CAM predicted a stronger visual or working narrative encoded in STM when the comedic trailer was paired with happy music, we found no such performance advantage in the congruent group. Instead, our significant group by test item interaction showed that participants in the incongruent condition displayed a recognition memory advantage for visual items over the congruent and control conditions. These findings support and extend previous research on the variable nature of affect-congruent memory effects and highlight the circumstances in which affect-incongruent effects may emerge in music-based manipulations of this kind (see Boltz et al., 1991; Vuoskoski & Eerola, 2012). Specifically, designs that encourage participants to encode and recall audio-visual pairings in an affective way may elicit stronger mood-congruency effects than designs where such strategies are not as salient (e.g., Boltz, 2004). In the current study, the level of affective cuing was relatively passive as participants watched the video without engaging in an encoding or selective attention activity in any way. While the recognition memory advantage observed in the music-incongruent group was initially surprising, it is important to note that the CAM can accommodate such effects via the expectancies generated in LTM (e.g., Hargreaves, 2012; Schaefer, 2017; Tan, 2017; Wingstedt et al., 2008). For instance, accessing knowledge of musical conventions and film grammar can enhance the distinctiveness of the visual or working narrative in STM when violations occur (e.g., a wedding scene set to slow-paced, sad music).

With regard to changes in self-reported affect, while the null effects observed in the pretest PANAS ratings confirmed that there were no existing mood differences between groups for either PA or NA, comparisons with posttest measures revealed several important fluctuations in affective response as a function of musical congruency. Specifically, a significant increase in PA was observed in participants in the music-congruent group—a finding which could be attributed to the additive effects of pairing the comedic scene with “happy” music (Ellis & Simons, 2005).

For participants in the music-incongruent group, there was no significant change in PA from baseline. While most instances of ironic contrast combine emotionally negative scenes (e.g., sadness, fear, and anger) with emotionally positive music (e.g., happiness; Boltz et al., 1991), the current findings reveal that pairing sad music with positive scenes can just as readily result in the emotional neutralization of a film scene, rendering the viewer's own affective experience unchanged (Rosenfeld & Steffens, 2019). For participants in the control group, there was a significant drop in PA observed. Watching film scenes that are accompanied with long periods of silence are often associated with feelings of unease (e.g., Chion & Gorbman, 2019; Tan et al., 2007). This phenomenon appears to generalize to watching a silent movie trailer under laboratory conditions with participants reporting a reduction in positive mood. In terms of self-reported NA levels, only the congruent group displayed a marked change with levels significantly decreasing from baseline. Thus, music congruency effects as operationalized in the current study by pairing a comedic film scene with "happy" music appears to simultaneously influence levels of self-reported affect in two unique ways: (a) by creating an additive effect in increasing the levels of PA experienced and (b) by creating a suppressive effect in decreasing the levels of NA experienced (see also Damjanovic et al., 2018). The extent to which these potential additive/suppressive mechanisms might contribute to the broader debate of independence versus bipolarity of affect is beyond the scope of the current work (e.g., Diener & Emmons, 1984; Feldman Barrett & Russell, 1998; Tellegen et al., 1999). However, future designs that take into account other dimensional aspects of musical congruency pairings may offer new reconciliatory insights into this important theoretical issue.

While these changes of affect in the music congruent condition would indicate above-threshold activation of the PA node in semantic memory, the subsequent spread of activation within the cognitive system was not sufficiently strong enough to prime recognition memory for the filmed event. Accounting for a meager 0% to 2% of the variance, the weak to non-existent PANAS change scores correlations appear to align with such a theoretical interpretation. In contrast, stronger relationships were observed when the participant was asked to assess their perception of emotion-specific categories within the musical piece. Specifically, higher levels of perceived sadness are associated with higher levels of recognition accounting for 28% of the variance, whereas higher levels of perceived happiness are associated with lower levels of recognition, accounting for 23% of the variance. This pattern of results was only observed with visual test items (e.g., Evans & Schubert, 2008; Gabrielsson, 2002; Juslin & Västfjäll, 2008; Kallinen & Ravaja, 2006; Kivy, 1990; Timmers, 2017). Taken together, these findings suggest an affective component in the integration and representation of audiovisual action that is likely to emerge where a participant perceives or recognizes expressed emotions in music, without necessarily feeling an overall PA or NA (Juslin & Västfjäll, 2008).

The results suggest that perceptions of happiness and sadness in music make different contributions to the processing of filmed events, especially when memory is tested with visual items. According to the levels-of-focus hypothesis, happiness is associated with the preferential processing of global aspects of an image over its local features, whereas sadness is associated with the preferential focus of smaller, local features that make up an overall image (Gasper & Clore, 2002; see also Auer et al., 2012). Part of the recognition test developed for the current study involved the capturing and reformatting of a moving image into smaller static images. This would have likely encouraged a local rather than a global level of image processing, thus potentially contributing to the opposing correlations obtained in the current study—a positive one for sadness, and a negative one for happiness (i.e., global level of focus).

Certain musical characteristics appear to be particularly effective in stimulating vivid imagery in the listener. These include repetition, predictability in melodic, harmonic, and rhythmic

elements, and slow tempo (e.g., Day & Thompson, 2019; McKinney & Tims, 1995). Using content analysis, Vuoskoski and Eerola (2012) showed that the visual imagery category (combined across sad and other imagery) was the most frequently reported response upon listening to experimenter-selected sad music. While the prevalence of visual imagery was not directly measured in the current study, it is possible that the slow tempo of the sad musical piece may have promoted a more concrete visual representation of the trailer during encoding and in turn enhanced the saliency of the visual cues at test. In the context of the current study's findings, this would have likely occurred on the basis of perceiving sadness in the music, rather than feeling an emotional response (e.g., Evans & Schubert, 2008; Gabrielsson, 2002). Future research endeavors that combine memory performance measures with continuous self-report responses of visual imagery for music-induced emotions will help to establish the extent to which imagery plays a role in the encoding and development of the visual narrative in STM (e.g., Damjanovic et al., 2020; Day & Thompson, 2019; Küssner & Eerola, 2019; Paivio, 1969; Presicce & Bailes, 2019; Timmers et al., 2006).

In interpreting these results, some limitations regarding the nature of the audiovisual pairings must be considered. Specifically, while the musical pieces were selected from a validated set for their emotional attributes (Eerola & Vuoskoski, 2011), other aspects of their structural elements such as the musical phrasing, metrical structure, structural harmonic events, melodic and rhythmic patterns, among many other musical components may favor perception and memory. Nevertheless, even with more exhaustive consideration of such surface level characteristics during presentation and encoding does not necessarily mean that the processing of its content after the scene has ended will always follow in a linear and chronological manner (e.g., Tan et al., 2007). On these issues and others reviewed in this article, there is still much to explore and understand with respect to how music interacts with numerous audio and visual elements to shape viewers' experience of dynamic and complex film scenes.

In conclusion, our findings add to the relatively nascent, but increasing, body of work that points toward an affective component in the integration and representation of audiovisual action. We extend this knowledge by proposing that such a component is likely to operate on the basis of emotion-specific perception rather than through changes in the way positive and negative moods are felt. From a practical perspective, we have been able to offer new empirical support to some of the intuitions proposed by film theorists and filmmakers about the effects of the soundtrack on the film audience. Pairing a comedic film scene with sad music renders the audience's response emotionally neutral—a finding consistent with the purported effects of ironic contrast techniques. In addition, we have been able to demonstrate that the ironic contrast effect as tested under laboratory conditions can result in the enhanced recognition of visual test items compared to viewing the same scene under congruent music or no music conditions. Future experimental designs that offer several ironic contrast pairings for music-induced emotions has the potential to discover new and distinct aspects of the multimodal cinematic experience. In film theorist Gorbman's (1987) words: "Change the score on the soundtrack, and the image-track can be transformed" (p. 30).

Data availability statement

Participants of this study did not agree for their data to be shared publicly, so supporting data are not available. Recognition memory test materials can be made available upon reasonable request from the first author.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Ljubica Damjanovic  <https://orcid.org/0000-0003-4971-0399>

References

- Auer, K., Vitouch, O., Koreimann, S., Pesjak, G., Leitner, G., & Hitz, M. (2012, July). When music drives vision: Influences of film music on viewers' eye movements [Paper presentation]. Paper presented at the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, Thessaloniki, Greece.
- Baddeley, A. (1986). *Working memory*. Oxford University Press.
- Barsics, C., & Brédart, S. (2011). Recalling episodic information about personally known faces and voices. *Consciousness and Cognition*, *20*, 303–308. <https://doi.org/10.1016/j.concog.2010.03.008>
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, *58*, 47–73. <https://doi.org/10.1146/annurev.psych.57.102904.190152>
- 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. <https://doi.org/10.1037/h0094102>
- Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, *18*, 427–454. <https://doi.org/10.1525/mp.2001.18.4.427>
- Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory and Cognition*, *32*, 1194–1205. <https://doi.org/10.3758/BF03196892>
- Boltz, M. G. (2017). Facial biases on vocal perception and memory. *Acta Psychologica*, *177*, 54–68. <http://doi.org/10.1016/j.actpsy.2017.04.013>
- Boltz, M. G., Schulkind, M., & Kantra, S. (1991). Effects of background music on the remembering of filmed events. *Memory and Cognition*, *19*, 593–606. <https://doi.org/10.3758/BF03197154>
- Bordwell, D., & Thompson, K. (1979). *Film art: An introduction*. Addison-Wesley.
- Bouhuys, A. L., Bloem, G. M., & Groothuis, T. G. G. (1995). Induction of depressed and elated mood by music influences the perception of facial emotional expressions in healthy subjects. *Journal of Affective Disorders*, *33*, 215–226. [https://doi.org/10.1016/0165-0327\(94\)00092-N](https://doi.org/10.1016/0165-0327(94)00092-N)
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148. <https://doi.org/10.1037/0003-066X.36.2.129>
- Bower, G. H., & Cohen, P. R. (1982). Emotional influences in memory and thinking: Data and theory. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition* (pp. 291–331). Erlbaum.
- Brédart, S., Barsics, C., & Hanley, J. R. (2009). Recalling semantic information about personally known faces and voices. *European Journal of Cognitive Psychology*, *21*, 1013–1021. <https://doi.org/10.1080/09541440802591821>
- Bullerjahn, C., & Guldénring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, *13*, 99–118. <https://doi.org/10.1037/h0094100>
- Carroll, N. (1988). *Mystifying movies: Fads and fallacies in contemporary film*. Columbia University Press.
- Chion, M., & Gorbman, C. (2019). *Audio-vision: Sound on screen*. Columbia University Press.
- Clark, D. M., & Teasdale, J. D. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology*, *48*, 1595–1608. <https://doi.org/10.1037/0022-3514.48.6.1595>
- Cohen, A. J. (2001). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). Oxford University Press.
- 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 University Press.

- Collingnon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., & Lassonde, M. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126–135. <https://doi.org/10.1016/j.brainres.2008.04.023>
- Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception & Performance*, 9, 864–881. <https://doi.org/10.1037//0096-1523.9.6.864>
- Damjanovic, L. (2011). The face advantage in recalling episodic information: Implications for modeling human memory. *Consciousness and Cognition*, 20, 309–311. <https://doi.org/10.1016/j.con-cog.2010.06.002>
- Damjanovic, L., & Hanley, J. R. (2007). Recalling episodic and semantic information about famous faces and voices. *Memory and Cognition*, 35, 1205–1210. <https://doi.org/10.3758/BF03193594>
- Damjanovic, L., Meyer, M., & Sepulveda, F. (2017). Raising the alarm: Individual differences in the perceptual awareness of masked facial expressions. *Brain and Cognition*, 114, 1–10. <https://doi.org/10.1016/j.bandc.2017.01.010>
- Damjanovic, L., Pinkham, A. E., Clarke, P., & Phillips, J. (2014). Enhanced threat detection in experienced riot police officers: Cognitive evidence from the face-in-the-crowd effect. *Quarterly Journal of Experimental Psychology*, 67, 1004–1018. <https://doi.org/10.1080/17470218.2013.839724>
- Damjanovic, L., Roberson, D., Athanasopoulos, P., Kasai, C., & Dyson, M. (2010). Searching for happiness across cultures. *Journal of Cognition and Culture*, 10, 85–107. <https://doi.org/10.1163/156853710X497185>
- Damjanovic, L., & Santiago, J. (2016). Contrasting vertical and horizontal representations of affect in emotional visual search. *Psychonomic Bulletin & Review*, 23, 62–73. <https://doi.org/10.3758/s13423-015-0884-6>
- Damjanovic, L., Wilkinson, H., & Lloyd, J. (2018). Sweet emotion: The role of odor-induced context in the search advantage for happy facial expressions. *Chemical Senses*, 43, 139–150. <https://doi.org/10.1093/chemse/bjx081>
- Damjanovic, L., Williot, A., & Blanchette, I. (2020). Is it dangerous? The role of an emotional visual search strategy and threat-relevant training in the detection of guns and knives. *British Journal of Psychology*, 111, 275–296. <https://doi.org/10.1111/bjop.12404>
- Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology*, 29, 75–89. <https://doi.org/10.1037/pmu0000220>
- De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and eye. *Cognition & Emotion*, 14, 289–311. <https://doi.org/10.1080/026999300378824>
- Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47, 1105–1117. <https://doi.org/10.1037//0022-3514.47.5.1105>
- Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39, 18–49. <https://doi.org/10.1177/0305735610362821>
- Eisenstein, S. M. (1949). *Film form* (J. Leyda, Ed. & Trans.). Harcourt, Brace & Jovanovich.
- Ellis, R. J., & Simons, R. F. (2005). The impact of music on subjective and physiological indices of emotion while viewing films. *Psychomusicology*, 19, 15–40. <https://doi.org/10.1037/h0094042>
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royden, L., Saur, R., Reiter, S., Grodd, W., & Wildgruber, D. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI Study. *Human Brain Mapping*, 27, 707–714. <https://doi.org/10.1002/hbm.20212>
- Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, 12, 75–99. <https://doi.org/10.1177/102986490801200105>
- Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74, 967–984. <https://doi.org/10.1037/0022-3514.74.4.967>
- Finsterwalder, J., Kuppelwieser, V. G., & de Villiers, M. (2012). The effects of film trailers on shaping consumer expectations in the entertainment industry: A qualitative analysis. *Journal of Retailing and Consumer Services*, 19, 589–595.
- Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, 5, 123–147. <https://doi.org/10.1177/10298649020050S105>

- Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science, 13*, 34–40. <https://doi.org/10.1111/1467-9280.00406>
- Giannetti, L. D. (1982). *Understanding movies* (3rd ed.). Prentice-Hall.
- Gorbman, C. (1987). *Unheard melodies: Narrative film music*. Indiana University Press.
- Hanley, J. R., & Damjanovic, L. (2009). It is more difficult to retrieve a familiar person's name and occupation from their voice than from their blurred face. *Memory, 17*, 830–839. <https://doi.org/10.1080/09658210903264175>
- Hansen, C. H., & Hansen, R. D. (1988). How rock music videos can change what is seen when boy meets girl: Priming stereotypic appraisal of social interactions. *Sex Roles, 19*, 287–316. <https://doi.org/10.1007/BF00289839>
- Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music, 40*, 539–557. <https://doi.org/10.1177/0305735612444893>
- Herget, A. (2021). On music's potential to convey meaning in film: A systematic review of empirical evidence. *Psychology of Music, 49*, 21–49. <https://doi.org/10.1177/0305735619835019>
- Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48*, 246–269. <https://doi.org/10.2307/1415746>
- 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, 5*, 146–153. <https://doi.org/10.1037/a0021544>
- Holmes, N. P., Calvert, G. A., & Spence, C. (2009). Multimodal integration. In M. D. Binder, N. Hirokawa & U. Windhorst (Eds.), *Encyclopedia of neuroscience* (pp. 2457–2461). Springer. https://doi.org/10.1007/978-3-540-29678-2_3640
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559–575. <https://doi.org/10.1017/S0140525X08005293>
- Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae, 10*, 191–213. <https://doi.org/10.1177/102986490601000203>
- Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Cornell University Press.
- Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology, 29*, 90–99. <https://doi.org/10.1037/pmu0000238>
- Levi, D. S. (1982). The structural determinants of melodic expressive properties. *Journal of Phenomenological Psychology, 13*, 19–40.
- Levy, S., McNulty, T., Roberts, M., Dana, P. J. (Producers), & Blitz, J. (Director). (2017). *Table 19* [Motion picture]. Fox Searchlight Pictures.
- Marshall, S., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception, 6*, 95–112. <https://doi.org/10.2307/40285417>
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748. <https://doi.org/10.1038/264746a0>
- McKinney, C. H., & Tims, F. C. (1995). Differential effects of selected classical music on the imagery of high versus low imagers: Two studies. *Journal of Music Therapy, 32*, 22–45. <https://doi.org/10.1093/jmt/32.1.22>
- Montepare, J. M., Goldstein, S. B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior, 11*, 33–42. <https://doi.org/10.1007/BF00999605>
- Nosal, A. P., Keenan, E. A., Hastings, P. A., & Gneezy, A. (2016). The effect of background music in shark documentaries on viewers' perceptions of sharks. *PLOS ONE, 11*, Article e0159279. <https://doi.org/10.1371/journal.pone.0159279>
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*, 241–263. <https://doi.org/10.1037/h0027272>
- Pan, F., Zhang, L., Ou, Y., & Zhang, X. (2019). The audio-visual integration effect on music emotion: Behavioral and physiological evidence. *PLOS ONE, 14*, Article e0217040. <https://doi.org/10.1371/journal.pone.0217040>

- Parfitt, D., Gigliotti, D., Weinstein, H., Zwick, E., Norman, M. (Producers), & Madden, J. (Director). (1998). *Shakespeare in Love* [Motion picture]. Miramax Films.
- Presicce, G., & Bailes, F. (2019). Engagement and visual imagery in music listening: An exploratory study. *Psychomusicology*, 29, 136–155. <https://doi.org/10.1037/pmu0000243>
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–154. <https://doi.org/10.1080/713752551>
- Rigg, M. G. (1964). The mood effects of music: A comparison of data from four investigators. *Journal of Psychology*, 58, 427–438. <https://doi.org/10.1080/00223980.1964.9916765>
- Rosenfeld, N., & Steffens, J. (2019). Effects of audiovisual congruency on perceived emotions in film. *Psychomusicology*, 29, 200–208. <https://doi.org/10.1037/pmu0000242>
- Schaefer, R. S. (2017). Music in the brain: Imagery and memory. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* [Routledge Music Companions] (pp. 25–25). Taylor & Francis.
- Scherer, K. (1979). Acoustical concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.), *Nonverbal communication* (pp. 249–253). Oxford University Press.
- Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52, 1061–1086. <https://doi.org/10.1037//0022-3514.52.6.1061>
- Steffens, J. (2020). The influence of film music on moral judgments of movie scenes and felt emotions. *Psychology of Music*, 48, 3–17. <https://doi.org/10.1177/0305735618779443>
- Strobin, A. A., Hunt, J. B., Spencer, F. J., & Hunt, T. G. (2015). The role of music in motion picture advertising and theatrical trailers: Altering music to modify emotional response and genre expectations. *Academy of Marketing Studies Journal*, 19, R244–R260.
- Tan, S. L. (2017). Scene and heard: The role of music in shaping interpretations of film. Multimodal integration. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* [Routledge Music Companions] (pp. 363–377). Taylor & Francis.
- Tan, S. L., Spackman, M. P., & Bezdek, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception*, 25, 135–152. <https://doi.org/10.1525/MP.2007.25.2.135>
- Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, 10, 297–303. <https://doi.org/10.1111/1467-9280.00157>
- Timmers, R. (2017). Emotion in music listening. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* [Routledge Music Companions] (pp. 489–501). Taylor & Francis.
- Timmers, R., Marolt, M., Camurri, A., & Volpe, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: An explorative case study. *Psychology of Music*, 34, 481–510. <https://doi.org/10.1177/0305735606067165>
- Vuoskoski, J. K., & Eerola, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 204–213. <https://doi.org/10.1037/a0026937>
- Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception and Psychophysics*, 76, 591–604. <https://doi.org/10.3758/s13414-013-0582-2>
- Warbeck, S. (1998). *The Beginning of the Partnership. On Shakespeare in love* [CD]. Sony Classical.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Weijkamp, J., & Sadakata, M. (2017). Attention to affective audio-visual information: Comparison between musicians and non-musicians. *Psychology of Music*, 45, 204–215. <https://doi.org/10.1177/0305735616654216>
- Wingstedt, J., Brändström, S., & Berg, J. (2008). Narrative music, visuals and meaning in film. *Visual Communication*, 9, 193–210. <https://doi.org/10.1177/1470357210369886>
- Yared, G. (1996). *Black Nights. On The English Patient* [CD]. Fantasy Records.
- Zaentz, S. (Producer), & Minghella, A. (Director). (1996). *The English Patient* [Motion picture]. Miramax Films.