Psychological Topics, 32 (2023), 1, 57-80
Original Scientific Paper UDC: 159.9:78
https://doi.org/10.31820/pt.32.1.4 159.942.072

The Influence of the Moving Image on Music-Induced Emotions

Elias Gram-Nilsen[1] and Jonna K. Vuoskoski[1,2,3]

[1] University of Oslo, Department of Musicology, Norway
[2] University of Oslo, Department of Psychology, Norway
[3] RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Norway

Abstract

The combination of music and the moving image is prevalent in our society, occurring in visual art forms and media such as movies, music videos, and commercials. The relationship between the two has received much interest in existing research, but the focus has mainly been on the influence that music may have on the experience of the moving image. The present study adopts the reverse approach, and examines the potential impact the moving image may have on emotions felt and perceived while listening to music. Sixty-six participants were presented with three music excerpts that were introduced either alone or paired with one of two videos aimed to elicit strong emotions (either awe or being moved). The music excerpts were played a second time (without visual accompaniment) to investigate whether a possible influence of previously presented visual information was persistent. The results revealed that the moving image did have an intensifying effect on emotions induced (both awe and being moved) while listening to music. A significant negative effect on perceived valence and enjoyment was found for the least congruent music & video pairing. Interestingly, this incongruent pairing had the most persistent influence, being the only significant carry-over effect during the second presentation of music alone.

Keywords: music, moving image, emotion, awe, being moved, cross-modal effects

Introduction

Visual information plays an important role in music listening, regardless of whether it is directly related to the source of the music (e.g., live performances or concert videos), or merely accompanying the music (e.g., movies, music videos, or commercials). Importantly, experience-sampling studies have estimated that approximately 10% of emotional episodes with music occur while watching TV or movies (Juslin et al., 2008). Despite the prevalence of multimodal musical experiences, little is known about how visual information affects music-induced

✉ Jonna K. Vuoskoski, RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion, P.O. Box 1133, 0318 Oslo, Norway. E-mail: *j.k.vuoskoski@imv.uio.no*

emotions – and whether these effects might carry over to subsequent encounters with music in isolation.

Research on multisensory perception has demonstrated that visual and auditory inputs are integrated to form a unified experience, and interactions between these modalities may in some cases occur (Lalanne & Lorenceau, 2004). Multisensory integration is facilitated by features like the 'unity assumption' (i.e., a belief that two or more unisensory cues belong together; Chen & Spence, 2017) and semantic congruency (Doehrmann & Naumer, 2008). In the context of musical performance, bodily (Krahé et al., 2015; Vines et al., 2006, 2011; Vuoskoski et al., 2013, 2015) and facial gestures of the performer (Livingstone et al., 2015; Thompson et al., 2005, 2008) have been found to alter the perception of musical features, and to modulate perceived and felt emotions.

Possible mechanisms through which the multimodal experience of music may have carry-over effects to subsequent experiences of the music (in isolation) include episodic memories and evaluative or emotional conditioning (e.g., Juslin & Västfjäll, 2008). In their theoretical account of different psychological mechanisms underlying music's emotional effects, Juslin and Västfjäll (2008) propose that music may sometimes induce particular emotions because it has previously been experienced in association with other emotion-inducing stimuli or events. However, while episodic memories are consciously accessible, evaluative conditioning involves unconscious and effortless processes (e.g., Juslin & Västfjäll, 2008). Although studies on evaluative and emotional conditioning typically show stronger effects after repeated exposure to pairs of conditioned and unconditioned stimuli (see e.g., Hofmann et al., 2010), single-exposure paradigms have also demonstrated significant effects (e.g., Forester et al., 2020; Vermeulen & Beukeboom, 2016).

Awe and *being moved/touched* are often considered as the central 'aesthetic' emotions evoked by music (e.g., Cotter et al., 2018; Konecni, 2005). Although these emotions are sometimes reported to co-occur (e.g., Cotter et al., 2018), the eliciting appraisals and theoretical frameworks associated with awe and being moved/touched have notable differences. Recently, the experience commonly labelled as 'being moved or touched' has been conceptualized as the subjective feeling component of a social-relational (rather than aesthetic) emotion that is evoked by situations involving interpersonal closeness and/or moral/prosocial acts (Fiske et al., 2019; Seibt et al., 2017). This emotion has been termed *kama muta* to avoid equating the emotion construct with one of its colloquial labels (i.e., being moved; Fiske et al., 2019). Also, in the context of music listening, experiences of being moved/*kama muta* are associated with increased feelings of connectedness (Vuoskoski et al., 2022; Swarbrick et al., 2021). Awe, in contrast, is associated with perceived vastness (i.e., perceiving something to be larger or grander than oneself), and is characterized by a need for cognitive accommodation (Keltner & Haidt, 2003). While experiences of music-induced *kama muta* appear to be overwhelmingly positive, experiences of awe tend to be more mixed (e.g., Swarbrick & Vuoskoski, submitted).

In the context of film and cinema, a number of studies have investigated how music might modulate the interpretation of (and emotional reactions to) the moving image (for example, Bolivar et al., 1994; Boltz, 2001; Bullerjahn & Güldenring, 1994; Cohen, 2001). This line of research has shown that music can bias the understanding of the plot, intention and motivation of characters, as well as the perception of mood in the moving image. In contrast, how the moving image might affect the experience of music, has only been explored to a limited extent (however, see Bannister & Eerola, 2021; Boltz et al., 2009; Eldar et al., 2007; Geringer et al., 1996, 1997; Goldberg et al., 1993).

A small number of previous studies have nevertheless investigated the potential effect of the moving image on emotional responses to music. Geringer and colleagues (1996, 1997) found that, compared to music alone, the combination of music and video led to increases in involvement, emotional reactions and preference, but these effects were small. Focusing on perceived emotions and potential memory biases, Boltz and colleagues (2009) demonstrated that the perceived affective character of neutral music was influenced by the accompanying videos in an affect-congruent manner. Boltz and colleagues also found that, when subsequently presented with the original as well as two manipulated versions (either more positive or negative) of each musical stimulus, participants' recognition memory was biased by the previously presented video accompaniment in a mood-congruent fashion.

Most relevantly for the present study, Bannister and Eerola (2021) demonstrated that the moving image can have an influence on the emotions and physiological responses induced by music. In their study, Bannister and Eerola distinguished music-induced chills into two categories based on possible underlying mechanisms. One type, labelled *vigilance chills,* was hypothesised to be associated with the experience of awe, while the other type, *social chills*, was predicted to be associated with being moved/*kama muta*. Animations presented with music, as well as extra-musical information given before listening, were employed to trigger one of the two distinct types of chills. For *vigilance chills*, stimuli (animations and extra-musical information) that emphasised the structure of the music was used (vigilance condition), while for *social chills*, the stimuli contained bittersweet and moving narratives (social condition). As hypothesised, the vigilance condition was associated with heightened experience of awe, while the social condition was associated with heightened feelings of being moved/kama muta.

Similar findings have also been provided by studies that have investigated the impact of extramusical information, such as written narratives rather than visual information, on the experience of music. As the moving image is not only a visual experience but also a narrative one, these studies have some relevance for the present study. Lavy (2001) found that reading a short story while listening to music had an influence on the perceived valence of music. Similarly, Vuoskoski and Eerola (2015) demonstrated that narrative descriptions of the music helped to intensify the emotions that were induced during music listening. Before listening to an excerpt of sad-

sounding music, participants were given descriptions of either a sad narrative of an encounter in a holocaust camp, or a neutral narrative related to nature in Yellowstone national park. These conditions were compared to music listening without any description, and the induced emotions were measured indirectly through cognitive biases in memory and judgment. The results revealed that the sad narrative seemed to enhance the emotional effect of the music, leading to a stronger memory bias.

The Present Study

The aim of the current study was to investigate the possible effect of moving image on the emotions induced and perceived in the context of music listening. Furthermore, the aim was to explore whether these effects might carry over to a subsequent presentation of the music in isolation. While it has been theorized that extra-musical sources of emotion can influence the emotions induced by music listening (e.g., via the episodic memory and evaluative conditioning mechanisms; Juslin & Västfjäll, 2008), it is currently not well understood whether a single prior exposure in combination with an emotion-inducing video could have a significant effect on the emotions induced by a piece of music during subsequent encounters (in isolation). This would have significant implications for understanding our musical experiences in everyday life, where music is often experienced in the context of audio-visual media.

Participants were randomly assigned to one of three conditions, where three music excerpts were presented either alone or paired with one of two videos expected to elicit either awe or being moved/kama muta. Rather than operating with one control condition and two experimental conditions, different audio-visual (and audio-only) combinations were balanced across the three conditions. In each of the conditions, one of the three music excerpts was presented alone, while the two remaining pieces were paired with the two videos. In addition to this, participants were randomly presented with one of the two videos alone (randomization occurred within each condition). The presentation of the video alone was done to account for responses to the moving image alone.

The selection of videos was based on their predicted ability to induce either *awe* or *being moved*, as these emotions may be more effectively and intensely induced by music compared to basic emotions (e.g., Konecni, 2005). For instance, these emotions have been found to be associated with physiological responses like chills, warm sensations and tears (Bannister & Eerola, 2021; Braud, 2001; Konecni, 2005; Menninghaus et al., 2015; Schubert et al., 2018; Schurtz et al., 2012; Wassiliwizky et al., 2015). Felt emotions tend to be associated with the enjoyment aspect of music listening (Thompson, 2006; Vuoskoski et al., 2011), and the latter was, therefore, also investigated in the present study. Feelings of being moved have, for instance, been found to mediate the enjoyment of sad music, resulting in what is often characterized as pleasurable sadness (Vuoskoski & Eerola, 2017). The moving image

was hypothesised to have a positive influence on enjoyment, potentially by intensifying induced emotions.

In order to explore a potential carry-over effect of the emotions induced by the videos on the experience of music, the music excerpts were presented (without video) for a second time after a filler task. In line with predictions arising from the episodic memory and evaluative/emotional conditioning mechanisms proposed by Juslin and Västfjäll (2008), it was hypothesised that the initial video pairing would have a lasting effect on emotional responses to music. Finally, the present research also set out to investigate whether congruency between the music and the moving image would modulate the effects of the moving image. Since semantic congruency and 'unity assumption' have previously been found to facilitate multisensory integration and the magnitude of cross-modal effects (e.g., Doehrmann & Naumer, 2008; Chen & Spence, 2017), it was hypothesised that the effect of the moving image on music-evoked emotions would be greater when the two were perceived to be congruent rather than incongruent.

Method

Participants

Sixty-six participants between the ages of 20 and 67 (M = 30.8, SD = 11.56) took part in an online experiment. Originally, 68 participants were recruited, but two participants were removed from the data due to empty responses to several questions. Thirty-three (50%) participants were females, 32 were males and 1 was non-binary. Participants were mainly recruited by distributing a link to the experiment on social media (Facebook and Twitter). The study included participants from numerous countries, but the majority of the sample came from Norway (48). The other nationalities that were represented in this study were Denmark (7), Sweden (3), United Kingdom (3), France (1), USA (1), Mexico (1), New Zealand (1), and Japan (1). Regarding musical background, 21 participants described themselves as musicians, 16 participants reported playing an instrument at a competent level, 14 participants reported playing an instrument at an amateur level, 14 participants chose the option "*I like music a lot*", while one participant chose the option "*I think music is okay*". Twenty-one participants were assigned to the first condition, 20 to the second condition, and 25 to the third condition.

Stimulus Materials

There were three musical stimuli, and each had the duration of approximately 1 minute and 20 seconds (see https://osf.io/2qcps/ for the stimuli). The stimuli were extracts from a slowed down version (800%) of Section 1 from *Music for 18 Musicians* by Steve Reich (this excerpt was not manipulated for the experiment, but

the version was found on YouTube), the second movement of *Company* by Philip Glass, and *Dream* by Ian Post. In the selection of music, vocal music with lyrics was excluded due to the possible impact the content of the lyrics could have on emotional reactions, and well-known music was avoided due to the possibility of it evoking emotions through episodic memories, or other idiosyncratic associations. The excerpt from *Music for 18 Musicians* has an ambient vibe (and will for that reason be labelled as the "ambient track" further in this paper) and this was hypothesised to be perceived as ambiguous, in particular in relation to valence. *Company* by Philip Glass is a classical piece for the string quartet ("string quartet"), and it was predicted to evoke awe, and to be perceived as negative in valence and high in arousal. *Dream* by Ian Post is a pop tune with a positive, energetic feeling ("pop tune"), and it was hypothesised to facilitate being moved, and to be perceived as conveying positive valence and high arousal.

The two videos that were chosen for the study were a video depicting a person surfing on a large wave, and a video of the reunion between a lion called Christian and his former owners (see https://osf.io/2qcps/ for the stimuli). Each video lasted, as the music excerpts, for approximately 1 minute and 20 seconds. The selection of the videos was grounded in their expected ability to induce strong emotions (awe and being moved) and physiological responses in participants. The "surfing video" was hypothesised to evoke awe, while the "Christian the lion video" was hypothesised to elicit feelings of being moved. These hypotheses were based on theoretical predictions (see e.g., Keltner & Haidt, 2003; Fiske et al., 2019; Menninghaus et al., 2015), as well as the use of "Christian the lion video" in previous research to evoke being moved/kama muta (Zickfeld et al., 2017). In relation to perceived emotions, the "surfing video" was predicted to be rated on the negative side of the valence dimension, due to the dramatic content, and on the higher side of the arousal dimension. The "Christian the lion video" was predicted to be given high ratings of both valence and arousal.

All the videos were edited and paired with music in iMovie. Sounds coming from the natural environment (diegetic sounds) of the videos were erased and they were shown with no sound when presented alone. This was done to fully isolate the moving image variable from the music variable and prevent any unwanted interactions from occurring. The selection of videos and music was guided by both structural and emotional aspects. The "string quartet" and "surfing video" were affectively congruent in terms of evoking awe, while the "pop tune" and "Christian the lion video" were affectively congruent in terms of evoking being moved/kama muta. Furthermore, the expected peaks in the music and the predicted climaxes in the videos were aligned within these pairings, promoting structural as well as affective congruence. The "ambient track" was considered ambiguous, and was structurally aligned with both videos. The emotional contrast between music and video was expected to be the strongest in the pairing of "string quartet" and "Christian the lion video", where the expected peaks were also unaligned.

Procedure

The experiment was conducted online on the Qualtrics platform. The experiment took approximately 15 minutes to complete, and participants had the opportunity of completing the survey in either Norwegian or English. Informed consent was obtained from all participants, and those who completed the experiment were given the opportunity to take part in a lottery to win two movie tickets. To reduce unwanted influences from surroundings and to ensure better sound quality, participants were asked to use headphones throughout the survey (however, adherence to this request was not checked).

Participants were randomly assigned to one of three conditions (see Table 1 for a detailed description of the stimuli included in different conditions), and completed the experimental blocks in the following order:

1. Video without music (one of two elements randomly presented; 31 of the participants were shown the "surfing video", while 35 participants were shown the "Christian the lion video")
2. The three musical excerpts in random order (two of the excerpts were accompanied by videos)
3. Filler task (a question taken from an IQ test; see Appendix)
4. The same musical excerpts without visuals, presented in a random order.

Table 1

Stimuli Presented In Different Conditions

Details	**Condition 1**	**Condition 2**	**Condition 3**
One of two videos randomly presented	Video 1 or Video 2	Video 1 or Video 2	Video 1 or Video 2
Stimuli presented in a randomized order	Music 1 & Video 1 Music 2 & Video 2 Music 3 / No Video	Music 1 & Video 2 Music 2/ No Video Music 3 & Video 1	Music 1/ No Video Music 2 & Video 1 Music 3 & Video 2
IQ question	Filler task	Filler task	Filler task
Stimuli presented in a randomized order	Music 1 Music 2 Music 3	Music 1 Music 2 Music 3	Music 1 Music 2 Music 3

Note. Video 1 = "surfing video"; Video 2 = "Christian the lion video"; Music 1 = ambient track; Music 2 = string quartet; Music 3 = pop tune.

The purpose of the filler task was to take the focus away from the stimuli that participants had recently been exposed to, and to prevent the repetition of the music from being experienced too soon after the first presentation.

After being presented with a stimulus (either video alone, music alone, or music & video together), participants were requested to rate how much they enjoyed the music (only music alone and music & video), perceived fit/congruency between

music and visuals (only music & video), the occurrence of chills, the experience of awe and being moved, and perceived valence and perceived arousal. Note that when audio-visual stimuli were presented, the questions related to enjoyment, perceived valence and perceived arousal were always directed at the music, while the questions related to induced emotions were focused on the subjective emotion experienced (see the Appendix for exact questions and scales). All ratings were collected with slider scales (output range: 0-100), as they provide numeric data that can be statistically analysed. There was also one open question about other potential physiological reactions and sensations that participants may have experienced. After the presentation of all the stimuli, participants were asked to answer demographic questions related to their age, gender, nationality, musical background, as well as familiarity with the stimuli. In the end they were also asked to report if they experienced any problems with playing the excerpts, and they were given the opportunity to give general feedback on the study.

Results

The data were imported to RStudio for analysis. Since the main interest of the study lay in the cross-modal effects of music and video that are contingent on the specific audio-visual pairings, separate one-way ANOVAs were carried out for each music excerpt. The following dependent variables were analysed: enjoyment, occurrence of chills, awe, being moved, perceived valence, and perceived arousal. Presentation condition (i.e., audio-visual pairing with the "surfing video", "Christian the lion video", or no video) was the independent variable. In addition, independent t-tests were carried out to investigate the congruency between visuals and music, and to compare the two videos presented alone. No statistical analysis was conducted on the open question related to physiological reactions, since the number of responses from each condition was much lower than for the slider ratings. Responses were nevertheless systematically mapped and responses from different stimulus categories compared with each other.

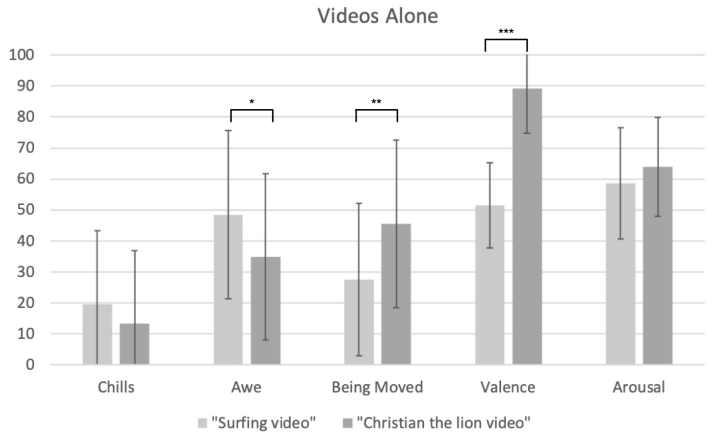
Videos Alone

The mean ratings (and standard deviations) for both videos are displayed in Figure 1. As predicted, the ratings of experienced awe were higher for the "surfing video", compared to the "Christian the lion video"; $t(62.75) = 2.044$, $p = .04$, $d = 0.49$ (small effect). In contrast, "Christian the lion video" had significantly higher ratings of being moved compared to the "surfing video"; $t(63.95) = 2.821$, $p = .006$, $d = 0.65$ (medium effect). The largest effect was found for perceived valence; $t(63.66) = 10.857$, $p < .001$, $d = 1.60$ (large effect), with "Christian the lion video" being given substantially higher ratings than the "surfing video". No significant differences were

found between the videos for either chills; $t(62.89) = 1.059$, $p = .29$, or perceived arousal; $t(60.17) = 1.266$, $p = .21$.

Figure 1

Mean Ratings for the Two Videos in the Video Alone Rating Condition



Note. The error bars represent standard deviations. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

First Presentation of Music and Audio-Visual Stimuli

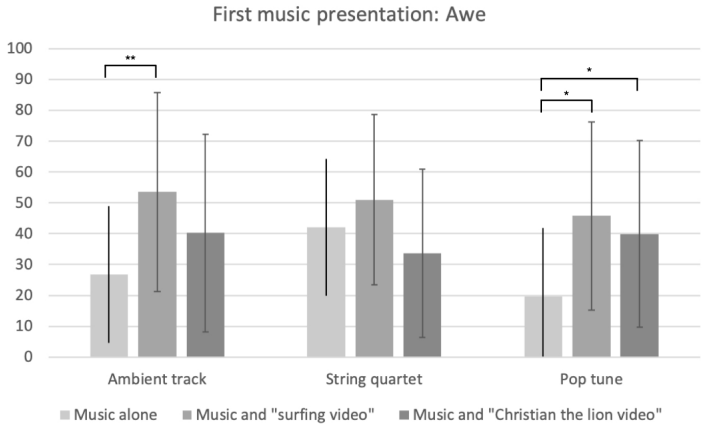
To investigate the potential cross-modal effects of the different audio-visual pairings compared to music alone (i.e., Presentation condition), a series of one-way ANOVAs were carried out for each music excerpt.

Awe

For the experience of awe (see Figure 2), a significant effect of Presentation condition was found for both the "ambient track"; $F(2,63) = 4.995$, $p = .009$, $\eta^2 = .14$, and the "pop tune"; $F(2,63) = 4.877$, $p = .01$, $\eta^2 = .13$. No significant effect was found for the "string quartet" excerpt. A post-hoc analysis (Tukey HSD) revealed that for the ambient track, the music in combination with the "surfing video" led to significantly higher ratings of experienced awe compared to the music alone, $p = .007$, $d = 0.89$ (large effect). For the "pop tune", pairing the music with both the "surfing video" and the "Christian the lion video" led to significantly higher awe ratings compared to the music alone; $p = .01$, $d = 0.87$ (large effect) and $p = .05$, $d = 0.70$ (medium effect), respectively.

Figure 2

Mean Ratings of Experienced Awe in Response to the Three Musical Stimuli either Alone or In Combination with the Two Videos



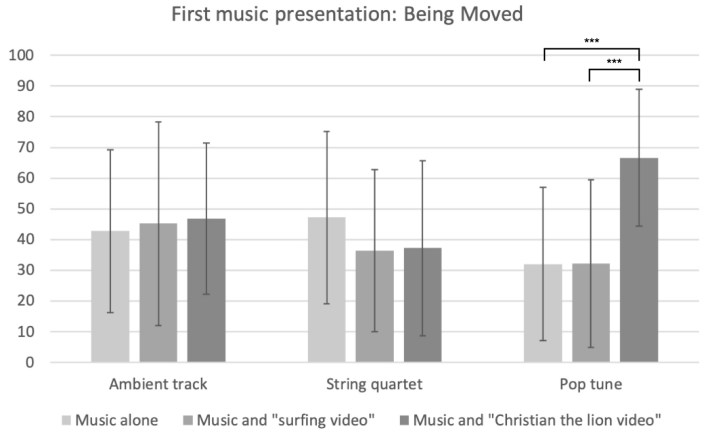
Note. The error bars represent standard deviations. $^{*}p < .05$, $^{**}p < .01$.

Being Moved

In relation to ratings of being moved (see Figure 3), a significant main effect of Presentation condition was only found for the "pop tune"; $F(2,63) = 15.14$, $p < .001$, $\eta^2 = .32$. Post-hoc comparisons revealed that when the music was paired with the "Christian the lion video", participants experienced significantly stronger feelings of being moved compared to both music alone; $p < .001$, $d = 1.19$ (large effect), and music paired with the "surfing video"; $p < .001$, $d = 1.15$ (large effect).

Figure 3

Mean Ratings of Being Moved In Response to the Three Musical Stimuli Either Alone or In Combination with the Two Videos



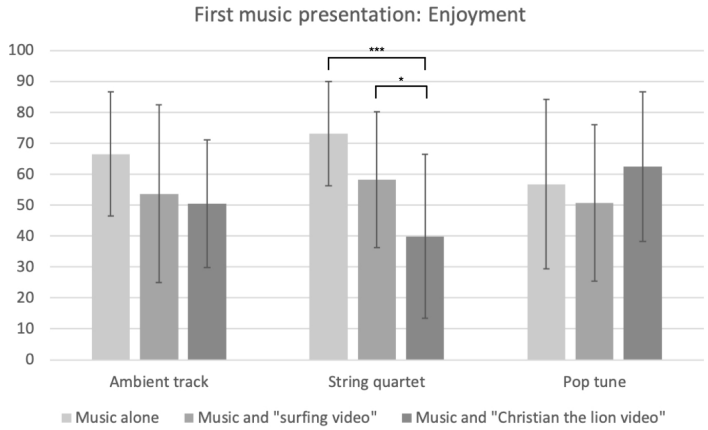
Note. The error bars represent standard deviations. $^{***}p < .001$.

Enjoyment & Chills

For the enjoyment ratings (see Figure 4), a significant main effect of Presentation condition was found only for the "string quartet" excerpt; $F(2,63) = 11.57$, $p < .001$, $\eta^2 = .27$. Post-hoc tests revealed that, compared to music alone, pairing the music with the "Christian the lion video" had a negative impact on the enjoyment of the music; $p < .001$, $d = 1.19$ (large effect). Pairing the "string quartet" excerpt with the "Christian the lion video" also led to significantly lower ratings of music enjoyment compared to the "surfing video"; $p = .02$, $d = 0.71$ (medium effect). No significant effect of Presentation condition was found for the chills variable.

Figure 4

Mean Ratings of Enjoyment in Response to the Three Musical Stimuli either Alone or In Combination with the Two Videos



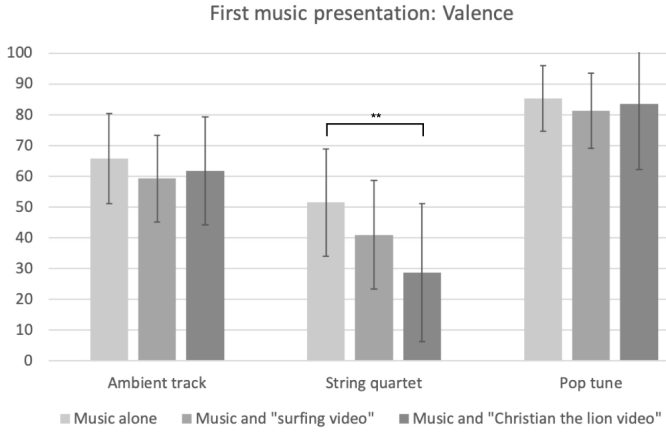
Note. The error bars represent standard deviations. $^{*}p < .05$, $^{***}p < .001$.

Perceived Valence & Arousal

With regard to the ratings of perceived valence in the music (see Figure 5), a significant main effect of Presentation condition was found only for the "string quartet" excerpt; $F(2,63) = 7.159$, $p = .001$, $\eta^2 = .18$. Similar to the enjoyment ratings, the "Christian the lion video" had a significant negative (mood-incongruent) effect on valence ratings when compared to the music alone; $p = .001$, $d = 0.99$ (large effect).

Figure 5

Mean Ratings of Valence Perceived In the Three Musical Stimuli, Presented either Alone or In Combination with the Two Videos

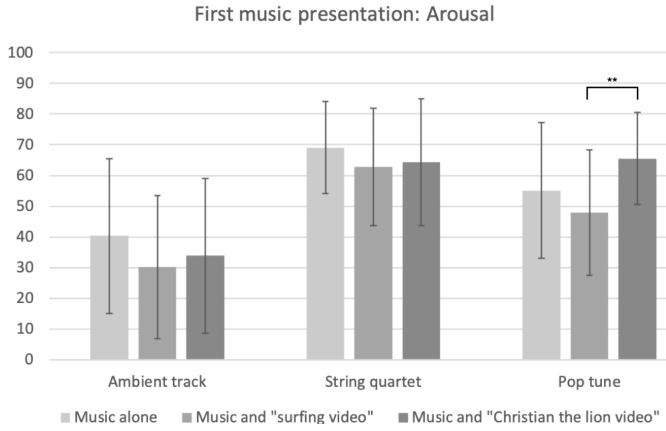


Note. The error bars represent standard deviations. $^{**}p < .01$.

A significant effect of Presentation condition on ratings of perceived arousal was found only for the "pop tune"; $F(2,63) = 4.836$), $p = .01$, $\eta^2 = .13$. This effect reflected a significant difference between the "surfing video" and the "Christian the lion video" when accompanying the "pop tune", $p = .008$, $d = 0.90$ (large effect), where the combination with "Christian the lion video" was rated as more arousing. Although the videos alone were rated as similar in terms of perceived arousal, this significant difference could be related to the considerably stronger feelings of being moved evoked by the "Christian the lion video" (compared to the "surfing video") in combination with the "pop tune". The mean arousal ratings are displayed in Figure 6.

Figure 6

Mean Ratings of Arousal Perceived In the Three Musical Stimuli, Presented either Alone or In Combination with the Two Videos



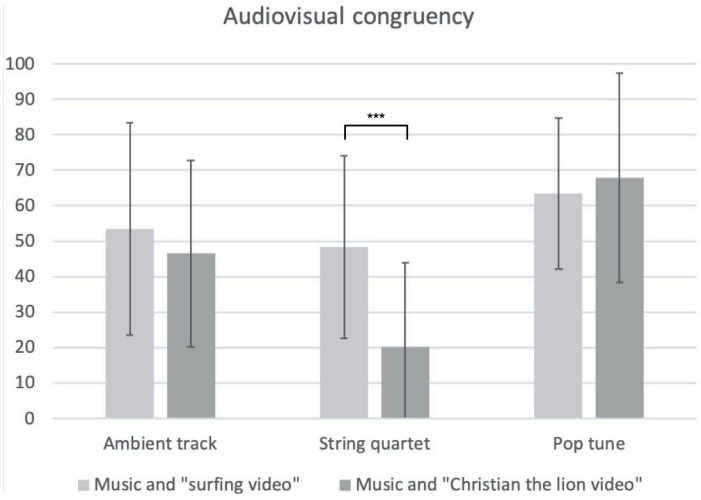
Note. The error bars represent standard deviations. $^{**}p < .01$.

Perceived Fit/Congruency

The significant effects found in the enjoyment and valence ratings for the "string quartet" excerpt were also reflected in the ratings of perceived fit/congruency, as the pairing of the "string quartet" excerpt and the "Christian the lion video" was given the lowest mean congruency rating of all the video pairings (see Figure 7), being significantly lower than the pairing of the "string quartet" and the "surfing video"; $t(43.55) = 3.84$, $p < .001$, $d = 0.99$ (large effect). This suggests that the aforementioned negative effects on enjoyment and perceived valence may be related to a perceived incongruence between the music and visuals.

Figure 7

Mean Ratings of Perceived Fit/Congruency between the Three Musical Stimuli and the Two Videos



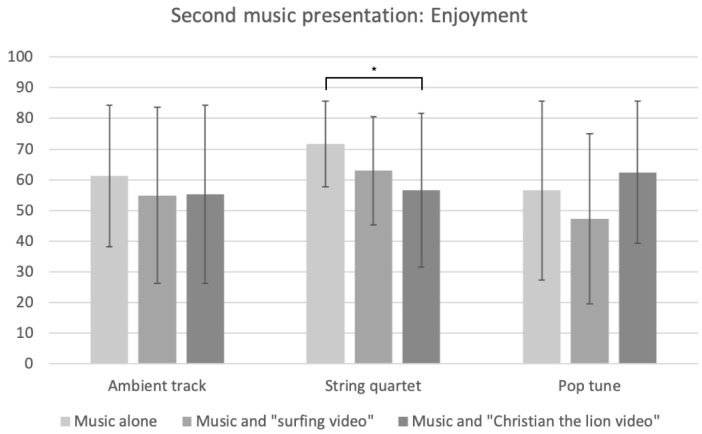
Note. The error bars represent standard deviations. $^{***}p < .001$.

Second Presentation of Music Excerpts

In order to investigate whether the video accompaniment (compared to no video) during the first presentation of the music excerpts had any carry-over effects to the second presentation (where music excerpts were presented alone), one-way ANOVAs with First presentation condition as the independent variable were carried out. For music enjoyment ratings (see Figure 7), a significant main effect of First presentation condition was found for the "string quartet"; $F(2,63) = 3.071$, $p = .05$, $\eta^2 = .09$. Post-hoc analyses revealed that, compared to hearing music alone during the first presentation, having previously been exposed to the excerpt in combination with the "Christian the lion video" led to significantly lower enjoyment ratings during the second presentation; $p = .04$, $d = 0.69$ (medium effect).

Figure 8

Mean Ratings of Enjoyment in Response to the Second Presentation of the Three Musical Stimuli, Grouped According To the Presentation Condition of the First Exposure

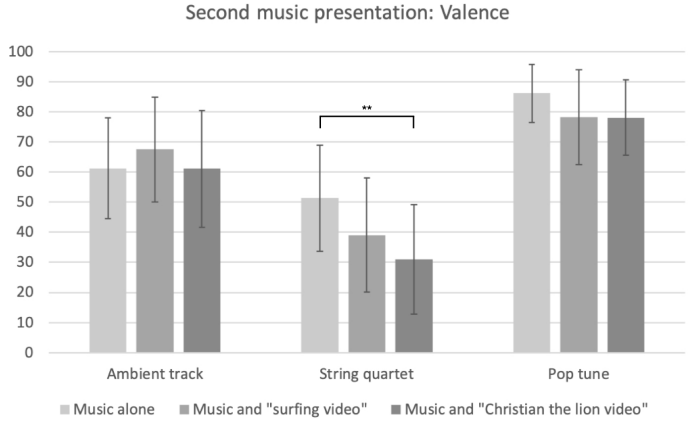


Note. The error bars represent standard deviations. $^{*}p < .05$.

A similar pattern emerged for perceived valence (see Figure 8), suggesting a carry-over effect from the first presentation. A significant main effect of First presentation condition was again found for "string quartet"; $F(2,63) = 6.32$, $p = .003$, $\eta^2 = .17$. Post-hoc tests revealed that when "string quartet" had previously been paired with the "Christian the lion video", it was perceived as significantly more negative than when it had been presented alone; $p = .002$, $d = 0.99$ (large effect).

Figure 9

Mean Ratings of Perceived Valence in Response to the Second Presentation of the Three Musical Stimuli, Grouped According To the Presentation Condition of the First Exposure



Note. The error bars represent standard deviations. $^{**}p < .01$.

For the ratings of awe and being moved, none of the significant effects form the first presentation carried over to the second presentation. No significant effects were present for chills or arousal ratings either.

Free Descriptions of Physiological Responses

The number of reported experiences of physiological responses other than chills are summarised in Table 2. Due to the open, exploratory, nature of this question (and the reliance on participants' own subjective experience of physiological sensations), the interpretation of the findings must remain tentative. The open descriptions may nevertheless give some insight into the overall pattern of physiological responses in relation to the different audio-visual combinations. The findings suggest that when music was paired with the "surfing video", participants reported increases in heart rate more often, potentially indicating higher arousal. In contrast, when music was paired with "Christian the lion video", participants reported smiling and experiencing a warm feeling in the chest, consistent with the pattern of sensations typically associated with being moved/kama muta (e.g., Schubert et al., 2018).

Table 2

Summary of the Physiological Sensations Reported In the Open Descriptions in Response to the First Presentation of the Experimental Stimuli

Presentation condition	Ambient track	Pop tune	String quartet
Music alone	Increased HR (1)	Warmth in chest (2)	Increased HR (1) Goosebumps (1) Warmth in chest (1)
Surfing video	Warmth in chest (2) Increased HR (1)	Increased HR (3)	Increased HR (4) Warmth in chest (1) Lump in throat (1)
Christian the lion video	Warmth in chest (5) Watery eyes (2) Increased HR (1)	Warmth in chest (6) Smiling (6) Lump in throat (1) Increased HR (1)	

Note. HR = heart rate.

A common physiological sensation reported in response to the "surfing video" on its own was an increase in heart rate (experienced by 8 out of 31 participants), while "Christian the lion video" alone evoked smiles (experienced by 8 out of 35 participants) and a warm feeling in the chest (6 participants).

During the second presentation of the music excerpts, 3 participants reported experiencing an increase in heart rate when presented with the "string quartet" previously paired with the "surfing video", and 4 participants reported a warm sensation in the chest in response the "pop tune" (previously paired with the "Christian

the lion video"). Smiles were not reported for the second presentation of the "pop tune".

Discussion

The results of this study suggest that the moving image can intensify emotions induced while listening to music. The videos had, in accordance with the hypothesis, an intensifying influence on the experience of awe (for the "ambient track" and the "pop tune") and being moved (for the "pop tune"). Ratings of awe were, as predicted, mainly influenced by the "surfing video", while an impact on ratings of being moved was only found for the "Christian the lion video". Despite this significant influence on induced emotions, ratings of experienced chills were unaffected by the presence of the moving image. This finding was rather surprising, but the lack of influence from the videos may partially be explained by chill-triggering features in the music excerpts. A high-pitched tone and an increase in dynamics, present in the music excerpts, have been found to elicit chills in several studies (Bannister, 2020; Grewe et al., 2007; Panksepp, 1995; Sloboda, 1991). Free descriptions, nevertheless, seemed to indicate that the videos had some influence on physiological responses, like an increase in heart rate and a warm sensation in the chest during music listening. The intensifying effects on awe and being moved did not, contrary to predictions, lead to increased enjoyment of the music, and the increases in awe and being moved were not present during the second presentation of the music. These findings seem to suggest that the impact that the videos had on the experience of the music excerpts was mainly limited to the first, bimodal presentation condition.

The moving image did not have notable mood-congruent effects on emotions perceived in the music. In relation to perceived valence, no significant mood-congruent effects were found, and in relation to perceived arousal, the significant difference between the two video conditions of the "pop tune" was not consistent with the arousal ratings of the videos alone. One possible explanation for this discrepancy is that the higher arousal ratings for the "pop tune-Christian the lion video" pairing reflected more intense felt emotions. The pattern of ratings suggests that while the "Christian the lion video" significantly intensified the feelings of being moved evoked by the "pop tune", the "surfing video" did not intensify awe to a similar extent in this pairing. In addition, the limited ambiguity in the music excerpts may also explain why no mood-congruent effects were found, since an influence on perceived emotions may depend on a certain level of ambiguity.

Compared to the positive effects on awe and being moved, the negative effects found in relation to enjoyment and valence ratings seemed to be more robust. Questions related to these variables were more explicitly directed at the music, and the negative effects were also present during the second presentation of the music excerpts, being the only significant carry-over effects. The negative effects of video

accompaniment on valence and enjoyment ratings may be related to a lack of congruency between music and visuals, as the negative effects seemed to occur for those pairings that had the lowest congruency ratings. This finding goes against our original hypothesis, which predicted that the effect of the moving image would be strongest in the case of the congruent pairings, facilitated by the 'unity assumption' (e.g., Chen & Spence, 2017). However, it is conceivable that a mismatch between visuals and music made a bigger impact than the other pairings, as it may have conflicted with participants' expectations about how the two stimuli would fit together. The fact that only the negative effects of the incongruent paring were sustained during the second presentation of the music was rather surprising. However, the recollection of stimuli associated with negative emotions (which is consistent with the low rating of enjoyment) has been found to be less dependent on attention than stimuli associated with positive emotions (Talmi et al., 2007). Negatively valenced stimuli may therefore be remembered more easily, and this may explain the sustained negative effects that were observed in the present study.

The present research does entail some important limitations that need to be discussed. The presentation of the videos on their own was important to better account for the responses that they alone could account for, but this decision came with some limitations. The lack of sound in the videos, although necessary in order to fully isolate visuals and sounds, may have reduced felt emotions. Videos and films are almost always accompanied by either diegetic, or non-diegetic sounds (often in the form of music), and this further strengthens an expectation of sound in encounters with the moving image. The violation of this expectation may have had a negative impact on the experience of the videos, and potentially also affected the second presentation of the video (in combination with music). However, we tried to mitigate the potential impact of this aspect by only presenting one of the videos alone to each participant, and by balancing the silent videos ("surfing video" or "Christian the lion video") within each condition to avoid any systematic differences between conditions.

It should also be noted that the dependent variables used in the present study were based on single self-report items. Single items can be less reliable than multi-item measures, and they are more fragile and more sensitive to noise (e.g., Allen et al., 2022). Future studies could try to implement multi-item measures of perceived and/or felt emotions in a similar design to corroborate present findings. Furthermore, translation issues may have had, to some extent, an unwanted influence on the results. In particular, the translation of awe into Norwegian is not completely unproblematic, as the Norwegian translation of awe, æ*refrykt* (comprising the words *honour* and *fear*), has connotations that are of a less positive valence. *Ærefrykt* tends to be more associated with deep reverence and fear, and is less closely related to wonder and admiration compared to the English term (awe). In addition to this, the Norwegian term is less commonly used in daily language compared to awe (an explanation of the term was given in the survey for this reason) and therefore not as intuitive.

Another limitation of the current study is the lack of distinction between structural and semantic congruency in the question related to how well the music fits together with the film. The pairing of the "string quartet" and the "Christian the lion video" was, for instance, predicted to be perceived as incongruent both in terms of semantic and temporal congruency, but it is impossible to determine the relative contributions of the two forms of congruency on participants' ratings. The current study cannot, therefore, fully explain whether the negative impact on enjoyment and valence was mostly due to semantic incongruence (i.e., mismatch in affect), or temporal incongruence (lack of synchrony between the peak moments in the two stimuli). The fact that none of the excerpts were pilot-tested is also likely to have put some limitations on the present research. By piloting a large sample of music excerpts and videos in an initial phase of the study, it may have been possible to select excerpts that would have been even more effective in conveying the intended emotions, while controlling for other attributes such as chill-triggering features.

Future investigations of the influence of the moving image on music may consider looking into other felt emotions besides awe and being moved. In particular, basic emotions (such as sadness, anger and happiness) may correspond more closely with the emotional expression of the music (such as basic emotions), and the relation between the two may, therefore. be more straightforward. Although basic emotions induced by music may be less intense in comparison to awe or being moved, for example, they should in general be given more attention in future research.

The negative influence of the incongruent pairing in particular is an area that should be given more attention in future research. In the present study, this negative influence provided the most robust results, being present also for the second presentation, and this is, to the best of our knowledge, the first study to examine this. The impact of congruency should be examined more systematically with a clearer division between semantic and formal congruency. It would for instance be interesting to see if pairings that are semantically incongruent but temporally congruent, have the same negative influence as music excerpts that are semantically congruent but lacking in terms of temporal synchrony. This would give further insight into the influence of different forms of congruency.

References

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. *European Journal of Psychological Assessment, 38*(1), 1–5. https://doi.org/10.1027/1015-5759/a000699

Bannister, S. (2020). A survey into the experience of musically induced chills: Emotions, situations and music. *Psychology of Music*, *48*(2), 297–314. https://doi.org/10.1177/0305735618798024

Bannister, S., & Eerola, T. (2021). Vigilance and social chills with music: Evidence for two types of musical chills. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000421

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: A Journal of Research in Music Cognition*, *13*(1-2), 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*(4), 427–454. https://doi.org/10.1525/mp.2001.18.4.427

Boltz, M. G., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception*, *27*(1), 43–59. https://doi.org/10.1525/mp.2009.27.1.43

Braud, W. (2001). Experiencing tears of wonder-joy: Seeing with the heart's eye. *Journal of Transpersonal Psychology*, *33*(2), 99–112.

Bullerjahn, C., & Güldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology: A Journal of Research in Music Cognition*, *13*(1-2), 99–118. https://doi.org/10.1037/h0094100

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'unity assumption' on multisensory integration: A review. *Frontiers in Psychology*, *8*, 445. https://doi.org/10.3389/fpsyg.2017.00445

Cohen, A. J. (2001). Music as a source of emotion in film. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 879–908). Oxford University Press.

Cotter, K. N., Silvia, P. J., & Fayn, K. (2018). What does feeling like crying when listening to music feel like?. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(2), 216–227. https://doi.org/10.1037/aca0000108

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, *1242*, 136–150. https://doi.org/10.1016/j.brainres.2008.03.071

Eldar, E., Ganor, O., Admon, R., Bleich, A., & Hendler, T. (2007). Feeling the real world: Limbic response to music depends on related content. *Cerebral Cortex*, *17*(12), 2828–2840. https://doi.org/10.1093/cercor/bhm011

Fiske, A. P., Seibt, B., & Schubert, T. (2019). The sudden devotion emotion: Kama muta and the cultural practices whose function is to evoke it. *Emotion Review*, *11*(1), 74–86. https://doi.org/10.1177/1754073917723167

Forester, G., Halbeisen, G., Walther, E., & Kamp, S. M. (2020). Frontal ERP slow waves during memory encoding are associated with affective attitude formation. *International Journal of Psychophysiology*, *158*, 389–399. https://doi.org/10.1016/j.ijpsycho.2020.11.003

Geringer, J. M., Cassidy, J. W., & Byo, J. L. (1996). Effects of music with video on responses of nonmusic majors: An exploratory study. *Journal of Research in Music Education*, *44*(3), 240–251. https://doi.org/10.2307/3345597

Geringer, J. M., Cassidy, J. W., & Byo, J. L. (1997). Nonmusic majors' cognitive and affective responses to performance and programmatic music videos. *Journal of Research in Music Education*, *45*(2), 221–233. https://doi.org/10.2307/3345582

Goldberg, M. E., Chattopadhyay, A., Gorn, G. J., & Rosenblatt, J. (1993). Music, music videos, and wear out. *Psychology & Marketing*, *10*(1), 1–13. https://doi.org/10.1002/mar.4220100102

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception*, *24*(3), 297–314. https://doi.org/10.1525/mp.2007.24.3.297

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. https://doi.org/10.1037/a0018916

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, *8*(5), 668–683. https://doi.org/10.1037/a0013505

Juslin, P. N., & Vastfjall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, *17*(2), 297–314. https://doi.org/10.1080/02699930302297

Konecni, V. J. (2005). The aesthetic trinity: Awe, being moved, thrills. *Bulletin of Psychology and the Arts*, *5*(2), 27–44. https://doi.org/10.1037/e674862010-005

Krahé, C., Hahn, U., & Whitney, K. (2015). Is seeing (musical) believing? The eye versus the ear in emotional responses to music. *Psychology of Music*, *43*(1), 140–148. https://doi.org/10.1177/0305735613498920

Lalanne, C., & Lorenceau, J. (2004). Crossmodal integration for perception and action. *Journal of Physiology-Paris*, *98*(1-3), 265–279. https://doi.org/10.1016/j.jphysparis.2004.06.001

Livingstone, S. R., Thompson, W. F., Wanderley, M. M., & Palmer, C. (2015). Common cues to emotion in the dynamic facial expressions of speech and song. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *68*(5), 952–970. https://doi.org/10.1080/17470218.2014.971034

Lavy, M. M. (2001). *Emotion and the experience of listening to music: A framework for empirical research.* [Unpublished doctoral dissertation]. University of Cambridge.

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PloS One*, *10*(6), e0128451. https://doi.org/10.1371/journal.pone.0128451

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception*, *13*(2), 171–207. https://doi.org/10.2307/40285693

Schubert, T. W., Zickfeld, J. H., Seibt, B., & Fiske, A. P. (2018). Moment-to-moment changes in feeling moved match changes in closeness, tears, goosebumps, and warmth: Time series analyses. *Cognition and Emotion*, *32*(1), 174–184. https://doi.org/10.1080/02699931.2016.1268998

Schurtz, D. R., Blincoe, S., Smith, R. H., Powell, C. A., Combs, D. J., & Kim, S. H. (2012). Exploring the social aspects of goose bumps and their role in awe and envy. *Motivation and Emotion*, *36*(2), 205–217. https://doi.org/10.1007/s11031-011-9243-8

Seibt, B., Schubert, T. W., Zickfeld, J. H., & Fiske, A. P. (2017). Interpersonal closeness and morality predict feelings of being moved. *Emotion, 17*(3), 389–394. https://doi.org/10.1037/emo0000271

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*(2), 110–120. https://doi.org/10.1177/0305735691192002

Swarbrick, D., Seibt, B., Grinspun, N., & Vuoskoski, J. K. (2021). Corona concerts: The effect of virtual concert characteristics on social connection and Kama Muta. *Frontiers in Psychology*, *12*, 648448. https://doi.org/10.3389/fpsyg.2021.648448

Swarbrick, D., & Vuoskoski, J. K. (submitted). Collectively classical: Connectedness, awe, feeling moved, and motion at a live and livestreamed concert. Manuscript submitted for publication.

Talmi, D., Schimmack, U., Paterson, T., & Moscovitch, M. (2007). The role of attention and relatedness in emotionally enhanced memory. *Emotion*, *7*(1), 89–102. https://doi.org/10.1037/1528-3542.7.1.89

Thompson, S. (2006). Audience responses to a live orchestral concert. *Musicae Scientiae*, *10*(2), 215–244. https://doi.org/10.1177/102986490601000204

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *2005*(156), 203–227. https://doi.org/10.1515/semi.2005.2005.156.203

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, *22*(8), 1457–1470. https://doi.org/10.1080/02699930701813974

Vermeulen, I., & Beukeboom, C. J. (2016). Effects of music in advertising: Three experiments replicating single-exposure musical conditioning of consumer choice (Gorn 1982) in an individual setting. *Journal of Advertising*, *45*(1), 53–61. https://doi.org/10.1080/00913367.2015.1088809

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170. https://doi.org/10.1016/j.cognition.2010.11.010

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177/0305735613502373

Vuoskoski, J. K., & Eerola, T. (2017). The pleasure evoked by sad music is mediated by feelings of being moved. *Frontiers in Psychology*, *8*, 439. https://doi.org/10.3389/fpsyg.2017.00439

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2013). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics, 76*(2), 591–604. https://doi.org/10.3758/s13414-013-0582-2

Vuoskoski, J. K., Thompson, M. R., Spence, C., & Clarke, E. F. (2015). Interaction of sight and sound in the perception and experience of musical performance. *Music Perception*, *33*(4), 457–471. https://doi.org/10.1525/mp.2016.33.4.457

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2011). Who enjoys listening to sad music and why?. *Music Perception*, *29*(3), 311–317. https://doi.org/10.1525/mp.2012.29.3.311

Vuoskoski, J. K., Zickfeld, J. H., Alluri, V., Moorthigari, V., & Seibt, B. (2022). Feeling moved by music: Investigating continuous ratings and acoustic correlates. *PloS One*, *17*(1), e0261151. https://doi.org/10.1371/journal.pone.0261151

Wassiliwizky, E., Wagner, V., Jacobsen, T., & Menninghaus, W. (2015). Art-elicited chills indicate states of being moved. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 405–416. https://doi.org/10.1037/aca0000023

Zickfeld, J. H., Schubert, T. W., Seibt, B., & Fiske, A. P. (2017). Empathic concern is part of a more general communal emotion. *Frontiers in Psychology*, *8*, 723. https://doi.org/10.3389/fpsyg.2017.00723

Received: November 9, 2022

Appendix

1. Questions asked after audio-visual stimuli

Did you enjoy the music?

Not at all — Very much

Did the music fit together with the film?

Not at all — Very well

Did you experience chills/goosebumps?

Not at all — Very much

Did you experience any other form of bodily reaction (warmth in chest, increase in heart rate, lump in throat, tears etc.)?

- ◯ Yes
- ◯ No

If yes: Please describe your reaction.

Did you experience a feeling of awe (mixture of admiration, surprise and fear)?

Not at all — Very much

To what extent were you moved/touched?

Not at all — To a large extent

How would you describe the mood of the music?

Negative Neutral Positive

How would you describe the energy of the music?

Very low (calm) Neutral Very high (energetic)

2. Filler task presented in the middle of the survey

Before you'll be presented with new examples, you need to do this task to refocus your mind.

What is the next number in the line?

- ◯ A= 26
- ◯ B= 21
- ◯ C= 22
- ◯ D= 23
- ◯ E= 24
- ◯ F= 25