



This is a repository copy of *The role of embodied simulation and visual imagery in emotional contagion with music.*

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/186858/>

Version: Published Version

Article:

Cespedes-Guevara, J. and Dibben, N. orcid.org/0000-0002-9250-5035 (2022) The role of embodied simulation and visual imagery in emotional contagion with music. *Music & Science*, 5. ISSN 2059-2043

<https://doi.org/10.1177/20592043221093836>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The Role of Embodied Simulation and Visual Imagery in Emotional Contagion with Music

Music & Science
Volume 5: 1–27
© The Author(s) 2022
DOI: 10.1177/20592043221093836
journals.sagepub.com/home/mns

Julian Cespedes-Guevara¹ and Nicola Dibben²

Abstract

Emotional contagion has been explained as arising from embodied simulation. The two most accepted theories of music-induced emotions presume a mechanism of internal mimicry: the BRECVEMA framework proposes that the melodic aspect of music elicits internal mimicry leading to the induction of basic emotions in the listener, and the Multifactorial Process Model proposes that the observation or imagination of motor expressions of the musicians elicits muscular and neural mimicry, and emotional contagion. Two behavioral studies investigated whether, and to what extent, mimicry is responsible for emotion contagion, and second, to what extent context for affective responses in the form of visual imagery moderates emotional responses. Experiment 1 tested whether emotional contagion is influenced by mimicry by manipulating explicit vocal and motor mimicry. In one condition, participants engaged in mimicry of the melodic aspects of the music by singing along with the music, and in another, participants engaged in mimicry of the musician's gestures when producing the music, by playing along ("air guitar"-style). The experiment did not find confirmatory evidence for either hypothesized simulation mechanism, but it did provide evidence of spontaneous visual imagery consistent with the induced and perceived emotions. Experiment 2 used imagined rather than performed mimicry, but found no association between imagined motor simulation and emotional intensity. Emotional descriptions read prior to hearing the music influenced the type of perceived and induced emotions and support the prediction that visual imagery and associated semantic knowledge shape listeners' affective experiences with music. The lack of evidence for the causal role of embodied simulation suggests that current theorization of emotion contagion by music needs refinement to reduce the role of simulation relative to other mechanisms. Evidence for induction of affective states that can be modulated by contextual and semantic associations suggests a model of emotion induction consistent with constructionist accounts.

Keywords

contagion, emotion, embodiment, mimicry, simulation, visual imagery

Submission date: 24 March 2021; Revised date: 21 January 2022; Acceptance date: 16 March 2022

Emotions are contagious: on many occasions, we "catch" other people's emotional states, reacting to their emotions by feeling the same emotions. Emotional contagion also seems to occur as a response to music: frequently, when we perceive that music expresses a particular emotion, we feel the same emotion aroused in ourselves. Both have been explained as sharing the same causal mechanism: embodied simulation (Cochrane, 2010; Juslin & Västfjäll, 2008; Molnar-Szakacs & Overy, 2006; Scherer & Coutinho, 2013). However, there are few empirical studies and scant direct evidence of the causal role of embodied simulation in emotion induction when listening to music. The two experiments reported here examined the causal role of embodied simulation in emotional contagion with music, its strength relative to another proposed

mechanism of emotion induction—visual imagery, and the role that contextual information has in the affective experiences induced by contagion.

Several theories have proposed that there is an overlap between the neural systems for planning, executing

¹ Departamento de Estudios Psicológicos, Universidad Icesi, Cali, Colombia

² Department of Music, University of Sheffield, Sheffield, UK

Corresponding author:

Julian Cespedes-Guevara, Departamento de Estudios Psicológicos, Universidad Icesi, Calle 18 No. 122-135, Cali, Colombia.
Email: jcespedes@icesi.edu.co



actions, perceiving other people's actions, and inferring their states of mind and feelings. According to those theories, social cognition is based on *embodied simulation*, whereby we achieve a direct understanding of another person's thoughts, feelings, and intentions by internally mirroring or reenacting their mental states and actions (Barsalou, 2008; Gallese & Sinigaglia, 2011; Glenberg, 2010; Iacoboni, 2009; Springer et al., 2013). This internal mimicry would be activated when people respond to emotional stimuli, and even when they *think* about emotional concepts (Barrett, 2006; Niedenthal et al., 2005; Wilson-Mendenhall et al., 2011). Importantly, these theories also predict that inhibiting simulation has the opposite effect on social cognition. For example, Michael and colleagues (Michael et al., 2014) found that using transcranial magnetic stimulation (TMS) in the participants' hand area of the premotor cortex, resulted in difficulty understanding pantomimed hand gestures.

Recent music cognition theories have proposed that music perception is based on embodied simulation of the motor actions executed by the musicians, supported by the mirror neuron system (Colling & Thompson, 2013; Cox, 2011; Godøy & Leman, 2010; Leman & Maes, 2014; Overy & Molnar-Szakacs, 2009) and several studies provide indirect evidence. For example, listeners' arm movements while listening to music played on a traditional Chinese plucked string instrument matches the pattern of movements of the performers' shoulders (Leman et al., 2009). Several brain areas associated with music perception are also associated with vocalizing and motor planning (Brown & Martinez, 2007; Callan et al., 2006; Chen et al., 2008; Miller, 2016); and brain motor planning areas are activated during music listening (Gordon et al., 2018; Wallmark et al., 2018), especially in musicians (Alluri et al., 2017). Subvocalization (subtle movements of muscles related to vocal production) has been observed during musical mental imagery tasks (Pruitt et al., 2019;—but not during listening: Bruder & Wöllner, 2021). And TMS of pianists' left-hand brain areas hampers their ability to play along with their right hand while listening to music that they previously learned to play with both hands (Novembre et al., 2014), or while synchronizing their right hand with a left-hand part recorded by a different pianist (Timmers et al., 2020). Such evidence gives credence to the hypothesis that simulation can take varied forms: intramodal (e.g., imitating the finger movements of a pianist), crossmodal (e.g., subvocalizing the melody of a piano piece) or even amodal, symbolic, or metaphorical (“abdominal imitation of the exertion dynamic evident in sounds”) (Cox, 2011, p. 37).

Drawing from such theories and evidence, several music cognition theorists have concurred that: (a) some features of musical sounds resemble the expression of emotions (or even the proprioceptive feelings) in the voice and the body and (b) perceiving those expressive musical features elicits internal mimicry, which in turn, leads to emotional contagion with the emotion expressed by the music

(Cochrane, 2010; Davies, 2013; Jackendoff & Lerdahl, 2006; Juslin & Västfjäll, 2008; Molnar-Szakacs & Overy, 2006; Scherer & Coutinho, 2013). These theories differ in regards to the aspect of music proposed to elicit simulation, and therefore, in the type of internal mimicry they involve. For the BRECVEMA framework (Juslin, 2013a; Juslin et al., 2010; Juslin & Västfjäll, 2008), the *melodic* aspect of music can elicit internal mimicry when particular instrumental timbres (such as the violin or the cello) are heard as “super-expressive voices” resembling the expression of basic emotions in vocalizations, which in turn, activates subvocalization and induces the corresponding basic emotions in the listener. In contrast, the Multifactorial Process Model (Scherer & Coutinho, 2013; Scherer & Zentner, 2001), claims that the observation or imagination of *motor expressions* of the musicians elicits muscular and neural mimicry, and subsequently, emotional contagion. It is plausible that these theories are not mutually exclusive, since both types of simulation can operate simultaneously: whereas simulation via subvocalization may be responsible for the perception of emotional qualities associated with variations in timbre, pitch, and melodic contour, simulation of motor gestures may be responsible for the perception of emotional qualities associated with bodily movements, such as strength, speed, and energy, as indeed predicted by Overy and Molnar-Szakacs' SAME model of affective experience (Molnar-Szakacs & Overy, 2006; Overy & Molnar-Szakacs, 2009).

In two experimental studies that aimed to test the BRECVEMA framework mechanisms, Juslin and colleagues found that participants who listened to music expressive of sadness reported induced sadness, and interpreted these findings as evidence of emotional contagion (Juslin et al., 2014, 2015). However, neither of these studies provided evidence that the correspondence between perceived and induced sadness was due to the proposed internal mimicry mechanism. Several other studies have also interpreted the correspondence between perceived and induced emotions as emotional contagion (Egermann & McAdams, 2013; Garrido & Macritchie, 2020; Peltola & Eerola, 2016; Van den Tol & Edwards, 2011), but to our knowledge, no empirical research has attempted to test the internal mimicry hypotheses for music directly. However, three lines of research have provided indirect evidence for the involvement of embodied simulation in emotional experiences with music. First, listening to pleasant pieces of music (compared to dissonant, unpleasant versions of the same pieces) activates brain areas associated with the formation of premotor representations of vocal sound production (Koelsch et al., 2006). Second, listening to music can elicit a pleasant motivation to mimic some aspect of it, such as moving along to rhythmic music (Janata et al., 2012; Labbé & Grandjean, 2014; Witek et al., 2014); singing along (DeNora, 2000; Dibben & Williamson, 2007); and mimicking singers' facial expressions (Chan et al., 2013). And finally, people's movements while listening to music affect their preferences (Sedlmeier et al., 2011), and their perception of music-expressed emotions (Maes & Leman, 2013).

Significantly, most of the evidence so far suggests that adopting or mimicking emotional postures, gestures, vocalizations, etc., alters participants' affective state by inducing changes in experienced pleasantness (valence), but it does not produce new particular (i.e., discrete) emotional states by itself. More precisely, mimicking a positive or negative stimulus (e.g., an expression of joy, fear, anger, disgust, etc.) leads to changes in experienced valence in a manner congruent with the valence of the stimulus, but not necessarily to the induction of the *same discrete emotion* as the one observed (Flack, 2006; Hatfield et al., 1995; Hess & Blairy, 2001; McIntosh, 2006; Neumann & Strack, 2000). It's notable, for instance, that participants' ratings of induced emotion in these studies showed a "bleeding effect" at odds with the notion of discrete emotions: when they reported feeling a negative emotion (e.g., disgust) they also reported feeling other negative ones simultaneously (e.g., anger and fear). To our knowledge, only one study (Hawk et al., 2012) has found that *hearing* emotional vocalizations of discrete emotions leads to both mimicry and *induction* of the corresponding discrete emotions in the participants. Taken together, these findings cast doubt on the notion that engaging in behaviors that facilitate embodied simulation can *by itself* lead to the induction of discrete emotions with music.

If, as the evidence above suggests, manipulating people's facial, vocal, or bodily expressions can *bias* their affective state (i.e., their experienced valence), but does not lead to the induction of discrete, full-blown emotions, then it is unlikely that engaging in any type of mimicry of the music expressive qualities *by itself* can lead to a contagion of discrete emotions. At first glance this observation is compatible with current theory; two of the most influential contemporary theories of musical emotions, the BRECVEMA framework (Juslin, 2013a), and the Multifactorial Process Model (Scherer & Coutinho, 2013) propose that emotional responses to music emerge from the interaction of factors in the music, the listening situation, and the individual.

Juslin has recently noted that extramusical information such as biographical information about the composer or song lyrics can influence emotions induced by music (Juslin, 2019). However, in his view, these mechanisms only influence listeners' emotions via the activation of one of the mechanisms in the BRECVEMA framework. In other words, contextual information only has an effect because it evokes personal memories (and activates the *episodic memories mechanism*), because it evokes associations with past experiences (and triggers the *evaluative conditioning* mechanism) or because it stimulates visual mental images (i.e., the *visual imagery* mechanism) (Juslin, 2019, p. 379; Taruffi & Küssner, 2019). To our knowledge, very few studies have explored the effect of extramusical information on the induction of musical emotions, albeit there is evidence for the effects of context on perception of emotion in music (Margulis et al., 2017) and on aesthetic judgments (Anglada-Tort et al., 2019). A notable exception

is a study by Vuoskoski and Eerola (2015), in which three groups of participants listened to the same sad-sounding piece of music after reading a description of the music as depicting a sad narrative, reading a neutral narrative, or not having read any previous information about the music, correspondingly. The results showed that, compared to the other groups, participants who read the sad narrative experienced more induced sadness, and evoked more sad imagery while listening to the music. The authors interpreted these results as stemming from the activation of the visual imagery mechanism. These results suggest that extramusical information affects listeners' emotions by intensifying emotional responses that they would have felt even in the absence of such information. However, it remains unclear whether this contextual effect can be observed in music that expresses other emotions, whether contextual information can modulate the *type* of induced emotion (e.g., music expressive of happiness heard at the funeral of a loved one may be experienced as sad), and whether the presence of mental visual imagery is a necessary condition for these effects to happen.

In contrast to the BRECVEMA framework, other theorists have proposed that extramusical information, and even musical sounds by themselves can evoke *semantic associations* (i.e., meaningful connections between music and cultural contents, physical events, biological entities, etc.); and that when these associations have emotional connotations, they contribute to the induction of musical emotions as a distinct mechanism (Clarke et al., 2010; Fritz & Koelsch, 2008). However, according to Juslin, these associations can only contribute to perceiving emotions expressed by music, and do not constitute an induction mechanism (Juslin, 2019, p. 377). Moreover, since according to the BRECVEMA framework, music expresses basic emotions (Juslin, 2013b), then it predicts that the contagion mechanism induces basic emotions in the listener as well (Juslin, 2019, p. 299), and it does not specify if, or how contextual and personal factors interact with the contagion response making it more or less likely that the listener experiences a basic emotion or a different, nonbasic one.

The assumption that emotional phenomena are organized around biologically predetermined categories ("basic emotions"), that tend to be expressed, perceived, and experienced cross-culturally (Ekman, 1992b), is refuted by constructionist approaches to emotion such as Russell's Core Affect model (Russell, 2003) and Barrett's Theory of Constructed Emotion (Barrett, 2006, p. 2017). Instead, constructionist theories have proposed that contextual factors such as the meaning of the present situation and the use of culturally specific linguistic categories are critically involved in emotion perception and elicitation (Barrett et al., 2007, 2010; Carroll & Russell, 1996; Lindquist & Barrett, 2008). This has implications for theorization of emotion contagion. Drawing from these theories, Cespedes-Guevara and Eerola (2018) challenged the commonly held assumption that the expression and perception of emotion are constrained to a set of basic emotions, and

suggested that musical expressivity is grounded on core affect (arousal and valence) instead. And since music expresses fluctuations in core affect (i.e., arousal and valence) rather than basic emotions, it follows that engaging in an embodied simulation of the emotions expressed by music can only lead to the corresponding experienced changes in core affect, not to the induction of discrete emotions in the listeners. Experiencing the induction of discrete emotions also requires the confluence of factors in the music, the person, and the context that help shape those changes of core affect into a discrete, full-blown emotion (Cespedes-Guevara, 2021). Therefore, changing contextual information about the music changes the quality and intensity of the emotions that a listener perceives and experiences (Cespedes-Guevara & Eerola, 2018). So far, the mechanism closest to this is the visual imagery mechanism of BRECVEMA: it is possible that visual imagery which arises from the personal and cultural associations evoked while listening to music works as one of the “contextual factors” that helps shape core affect into specific emotions.

Overview of the Present Studies

As argued above, although there is some evidence that embodied simulation is involved in emotional processing and in music perception, there is no evidence that it plays a causal role in emotional contagion with music, nor the extent to which it might be mediated by contextual information. Consequently, we conducted two empirical studies to investigate these questions. In the first experiment, we tested two hypotheses about the role of mimicry in emotional contagion: the hypothesis derived from BRECVEMA (Juslin, 2013a), which claims that internal mimicry of *the melodic aspects of the musical material* leads to emotional contagion, and that from the Multifactorial Process Model (Scherer & Coutinho, 2013), which claims that internal mimicry of *the musician’s gestures when producing the music* drives the contagion response. In the second experiment, based on the constructionist theory proposed by Cespedes-Guevara (2021), we studied the role of contextual information in the emotional responses evoked by contagion. Therefore, we tested the hypothesis that, in the absence of any contextual information about the music, engaging in an embodied simulation would facilitate contagion with the core affect expressed by the music, rather than induce specific basic emotions. In other words, we expected that the participants would experience a variety of emotions that are congruent with the music’s valence and arousal characteristics, but they would not experience only a limited set of basic emotions such as joy, sadness, or fear; we expected that they would also report emotions such as pride, nostalgia, tenderness, and determination. Furthermore, we predicted that providing participants with contextual information about the music would modulate the core affect induced by embodied simulation, producing the induction of discrete emotions whose quality would be coherent with the provided contextual information. Additionally, in both experiments, we

collected qualitative data about the participants’ mental experiences while listening to the music. This allowed us to explore if the biasing effect of contextual information could be attributed to the activation of the visual imagery mechanism, or if the biasing effect could be observed even in the absence of such imagery. The inclusion of a new Action Tendencies questionnaire enabled us to capture this established component of emotional response which the majority of studies in music and emotion have ignored.

Experiment I

Participants were asked to listen to music and to perform behavioral tasks that either facilitated or prevented simulation: in the vocal simulation condition (BRECVEMA), participants mimicked the music’s melody by singing along with the piece, and in the motor simulation condition (Multifactorial Process Model), participants mimicked the gestures implied in the production of the sounds by pretending to play the instruments. The third group performed a simulation-hampering, distracting task: they used their arms and their voice in a different task while the music played. Finally, the fourth group constituted a control condition: these participants remained quiet and still while the music played. We predicted that participants in the two simulation conditions (vocal, motoric) would experience more intense perceived and induced emotions than participants in the other groups (Hypothesis 1). Second, we predicted that the participants in the distracting task group would experience the least intense perceived and induced emotions (Hypothesis 2). We also asked the participants to report what went through their minds while listening to check whether mimicry could be detected in visual imagery experienced by the participants. Based on the hypothesis of the visual imagery mechanism from the BRECVEMA framework (Juslin, 2013a), we predicted that there would be a correspondence between the content of the participants’ imagery and their ratings of perceived and induced emotions (Hypothesis 3).

Method

Design. The experiment used a between-subjects design, with Type of Simulation as the between-subjects independent variable (IV) (four levels: Vocal Simulation/Motor Simulation/Distracting Task/Stationary); and Perceived and Induced affective states as dependent variables.

Participants. One hundred and twenty-seven participants were recruited from [name on acceptance] (age range = 19–66 years; Mean_{age} = 29.9, *SD* = 9.5, 79 women). The sample size was determined using an a priori power analysis run in the G-Power 3.1 software (Faul et al., 2007) (power = 0.95; alpha = 0.05; effect size $f = 0.4$). Participants were randomly assigned to four experimental conditions: vocal simulation, motor simulation, distracting task, and control task.

Stimuli. Participants listened to three instrumental pieces from the Movie Soundtrack Database (Eerola & Vuoskoski, 2011) that are perceived as expressive of emotions in the four quadrants of Two-Dimensional Affective Space (Russell & Barrett, 1999): Sadness/Tenderness (low arousal, negative/positive valence), Fear/Anger (high arousal, negative valence), Joy/Excitement (high arousal, positive valence). To select the stimuli, we asked 28 participants to rate the emotions expressed by 9 pieces from the same database (these participants did not take part in the main experiment). We chose three pieces rated as moderately expressive of the target emotions rather than pieces with the highest scores to avoid a ceiling effect, which would make the effects of the IV unobservable (Appendix A). The experiment used longer duration versions of the pieces ($M = 137.33$ s, $SD = 8.02$). Consistent with the original dataset construction (Eerola & Vuoskoski, 2011), participants were unfamiliar with the specific stimuli ($M = 0.25$ on a scale from 0 to 3), and only one of the 124 participants correctly guessed the movie to which one of the pieces belonged.

Measures. Participants' affective experience was measured using direct and indirect techniques. Induced affect was measured directly with a 14-item questionnaire including core affect adjectives (i.e., valence, tense arousal, and energetic arousal), discrete emotional adjectives taken from the GEMS-25 (Zentner et al., 2008), and items from a questionnaire used by Juslin and colleagues (Juslin et al., 2014). Each item of this Induced Emotions questionnaire included two or three adjectives related to the same emotional category. Induced affective state was also measured indirectly using a facial expression technique (Niedenthal et al., 2000) in which participants control a computerized video displaying a changing (positive to negative) facial expression to detect the offset of the initial expression. Participants in positive affective states (e.g., joy, tenderness) would see the offset of the happy expression at an earlier point than participants in negative affective states (e.g., fear, sadness, and anger), as calculated by frame number. Perception of emotions expressed by the music was measured using a 15-item questionnaire containing pairs of adjectives with the same semantic content as the adjectives in the Induced Emotions questionnaire. An additional 15-item questionnaire (Subjective Feelings and Action Tendencies) developed by the first author measured subjective action tendencies (Frijda et al., 1989)—an important but often-overlooked component of emotion in music studies (see Table S1 in Supplemental materials for the questionnaire item list).

The participants were also asked to report how much they *liked* the piece they listened to, how *familiar* they were with it, how *difficult* they found the task, and how *embarrassing* they found performing the task. (The participants in the stationary condition were asked whether they felt uncomfortable with the presence of the experimenter behind them

during the experiment). Participants were unobtrusively observed to check they complied with instructions, including the need to stay still during the stationary task.

The participants' thoughts while listening to the music were explored with a short interview at the end of the experiment, to verify the emotion induction and perception ratings, to check whether mimicry features in imagery, and to examine the extent to which imagery and emotional experiences coincided. Theories of embodied simulation assume that simulation is an implicit process (Barsalou, 2008), nevertheless it could be speculated that the more embodied simulation is active in listeners' brains, the more they may evoke images of the musicians (or themselves) playing the music while listening to it (Holmes et al., 2008; Kan et al., 2003; Wu & Barsalou, 2009). Additional measures consisted of a questionnaire of demographic information, the Engagement and Musical Training scales from the Goldsmiths Musical Sophistication Index [Gold-MSI] (Müllensiefen et al., 2013). Finally, since recent evidence suggests that expertise is positively correlated with neural activation of motor brain areas (e.g., Alluri et al., 2017) and negatively correlated with subvocalization (Pruitt et al., 2019), we also asked participants which musical instruments they could play, if any, to explore the potential mediating role of this covariate.

Procedure. Participants were tested individually. They were informed that the experiment concerned the psychological effects of listening to instrumental music. Once the participants had satisfactorily practiced the indirect measurement task, a computer screen displayed the instructions of their corresponding experimental condition. The participants in the vocal simulation condition were asked to "sing or hum along to the melody while the music unfolds". The instructions emphasized that it was not important if they could sing in tune, or if they shifted their attention from one instrument to another while singing. The participants in the motor simulation condition were asked to "pretend they were playing the instruments that they heard" by making the movements they thought the musicians would make while playing. Again, the instructions emphasized that it was not important if they knew how to play the instruments, or if they switched from one instrument to another along the way. The participants in the distracting task condition were instructed to move a group of cubes from a box in front of them, to two other boxes at their sides, one at a time and counting out loud each time they moved a cube. Finally, the participants in the control condition were asked to "stay completely still and silent" while listening to the music, and to avoid humming the melody, tapping their feet, or swaying their body to the rhythm of the music. Participants practiced the corresponding task while listening to Satie's *Gymnopédie No.1*. (See Figure 1 for an illustration of these tasks).

The experiment followed the same procedure for each trial: listening to a piece of music twice while performing the experimental task, completing the indirect measurement technique, and filling the self-report questionnaires of

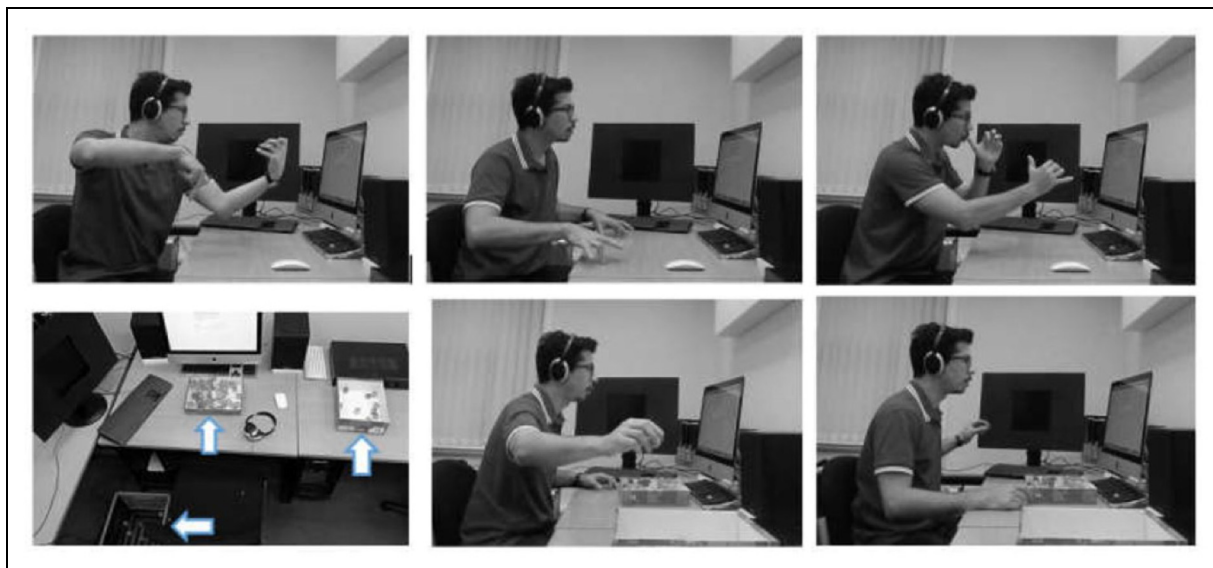


Figure 1. Enactment of the motor simulation and distracting task conditions in Experiment 1. The panels represent the typical gestures made by participants while listening to the music in the motor simulation condition. The panel on the bottom left corner shows the setting that participants in the distracting task condition found: notice the box with cubes in the middle, and the two boxes at the right and left of the computer. The middle and bottom right panels represent the way participants allocated the cubes in the boxes while counting out loud and listening to the music.

induced and perceived affective states. Stimuli were presented twice to make it easier for the participants to predict how the music was unfolding, and in consequence, to perform the experimental task more effectively. The musical stimuli were presented in a counterbalanced order through headphones (Bose AE2), adjusted at a comfortable sound level.

After finishing all the trials, participants were asked to listen to a fragment of each stimulus once more, and to say “what went through their minds” while they listened to this music in the experiment. The experiment ended by filling questionnaires about demographics, and musicality. Once debriefed, the participants were offered a chocolate bar as a reward for their participation. The procedure took 50 min on average.

Results

Three participants were excluded from the analysis because they did not follow the instructions correctly, resulting in a sample of 124 people. Participants had an intermediate level of musicianship as inferred from comparison of the mean scores from our sample to the data norms of the Musical Sophistication Index test [Gold-MSI] (Müllensiefen et al., 2013): Musical Engagement $M = 38.7$, $SD = 8.18$ (Gold-MSI v.1.0 $M = 41.52$, $SD = 10.36$); Musical Training $M = 25.74$, $SD = 6.65$ (Gold-MSI v.1.0 $M = 26.52$, $SD = 11.44$).

Validity and Manipulation Check

Validity of Perception of Morphing Faces Technique. The correlations between z-scores (used to determine variability in

the use of the scale) and induced affect (self-report ratings from the questionnaires) were not compatible with our prediction that participants who were in a positive affective state would perceive the change from a positive to a negative emotional expression earlier than those who were in a more negative state (Niedenthal et al., 2000, 2001): only the data from the morphing face task after listening to the Joy/Excitement piece showed the expected correlations, and none of them was higher than .28 (Tables S2 and S3 in Supplemental materials). It is likely that variability in the way the participants approached the morphing faces task, and individual differences in perception of emotional expressions produced the observed results. In the context of this experiment, the morphing face technique cannot be assumed to be a reliable measurement of the participants’ affective state, and therefore, is not included in the subsequent analyses.

Validity of the Action Tendencies Questionnaire. The correlation analysis between participants’ ratings in the Subjective Feelings and Action Tendencies questionnaire and their ratings in the Induced Emotions questionnaire reveals a coherent pattern of correlation indicating it is a valid measure of how emotionally moved the participants felt. There were moderate to high positive correlations between scores of needing-to-be-comforted and of feeling-like-crying and induced sadness ($r = .41$; $r = .59$, respectively); between scores of feeling-like-laughing and induced happiness ($r = .49$); between scores of feeling-in-command-of-the-situation and induced triumph ($r = .54$); of wanting-to-attack-something and induced irritation ($r = .47$); and between scores of

wanting-to-hide and induced anxiety ($r = .66$) (all p values $< .001$). (See Table S4 in Supplemental materials).

Manipulation Check. The stimuli were successful at expressing and inducing the target emotions as inferred from the highest mean scores being given to the related emotion adjectives in each case relative to the scores for the other adjectives (on a scale of 0–4 where 4 represents greatest intensity). Importantly, although the participants experienced perceived and induced emotions that are coherent with the arousal and/or valence characteristics of the music, their experiences were not limited to the induction of *basic emotions* (e.g., happiness, sadness, anger, fear, etc.), as predicted by the BRECVEMA framework. In contrast, our results are more consistent with the findings of previous experiments where participants who were asked to mimic nonmusical emotional stimuli reported experiencing a variety of emotions consistent with the valence of the observed stimuli (e.g., Flack, 2006; Hatfield et al., 1995; Hess & Blairy, 2001; McIntosh, 2006; Neumann & Strack, 2000). The Sadness/Tenderness piece was associated with the induction of bittersweet low arousal emotions such as feeling mellowed ($M = 2.09$) or soothed ($M = 2.16$) and with the perception of expressed tenderness ($M = 2.49$), peacefulness ($M = 2.46$) and nostalgia ($M = 1.93$); the Anger/Fear piece was associated with induced emotions of high arousal such as feeling anxious ($M = 1.60$) or triumphant ($M = 1.53$), and with the perception of expressed anger ($M = 2.19$), pride ($M = 2.14$) and fear ($M = 2.02$); and the Joy/Excitement piece was associated with the induction of high arousal and positive valence feelings of happiness ($M = 2.69$) or triumph ($M = 1.69$), and the perception of expressed joy ($M = 1.99$), peacefulness ($M = 2.07$) or tenderness ($M = 1.46$). Spearman correlation coefficients of corresponding perceived and induced emotions ranged from .20 to .67 (all p values $< .005$) which is consistent with the condition for emotional contagion with music to have occurred. (See Table S5 in Supplemental materials).

Hypotheses Testing

Testing Hypotheses 1 and 2 involves establishing which were the most intensely induced and perceived emotions for each piece; to this end, we calculated this set of dependent variables:

- Most Intense Induced Emotion: the highest score for each participant in the Induced Emotions questionnaire.
- Most Intense Action Tendency: the highest score for each participant in the questionnaire of Subjective Feelings and Action Tendencies.
- Most Intense Perceived Emotion: the highest score reported by each participant in the Perceived Emotions questionnaire.

The distributions of these dependent variables were somewhat skewed, but we did not perform any transformation given that ANOVA is robust to violations of the assumption of normality (Howell, 2002; Schmider et al., 2010). The Bonferroni correction was applied to the p values in all the post hoc comparisons.¹

Hypothesis 1 predicted that participants in the two simulation conditions would report higher perceived and induced emotions than participants in the other conditions. This prediction was not supported by the results, albeit, in most cases, the motor simulation group had higher, but non-significantly different scores, compared to the other conditions (see Figure 2; and Tables S6, S7, and S8 in Supplemental materials). Also, contrary to Hypothesis 1, the vocal simulation group had significantly lowest scores overall for every stimulus (all p 's $> .05$; except for induced mellowness in the sad/tenderness piece, $p = .019$, see Tables S7 and S9 in Supplemental materials). The analysis of the responses to the Action Tendencies questionnaire was consistent with this pattern of results: there were no significant differences in action tendencies reported by the simulation groups compared to the other conditions (see Table S10 in Supplemental materials).

Hypothesis 2 predicted that participants in the distracting task condition would report the least intense affective states. The results do not support this hypothesis: emotion intensity reported by participants in the distracting task condition did not differ significantly from those in the other conditions (Figures 2–4; and Tables S6–S12 in Supplemental materials).

Imagery consistent with embodied simulation (i.e., mental images of the music being performed/played) was present in answers to the question “what went through your mind while you were listening to the music?” (see Appendix B, Table B1), but it did not represent the majority of cases. We found descriptions of this type of imagery in 14.52% of the narratives associated with the Sadness/Tenderness piece, in 8.87% of the narratives associated with the Fear/Anger piece; and in 16.13% of the narratives associated with the Joy/Excitement piece.

Hypothesis 3 predicted that the content of the participants' answers to the question “what went through your mind while you were listening to the music?” would have the same emotional content as the perceived and induced emotions reported by participants. We analyzed the content of each participant's narrative, and created two new variables to indicate whether there was a coincidence between the content of the narrative, and the participant's highest scores of perceived, and induced emotions, correspondingly. We assigned a value of 1 if the content of the narrative matched the highest perceived/induced emotion, and a value of 0, if it did not. (The first author did the first analysis of the data and codified it, and subsequently, the second author checked the consistency of the code). The hypothesis was supported by the data. As expected from the ambiguous expressive character of the music, in every piece the contents of the participants' narratives can be categorized in terms of two or three emotions with similar levels of arousal,

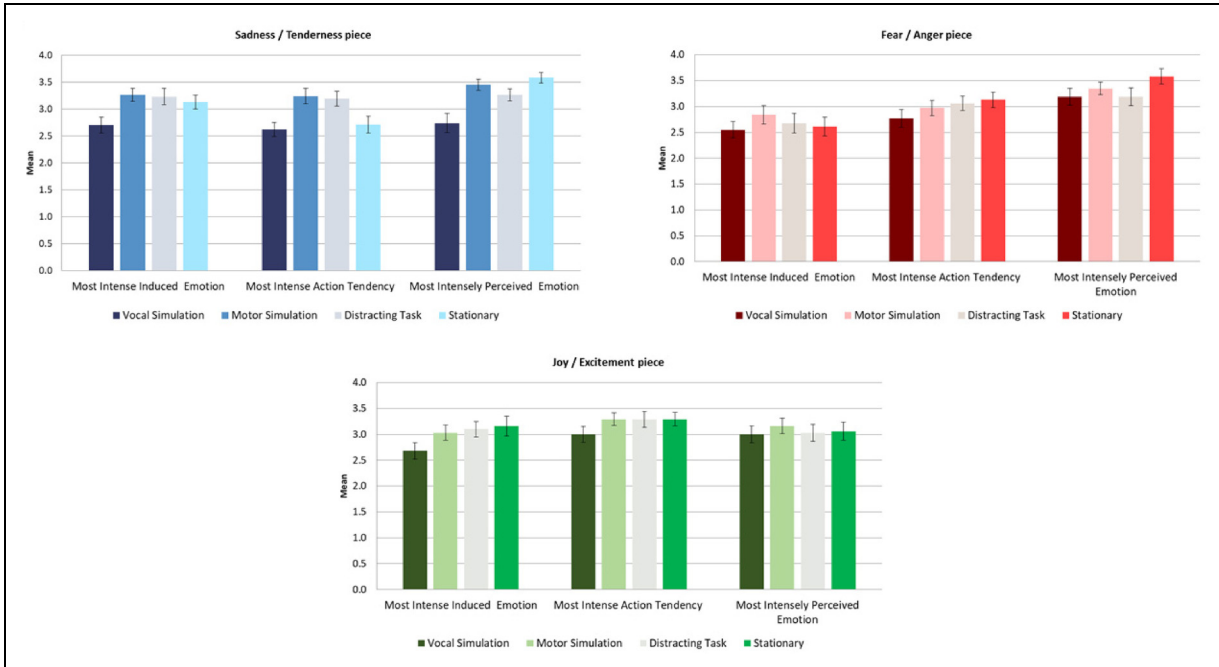


Figure 2. Experiment 1: Mean scores of most intense induced emotion, most intense action tendency, and most intensely perceived emotion as a function of experimental condition for each stimulus. (Error bars represent SE).

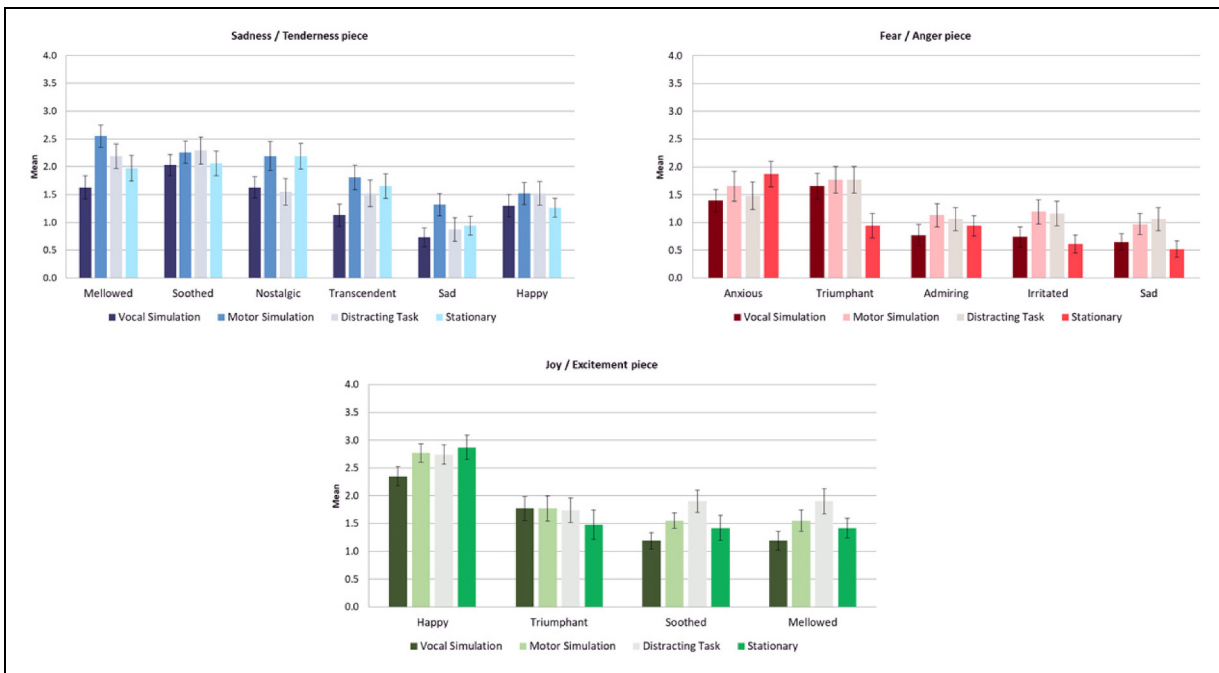


Figure 3. Experiment 1: Mean scores of most intense induced discrete emotions as a function of experimental condition for each stimulus. (Error bars represent SE).

and sometimes of valence. The coincidence rate (the frequency of narratives where the emotional content of the narrative matched the reported emotions) was: Sadness/Tenderness piece, perceived emotion = 84.68%, induced emotions 79.84%; Fear/Anger piece, perceived emotions = 91.13%, induced emotions = 72.58; Joy/Excitement piece,

perceived emotions = 77.42%, induced emotions = 69.35% (Appendix B, Table B2).

Finally, we ran additional regression analyses to examine whether covariates liking, musical engagement, and the ability to play a musical instrument mediated the results. The analyses indicated that in the Sadness/

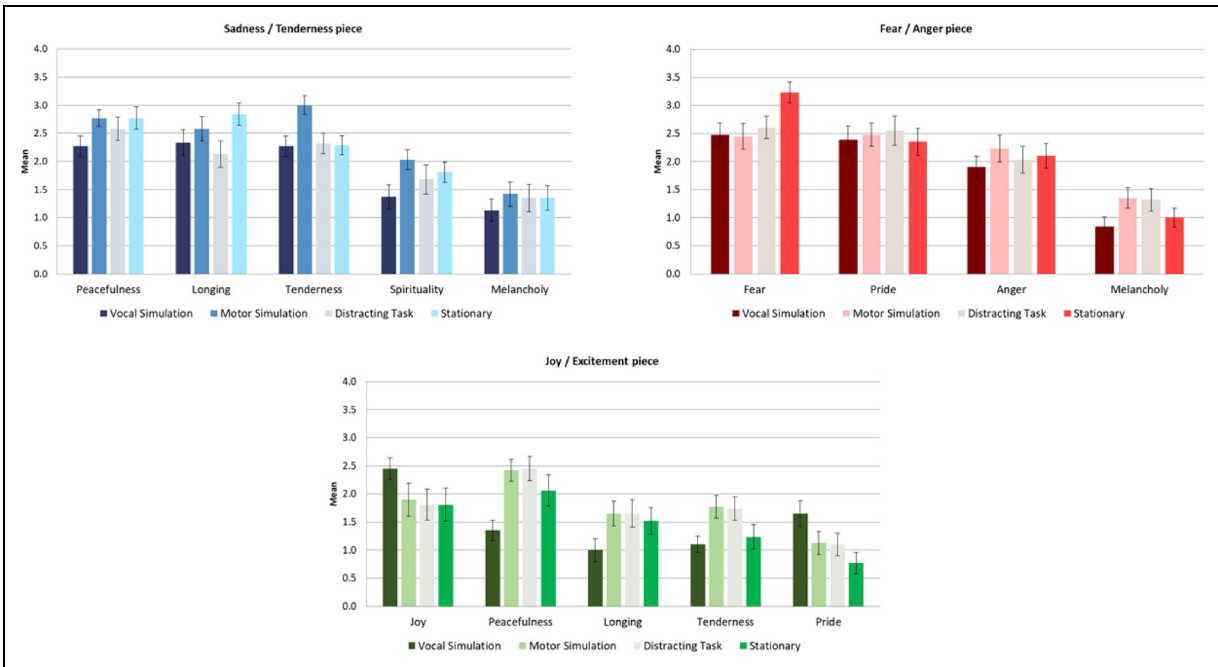


Figure 4. Experiment I: Mean scores of most intensely perceived discrete emotions as a function of experimental condition for each stimulus. (Error bars represent SE).

Tenderness and the Joy/Determination piece, the ratings of liking were significant, positive predictors of the intensity of the most intense perceived and induced emotions. In other words, the more participants liked the music, the higher their ratings of perceived and induced emotions. Curiously, in the Joy piece, the analyses indicated that the more the participants were able to play an instrument present in the piece, the less intense the induced emotions they reported. None of the other covariates introduced in the analyses (including ratings of perceived difficulty and embarrassment) were retained as significant predictors in any of the three pieces (see Supplementary Table S13).

Experiment I Discussion

The first hypothesis predicted that simulation involving mimicry of the melody, and/or of the musicians' gestures strengthens the contagion response. Additionally, it was predicted that performing a distracting activity that involved activation of motor and vocal brain areas would obstruct simulation mechanisms and lead to more subdued affective responses in comparison (Niedenthal et al., 2005). The results give little support to these hypotheses. There were very few significant differences between the groups, suggesting that the experimental manipulation did not have strong effects on the participants' emotional experience. This was particularly true in the case of the vocal simulation group, which in most of the dependent measures displayed an *opposite* trend to the predictions. We could possibly attribute this finding to the difficulty of the vocal simulation task, in which participants were

required to sing along with unfamiliar music. However, if the cognitive load of singing along was a factor this would likely have hindered enjoyment of music and should have prevented the induction of positive emotions such as joy and tenderness, but not negative ones such as anxiety, which was not the case here. Moreover, it would be counter to the fact that people commonly choose to sing along to music. Alternatively, it may be that this result reflects the fact that vocal mimicry (and other types of mimicry) are not sufficient to induce emotion. It is worth noting that two recent studies also suggest that vocal simulation does not play as big a role in music perception as proposed by embodied cognition theories: Bruder and Wöllner (2021) found that subvocalization effects were present when participants imagined music but not when they listened to it; and Weiss et al. (2020) found that engaging the vocal muscles in distracting tasks (such as chewing gum, or singing) did not have a detrimental effect on the participants' recall for vocal melodies. Similarly, our results do not support the prediction that engaging in a distracting task would have a hampering effect on the participants' affect. Moreover, individual differences in musical skills (such as the ability to play a musical instrument) did not play a significant mediating role in the dependent measures.

Does this mean that the participants' bodily behavior had nothing to do with their emotional experience? Probably not. The participants' responses reveal, for example, that in the Fear/Anger piece, the participants in the stationary condition (who were asked to remain completely still while listening to this "threatening" music) felt significantly

more scared than the participants in the other conditions, perhaps because they felt more vulnerable to the powerful events portrayed by the music. At the same time, the participants in the motor simulation condition tended to feel more “triumphant, strong”, and to perceive the piece as more expressive of “pride, power” than the rest, suggesting that making energetic movements while listening to this music facilitates experiencing oneself as the agent of the power expressed by the music. Furthermore, participants experienced relevant bodily urges while listening to the music, as evidenced by the results of the Action tendencies and Subjective Feelings questionnaire.

A number of explanations can be proposed to account for these null results. First, as mentioned earlier, the tasks proposed to the participants were difficult, particularly in the simulation conditions: although the participants had the chance to listen to the piece twice, it is still difficult to follow a piece of music and to pretend to play it, or to sing along to it if one has never heard it before. This is potentially more cognitively demanding than analogous studies which have asked people to mimic vocal or facial expressions after they have seen/heard a stimulus rather than at the same time. Furthermore, perhaps singing along with the music prevented the participants in the vocal simulation group from hearing the piece very well. Second, in the case of the distracting task group, which did not exhibit the expected hampering effect, it is possible that the participants somehow entrained their movements to the music, and therefore the task facilitated their emotional engagement with the pieces, rather than prevented it. Third, we used explicit forms of mimicry whereas most theories talk about implicit or covert mimicry (a distinction noted by Cox’s (2016) theorization of “mimetic motor action” vs. “mimetic motor imagery”). Nonetheless, virtually all empirical studies of emotion contagion ask participants to mimic the emotional expressions they observed (Hatfield et al., 2014) and, on the basis of this previous research, asking participants to sing or to perform other kinds of music production movements should have an enhancing effect rather than the opposite.

The fourth, more optimistic interpretation is that at least in the case of the motor simulation condition, the null results were due to lack of statistical power (on average, observed power was 0.41, and ranged from 0.12 to 0.69). The participants in this group displayed the predicted trend in 52% of the evaluated variables, suggesting that pretending to play the musical instruments that we listen to has a positive effect on the intensity of our emotional responses, but that this effect is very small, and therefore the statistical tests did not detect it.

The fifth interpretation is that embodied simulation is a necessary, but not sufficient condition for the perception and induction of musical emotions. That is, even though it is probable that perceiving sounds as “music” involve embodied simulation, the effects of this internal mimicry are restricted to facilitating the *perceptual* experience of sounds as organized, intentional, human-produced musical sounds (Godøy & Leman, 2010; Launay, 2015; Leman &

Maes, 2014). These effects would not necessarily extend to producing affective responses to the music. Consequently, we propose that emotional responses would only happen when these (implicitly or explicitly) mimicked physical gestures and sounds have an emotional connotation or emotional relevance for the listener. The second experiment explores this possibility.

Analyses of the participants’ subjective experiences largely support **Hypothesis 3**, which predicted a correspondence between the contents of the participants’ narratives and their scores in the questionnaires. In other words, participants reported experiencing mental imagery, and evoking abstract concepts that matched the quality of the emotions they perceived as expressed by the music and the induced emotions they felt. A potential limitation of our design is that participants reported imagery at the end of all trials rather than after each in turn (to avoid biasing their attention) and may therefore have misremembered or confused imagery across stimuli. However, no participants reported such confusions, and only a few minutes had passed between the first stimulus and the gathering of data on imagery.

It is impossible to determine whether the narratives and imagery produced the observed emotional responses, or the aroused emotional states triggered the evoked imagery because we did not manipulate the participants’ imagery and associations with the music. However, a recent investigation found that participants perceived and experienced induced emotions *before* forming mental images while listening to music (Day & Thompson, 2019), and another found that a distractor task reduced the prevalence and vividness of imagery while listening to music, but had minimal impact on the emotion felt (Hashim et al., 2020). Therefore, our interpretation is that these narratives and imagery were components of the participants’ emotional reactions, not their primary cause (Clore & Ortony, 2013). We propose that both the participants’ emotional responses, and the imagery they evoked, were at least partially caused by another underlying mechanism: the activation of semantic associations while listening to the music. Support for this interpretation can be found in the observation that about 25% of the participants did not include any emotional terms nor connotations in their answers to the abovementioned open question, but they still chose the same emotional adjectives in the questionnaires as those participants who used these kinds of terms in their answers. The second experiment tested this conjecture by asking the participants to read narratives about the musical pieces before listening to them and investigating any associated differences in emotion reported.

Experiment 2

This experiment had two aims: first, it represented a further attempt to test the hypothesis that embodied simulation facilitates the perception and induction of musical emotions, while mitigating potential limitations of the previous design. We carried out this study as a web-based experiment, eliminating the potentially embarrassing or

distracting presence of the experimenter (Egermann et al., 2009; Tesoriero & Rickard, 2012; Witek et al., 2014), and using a task involving a more implicit version of mimicry (Barsalou, 2008); participants were instructed to *imagine* themselves performing one of two tasks—either to imagine themselves as musicians playing the instruments (simulation condition) or to imagine themselves as sound engineers required to check the quality of the recording (distracting task). This mental simulation is validated by neuroimaging studies that have concluded that imagining and planning motor actions activates the same brain areas as actually performing the movements (Bangert et al., 2006; Jeannerod, 1995; Jeannerod & Frak, 1999; Zatorre & Halpern, 2005).

Second, Experiment 2 examines the role that contextual information has in shaping listeners' perceived and induced emotions. According to a constructionist account, motor actions acquire emotional meaning by virtue of the context in which they occur: for example, a smile can be interpreted as a sign of joy at a birthday party, or a sign of contempt in the middle of a heated discussion (Barrett et al., 2011; Hoemann et al., 2019). Moreover, the constructionist account proposes that musical sounds are able to express core affect (i.e., valence and arousal), not basic emotions (Cespedes-Guevara & Eerola, 2018). Therefore, according to these constructionist theories, and in contrast to the BRECVEMA framework, the emotional states that embodied simulation induce in listeners can therefore be restricted to changes in feelings of pleasure/displeasure (valence) and activation (arousal), which in turn may be shaped into particular emotions according to the characteristics of the sociocultural and personal meaning of the context in which music listening happens (Cespedes-Guevara, 2021).

Therefore, we expect that any emotionally biasing effects of embodied simulation would be dependent on the presence of a relevant context. To our best knowledge, only a handful of experimental studies have examined the effect of contextual information on emotions induced by music. In one study, Miu and Balteş (2012) presented video clips of opera songs with subtitles of the lyrics to two groups of participants: one group was instructed to “imagine vividly how the performer felt”, and another was asked to adopt an “objective perspective towards what was described in the music” (p. 3). They found that participants in the first group had stronger physiological responses and more intense induced emotions than participants in the second one. In another study, Vuoskoski and Eerola (2015) asked three groups of participants to listen to music expressive of sadness; the first group listened to the music without any previous information, the second read an emotionally neutral narrative about the music, and the third read a sad narrative about it. The researchers found that, compared to participants in the other groups, participants who read the sad narrative experienced greater induced sadness and reported increased thoughts containing sadder imagery. The authors interpreted these results as stemming from the activation of the visual

imagery mechanism proposed in the BRECVEMA framework (Juslin & Västfjäll, 2008). Similarly, O'Neill and Egermann (2020) compared three groups of participants: one group read descriptions about the composer's feelings while writing the music, another read neutral descriptions about the music, and the third did not read any descriptions. The researchers found that, compared to the other two groups, participants who read the emotional descriptions reported higher levels of induced valence and arousal which coincided with the music's expressive qualities. For these researchers, this effect was evidence that the contextual information led participants in the first group to empathize with the composers.

Taken together, these findings suggest that providing contextual information about the music can *intensify* the listeners' affective responses to it. However, they do not indicate whether that information can modulate the *type of discrete emotions* experienced by listeners. In contrast, based on the assumption that music expresses valence and arousal but not discrete emotions (Cespedes-Guevara & Eerola, 2018), in Experiment 2, we predicted that providing contextual information would shape the listeners' affective responses, producing discrete emotions that match the emotional quality suggested by said context.

Additionally, in Experiment 1, we found that most participants evoked mental visual images while listening to the music, albeit unlike Vuoskoski and Eerola (2015), we suggest that this phenomenon is secondary to the activation of semantic knowledge while listening to the music. From our constructionist perspective, we infer from this that a relevant emotional context for simulation is needed and that providing such a context should help intensify the effects of any simulation in creating an emotion, enabling us to observe an effect even if small.

Hypotheses

We predicted that, compared to the participants who perform the non-simulation task, the participants in the simulation condition would experience more intense perceived and induced emotions while listening to the music (Hypothesis 4). In addition, this experiment aimed to explore the contribution of visual imagery/semantic associations to emotion induction by testing the effect of manipulating information about the music on emotion contagion; we wanted to check whether emotion contagion is enhanced if bodily movements are given a meaningful context. Following a similar procedure to that used by Vuoskoski and Eerola (2015) and O'Neill and Egermann (2020), before listening to each piece, the participants read a description intended to bias their perceived and aroused emotions in a coherent manner (Hypothesis 5). However, unlike those experiments, we tested this effect more extensively: for each piece, we compared the effect of a neutral description versus the effect of two different emotional descriptions, whose content matched the arousal level of the music.

The final hypothesis is based on the argument that the results from Experiment 1 were in part due to the fact that the mimicked gestures performed by the participants lacked an emotional connotation for the listeners. Thus, we predicted an interaction: those participants who perform the simulation task and read the emotional descriptions of the pieces would experience more intense emotions than those participants who perform the non-simulation task and who read the neutral descriptions (Hypothesis 6).

Method

Design. The experiment used a between-subjects design, with two dependent variables: perceived and induced affective states, and two IVs. The first IV was simulation (two levels: simulation/non-simulation), and the second IV was the type of description of each musical piece (three levels: neutral description/emotional description 1/emotional description 2). In the neutral description, the pieces were described using impersonal technical terms. In the first type of emotional description, the participants read stories about the composer's experience while writing each piece in terms that suggested sadness, fear, and joy, correspondingly. In the second type of emotional description, the stories portrayed the composer's experience in terms that suggested tenderness, pride, and determination, correspondingly. (See Appendix C for a full transcription of the descriptions).

Participants. Participants were recruited by personal invitation via email, by snowballing sampling, and by links to the study from social media websites. All participants could take part in a prize draw to win one £30 Amazon voucher.

Participants completed the experiment in their preferred language (English or Spanish), and were randomly allocated to one of six different experimental conditions. Of the 447 participants who completed the experiment, 212 were excluded from the analysis (47% of the initial sample), leaving a final sample of 235 individuals. This represents an exclusion rate comparable to other web experiments (Egermann et al., 2009). Exclusion was based on time taken to do the experiment (1 *SD* above or below the mean duration (<19 min or > 42 min) and performance in four self-report items designed to control for attention and commitment to the experiment.

Stimuli. We used shorter versions of the same three instrumental pieces used in Experiment 1² (c.60 s) to prevent participant dropout. The pieces were edited such that every participant heard them twice in a row. As mentioned above, we created three descriptions for each piece: two descriptions suggesting that the piece was composed during an emotionally important episode in the composer's life, and one describing the piece in emotionally neutral terms (Appendix C).

Measures. We used the same set of questionnaires as in Experiment 1 to measure the participants' affective responses. They also reported their liking, familiarity with the piece, and how difficult they found it to follow the experimental instructions. After each trial, the participants were asked to write down a summary of what went through their minds while listening to the music. At the end of the experiment, the participants completed a questionnaire about demographic information, and their musical engagement and training, including any musical instruments they could play.

Procedure. The procedure followed the same format as Experiment 1. The instructions for the participants in the simulation condition were the following:

Please listen to the piece while imagining that you are one of the musicians playing the music. (You can choose to imagine playing only one of the instruments, or if you prefer, you can imagine switching from one instrument to the other as the music progresses). Please avoid moving, tapping, dancing or singing while listening to the music.

The instructions for the participants in the non-simulation condition were:

Please listen to the piece while imagining that you are a sound engineer, who is in charge of checking that the recording does not contain any glitches or errors, before it is copied to a CD. Please avoid moving, tapping, dancing or singing while listening to the music.

Results

The 212 participants included in the analysis had a mean age of 28.8 years (*SD* = 9.43); (Female = 58.7%, Male = 40.4%, Other = 0.9%). They had 26 different mother tongues, but most of them had either Spanish (47.23%) or English (28.09%) as their first language. Participants came from 42 different nationalities. The English version was completed by 128 participants (54.5%) and the Spanish version by 107 (45.5%). The participants had a relatively high level of musical engagement and training: more than 65% reported spending a lot of free time in music-related activities, and "not being able to live without music"; more than half had received at least 3 years of musical training, and 78.7% of them reported being able to play at least one musical instrument.

Manipulation Check. The musical pieces elicited the intended perceived and induced emotions. The results from the open-ended questions about what the participants thought about while listening to the music indicated that very few participants (equivalent to 2.5% of the sample) considered the descriptions incongruent with the music. Moreover, most participants who read the neutral descriptions evoked narratives and mental images that were

compatible with the descriptions provided to the other two groups (see Appendix D). This suggests that the provided emotional descriptions matched the type of visual imagery and semantic associations that a listener might spontaneously evoke while listening to these pieces.

We observed high and significant correlations between perceived and induced emotions: overall, the Pearson correlation coefficients range from .20 to .67 (all p values $< .005$). (See Supplementary Table S14). These results are consistent with the condition for emotional contagion with music to have occurred. However, as in Experiment 1, we found that participants in all conditions reported a variety of emotions that were not restricted to basic emotions.

Hypotheses Testing. We used the same procedure described in Experiment 1 to calculate the most intense induced emotion, the most intense action tendency, and the most intense perceived emotion reported by the participants. None of these dependent variables was normally distributed. However, given the large size of the sample, it can be assumed that these characteristics are unlikely to be due to sampling errors. At the same time, the variables do not display the same degree or type of skewness; therefore applying data transformations to all the data was not viable. Given that ANOVA is a robust test when the normality assumption is not met (Finch, 2005), and that there are no nonparametric alternatives to two-way ANOVA, we ran two-way Factorial ANOVA tests to analyze the data, with simulation condition (non-simulation/simulation) and type of description (neutral/emotional description 1, emotional description 2) as IVs, and perceived and induced emotions as dependent variables

for each piece separately. Pairwise comparisons were Bonferroni-adjusted for multiple testing.

An analysis of the main effect of the type of simulation showed that, as predicted by **Hypothesis 4**, in the majority of cases, the participants in the simulation group reported more intense affective states than participants in the non-simulation group for all the pieces, but none of the differences was statistically significant (all p values $> .05$) with the exception of perceived emotion with the Joy/Determination piece (see Figure 5 and Supplementary Table S15).

Hypothesis 5 predicted a main effect of type of description, such that there would be a correspondence between the content of the description of the piece and the perceived and induced affective responses. Responses to the Sadness/Tenderness piece supported this prediction. Participants who read the sadness description experienced stronger induced sadness than participants who read the tenderness description ($p < .000$). The sadness description group also had higher ratings of perceived melancholy than the tenderness description group ($p < .000$), and than the neutral description group ($p = .033$). Also as predicted, the ratings of perceived tenderness were the highest in the tenderness description group when compared to the sadness description group ($p = .001$), and to the neutral description group ($p < .000$) (Figures 6 and 7; and Supplementary Tables S17, S19, and S20). Participants in the tenderness description group also experienced positive action tendencies more intensely, and negative ones less intensely than participants who read the other two descriptions: their scores of feeling-in-command-of-the-situation were higher compared to the sadness description group ($p = .002$) and

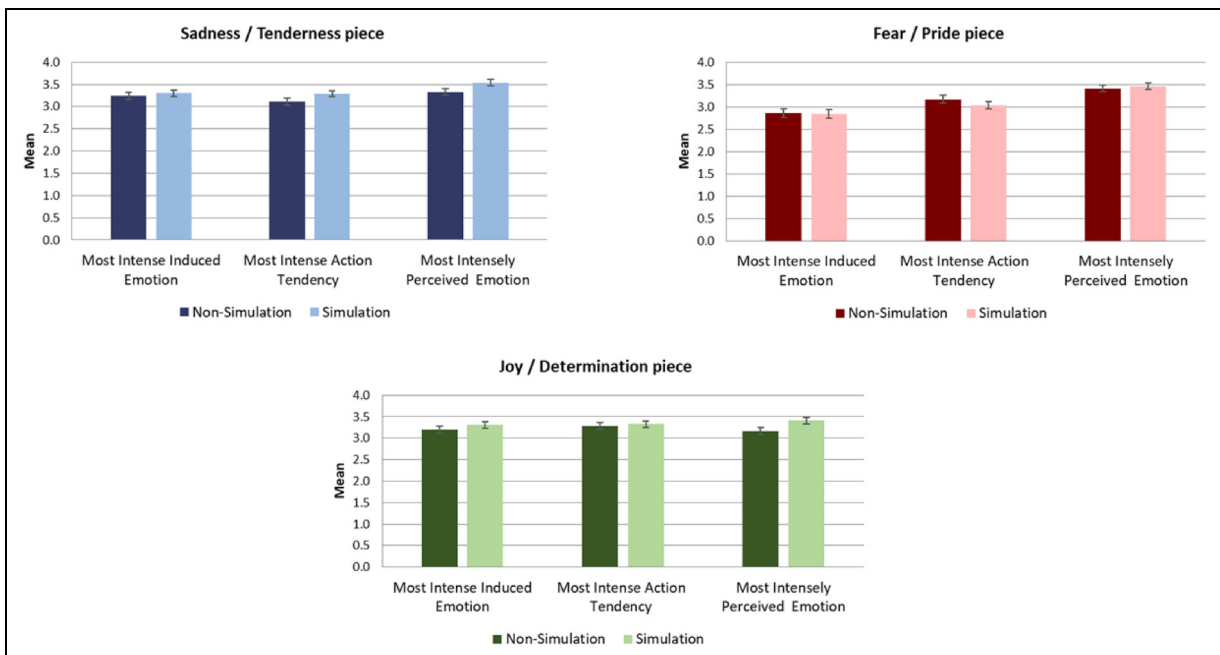


Figure 5. Experiment 2: Mean scores of most intense induced emotion, most intense action tendency, and most intensely perceived emotion as a function of experimental condition for each stimulus. (Error bars represent SE).

to the neutral description group ($p < .000$); they had higher scores of feeling-like-everything-is-fine than the sadness description group ($p < .000$); and than the neutral description group ($p = .001$). The tenderness description group also had lower scores of needing-to-be-comforted than the sadness description group ($p = .015$), and than the neutral description group ($p = .048$). However, contrary to the prediction, participants who read the sadness description did not report higher scores of feeling-like-crying than participants who read either of the other two descriptions (Tables S18 and S21 in Supplemental materials).

In the case of the Fear/Pride piece, the main effect of type of description was consistent with **hypothesis 5**: the participants who read the fear description reported the highest levels of induced anxiety and perceived fear; and the participants who read the pride description reported the highest ratings of feeling triumphant, strong and of perceived pride. However, the differences between the groups are only statistically significant in the case of perceived fear, where the fear description group had significantly higher scores than the pride description group ($p = .007$) (Figures 6 and 7; and Supplementary Tables S16 and S17). The results of the Action Tendencies questionnaire also support hypothesis 5: the fear description group reported higher ratings of needing-to-be-comforted than the pride description group ($p = .004$), and than the neutral description group ($p = .021$). The fear description group also had higher scores of feeling frozen than the neutral description group ($p = .012$); and higher ratings of feeling-like-crying than the pride description group ($p = .011$), and than the

neutral description group ($p = .02$) (see Supplementary Table S18). A “bleeding effect” was observed: the fear description group had significantly higher ratings of induced sadness than the pride description group ($p = .008$); and higher ratings of perceived melancholy than the neutral description group ($p = .031$) (Figure 7 and Supplementary Tables S16, S19, and S20).

The analysis of the main effect of the type of description for the Joy/Determination piece revealed that the trend predicted by **Hypothesis 5** was not clearly observed. Thus, although the determination description group had the highest ratings of feeling triumphant, and higher ratings of perceived pride than the joy description group, the determination description group also reported the highest ratings of induced happiness and of induced valence. Moreover, only in the ratings of valence did the determination description group have significantly higher scores than the neutral description group ($p = .025$) (Figures 6 and 7; and Supplementary Tables S15, S16, and S19–S21). As expected, the answers to the Action Tendencies questionnaire showed that, compared to the joy description group, the determination description group had slightly higher scores. But, unexpectedly, the determination description group also had significantly higher ratings of wanting-to-make-the-experience-longer when compared to the neutral description group ($p = .022$); and of feeling-like-everything-is-fine when compared to the neutral description group ($p = .019$). These results suggest that the participants who read the determination description had the most positive affective experience while listening to this piece (see Supplementary Tables S18 and S21).

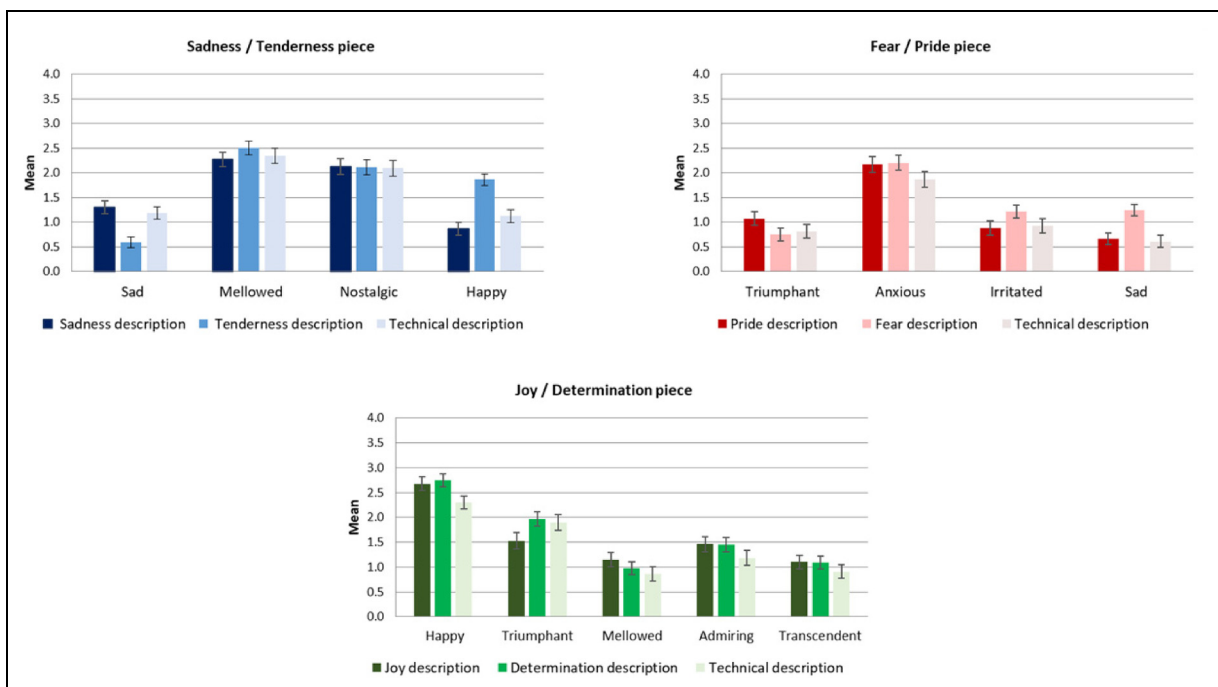


Figure 6. Experiment 2: Mean scores of most intense induced discrete emotions as a function of provided description for each stimulus. (Error bars represent SE).

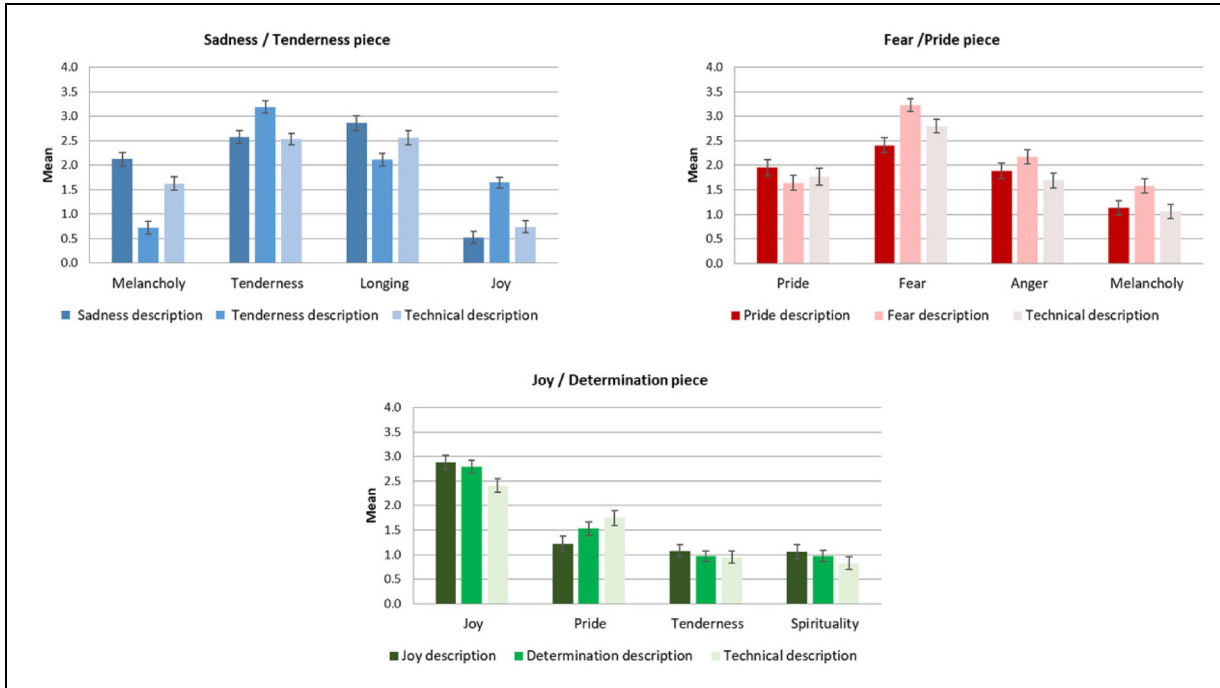


Figure 7. Experiment 2: Mean scores of most intensely perceived discrete emotions as a function of provided description for each stimulus. (Error bars represent SE).

Hypothesis 6 predicted an interaction whereby the participants in the simulation condition who read an emotional description would experience higher perceived and induced affective states than participants in the non-simulation condition who read a neutral description. The ANOVA tests showed no significant interactions between these two IVs, for any of the pieces.

Finally, as in Experiment 1, we ran additional regression analyses to examine whether covariates liking, musical engagement, and the ability to play a musical instrument mediated the results. The results indicated that, in the Sadness/Tenderness and the Joy/Determination pieces, the ratings of liking were significant, positive predictors of the intensity of the most intense perceived and induced emotions, such that the more participants liked the musical stimuli, the more intense the emotions they perceived in the music and experienced in themselves. Additionally, in the Joy/Determination piece, the ratings of musical engagement were also positive predictors of the intensity of induced emotions. None of the other covariates introduced in the analyses were retained as significant predictors in any of the three pieces (see Supplementary Table S22).

Visual Imagery and Semantic Associations Evoked by the Music and the Descriptions. We used the same procedure as in Experiment 1 to analyze the extent to which the participants' descriptions of what went through their minds while listening to the music coincided with the emotion they reported in the questionnaires. (See Appendix D, e.g., from the participants' narratives.)

In the Sadness/Tenderness piece, the coincidence rate between the participants' ratings of perceived emotions and their narratives was 65.88% and between ratings of induced emotions and narratives, the rate was 60.19%. These rates are lower than those observed in Experiment 1 (84.68% and 79.84%, correspondingly). The analysis of the presence of elements from the description in the participants' narratives (restricted to description groups 1 and 2) shows that only a third of the participants (32.62%) explicitly mentioned an element of the description in their narratives.

In the Fear/Pride piece, the coincidence rate between the participants' narratives and their scores of perceived and induced emotions was 71.86%, and 69.85%, respectively. Again, these rates are lower than those of Experiment 1 for the same piece (91.13% and 72.58%, correspondingly). The presence of elements from the provided description in the participants' narratives (description groups 1 and 2) was observed in 39.55% of the cases.

In the Joy/Determination piece, the coincidence rate between the participants' ratings of perceived emotions and their narratives was of 61.50%, and of 63% for their ratings of induced emotions. (These percentages are lower than the percentages observed in Experiment 1, which were 77.42%, and 69.35%, respectively). Elements from the provided description in the participants' narratives were present in 40.14% of the cases (description groups 1 and 2).

As in Experiment 1, many participants commented that the pieces reminded them of movie soundtracks and their

film genres rather than specific films, and only two participants correctly identified the movie source of two of the pieces.

Discussion of Results From Experiment 2

Just as in the first experiment, the results did not support the hypothesis that engaging in a mental task associated with motor simulation would lead to greater intensity of a target emotion than engaging in a mental task that prevented simulation. Nor did we find evidence supporting an interaction. In fact, while we found that reading an emotional description provided to participants had an effect on the *type* of perceived and induced emotions, in contrast to Vuoskoski and Eerola's (2015) study, it did not have a significant effect on the intensity of the emotions reported by the participants (when compared to reading a neutral description). This finding suggests that factors other than the presence or absence of emotional elements in the descriptions were more important in determining the intensity of the participants' emotional reactions (e.g., idiosyncratic associations of the music with episodic memories).

It is unlikely that these null results can be attributed to the tasks having failed in generating the intended simulation and non-simulation conditions. Previous research has shown that engaging in imagery of motor actions correlates with the activation of the same brain areas involved in performing those actions (Bangert et al., 2006; Jeannerod, 1995; Jeannerod & Frak, 1999; Zatorre & Halpern, 2005), suggesting that the task designed to facilitate simulation was valid. Moreover, many of the free descriptions provided by the participants of the non-simulation condition include comments about having noticed small glitches in the musical pieces, suggesting that they assumed the intended third-person perspective during the listening task. Moreover, it is unlikely these results are due to the participants having failed to perform the mental task adequately, because we only included in the analyses those participants who reported having followed the instructions correctly, without distractions. One limitation is that we manipulated both the presence of contextual information and changed the simulation task in this second experiment which means we can't infer which of these might individually have been responsible for the null result.

The results support the prediction that providing participants with descriptions about the composer's emotional state while writing the music would have a coherent effect on the quality of the perceived and induced emotions reported by the participants. This finding advances previous knowledge: it extends insights from Vuoskoski and Eerola (2015) by showing that contextual information can shape the type of emotions that listeners experience, and not only their intensity; and furthers the findings of O'Neill and Egermann (2020), by showing that contextual information can bias the listeners' affective responses producing discrete emotions, not only changes in valence and arousal. However, unlike O'Neill and Egermann's study, our experiment did

not find that providing emotional descriptions about the composer's feelings was associated with more intense induced emotions. These different findings may be due to the fact that O'Neill and Egermann used longer and more emotionally explicit narratives about the composer's feelings than those we used. Additionally, whereas those researchers interpreted their results as evidence of participants' empathy for the composer's feelings, we prefer to interpret ours as evidence of contagion, because fewer than half of the participants who read the emotional narratives explicitly mentioned elements from them in the descriptions of their thoughts while listening to the music (see Table B2 in the appendix for a sample of the participants' responses).

The biasing effect of context was most clearly observed in the Sadness/Tenderness piece; in the Fear/Pride and Joy/Determination pieces, the differences between the groups were not always statistically significant. This mixed pattern of results can be attributed to a range of factors: it is possible that the Sadness/Tenderness piece was more expressively ambiguous than the other two, that individual differences influenced the enjoyment of "scary" esthetic stimuli, and that the absence of an adjective in the questionnaire for "determined" meant responses were not as well captured for this item (the closest was "triumphant, strong"). These findings are consistent with the fact that positive emotional states tend to be less differentiated, and that in most languages (including English) there are fewer emotional adjectives to describe nuances in positively valenced and high-arousal emotions than to describe negatively valenced emotions (Rozin et al., 2010).

It is also worth noting that, in general terms the observed effect sizes were small (on average $\eta^2 = .03$), and that the significant differences were not always found between the neutral description group and the emotional description groups, but between the two emotional description groups. Hence, a better interpretation of the results may be that due to the artificial listening circumstances of the experiment, the descriptions had little power, but in real-life listening circumstances, the listener's motivation and opportunity to find meaning for the music is likely to be stronger due to the many contextual components of the listening experience (e.g., the paratextual information provided with musical artifacts, including cover imagery, booklets, lyrics, details about the musician's biography, meaning of the social occasion, etc.).

General Discussion

Embodied Simulation as a Weak Emotion Induction Mechanism

Taken together, the results of these two experiments suggest that if embodied simulation plays a role in emotional experiences with music, this role is small, and probably masked by the effect of other mechanisms simultaneously activated while listening. Indeed, the most influential contemporary theories of the induction of musical emotions, namely, the

BRECVEMA framework (Juslin, 2013a) and the Multifactorial Process Model (Scherer & Coutinho, 2013) propose that there are multiple routes to emotion induction, but they have not fully specified under which circumstances each mechanism has predominant effects over others in the emotion-eliciting process. Our results give some indication of the strength of visual imagery (and/or semantic associations) relative to embodied simulation, using self-report questionnaires aligned with current theories of musical emotions: established measures of induced and perceived emotions, capturing both discrete and dimensional emotion constructs, and with the addition of a novel questionnaire to measure action tendencies. Future studies should investigate which factors in the music, the person, and the situation make embodied simulation the maximal mechanism for emotion induction. This would enhance our understanding of why, for example, some people seem to derive great pleasure from “air-playing” while listening to music. Is this mimicry the cause, the consequence of the emotional reaction, or both? Is this effect only possible when music evokes positive emotional states? Is it due to a sense of enhanced personal agency, to a sense of synchronization with another real or virtual human being (Launay et al., 2013), or even to a sense of merged subjectivity with the music (Clarke, 2014)? To what extent does this pleasurable experience relate to the listener’s real ability to play the instruments? These are all questions that await an empirical answer.

Our claim is not that embodied simulation is not involved in music perception. The close link between perceiving, predicting, and executing motor actions has been demonstrated in several behavioral and neuroimaging studies (Chen et al., 2008; Leman et al., 2009; Stupacher et al., 2013), suggesting that perceiving musical sounds involves the activation of internal mimicry mechanisms manifesting an implicit notion of music as an activity produced by human agency (Launay, 2015). Indeed, this is not the only way to conceive of the role of overt and covert mimicry but this lies beyond scope of this paper which aimed to evidence the emotion contagion mechanism construct. While it is probable that embodied simulation of motor actions plays a central role in music perception, from our results, we can conclude that if this mechanism has any consequences for the elicitation of emotional experiences with music at all, its effects are very small. We articulate two arguments for this conclusion in the following paragraphs.

The first argument is that having a first-person notion of the motor actions involved in playing a musical instrument does not involve perceiving those actions and their associated sounds as embedded with emotional meanings. For example, it is conceivable that embodied simulation helps us understand that to make a musical instrument sound loud, the musician has to make a powerful bodily movement, but this implicit understanding of the immediate goal behind the musician’s action does not equate to inferring that the loud sound intends to communicate an emotion of anger, joy, despair, fear, or hope, etc. According to a

constructionist account, motor actions, like vocalizations and facial gestures, only acquire emotional meaning when they are placed in relation to the wider context in which they are observed and produced. There is empirical evidence about this in the case of perception of faces (Barrett et al., 2011; Carroll & Russell, 1996) and emotion perception of vocal sounds (Liuni et al., 2020). Moreover, simply doing a mental experiment similar to the one proposed by Jacob and Jeannerod (2005) demonstrates that the same principle can be applied to the case of bodily gestures and musical actions: consider how the same observed action (e.g., frowning, cutting someone’s abdomen with a scalpel, playing a trumpet loudly) has very different emotional goals and meanings according to the context in which they occur (frowning can communicate anger to an adversary, or physical exertion when lifting a heavy weight; cutting someone with a scalpel can be done by a psychopath torturing another person, or by a surgeon performing a surgery; playing a loud note in a trumpet loudly can have the intention of communicating joy, but also anger, etc.). This issue of whether bodily mimicry is sufficient to induce emotion is central to the design of behavioral studies of emotion, whether that be singing or playing along with music as in our study, or reproducing facial or vocal expressions as done in other studies. Although embodied simulation explains how we perceive actions as produced by human agents using particular bodily movements, it cannot *by itself* explain how we perceive the *emotional intentions* behind the actions. Inferring that intention requires processing more information about the observed person (or about the music) and the context where the action (or music) takes place.

In the context of music listening, this argument has to be refined even further. The results from the second experiment suggest that the presence of a relevant emotional context, such as learning about the emotional intentions of the composer, is not enough to reinforce any effects of simulation. The participants who engaged in simulation and read emotional descriptions of the pieces did not report more intense emotions than the participants who did not engage in simulation and read neutral descriptions. It may be that if simulation is to have a reinforcing effect on the listeners’ emotional experience, it is necessary that they map the simulated movements and melodies onto emotional meanings. For instance, it is not enough that the listeners perceive that the piece expresses anger in general, it is necessary that they associate the specific movements that make the instruments sound loud and fast with the experience of producing an aggressive discharge of physical power against a rival.

The second argument for the claim that motor simulation plays a small role in musical emotions is that theories like that of Davies (1994, 2013), Jackendoff and Lerdahl (2006), Molnar-Szakacs and Overy (2006), and Overy and Molnar-Szakacs (2009) reduce emotional experiences to behaviors or gestures, arguably mistaking the part for the whole. Gestures and expressive behaviors are merely *one* of the components of emotional experiences, which always include the evaluation

of an event as personally relevant for our goals within a given context (Clore & Ortony, 2013; Scherer, 2005). It follows that simulating or mimicking gestures or behaviors can only be at most one contributing mechanism to the perception and induction of musical emotions among others.

This reasoning can also help explain the finding that mimicry has limited effects on emotional elicitation. Both the experiments reported here, and previous studies in which participants mimicked observed emotional expressions, have found that this manipulation facilitates and biases the perception and induction of coherent affective states (i.e., changes in valence and arousal), but it does not lead to the induction of full-blown, discrete emotional experiences (Flack, 2006; Hatfield et al., 1995; Hess & Blairy, 2001; McIntosh, 2006; Neumann & Strack, 2000). This is true even in studies where participants observed and mimicked facial expressions, which are the type of stimuli with the greatest ability to communicate affective states (Russell et al., 2003). Hence it can be expected that more ambiguous stimuli like vocalizations, bodily gestures, and musical sounds, should be even less effective in communicating and inducing discrete emotions via mimicry mechanisms (Cespedes-Guevara & Eerola, 2018).

Visual Imagery as an Outcome of Semantic Associations

These studies also shed light on the role of visual imagery in emotion induction by music. In line with previous studies (Hashim et al., 2020; Küssner & Eerola, 2019; Taruffi et al., 2017; Vuoskoski & Eerola, 2015), in both our experiments, we found that the participants' descriptions of their thoughts while listening to the music indicate that the majority of them experienced visual images, whose content coincided with the emotions they experienced. For Vuoskoski and Eerola (2015), this visual imagery mechanism was activated by an interaction of the music and the provided narratives, and led to the induction of emotional responses in the participants. In our view, the results of the present experiments make this explanation plausible, but incomplete.

Our interpretation is that while the visual imagery evoked by the participants might have contributed to their emotional experiences, this visual imagery was a consequence of a more basic process: the activation of semantic information primed by the musical materials, and by the descriptions (Dibben, 2003; Herget, 2021; Koelsch et al., 2004). We base this conclusion on the following observations: first, it is possible that the visual imagery that the participants experienced was simply a consequence of the fact that the experiments did not provide them with any meaningful visual stimulation while listening to the music (Thompson & Coltheart, 2008). Second, there is evidence that listeners perceive music-expressed emotions and experience music-induced emotions *before* they experience visual imagery (Day & Thompson, 2019). Third, although there were high levels of agreement in the emotions they reported,

both experiments revealed individual differences within the narratives and imagery. Moreover, in the second experiment, less than 41% of the sample explicitly mentioned (at least some) elements from the descriptions provided. This variation suggests that personal associations were powerful factors that drove the participants' affective experience (Küssner & Eerola, 2019). However, at the same time, we interpret the high levels of agreement in the reported emotions as a consequence of two factors: first, the affective qualities of the musical stimuli (Taruffi & Küssner, 2019), and second, the priming of idiosyncratic narratives and imagery by shared cultural knowledge activated by each piece (and by the provided descriptions, in the second experiment).

What Kind of Affective States can be Induced by Music Through Contagion?

According to the BRECVEMA framework (Juslin, 2013a), observing a correspondence between perceived and induced emotions is explained by an automatic process of internal mimicry whereby perceiving an emotion expressed by the music leads to the induction of the same discrete emotion. This framework proposes that this mechanism leads to the induction of *basic emotions*, because the phylogenetic importance that these emotions have had in our evolution as a species is reflected in the way we perceive emotions in vocalizations and in music (Juslin, 2013b). If this account is correct, it follows that basic emotions should be more easily expressed, perceived (Juslin, 2019, p. 182), and aroused by music than other emotions, and that it should be less possible to change the type of emotion perceived and induced by manipulating the contextual information that the listener has access to.

The results of the present experiments do not support these predictions, for several reasons. First, both experiments showed a "bleeding" effect: most participants did not choose only one category to rate their perceived and induced emotions; they chose several categories which were always compatible in terms of valence and/or arousal. While these results could be attributed to the ambiguous character of the stimuli used in these experiments, it should be noted that this effect has been found in many experiments on mimicry of facial and vocal expressions (Flack, 2006; Hatfield et al., 1995; Hess & Blairy, 2001; McIntosh, 2006; Neumann & Strack, 2000), and in previous experiments on induction of musical emotions in which the researchers used less ambiguous stimuli (e.g., Juslin et al., 2014; Kallinen & Ravaja, 2006; Lundqvist et al., 2009).

The second argument against the idea that basic emotions can be induced through emotional contagion is that the participants readily chose nonbasic emotions to describe both the affective states expressed by the music, and aroused in themselves (e.g., nostalgia, pride, transcendence, and mellowness). This cannot be attributed to the effect of the provided emotional descriptions in Experiment 2, because this effect was also observed in the group of

participants who read the neutral descriptions, and in Experiment 1, no descriptions were provided. It seems that just as it has been found in research with the perception of faces and voices (Russell et al., 2003), concluding that basic emotions are privileged in the participants' answers depends largely on using only basic emotions in the list of adjectives they have to choose from.

Third, the results of these two studies indicate that appraisal mechanisms mediated the participants' responses, producing emotions that did not correspond to the perceived emotion. The best example of this was the finding that many participants perceived the Fear/Pride piece to be expressive of "anger", and reported feeling "anxious". These responses cannot be attributed to mimicry (which would lead to induced "irritation" or feelings of "power, pride") but to an evaluation of the musical stimulus as "threatening", which elicited a defensive response of fear or anxiety.

One counterargument to this from the perspective of the BRECVEMA framework, as with the Basic Emotion theory generally (Ekman, 1992a; Izard, 2009; Panksepp, 1992), would be that this type of variability in arousal of participants' felt responses is to be expected because musical emotions (like emotions in general) interact with other mechanisms and contextual factors producing nuances in what was originally a basic emotion (as argued by Juslin (2013b) in relation to expressed and perceived emotion). However, unless these interactions are fully specified, one would be in danger of resorting to an *ad hoc hypothesis* (Popper, 1935/1992).

An alternative, more parsimonious alternative is to assume that music is only capable of expressing diffuse affective states and therefore, if emotional contagion with music happens, we should expect to observe correspondence of expressed and induced emotions in terms of variations in core affect (i.e., valence and arousal), but not of discrete emotions (Cespedes-Guevara & Eerola, 2018). In fact, several authors have used the term "contagion" in this sense in their research (e.g., Egermann & McAdams, 2013; Evans & Schubert, 2008). In our view to advance research of this phenomenon, it is necessary that theoretical assumptions are explicit; researchers should specify if they are using the term "emotional contagion" in a merely descriptive sense (without any commitment to explaining how the observed correspondence is produced), or if they are using the term in a theoretical sense.

Importantly, assuming that music can only express fluctuations of core affect does not imply that when people listen to emotionally expressive music, they only experience general affective states such as feeling pleasure/displeasure and calm/relaxation; the results of our experiments clearly show that this was not the case. Our proposal, based on constructionist theories such as Barrett's (2006), is that listening to expressive music often changes the listener's core affect (making the individual feel more or less pleasure and relaxation), and these changes in core affect can become a variety of perceived and induced emotions given the right circumstances. More precisely, on those occasions when the listener

meaningfully associates the changes in core affect with contextual and/or personal information, the person has the experience of perceiving discrete emotions in the music, and discrete induced emotions as well. In the case of our experiments, that contextual information was provided by the response options that the participants found in the questionnaires, the imagined visual narratives that they evoked while they listened to the music, the information we gave them about the composer's intentions when writing the music (Experiment 2), and by semantic associations that the music primed for participants.

This approach helps explain why, although we observed "emotional contagion" in a general sense (i.e., a correspondence between perceived and induced emotions), we also observed a variety of perceived and induced affects which, while not restricted to a set of basic emotions, was not idiosyncratic or random. On the contrary, our results indicate that the emotional experiences of our participants were systematically constrained both by the variations of valence and arousal that music could express, and by the abovementioned contextual factors.

We propose that this constructionist theoretical proposal can be useful for devising further studies that explore the relative contribution of acoustic factors in the music (such as the resemblance of its structure to the expression of affective vocalizations), psychological mechanisms in the listener (such as personal memories and culturally shared semantic associations), and the sociocultural context (such as the cultural meaning of a musical event) in the phenomena of perception and induction of musical emotions. Future research from this constructionist perspective could trace the ways in which these factors interact with core affect to produce particular affective experiences.

Conclusion

The results of these two experiments suggest that embodied simulation of melodies and of the implied motor actions performed by the musicians do not play a substantial role in the phenomenon of emotional contagion with music. They indicate that emotional contagion with music is a mediated phenomenon where factors such as contextual information about the composer, personal associations, and socially shared semantic concepts shape the quality of the perceived and induced emotions that listeners experience. Moreover, we interpret these results to propose that rather than being a direct mechanism of emotion induction, visual imagery experienced while listening to music (and arising from semantic associations primed by musical materials and accompanying textual descriptions in our experiments) acts as a contextual factor helping shape core affect into particular emotions.

Action Editor

Jin Hyun Kim, Humboldt-Universität zu Berlin, Musicology and Media Studies.

Peer Review

Mats Küssner, Humboldt-Universität zu Berlin, Musicology and Media Studies.

Two anonymous reviewers.

Author Contributions

JC and ND researched literature, and conceived the study. JC designed the study, gained ethical approval, recruited participants, ran the study, and ran data analyses. JC wrote the first draft of the manuscript. Both authors reviewed and edited the manuscript and approved the final version of the manuscript.

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) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a doctoral scholarship from the University of Sheffield (UK); and by a doctoral scholarship from Colombia's Departamento Administrativo de Ciencia, Tecnología e Innovación, COLCIENCIAS.

Supplemental Material

Supplemental material for this article is available online.

ORCID iD

Julian Cespedes-Guevara  <https://orcid.org/0000-0002-8816-3650>

Notes

1. To avoid making Type I errors, we also analyzed the data using nonparametric tests, and we found the same pattern of results. We decided to report the results of the parametric tests since they have more statistical power.
2. The pieces are the same as in Experiment 1, but due to the content of the descriptions that the participants read, in this experiment, we call the second and third pieces by slightly different names: "Fear/Pride" (instead of "Fear/Anger") and "Joy/Determination" (instead of Joy/Excitement).

References

- Alluri, V., Toiviainen, P., Burunat, I., Kliuchko, M., Vuust, P., & Brattico, E. (2017). Connectivity patterns during music listening: Evidence for action-based processing in musicians. *Human Brain Mapping, 38*(6), 2955–2970. <https://doi.org/10.1002/hbm.23565>
- Anglada-Tort, M., Steffens, J., & Müllensiefen, D. (2019). Names and titles matter: The impact of linguistic fluency and the affect heuristic on aesthetic and value judgements of music. *Psychology of Aesthetics, Creativity, and the Arts, 13*(3), 277. <https://doi.org/10.1037/aca0000172>
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H. J., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage, 30*(3), 917–926. <https://doi.org/10.1016/j.neuroimage.2005.10.044>
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2
- Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1–23. <https://doi.org/10.1093/scan/nsw154>
- Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11*(8), 327–332. <https://doi.org/10.1016/j.tics.2007.06.003>
- Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286–290. <https://doi.org/10.1177/0963721411422522>
- Barrett, L. F., Mesquita, B., Gendron, M., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science, 21*(4), 595–599. <https://doi.org/10.1177/0956797610363547>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Brown, S., & Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition, 63*(1), 59–69. <https://doi.org/10.1016/j.bandc.2006.08.006>
- Bruder, C., & Wöllner, C. (2021). Subvocalization in singers: Laryngoscopy and surface EMG effects when imagining and listening to song and text. *Psychology of Music, 49*(3), 567–580. <https://doi.org/10.1177/0305735619883681>
- Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: Brain regions involved with perception and covert production. *NeuroImage, 31*(3), 1327–1342. <https://doi.org/10.1016/j.neuroimage.2006.01.036>
- Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205–218. <http://www.ncbi.nlm.nih.gov/pubmed/8636880>. <https://doi.org/10.1037/0022-3514.70.2.205>
- Cespedes-Guevara, J. (2021, November 7). Cespedes-Guevara, J. - A constructionist theory of music induced emotions. <https://doi.org/10.31234/osf.io/sfzm2>
- Cespedes-Guevara, J., & Eerola, T. (2018). Music communicates affects, not basic emotions—a constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology, 9*, 215. <https://doi.org/10.3389/fpsyg.2018.00215>
- Chan, L. P., Livingstone, S. R., & Russo, F. A. (2013). Facial mimicry in response to song. *Music Perception, 30*(4), 361–367. <https://doi.org/10.1525/mp.2013.30.4.361>
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854. <https://doi.org/10.1093/cercor/bhn042>
- Clarke, E., Dibben, N., & Pitts, S. (2010). *Music and mind in everyday life*. Oxford University Press.

- Clarke, E. F. (2014). Lost and found in music: Music, consciousness and subjectivity. *Musicae Scientiae*, 18(3), 354–368. <https://doi.org/10.1177/1029864914533812>
- Clore, G. L., & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review*, 5(4), 335–343. <https://doi.org/10.1177/1754073913489751>
- Cochrane, T. (2010). A simulation theory of musical expressivity. *Australasian Journal of Philosophy*, 88(2), 191–207. <https://doi.org/10.1080/00048400902941257>
- Colling, L. J., & Thompson, W. F. (2013). Music, action, and affect. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music, multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 197–212). Oxford University Press.
- Cox, A. (2011). Embodying music: Principles of the mimetic hypothesis. *Music Theory Online*, 17(2), 1–24. <https://doi.org/10.30535/mto.17.2.1>
- Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Indiana University Press.
- Davies, S. (1994). *Musical meaning and expression*. Cornell University Press.
- Davies, S. (2013). Music to listener emotional contagion. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music, multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 169–176). Oxford University Press.
- Day, R. A., & Thompson, W. F. (2019). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*, 29(2–3), 75–89. <https://doi.org/10.1037/pmu0000220>
- DeNora, T. (2000). *Music in everyday life*. Cambridge University Press.
- Dibben, N. (2003). Musical materials, perception, and listening. In M. Clayton, T. Herbert, & R. Middleton (Eds.), *The cultural study of music* (pp. 193–203). Routledge.
- Dibben, N., & Williamson, V. J. (2007). An exploratory survey of in-vehicle music listening. *Psychology of Music*, 35(4), 571–589. <https://doi.org/10.1177/0305735607079725>
- Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1), 18–49. <https://doi.org/10.1177/0305735610362821>
- Egermann, H., & McAdams, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception*, 31(2), 139–156. <https://doi.org/10.1525/gfc.2008.8.4.24>
- Egermann, H., Nagel, F., Altenmüller, E., & Kopiez, R. (2009). Continuous measurement of musically-induced emotion: A web experiment. *International Journal of Internet Science*, 4(1), 4–20.
- Ekman, P. (1992a). An argument for basic emotions. *Cognition & Emotion*, 6(3), 169–200. <https://doi.org/10.1080/02699939208411068>
- Ekman, P. (1992b). Are there basic emotions? *Psychological Review*, 99(3), 550–553. <https://doi.org/10.1037/0033-295x.99.3.550>
- Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, 12(1), 75–99. <https://doi.org/10.1177/102986490801200105>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: Flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Finch, H. (2005). Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology*, 1(1), 27–38. <https://doi.org/10.1027/1614-1881.1.1.27>
- Flack, W. (2006). Peripheral feedback effects of facial expressions, bodily postures, and vocal expressions on emotional feelings. *Cognition & Emotion*, 20(2), 177–195. <https://doi.org/10.1080/02699930500359617>
- Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2), 212–228. <https://doi.org/10.1037/0022-3514.57.2.212>
- Fritz, T., & Koelsch, S. (2008). The role of semantic association and emotional contagion for the induction of emotion with music. *Behavioral and Brain Sciences*, 31(5), 579–580. <https://doi.org/10.1017/s0140525x08005347>
- Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences*, 15(11), 512–519. <https://doi.org/10.1016/j.tics.2011.09.003>
- Garrido, S., & Macritchie, J. (2020). Audience engagement with community music performances: Emotional contagion in audiences of a ‘pro-am’ orchestra in suburban Sydney. *Musicae Scientiae*, 24(2), 155–167. <https://doi.org/10.1177/1029864918783027>
- Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(4), 586–596. <https://doi.org/10.1002/wcs.55>
- Godøy, R. I., & Leman, M. (Eds.). (2010). *Musical gestures: Sound, movement, and meaning*. Routledge.
- Gordon, C. L., Cobb, P. R., & Balasubramaniam, R. (2018). Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS One*, 13(11), e0207213. <https://doi.org/10.1371/journal.pone.0207213>
- Hashim, S., Stewart, L., & Küssner, M. B. (2020, January). Saccadic eye movements suppress visual mental imagery and partly reduce emotional response during music listening. *Music & Science*. <https://doi.org/10.1177/2059204320959580>
- Hatfield, E., Bensman, L., Thornton, P. D., & Rapson, R. L. (2014). New perspectives on emotional contagion: A review of classic and recent research on facial mimicry and contagion. *Interpersona: An International Journal on Personal Relationships*, 8(2), 159–179. <https://doi.org/10.5964/ijpr.v8i2.162>
- Hatfield, E., Hsee, C. K., & Costello, J. (1995). The impact of vocal feedback on emotional experience and expression. *Journal of Social Behavior and Personality*, 10, 293–313. <https://ssrn.com/abstract=1133985>
- Hawk, S. T., Fischer, A. H., & Van Kleef, G. A. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social*

- Psychology*, 102(4), 796–814. <https://doi.org/10.1037/a0026234>
- Herbert, R., & Dibben, N. (2018). Making sense of music: Meanings 10- to 18-year-olds attach to experimenter-selected musical materials. *Psychology of Music*, 46(3), 375–391. <https://doi.org/10.1177/0305735617713118>
- Herget, A.-K. (2021). On music's potential to convey meaning in film: A systematic review of empirical evidence. *Psychology of Music*, 49(1), 21–49. <https://doi.org/10.1177/0305735619835019>
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 40(2), 129–141. [https://doi.org/10.1016/S0167-8760\(00\)00161-6](https://doi.org/10.1016/S0167-8760(00)00161-6)
- Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, 55(9), 1830–1849. <https://doi.org/10.1037/dev0000686>
- Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion (Washington, D.C.)*, 8(3), 395–409. <https://doi.org/10.1037/1528-3542.8.3.395>
- Howell, D. C. (2002). *Statistical methods for psychology*. Duxbury Press.
- Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, 60, 653–670. <https://doi.org/10.1146/annurev.psych.60.110707.163604>
- Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology*, 60, 1–25. <https://doi.org/10.1146/annurev.psych.60.110707.163539>
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100(1), 33–72. <https://doi.org/10.1016/j.cognition.2005.11.005>
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Sciences*, 9(1), 21–25. <https://doi.org/10.1016/j.tics.2004.11.003>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54–75. <https://doi.org/10.1037/a0024208>
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33(11), 1419–1432. [https://doi.org/10.1016/0028-3932\(95\)00073-c](https://doi.org/10.1016/0028-3932(95)00073-c)
- Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*, 9(6), 735–739. [https://doi.org/10.1016/S0959-4388\(99\)00038-0](https://doi.org/10.1016/S0959-4388(99)00038-0)
- Juslin, P. N. (2013a). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235–266. <https://doi.org/10.1016/j.plrev.2013.05.008>
- Juslin, P. N. (2013b). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, 4, 1–14. <https://doi.org/10.3389/fpsyg.2013.00596>
- Juslin, P. N. (2019). *Musical emotions explained*. Oxford University Press.
- Juslin, P. N., Barradas, G., & Eerola, T. (2015). From sound to significance: Exploring the mechanisms underlying emotional reactions to music. *The American Journal of Psychology*, 128(3), 281–304. <https://doi.org/10.5406/amerjpsyc.128.3.0281>
- Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, 42(4), 599–623. <https://doi.org/10.1177/0305735613484548>
- Juslin, P. N., Liljeström, S., & Västfjäll, D. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford University Press.
- Juslin, P. N., & Västfjäll, 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/S0140525X08006079>
- Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, 10(2), 191–213. <https://doi.org/10.1177/102986490601000203>
- Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, 20(3–6), 525–540. <https://doi.org/10.1080/02643290244000257>
- Koelsch, S., Fritz, T., Cramon, D. Y. V., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, 27(3), 239–250. <https://doi.org/10.1002/hbm.20180>
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, 7(3), 302–307. <https://doi.org/10.1038/n1197>
- 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: Music, Mind, and Brain*, 29(2–3), 90–99. <https://doi.org/10.1037/pmu0000238>
- Labbé, C., & Grandjean, D. (2014). Musical emotions predicted by feelings of entrainment. *Music Perception*, 32(2), 170–185. <https://doi.org/10.1525/MP.2014.32.2.170>
- Launay, J. (2015). Musical sounds, motor resonance and agency detection. *Empirical Musicology Review*, 10(1), 30–40. <https://doi.org/10.18061/emr.v10i1-2.4579>
- Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization can influence trust following virtual interaction. *Experimental Psychology*, 60(1), 53–63. <https://doi.org/10.1027/1618-3169/a000173>
- Leman, M., Desmet, F., Styns, F., Van Noorden, L., Moelants, D., Noorden, L. V., & Moelants, D. (2009). Sharing musical expression through embodied listening: A case study based on Chinese Guqin music. *Music Perception*, 26(3), 263–278. <https://doi.org/10.1525/mp.2009.26.3.263>
- Leman, M., & Maes, P.-J. (2014). The role of embodiment in the perception of music. *Empirical Musicology Review*, 9(3–4), 236–246. <http://dx.doi.org/10.18061/emr.v9i3-4.4498>

- Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion, the experience of fear as a conceptual act. *Psychological Science, 19*(9), 898–903. <https://doi.org/10.1111/j.1467-9280.2008.02174.x>
- Liuni, M., Ponsot, E., Bryant, G. A., & Aucouturier, J. J. (2020). Sound context modulates perceived vocal emotion. *Behavioural Processes, 172*, 104042. <https://doi.org/10.1101/2020.01.08.898205>
- Lundqvist, L.-O. O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music, 37*(1), 61–90. <https://doi.org/10.1177/0305735607086048>
- Maes, P.-J., & Leman, M. (2013). The influence of body movements on children's perception of music with an ambiguous expressive character. *PLoS ONE, 8*(1), e54682. <https://doi.org/10.1371/journal.pone.0054682>
- Margulis, E. H., Levine, W. H., Simchy-Gross, R., & Kroger, C. (2017). Expressive intent, ambiguity, and aesthetic experiences of music and poetry. *PLoS One, 12*(7), e0179145. <https://doi.org/10.1371/journal.pone.0179145>
- Mcintosh, D. N. (2006). Spontaneous facial mimicry, liking, and emotional contagion. *Polish Psychological Bulletin, 37*(1), 31–42. <https://doi.org/10.1177/0269881106067097>
- Michael, J., Sandberg, K., Skewes, J., Wolf, T., Blicher, J., Overgaard, M., & Frith, C. D. (2014). Continuous theta-burst stimulation demonstrates a causal role of premotor homunculus in action understanding. *Psychological Science, 25*(4), 963–972. <https://doi.org/10.1177/0956797613520608>
- Miller, N. A. (2016). Musical listening and kinesthesia: Is there an audio-vocal tuning system? *Psychomusicology: Music, Mind, and Brain, 26*(2), 117–131. <https://doi.org/10.1037/pmu0000138>
- Miu, A. C., & Baltes, F. R. (2012). Empathy manipulation impacts music-induced emotions: A psychophysiological study on opera. *PLoS ONE, 7*(1), 1–6. <https://doi.org/10.1371/journal.pone.0030618>
- Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: From motion to emotion. *Social Cognitive and Affective Neuroscience, 1*(3), 235–241. <https://doi.org/10.1093/scan/nsl029>
- Müllensiefen, D., Gingras, B., Stewart, L., & Ji, J. (2013). Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0: Technical report and documentation revision 0.3. Goldsmiths, University of London.
- Neumann, R., & Strack, F. (2000). “Mood contagion”: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology, 79*(2), 211–223. <https://doi.org/10.1037/0022-3514.79.2.211>
- Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S. S., & Ric, F. F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184–211. https://doi.org/10.1207/s15327957pspr0903_1
- Niedenthal, P. M., Brauer, M., Halberstadt, J. B., Se, AÊ, Innes-Ker, H., & Innes-Ker, ÅH (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition & Emotion, 15*(6), 853–864. <https://doi.org/10.1080/02699930143000194>
- Niedenthal, P. M., Halberstadt, J. B., & Margolin, J. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology, 30*(2), 211–223.
- Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2014). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience, 9*(8), 1062–1068. <https://doi.org/10.1093/scan/nst086>
- O’Neill, K., & Egermann, H. (2020, December). Induced empathy moderates emotional responses to expressive qualities in music. *Musicae Scientiae*. <https://doi.org/10.1177/1029864920974729>
- Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*(5), 489–504. <https://doi.org/10.1525/mp.2009.26.5.489>
- Panksepp, J. (1992). A critical role for “affective neuroscience” in resolving what is basic about emotions. *Psychological Review, 99*(3), 554–560. <https://doi.org/10.1037/0033-295x.99.3.554>
- Peltola, H. R., & Eerola, T. (2016). Fifty shades of blue: Classification of music-evoked sadness. *Musicae Scientiae, 20*(1), 84–102. <https://doi.org/10.1177/1029864915611206>
- Popper, K. (1992). *The logic of scientific discovery*. Routledge.
- Pruitt, T. A., Halpern, A. R., & Pfordresher, P. Q. (2019). Covert singing in anticipatory auditory imagery. *Psychophysiology, 56*(3), e13297. <https://doi.org/10.1111/psyp.13297>
- Roizin, P., Berman, L., & Royzman, E. (2010). Biases in use of positive and negative words across twenty natural languages. *Cognition & Emotion, 24*(3), 536–548. <https://doi.org/10.1080/02699930902793462>
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329–349. <https://doi.org/10.1146/annurev.psych.54.101601.145102>
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805–819. <https://doi.org/10.1037/0022-3514.76.5.805>
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695–729. <https://doi.org/10.1177/0539018405058216>
- Scherer, K. R., & Coutinho, E. (2013). How music creates emotion: A multifactorial process approach. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music, multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 121–145). Oxford University Press.
- Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361–392). Oxford University Press.

- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Böhner, M. (2010). Is it really robust? *Methodology*, 6(4), 147–151. <https://doi.org/10.1027/1614-2241/a000016>
- Sedlmeier, P., Weigelt, O., & Walther, E. (2011). Music in the muscle: How embodied cognition may influence music preferences. *Music Perception*, 28(3), 297–306. <https://doi.org/10.1525/mp.2011.28.3.297>
- Springer, A., Parkinson, J., & Prinz, W. (2013). Action simulation: Time course and representational mechanisms. *Frontiers in Psychology*, 4, 1–20. <https://doi.org/10.3389/fpsyg.2013.00387>
- Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, 82(2), 127–136.
- Tagg, P., & Clarida, B. (2003). *Ten little tunes: Towards a musicology of the mass Media*. Mass Media Music Scholars' Press.
- Taruffi, L., & Küssner, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain*, 29(2-3), 62. <https://doi.org/10.1037/pmu0000226>
- Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, 7(1), 1–10. <https://doi.org/10.1038/s41598-017-14849-0>
- Tesoriero, M., & Rickard, N. S. (2012). Music-enhanced recall: An effect of mood congruence, emotion arousal or emotion function? *Musicae Scientiae*, 16(3), 340–356. <https://doi.org/10.1177/1029864912459046>
- Thompson, W. F., & Coltheart, M. (2008). The role of signal detection and amplification in the induction of emotion by music. *Behavioral and Brain Sciences*, 31(5), 597–598. <https://doi.org/10.1017/S0140525X08005529>
- Timmers, R., MacRitchie, J., Schabrun, S. M., Thapa, T., Varlet, M., & Keller, P. E. (2020). Neural multimodal integration underlying synchronization with a co-performer in music: Influences of motor expertise and visual information. *Neuroscience Letters*, 721, 134803. <https://doi.org/10.1016/j.neulet.2020.134803>
- Van den Tol, A. J. M., & Edwards, J. (2011). Exploring a rationale for choosing to listen to sad music when feeling sad. *Psychology of Music*, 41(4), 440–465. <https://doi.org/10.1177/0305735611430433>
- 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>
- Wallmark, Z., Iacoboni, M., Deblieck, C., & Kendall, R. A. (2018). Embodied listening and timbre: Perceptual, acoustical, and neural correlates. *Music Perception: An Interdisciplinary Journal*, 35(3), 332–363. <https://doi.org/10.1525/mp.2018.35.3.332>
- Weiss, M. W., Bissonnette, A. M., & Peretz, I. (2020). The singing voice is special: Persistence of superior memory for vocal melodies despite vocal-motor distractions. *Cognition*, 213, 104514. <https://doi.org/10.1016/j.cognition.2020.104514>
- Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, 49(5), 1105–1127. <https://doi.org/10.1016/j.neuropsychologia.2010.12.032>
- Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, 9(4), e94446. <https://doi.org/10.1371/journal.pone.0094446>
- Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, 132(2), 173–189. <https://doi.org/10.1016/j.actpsy.2009.02.002>
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47(1), 9–12. <https://doi.org/10.1016/j.neuron.2005.06.013>
- Zentner, M. R., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion (Washington, D.C.)*, 8(4), 494–521. <https://doi.org/10.1037/1528-3542.8.4.494>

Appendix

Appendix A: Experiment I stimuli

- “Kip’s lights” from the movie *The English Patient* (1997). This piece is perceived as expressive of **sadness** or **tenderness**. The piece consists of a piano playing a slow, right-hand melody accompanied by a small ensemble of strings, clarinet, and celesta. Major mode. Tempo = 60 BPM.
- “Max” From the movie *Cape Fear* (1991). Perceived as expressive of **fear** or **anger**. The piece starts with the brass playing two loud semibreve notes with a seventh major interval. After a quiet moment played by the strings, the brass instruments play loud descending phrases, including a tri-tone interval (Bb to E). These phrases are then answered by the strings playing descending chromatic scales with fast tremolo. Minor mode. Tempo = 60 BPM.
- “Oliver learns the hard way” from the movie *Oliver Twist* (2005). Perceived as expressive of **joy** or **excitement**. The piece is characterized by a solo clarinet repeatedly playing a simple, *cantabile* melody accompanied by a small ensemble of strings playing a syncopated rhythm. A percussion instrument plays a small role at the end of each phrase from the clarinet. Minor mode. Tempo = 120.

Appendix B: Experiment I: Examples of visual imagery data

Table B1. Examples of imagery consistent with embodied simulation (images of the music being performed/played) present in answers to the question “what went through your mind while you were listening to the music”.

Quotation from participant	Highest ratings given by participant for the Fear/Anger piece
<i>The orchestra, their movements, the tension in the orchestra.</i>	Pride = 3 Induced Triumphant = 3
<i>Someone at the piano, and strings players around in a circle, in a practice space.</i>	Perceived Longing = 3 Peacefulness = 3 Positive feelings = 3 Induced Relaxed = 3
<i>Cheerful, I imagined myself in a theatre, watching the orchestra playing the tune... [Also] A movie of a particular historical period.</i>	Perceived Joy = 4 Induced Mellowed = 4 Induced Happy = 4

Table B2. Experiment I: Themes and illustrative examples from participant’s narratives^a.

Piece	Most frequent themes	Examples
Sadness/ Tenderness	Images of nostalgia, sadness and romanticism, (several mentioned Jane Austen’s books), and calm time spent in the countryside.	<p>“I had a very vivid visual image of revisiting a time in my childhood, I felt transported to a time before my mother died. I am in the garden, playing, exploring the garden. [...] Even though I could identify the notes and the key, I was concentrated in these images.”</p> <p>Participant whose highest ratings were perceived Longing = 4, perceived Spirituality = 4, perceived Tenderness = 4, Induced Sad = 4 and Induced Nostalgic = 4</p> <p>“I’m on a field with flowers, and the weather is nice. A soothing feeling. <i>Pride & Prejudice</i> [The movie]”</p> <p>Participant whose highest ratings were perceived Tenderness = 3, perceived Peacefulness = 3, induced Mellowed = 3, and induced Soothed = 3</p>
Fear/Anger	Horror movies (e.g., the suspense of something bad going to happen, or images of being chased by an evil character), and epic movie scenes of war.	<p>“I didn’t like it. It was something was trying to chase me [sic], to catch me, to hurt me. I saw the image of a murderer... An episode from <i>The Simpsons</i>, when Bart is going to be killed.”</p> <p>Participant whose highest ratings were perceived Fear = 4, and induced Anxious = 4</p> <p>“I imagined volcanoes, dramatic scenes on a mountain top, the climax of a fight between superheroes, thunder in the background... A Wagner opera.”</p> <p>Participant whose highest ratings were perceived Pride = 3, and induced Triumphant = 3</p>
Joy/Excitement	Narratives featured three themes: fantastic characters dancing merrily; excitement, purposefulness or feelings of determination before an adventure; and medieval scenes of a community of busy people.	<p><i>The main instrument being played by an animal, a pig, super happy, jumping around, a bit childish and unreal... Fantasia</i> [the movie], there’s a scene with dancing mushrooms and flowers. Going to carnivals as a little kid.</p> <p>Participant whose highest ratings were Positive Feelings = 3, and induced Happy = 3</p> <p><i>It felt quite adventurous, going to an adventure in a positive way. A group of people going out, going hiking. The Hobbit.</i></p> <p>Participant whose highest rating were perceived Positive feelings = 3, induced Triumphant = 4, and induced Happy = 3</p>

Note: ^aThe narratives and images were often associated with screen music (movie genres, particular films or TV dramas). While this could be due to the fact that the stimuli derived from movie soundtracks, other research has found that cinematic references are a common way in which people report their experiences with music (Herbert & Dibben, 2018; Tagg & Clarida, 2003).

Appendix C: Experiment 2: Descriptions of pieces intended to bias induced and perceived affect

Piece	Neutral description	Emotion description 1	Emotion description 2
“Kip’s lights”	You will now listen to a short extract from a piece from the end of the last century. The piece is an adagio, characterized by a simple melody and a slow accompaniment.	Sadness description: The composer wrote this piece towards the middle of his career, in what proved to be one of the most difficult years of his life. His young wife died while giving birth to their first child, and a few months later he had to leave his position as lecturer of composition in a prestigious university and travel to a different country as a political refugee, because the start of war in his home country made the political climate too dangerous to remain there.	Tenderness description: The composer wrote this piece towards the middle of his career, in what proved to be one of the most productive years of his life. He was appointed as lecturer of composition in a prestigious university, and a few months later, his wife gave birth to their first son. The composer wrote this piece during the spring of that year, when he moved with his new family to a country house to spend as much time as possible with their new-born child.
“Max”	You will now listen to an extract from a 20th century symphony. It is characterized by the alternation of loud and descending sounds with quieter and ascending ones.	Pride description: The composer wrote this piece to commemorate the tenth anniversary of the victory of his country against an invading army. He took inspiration from his memories of the bravery and sacrifice that his countrymen displayed during the decisive battle.	Fear description: The composer wrote this music to commemorate the difficult times that his country lived during the war. He was inspired by his childhood memories of having his sleep interrupted by the sounds of the enemy’s airplanes dropping the bombs that destroyed large parts of his home town.
“Oliver learns the hard way”	This piece of music has a moderate speed and a syncopated rhythm, and features a simple and repetitive melody.	Joy description: The composer wrote this piece for the scene of a ballet that portrays folk traditions in his home country. Inspired by traditional tales, he wrote this piece for a scene full of fantastic animals dancing in the woods.	Determination description: Inspired by fantastic novels, the composer wrote this musical piece for the scene of a ballet in which the bold characters prepare to embark on the epic adventure that will be portrayed during the rest of the work.

Appendix D: Experiment 2: Examples of visual imagery data

Intended Music Emotion	Quotation from participant	Participant and condition
Sadness/ Tenderness	<i>I imagined myself playing the instruments, one by one. And then I saw images of a train, and a couple who says good-bye, a painful and melancholic scene.</i>	Participant in the simulation, neutral description condition, who perceived the Sadness/Tenderness piece as expressive of longing and melancholy, and experienced induced nostalgia.
	<i>Actually I imagined a child being wrapped up in a quilt, and being cuddle [sic] by his mom.</i>	Participant in the non-simulation, tenderness description condition, who perceived the Sadness/Tenderness piece as expressive of joy and tenderness, and experienced induced mellowness.
	<i>I felt very far away from something I loved very much. I wanted to be soothed and told it would be ok. Both of these however were quite enjoyable feelings.</i>	Participant in the simulation, sadness description condition, who perceived the Sadness/Tenderness piece as expressive of longing and melancholy, and who experienced induced mellowness and tenderness.
	<i>It started sad, but seemed to become more uplifted. Reminded me of spring and the seasons changing.</i>	Participant in the non-simulation, sadness description condition, who perceived the Sadness/Tenderness piece as expressive of peacefulness, longing and tenderness, and who experienced induced mellowness, nostalgia and relaxation.
Fear/ Pride	<i>Death. Things on fire. Run away, now run away faster [...] on the second go-around, where I found it easier to think more about being a performer.</i>	Participant in the simulation, neutral description condition, who perceived the Fear/Pride piece as expressive of fear, and experienced induced anxiety
	<i>This piece of music elicited intense emotions in me, of dread and anxiety. I felt tense and I could readily imagine the experiences which inspired this music- bombs dropping, war, fire... I think the composer certainly achieved his aim in that respect.</i>	Participant in the simulation, fear description condition, who perceived the Fear/Pride piece as expressive of fear, melancholy and anger, and experienced induced anxiety
	<i>"This track referred to me pictures of ancient or future battles. Like I was ready for them, like I needed to save someone."</i>	Participant in the simulation, pride description condition, who perceived the Fear/Pride as expressive of pride, and who experienced induced triumphant and transcendent feelings.
	<i>It seemed like someone was trying to kill me with the knife [sic]</i>	Participant in the non-simulation, pride description condition, who perceived the Fear/Pride piece as expressive of anger and fear, and who experienced induced anxiety.
Joy/ Determination	<i>The piece of music reminded me of adventure and travelling.</i>	Participant in the simulation, neutral description condition, who perceived the Joy/Determination piece as expressive of pride and longing, and experienced induced feelings of triumph.
	<i>A forest, animals trotting about together but also a bit of impending doom.</i>	Participant in the non-simulation, joy description condition, who perceived the Joy/Determination piece as expressive of tenderness and experienced induced happiness, mellowness and admiration.
	<i>I thought of myself on an adventure. I felt upbeat, excited and ready to explore. I imagined myself in woodlands, marching along with purpose.</i>	Participant in the non-simulation, determination description condition who perceived the Joy/ Determination piece as expressive of pride and joy, and who experienced induced happiness and triumphant feelings.
	<i>It seemed like a motivational song, like after listening you wanted to get to work. It was very uplifting.</i>	Participant in the simulation, determination description condition, who perceived the Joy/Determination piece as expressive of pride and joy, and who experienced induced happiness and triumphant feelings).