

An Interactive Approach to Emotional Expression Through Musical Cues

Annaliese Micallef Grimaud¹ and Tuomas Eerola¹

Abstract

Previous literature suggests that structural and expressive cues affect the emotion expressed in music. However, only a few systematic explorations of cues have been done, usually focussing on a few cues or a limited amount of predetermined arbitrary cue values. This paper presents three experiments investigating the effect of six cues and their combinations on the music's perceived emotional expression. Twenty-eight musical pieces were created with the aim of providing flexible, ecologically valid, unfamiliar, new stimuli. In Experiment 1, 96 participants assessed which emotions were expressed in the pieces using Likert scale ratings. In Experiment 2, a subset of the stimuli was modified by participants (N = 42) via six available cues (tempo, mode, articulation, pitch, dynamics, and brightness) to convey seven emotions (anger, sadness, fear, joy, surprise, calmness, and power), addressing the main aim of exploring the impact of cue levels to expressions. Experiment 3 investigated how well the variations of the original stimuli created by participants in Experiment 2 expressed their intended emotion. Participants (N = 91) rated them alongside the seven original pieces, allowing the exploration of similarities and differences between the two sets of related pieces. An overall pattern of cue combinations was identified for each emotion. Some findings corroborate previous studies: mode and tempo were the most impactful cues in shaping emotions, and sadness and joy were amongst the most accurately recognised emotions. Novel findings include soft dynamics being used to convey anger, and dynamics and brightness being the least informative cues. These findings provide further motivation to investigate the effect of cues on emotions in music as combinations of multiple cues rather than as individual cues, as one cue might not give enough information to portray a specific emotion. The new findings and discrepancies are discussed in relation to current theories of music and emotions.

Keywords

music, cues, perceived, emotion, expression

Submission date: 5 October 2020; Acceptance date: 12 October 2021

Introduction

Previous literature suggests that emotions can be successfully conveyed through music and recognised by the listeners (Juslin, 1997a, 2013). This notion allows music to be utilised as a means of emotional communication in different scenarios, such as an aid for non-verbal patients (Silverman, 2008), a method for emotional recognition development in children and young adults (Saarikallio et al., 2014; Saarikallio et al., 2019), and a tool for mood regulation (Lyvers et al., 2018). Due to music's ability to convey emotion and have an effect on an individual's emotional response, it is of great importance to understand how this is attained.

A distinction is made between the two kinds of emotional processes that can occur during a musical experience: perceived emotion and felt (or induced) emotion. Perceived

emotion refers to the listeners' perception of the emotional expression the music is supposed to convey, whilst felt emotion refers to the listener's emotional response to the music. The variance between the two types of emotion might be differentiated by a rather fine line, however, they are considered as different modes of emotional responses, which may produce contrasting results (Gabrielsson, 2002). This study focusses on the

¹ Music Department, Durham University, Durham, UK

Corresponding author:

Annaliese Micallef Grimaud, Music Department, Durham University, Durham, United Kingdom.

Email: annaliese.micallef-grimaud@durham.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page

(<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

communication of *perceived* emotional expressions in music and investigates how different emotions are successfully conveyed to the listener through music.

Following an expanded version of Brunswik's lens model (Brunswik, 1956), previous research suggests that musical cues utilised by composers and performers aid in encoding emotions in music and helps listeners successfully decode and recognise the intended emotions (Juslin, 1997a). Musical cues can be loosely divided into two groups: structural cues and expressive cues. Structural cues refer to properties of the music that relate to the score, such as tempo and mode, whilst expressive cues are features utilised by performers, such as articulation and timbre (Gabrielsson, 2002). Although a lax distinction is made between these two groups, which cues belong in which group is still debatable, as some cues such as dynamics can be altered by both composers and performers (Livingstone et al., 2010). In this paper, tempo, mode, pitch, and dynamics will be referred to as structural cues, whilst articulation and brightness will be regarded as expressive cues.

Over the last 90 years, various methodologies have been utilised to investigate the role of different cues in conveying emotion through music. Hevner introduced systematic manipulation of structural cues in short pieces of tonal music, by creating versions of the same musical samples that varied in cues such as mode (1935), rhythm, melodic line, harmony (1936), tempo and pitch level (1937). Participants then listened to the stimuli variations and chose appropriate terms to describe the emotion conveyed by the music, thus identifying how the different cue levels affected the communicated emotion. Since then, several scholars inspected the properties of specifically composed music in relation to the intended conveyed emotion (e.g., Thompson and Robitaille, 1992) or measured the acoustical properties of the music (e.g., Juslin, 1997b; Schubert, 2004). Certain cue combinations have been linked to specific emotions. For example, fast tempo and high levels of loudness and pitch are associated with high arousal emotions like happiness or anger. A slow tempo, *legato* articulation, and soft timbre are associated with low arousal emotions like sadness and calmness (Juslin, 1997b; Scherer & Oshinsky, 1977; Watson, 1942). Many of these cues – such as loudness, timbre, tonality, and rhythm – seem to be operating similarly in different cultures (Balkwill et al., 2004; Laukka et al., 2013; Midya et al., 2019).

Despite the research on emotion cues over the years, there are only a few *systematic* explorations of cue combinations contributing to the expressed emotion using causal manipulation of cues. Early studies usually explored either a few cues such as tempo and mode (Dalla Bella et al., 2001) or tested a bigger number of cues with only two cue levels (Juslin & Lindström, 2010; Scherer & Oshinsky, 1977). Eerola et al. (2013) used a fractional factorial design to ambitiously combine six cues, each with three to six levels. Their findings reported that musical

cues for basic emotions tend to be additive and linearly contribute to emotional expression. On the other hand, there have been numerous attempts focussing on a particular cue, such as timbre (Eerola et al., 2012), harmony (Lahdelma & Eerola, 2016), mode (Kastner & Crowder, 1990), and harmonic intervals (Costa et al., 2000). However, the common shortcomings of all studies dealing with cue combinations are that they are limited in terms of how many cues can be realistically explored simultaneously and that the cue levels are arbitrary.

A strategy aimed to circumvent many of the limitations of systematic manipulations is allowing participants to create music expressing different emotional qualities. For instance, composers were given the task of creating music expressing different emotions. The efficacy of structural cues utilised in the compositions was then examined via an emotion recognition listening experiment (Thompson & Robitaille, 1992). Another method involved asking musicians to provide their interpretation of different emotions by performing a set-piece on their instruments (Gabrielsson & Juslin, 1996; Juslin, 1997b; Laukka et al., 2013). These approaches either focussed on structural cues or expressive cues. However, previous research suggests that a combination of structural and expressive cues should be investigated simultaneously as the two types of cues are known to interact together in an additive fashion (Eerola et al., 2013; Friberg & Battel, 2002; Gabrielsson, 2008).

A different approach to systematic manipulation studies, and score and performance analyses is an analysis-by-synthesis methodology (Friberg et al., 2014). This approach allows participants to manipulate a selection of cues of existing music using an interface that does not require musical expertise (Bresin & Friberg, 2011; Kragness & Trainor, 2019). Furthermore, a bigger cue space may be explored as cue levels and combinations do not need to be pre-determined and rendered. Bresin and Friberg's (2011) approach allowed participants to manipulate seven musical cues simultaneously (timbre, register, articulation, tempo, sound level, phrasing, and attack speed) with no arbitrary level restrictions. A few other studies have used this production approach to investigate how adolescents and children would change three to five cues via sliders to express three different emotions (happy, sad, and anger) in music (Saarikallio et al., 2019, 2014), or to compare how five emotions (happiness, anger, peacefulness, sadness, and fear) are expressed via five cues in music and movement across two cultures (Sievers et al., 2013). However, the musical materials in these studies were somewhat limited (Vieillard et al., 2008) since the cues manipulations either only affected the melodic component of the stimuli, or monophonic melodies were utilised as stimuli. Nevertheless, this was a viable way of probing how the cues work together to create the optimal desired emotional expression. Kragness and Trainor (2019) devised an experiment which utilised one key on a MIDI keyboard to control tempo, articulation, and dynamics of chord sequences

taken from Bach chorales. This methodology allowed users without any prior musical knowledge to perform different emotions through the stimuli with minimal task demands. However, the utilisation of one MIDI key to control three cues is a challenging interface to control the cues independently.

Most studies tend to focus on the communication of a limited selection of basic emotions, such as happiness, sadness, and anger (Warrenburg, 2020b), following the theory that basic emotions are the easiest and most accurately recognised emotions in music due to their existence in everyday life (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Kragness & Trainor, 2019; Mohn et al., 2011; Saarikallio et al., 2019). Other studies ask participants to rate musical pieces on valence and arousal dimensions (Costa et al., 2004; Morreale, Masu, Angeli, and Fava, 2013; Quinto et al., 2014). However, utilising valence and arousal dimensions may be somewhat ubiquitous and limiting, as some emotional expressions might not be captured by these dimensions (Collier, 2007). A different framework that presents perceived emotional expressions in music as a product of core affects and the listeners' contextual information is the constructionist approach, which proposes that different affect dimensions are recognised in music due to the abilities of speech and music to communicate levels of valence and arousal (Cespedes-Guevara & Eerola, 2018). Following the argument made by Laukka et al. (2013), committing to one framework of emotion theory might limit us to a number of discrete emotional expressions or affective dimensions, and hinder our aim to investigate a substantial number of different emotional expressions which have been reported as being expressed in music and perceived by listeners, which might not necessarily fit in one emotion framework, such as the combination of joy, sadness, love, calmness, longing, and humour emotions (see: Juslin and Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Zentner et al., 2008). Therefore, our aim is to give new insight on emotional expressions that exist in music. To this end, nine different emotional expressions, incorporating both basic emotions and other complex ones, which have been perceived in music were investigated in this paper: joy (Akkermans et al., 2019; Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Vieillard et al., 2008), sadness (Behrens & Green, 1993; Juslin & Laukka, 2004; Mohn et al., 2011), calmness (Juslin & Laukka, 2004; Laukka et al., 2013; Lindström et al., 2003; Thompson & Robitaille, 1992; Vieillard et al., 2008; Zentner et al., 2008), anger (Akkermans et al., 2019; Behrens & Green, 1993; Juslin & Laukka, 2004; Laukka et al., 2013; Mohn et al., 2011), fear (Behrens & Green, 1993; Kreutz, 2000; Mohn et al., 2011; Vieillard et al., 2008; Zentner et al., 2008), surprise (Juslin & Laukka, 2004; Lindström et al., 2003; Mohn et al., 2011; Scherer & Oshinsky, 1977), love (Juslin & Laukka, 2004; Lindström et al., 2003), longing (Juslin & Laukka, 2004; Laukka et al., 2013; Lindström et al., 2003), and power (Zentner et al., 2008).

Another limitation highlighted in current literature is that the majority of previous studies utilised commercial recordings of existing music as stimuli, mostly classical and popular music (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). When commercial music is utilised, it might create familiarity bias issues which cannot be controlled, as participants might have had prior exposure to the stimuli (Juslin & Västfjäll, 2008). Although using commercial music retains high ecological validity, control over the cues is limited, making recognition of their effects difficult (Gabrielsson & Lindström, 2010). Contrastingly, systematically manipulating cues affects the real music properties resulting in artificially sounding stimuli, forfeiting their ecological validity (Eerola et al., 2013; Juslin & Lindström, 2010). A solution to eliminate the overuse of commercial recordings and familiarity bias, whilst attending to the balance between ecological validity and experimental control would be to compose original music for the experiments.

In this paper, the main aim was to investigate how a number of structural and expressive cues and their combinations affected the communication of different emotional expressions through music. We strove to do this by moving away from a traditional, systematic manipulation methodology and using an interactive paradigm (analysis-by-synthesis methodology) where participants used cues to change the music to express different emotions in real-time, which allowed for a bigger cue space to be investigated. Furthermore, we wanted to explore a number of different emotions that have been said to be expressed in music (Juslin & Laukka, 2003; Kreutz, 2000; Lindström et al., 2003). To achieve this main aim of exploring a large cue space using a production approach, we also needed to address certain shortcomings mentioned above, and thus, we created a hierarchy of two secondary goals together with our main goal for this paper:

1. Our first sub-goal was to create a new set of musical stimuli that would be able to express a broad selection of nine emotional expressions (sadness, joy, calmness, anger, fear, power, and surprise) which may be conveyed by music (Juslin, 2013; Juslin & Laukka, 2004; Lindström et al., 2003), as existing music stimuli dealt with less emotions and mostly basic ones (Vieillard et al., 2008). This ensured that musical stimuli used were unfamiliar to participants, eliminating the issue of any familiarity bias that might stem when commercial music is used as stimuli. Additionally, we wanted to create polyphonic music which is flexible and allows for cue manipulations of all parts of the music, rather than just melodic manipulations as in previous studies (Bresin & Friberg, 2011; Kragness & Trainor, 2019; Saarikallio et al., 2019, 2014; Sievers et al., 2013). Creating new stimuli also allowed us to attend to the delicate ecological validity and experimental control balance, therefore simultaneously

tackling shortcomings mentioned in previous studies. To confirm whether the compositions were successful in expressing their predefined emotion, we asked participants to listen to the new musical excerpts and rate which emotions were being expressed in the music (Experiment 1).

2. To achieve our next goal and main aim of the paper, which was exploring how the cues contributed to the different emotions, we carried forward the musical pieces rated in Experiment 1 as the best exemplars of the pre-defined emotions and used them in the analysis-by-synthesis cue manipulation experiment (Experiment 2). Participants used an interactive interface called *EmoteControl* (Micallef Grimaud & Eerola, 2021) to change the musical pieces via six available cues (tempo, pitch, dynamics, brightness, and mode) to create different emotional expressions out of our selection of musical pieces. Using this production approach, a bigger number of cue combinations could be simultaneously explored, as unlike traditional, systematic manipulation experiments, cue levels and combinations did not need to be pre-defined and rendered. Therefore, Experiment 2 tackled the restricted number of cue levels limitation identified in previous studies. Furthermore, a combination of structural and expressive cues were used to manipulate polyphonic musical pieces, rather than monophonic melodies used in previous studies (Bresin & Friberg, 2011; Saarikallio et al., 2019, 2014; Sievers et al., 2013).
3. Finally, as the results of Experiment 2 created new versions of musical pieces expressing the different emotions, we took the opportunity to investigate a second sub-goal: how well these new participant-proposed pieces expressed their intended emotion. Therefore, we carried out another experiment (Experiment 3) where a new set of participants rated the emotion(s) expressed in the musical pieces' variations created by the participants in Experiment 2. Furthermore, in Experiment 3, participants also evaluated the already-rated musical pieces from Experiment 1 which were carried forward to Experiment 2. This gave us the opportunity to examine how two variations of the same musical pieces were perceived and look at the similarities and differences between the composer's and participants' musical interpretations of the emotions and cue combinations in related musical pieces.

Experiment 1: Evaluation of New Music Stimuli

Twenty-eight musical pieces were composed by the first author to be used as stimuli for music emotion research. The aims were to provide new, unfamiliar, polyphonic music that allows for experimental flexibility whilst also retaining ecological validity. Furthermore, the pieces were composed with the aim of conveying a particular emotion

to the listener in order to investigate how different emotions are communicated through the structural and expressive alterations of the musical pieces. Each piece was composed to convey one specific emotional expression from the following selection: joy, sadness, calmness, anger, fear, surprise, power, love, and longing. These nine emotion categories were selected based on previous literature suggesting that these emotions may be expressed through music and perceived by listeners (Juslin, 2013; Juslin & Laukka, 2004; Lindström et al., 2003; Turnbull et al., 2008; Zentner et al., 2008), and thus, this experiment aimed to provide new information on how the aforementioned emotions may be encoded in the music, and communicated to the listener. These emotions also cover a broad range on the emotion spectrum (Plutchik, 2001) and valence-arousal circumplex model (Russell, 1980). Furthermore, the composition of these musical pieces was an attempt to provide stimuli that represented other emotion terms apart from the most common ones which are sadness, happiness, and anger (Warrenburg, 2020a). To validate whether these 28 music compositions were able to convey their intended emotion, a rating study was carried out.

Method

Participants. Participants were recruited via social media and email notices. Ninety-six participants (40 men and 56 women) between 19 and 75 years of age ($M = 37.60$, $SD = 15.60$) took part in the study. A one-question version of the Ollen Music Sophistication Index (OMSI) (Ollen, 2006) was utilised to distinguish between the participants' levels of musical expertise (Zhang & Schubert, 2019). Sixty-five of the participants were non-musicians and 31 were musicians. Participants also provided information on their fluency in the English language on a five-point Likert scale (extremely limited, limited, modest, competent, and good/ fluent user), with 89 participants reporting they are fluent in the English language, five reporting they are competent and two participants rating themselves as modest users of the English language. Participation in the study was voluntary, and institutional ethics approval was obtained.

Material. The music material was composed by the first author who has nearly 10 years of experience in music composition. The musical excerpts were composed using both knowledge from existing literature on which musical features tend to express certain emotions (for an overview see: Cespedes-Guevara and Eerola, 2018; Juslin and Lindström, 2010), as well as the composer's own intuitions. Furthermore, to ensure compatibility with the *EmoteControl* interface (Micallef Grimaud & Eerola, 2021) described and used in Experiment 2, certain requirements were adhered to:

- The music should be composed for one instrument as the interface plays all parts in the music with the one chosen virtual instrument.
- Music notes should have note durations that allow for different articulation changes.

- The pitch range of the music should be compatible with the virtual instrument's register range to ensure all notes are played through the interface. In this case, as the interface uses a chamber strings virtual instrument, the pitch range was from B0 to C7.
- The musical pieces should have no modulations outside of the piece's key signature for the switch between major to harmonic minor mode to be successful.

In total, 28 musical pieces were composed, with three to four pieces composed for each of the nine selected emotions. All pieces were short polyphonic piano pieces, mostly adhering to a tonal framework and with durations ranging from 14 to 40 s. Further details on the musical pieces can be found in the supplementary material. All musical pieces are available on OSF repository¹.

Procedure. The study was carried out in English and administered online. Participants were instructed to wear headphones or use good quality speakers in a quiet environment due to the nature of the survey and test their sound. Instructions at the start of the survey explained to participants that they'll be listening to different musical pieces and rating on different emotion scales how much they thought the music was expressing each emotion. The instructions noted that participants will be asked to assess which *perceived* emotion they think the music is conveying, rather than their emotional response to the music. The full instructions and question template of Experiment 1 can be found in the Supplementary Material, together with additional procedure details. The 28 musical pieces were presented to the participants in random order. For each piece, participants rated how much of each of the nine emotions utilised to compose the pieces (joy, sadness, calmness, power, anger, fear, surprise, love, longing) they thought the music was conveying. Ratings were done on nine separate five-point Likert scales, one for each emotion, which were simultaneously presented to the participants in a matrix. A rating of 1 (none at all) indicated that the music did not convey any of the emotion. A rating of 5 (a lot) indicated that the music strongly conveyed the emotion. Participants carried out a practice trial which allowed them to familiarise themselves with the music listening task and rating scales. The study took approximately 25 min to complete.

Results

The consistency among participants in using the emotion rating scales was calculated using Cronbach's alpha (intra-class correlation coefficient) to examine the inter-rater agreement within each emotion scale across each participant and musical piece. High consistency of agreement between participants was observed for all rating scales, especially in the calmness $\alpha = 0.994$, sadness $\alpha = 0.992$, fear $\alpha = 0.992$ and anger $\alpha = 0.990$ emotion rating scales. The other rating scales also had high consistencies

(love $\alpha = 0.989$, joy $\alpha = 0.989$, longing $\alpha = 0.984$, surprise $\alpha = 0.979$) with the power emotion rating scale having the lowest consistency score $\alpha = 0.967$.

The data were then subjected to a one-way repeated measures ANOVA to investigate whether overall, participants rated the intended emotion significantly different to the other emotion rating scales, with the intended emotion being the independent variable and the dependent variable being the collapsed ratings across all other 'unintended' emotion scales. The main effect of the intended emotion on the emotion scale rating was significantly different, suggesting that in general, participants rated the intended emotion to be recognised by participants higher than the other eight emotions in the musical pieces, $F(1, 95) = 1173.00, p < .001$.

The mean ratings given by participants for all nine emotion scales were calculated for the 28 musical excerpts. Table 1 displays the mean ratings collapsed across participants and musical excerpts grouped in their respective intended emotion category. Rows in the table refer to the nine different types of intended emotions in the excerpts. Each row groups the excerpts intending to convey the respective emotion (e.g., anger row groups the three excerpts aiming to convey anger). Columns in the table refer to the nine emotion scales rated for each excerpt, to establish how much of each emotion participants thought the excerpts were conveying. The ratings along the diagonal in bold are expected to be higher than the other ratings in their relative row, following the hypothesis that a composer can effectively communicate the intended emotion to the listeners. However, this was not the case for all intended emotions. Overall, the excerpts composed to convey calmness, fear, joy, power, sadness, and surprise were given the highest ratings for their intended emotion, whilst excerpts composed to convey anger, longing, and love were rated highest for other emotions.

A one-way repeated measures ANOVA was executed on the excerpts grouped in their respective emotion category to determine if the intended emotion was rated significantly higher than the other emotions across the pieces within the group. The asterisks following the mean emotion ratings in the columns in Table 1 represent how significantly different the intended emotion's rating was to the other emotions' ratings. All pieces composed to express calmness, fear, joy, power, and sadness were rated significantly higher than all other emotions. Although the mean rating of surprise candidates was overall the highest, it was not significantly higher than the joy mean rating. This result suggests that joy might have been rated higher than surprise in one or more excerpts in their group. The anger candidates were rated significantly different for their intended emotion in comparison with other emotions, apart from fear. Excerpts intending to convey longing and love were both rated highest for calmness. Excerpts in the longing and love categories had mixed emotion ratings which were not significantly different from the intended

Table 1. Mean Ratings of the Emotions Perceived by Participants Collapsed Across the Musical Pieces Within Their Respective Emotion Category. Standard Deviations for Each Mean Rating Are Given in Brackets. The Excerpts Row Indicates Which Musical Piece(s) From Each Category Were Rated Significantly Highest for Their Intended Emotion. Their Mean Ratings and Standard Deviations Are Denoted in Brackets.

Intended Emotion	Rating	Rated emotion								
		Anger	Calmness	Fear	Joy	Longing	Love	Power	Sadness	Surprise
Anger		3.47 (1.46)	1.07 *** (0.35)	3.68 (1.39)	1.61 *** (1.13)	1.36 *** (0.81)	1.13 *** (0.39)	2.83 *** (1.49)	1.55 *** (0.93)	2.52 *** (1.48)
Calmness		1.06 *** (0.31)	3.98 *** (1.26)	1.15 *** (0.49)	2.15 *** (1.23)	3.02 *** (1.33)	3.51 *** (1.27)	1.94 *** (1.28)	2.18 *** (1.33)	1.21 *** (0.58)
Fear		3.25 *** (1.41)	1.09 *** (0.36)	4.17 *** (1.18)	1.53 *** (0.98)	1.36 *** (0.76)	1.17 *** (0.48)	2.75 *** (1.47)	1.54 *** (0.97)	2.57 *** (1.51)
Joy		1.28 *** (0.76)	1.43 *** (0.84)	1.29 *** (0.76)	4.40 *** (0.99)	1.56 *** (0.99)	1.86 *** (1.12)	2.65 *** (1.37)	1.09 *** (0.36)	2.39 *** (1.36)
Longing		1.20 *** (0.63)	3.29 *** (1.46)	1.65 *** (1.14)	1.86 *** (1.20)	3.05 *** (1.43)	2.81 *** (1.41)	2.08 *** (1.27)	2.82 *** (1.51)	1.39 *** (0.85)
Love		1.05 *** (0.24)	3.82 *** (1.25)	1.11 *** (0.45)	2.56 *** (1.29)	2.78 *** (1.43)	3.53 *** (1.31)	2.19 *** (1.33)	1.84 *** (1.14)	1.33 *** (0.75)
Power		2.38 *** (1.39)	1.13 *** (0.47)	2.40 *** (1.40)	2.69 *** (1.57)	1.36 *** (0.79)	1.38 *** (0.78)	3.63 *** (1.34)	1.31 *** (0.67)	2.47 *** (1.44)
Sadness		1.23 *** (0.67)	3.14 *** (1.46)	1.76 *** (1.15)	1.12 *** (0.40)	3.32 *** (1.39)	2.32 *** (1.23)	1.77 *** (1.15)	4.27 *** (1.06)	1.18 *** (0.57)
Surprise		1.83 *** (1.27)	1.42 *** (0.90)	1.95 *** (1.33)	2.93 *** (1.51)	1.56 *** (0.99)	1.53 *** (0.94)	2.66 *** (1.42)	1.43 *** (0.90)	3.14 *** (1.44)
Excerpts		A1 (M=4.16, SD=1.11), -, -	C2 (M=4.06, SD=1.24), C3 (M=4.04, SD=1.18), -	F2 (M=4.30, SD=1.09), F1 (M=4.26, SD=1.08), F3 (M=3.96, SD=1.34)	J2 (M=4.56, SD=0.96), J1 (M=4.29, SD=0.99), J3 (M=4.34, SD=1.00)	-, -, -	-, -, -	P2 (M=3.72, SD=1.37), P3 (M=3.83, SD=1.21), P1 (M=3.33, SD=1.40)	S2 (M=4.49, SD=0.86), S1 (M=4.26, SD=1.04), S3 (M=4.06, SD=1.22)	SU2 (M=3.29, SD=1.43), -, -

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$.

Bonferroni corrected values from the one-way repeated measures ANOVA. Df (1, 95). The asterisks are indications of when the mean rating of the intended emotion was significantly different from the other rated emotions. Values without any asterisks (apart from the intended emotion) represent the emotion ratings which were not significantly different from the intended emotion's rating. The Excerpts row notes the tracks that are significantly conveying their intended emotion using a posthoc analysis described in the text. Tracks are presented in ranked order from highest to lowest, with '-' denoting tracks which were not significantly conveying the intended emotion. The different tracks are coded with the first letter of emotion and track number (e.g., Track A1 = A for Anger and track number 1 of the 3 Anger tracks). The mean rating of the intended emotion per track is shown in brackets.

emotion in their respective groups. Thus, excerpts composed to convey longing and love were not clear representatives of their intended emotion.

Posthoc comparisons were also carried out to explore the participants' ratings for the individual excerpts and identify which candidates in each group were the strongest conduit of the intended emotion. The 'Excerpts' row in Table 1 notes the musical pieces within each group that were rated significantly highest for their intended emotion. The pieces are listed in a ranked order starting with the strongest representative of the intended emotion. The mean rating of the individual pieces is denoted in brackets. A '-' in the table denotes when none of the pieces in the emotion group represented the intended emotion.

All excerpts in the fear, joy, power, and sadness groups were rated significantly higher for their intended emotions. This suggests that these excerpts are good representatives of their intended emotion. All calmness pieces were rated highest for calmness. However, only two of the three candidates' calmness ratings were significantly higher than the other emotions' ratings. Only one excerpt from the anger and surprise groups was a strong representative of its intended emotion. No longing and love excerpts were good indicators of their intended emotion due to the mixed ratings.

Discussion

In this experiment, 28 newly composed musical pieces with the aim of conveying one particular emotion from an array of emotions (sadness, joy, calmness, anger, fear, power, surprise, love, and longing) were rated by participants to determine whether the pieces were accurately communicating their intended emotion to the listeners. Sixteen out of the 28 pieces (57.14%) were correctly identified as conveying their intended emotion, which suggests that it is possible for listeners to correctly identify an intended emotion in a musical piece, despite the music being new and unfamiliar to them. This supports the notion that in general, musicians can encode certain emotions in the music by using musical cues, which in turn, listeners use to decode and identify the emotion communicated in the music (Akkermans et al., 2019; Juslin, 2000, 2013; Juslin & Lindström, 2010). However, it is important to note that this was not the case for all intended emotions in this experiment. All musical pieces representing fear, joy, power, calmness, and sadness were recognised as conveying their intended emotion. On the other hand, only one of the three anger and surprise excerpts were rated as conveying their intended emotion, whilst none of the longing and love excerpts were perceived as expressing their desired emotion. Instead, all love and longing excerpts were rated highest for calmness. The fact that love, longing, and calmness have similar musical features might explain why these three emotions tended to be confused. Previous research has suggested that music expresses basic emotions, i.e., happiness, anger, sadness, fear, surprise, and disgust (Ekman,

1992), and that basic emotions are easier to communicate in music and be recognised by listeners than other emotional expressions (Gabrielsson & Juslin, 1996; Juslin, 2000, 2013; Peretz et al., 1998). Although in Experiment 1, all musical pieces representing the sadness, joy, and fear basic emotions were correctly identified by participants, the anger emotion was correctly recognised in only one of the three anger excerpts. Furthermore, it is interesting to note that although calmness and power emotions are not considered as basic emotions, participants accurately identified the intended emotions in their respective musical pieces. This might be due to calmness and power being two emotions that have been frequently reported to be expressed by music and perceived by listeners (Juslin & Laukka, 2003; Lindström et al., 2003; Zentner et al., 2008). Furthermore, it is suggested that music can effectively communicate emotions that can be explained without an intentional situation context (Cespedes-Guevara & Eerola, 2018).

The musical pieces in each emotion category were composed with a range of cues that have been associated with their intended emotion in previous studies. Certain cue combinations also overlapped across emotion categories. Anger, fear, and power excerpts featured a fast tempo, minor mode, repetitive notes, dissonance, stepwise movement in the melodic line, and a constant rhythm (Costa et al., 2000; Ilie & Thompson, 2006; Juslin, 1997b; Krumhansl, 1997; Lindström, 2006; Scherer & Oshinsky, 1977). The excerpt rated as the best representative of fear had the most dissonance and most constant repetitive note pattern. Only one of the anger candidates was rated highest for anger, while the remaining two were rated highest for fear. Although previous studies suggest that anger is represented by a high pitch level and fast tempo (Juslin, 1997b; Scherer & Oshinsky, 1977), the strongest representative of anger had the lowest pitch level and slowest tempo from the three anger candidates. Furthermore, the piece rated highest for the intended emotion anger had the smallest pitch range of C1 to F2, while the other two pieces which were incorrectly rated highest for fear had pitch ranges spanning four octaves. This result is not surprising as other studies have also found that anger and fear do tend to be confused in music (Cunningham & Sterling, 1988; Kragness et al., 2021; Vidas et al., 2018), potentially due to them being both negative emotions and sharing multiple musical elements such as *staccato* articulation, minor mode, and a fast tempo (Mohn et al., 2011). Excerpts portraying power featured melodies with small intervals, mostly major thirds, perfect 4ths and 5ths (Smith & Williams, 1999), and a narrow melodic range (Gundlach, 1935). The strongest representative of power had the fastest tempo at 175 bpm and the piece had the smallest pitch range from the three excerpts.

Surprise and joy excerpts featured upward pitch leaps in the melodic line and variation (Scherer & Oshinsky, 1977). Joy excerpts were all in major mode (Peretz et al., 1998) whilst surprise excerpts varied in modes. The strongest

candidate for surprise was in minor mode and had the most rhythmic variation and rests, which perhaps aided in making the surprise element more defined. The best representative of joy was the fastest at 120 bpm and had the simplest harmonic complexity, which could potentially explain why it was preferred over the other pieces (Costa et al., 2000).

Excerpts composed to convey calmness, love, sadness, and longing all featured a slow to moderate tempo, smooth melodic progressions with stepwise or arpeggiated movement, a constant rhythm, and very similar pitch ranges (Gagnon & Peretz, 2003; Juslin, 1997b; Quinto et al., 2013; Thompson & Robitaille, 1992). Calmness and love pieces were consonant and in major mode, whilst sadness and longing pieces were in minor mode (except for one longing excerpt which was in major mode) (Hevner, 1936). Sad pieces featured low pitch levels (Hevner, 1937; Watson, 1942) and narrow melodic pitch ranges (Balkwill & Thompson, 1999). The best representative of sadness had the least movement and was the only piece with a descending stepwise melody rather than an arpeggiated one (Scherer & Oshinsky, 1977; Thompson & Robitaille, 1992). The highest-rated calmness excerpt had the most melodic movement, which was well-paced and held a steady rhythm. Interestingly, love was the second-highest rated emotion in the calmness excerpts, whilst all love excerpts were rated highest for calmness. All longing pieces got mixed ratings, with the highest ratings being for sadness, calmness, and longing emotions. This might be due to the heavy overlap in music features used to portray these emotions (Lindström, 2006) or simply due to the complexity and ambiguity that emotion poses for music (Gabrielsson & Juslin, 1996; Juslin, 2013).

Although overall, similar cues were utilised to portray the same emotion across different musical pieces, these results suggest that even small nuances affect the emotion being expressed by the music. This supports the notion that the different properties (cues) of the music work together to portray different, intended emotions (Argstatter, 2016; Eerola et al., 2013; Juslin & Timmers, 2010; Lindström et al., 2003), and thus, components of the music and their combinations should be investigated together to identify which specific cues and levels provide the determining factor in conveying one emotion rather than another.

Limitations of the experiment. A potential shortcoming of this experiment is that although participants were instructed to wear headphones or use good quality speakers in a quiet environment, the researchers do not have absolute control over the participants' environment due to the online nature of the study, and the requirements mentioned might not have been upheld by the participants. Furthermore, the instructions did not mention that participants should keep their volume constant, therefore, participants might have altered the volume level throughout the experiment, which could also affect results. Another

possible limitation of this experiment is the potential misunderstanding of terms and instructions due to modest language competence. Therefore, apart from enquiring about participants' English proficiency levels, a post-task question with regards to clarity of instructions and task would be helpful. It is good to note that the composer and the majority of participants (94.79%) that took part in this experiment are from a Western culture, and that the music composed and rated by participants was tonal, Western music. Thus, the results of this experiment represent a Western population sample and different results might be achieved in a cross-cultural setting, which would be an interesting avenue to pursue in future studies.

To investigate how the cues and their manipulations influence the emotions communicated through the music, only the best representative of each emotion (i.e., the piece rated highest for its intended emotion) was selected for the next experiments. Musical pieces that received mixed emotion ratings and were not successful in portraying their intended emotion (i.e., love and longing pieces) were not carried forward to the next experiments.

Experiment 2: Cue Manipulation Task

Experiment 2 addressed the main aim of this paper, which was to explore the role of six musical cues (tempo, articulation, pitch, dynamics, brightness, and mode) in conveying different emotional expressions through music and a large cue space by using an interactive paradigm which does not restrict us to a small number of predetermined cue levels and combinations. To achieve this, an analysis-by-synthesis method was utilised, where participants were presented with a selection from the newly composed musical pieces that were rated by participants in Experiment 1 as strongly conveying their intended emotion. Participants in Experiment 2 were then asked to alter these musical pieces in a computer interface called *EmoteControl* (Micallef Grimaud & Eerola, 2021) via the six available cues (tempo, articulation, pitch, dynamics, brightness, and mode) to change the emotion conveyed by the music. This approach allowed for an extensive exploration of cue levels and combinations to identify how the same six cues are altered to convey different emotions. The prediction was that across different musical pieces, the same cue combinations are used to convey the same emotion, and a unique pattern of cues will emerge for each emotion.

Method

Participants. Participants were recruited via social media and university communications. Forty-two participants (12 men, 29 women, one individual did not indicate their gender) between the ages of 18 and 58 years ($M = 26.17$, $SD = 8.17$) took part in the study. A one-question version of the OMSI (Ollen, 2006; Zhang & Schubert, 2019) was utilised to distinguish between the participants' levels of musical expertise. Twenty-two of the participants were

musicians, and 20 were non-musicians. Participants were compensated with chocolate for their time.

Material. Seven musical pieces previously validated in Experiment 1 as representing a specific emotion (joy, sadness, calmness, power, anger, fear, or surprise) were selected. Participants were asked to convey each of the seven emotions attributed to the musical pieces through all the excerpts.

Apparatus. *EmoteControl*, a graphical user interface created for music emotion research, was utilised for the study (Micallef Grimaud & Eerola, 2021). Figure 1 presents the *EmoteControl* user interface. *EmoteControl* allows users to input an instrumental musical piece in MIDI format in the interface and alter a combination of structural and expressive cues (tempo, articulation, pitch, dynamics, brightness, and mode) of the music file. A chamber strings sound synthesizer from Vienna Symphonic Library (VSL) is used as the default virtual instrument and sound output in the *EmoteControl* interface. When a music file is inputted in *EmoteControl*, the properties of the music are re-arranged depending on the initial values of the cue sliders. The cue values are initially set to the middle of the available range before playback starts, thus not exposing users to the ‘original’ version of the piece. Users can make cue changes via sliders for tempo, articulation, brightness, pitch, and dynamics, and a toggle button for the mode cue, while the music plays in real-time and the cue changes are instantly heard in the music. The interface records the cue changes at 10 Hz.

Cue details of *EmoteControl*. The *EmoteControl* interface allows participants to change a combination of four structural (tempo, mode, dynamics, and pitch) and two expressive cues (brightness and articulation) of the music, for a total of six cues. The tempo, mode, and dynamics cues have been reported as being the most contributing structural cues to the emotion communicated in music (Dalla Bella et al., 2001; Eerola et al., 2013; Kamenetsky et al., 1997; Morreale, et al., 2013), while the remaining three cues, pitch, articulation, and brightness, have been investigated to a lesser extent (Eerola et al., 2013; Juslin & Lindström, 2010; Quinto et al., 2014; Saarikallio et al., 2014). Therefore, investigating how tempo, mode, and dynamics are used by participants in an interactive setting is an opportunity to explore whether this current methodology will produce results that complement previous literature, whilst also providing a baseline for experiments utilising the *EmoteControl* interface. Furthermore, investigating a combination of influential cues together with less explored cues such as pitch, articulation, and brightness will allow for new data to be collected on these less explored cues, but most importantly, on their interaction with cues that have been established as strong contributing factors of emotion in music. Furthermore, both structural and expressive cues are responsible for the communication of emotion through music (Friberg & Battel, 2002) and should be investigated together due to their additive and interactive nature (Gabrielsson, 2008), which were two additional motivations taken into consideration when choosing the cues to be investigated in this experiment. The following subsections describe the ranges and levels used for each cue in the *EmoteControl* interface.

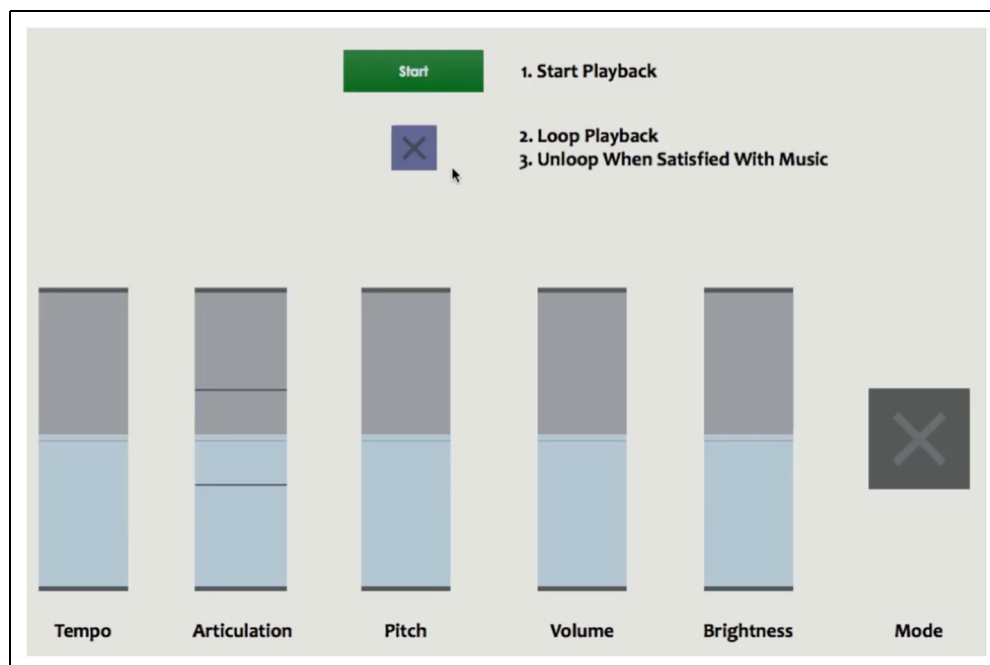


Figure 1. The *EmoteControl* user interface.

Tempo. The tempo cue is controlled via a slider and is measured in beats per minute (bpm). The tempo cue has a minimum value of 50 bpm and a maximum value of 160 bpm which covers a broad tempo range, and a step size of 1 bpm.

Articulation. The articulation cue consists of three levels of articulation: *legato*, *detaché*, and *staccato*. The different articulation levels are presented in a sequence from longest note-duration to shortest. Changes to the articulation cue are made by a use of a slider.

Pitch. The pitch slider controls a pitch shift range of ± 2 semitones from the default mid-point of the slider.

Dynamics. The dynamics slider controls the MIDI volume of the virtual instrument used as output rather than the overall volume via the dB level. The dynamics slider is set to a minimum MIDI volume of 30 and has a maximum value of 129, which translates to a range of 26 decibels that is known to have small non-linearities (Goebl & Bresin, 2001).

Brightness. The brightness slider controls the cut-off frequency value of a low-pass filter which affects how many harmonics sound. The low-pass filter has a cut-off frequency range of 305 Hz to 20,000 Hz.

Mode. The mode cue gives the participants the option to switch from major mode to harmonic minor mode (the third and sixth degrees of the scale are flattened) which is controlled via a toggle button.

Procedure. Ethical consent was obtained before testing and the experiment was carried out in the lab. The experiment was made up of two parts. In the first part, participants answered some demographic questions such as age, gender, and musical expertise. The full set of questions are presented in the Supplementary Material. In the musical cues task, participants were informed that they will be presented with different combinations of musical pieces and emotion terms. For each trial, their task was to alter how the music sounds to best represent what they think the intended emotion sounds like in music. They were instructed that they could change the music in real time using the six cues presented as five sliders and one toggle button. Each musical piece was looped so that the participants could keep on hearing it and making as many changes as they liked. When they were satisfied that the musical piece was best representing the intended emotion, a new musical piece was loaded, and a new emotion term was given. Changes to the cue values were recorded for each trial. It was explained that there was no time limit for the experiment. Prior to the musical task, the researcher gave a short demonstration of the interface, and participants were subjected to a practice trial. The full instructions as

well as details about the demonstration and practice trial are included in the Supplementary Material. At the end of the experiment, participants were presented with an optional open-ended question to leave feedback on their experience with the interface and the experiment in general. Overall, seven musical pieces were changed to convey seven different emotions, which yielded 49 different combinations. Participants were split into three groups of 14 participants to minimise fatigue. Each group carried out 21 combinations of musical pieces and intended emotion: conveying three different emotions through all pieces (3 emotions $\times$ 7 pieces). The experiment took approximately 30 min to complete.

Results

The consistency and reliability of the participants' cue usage was examined by calculating the inter-rater agreement using Cronbach's alpha, among participants within the 21 combinations of stimuli and emotions in each group. High consistency among participants was observed, particularly in Tempo ($\alpha = 0.943\text{--}0.964$), Articulation ($\alpha = 0.943\text{--}0.957$), Pitch ($\alpha = 0.939\text{--}0.956$), Dynamics ($\alpha = 0.832\text{--}0.865$), Brightness ($\alpha = 0.833\text{--}0.869$), and Mode ($\alpha = 0.950\text{--}0.960$).

Table 2 shows the overall, main effect of Emotion, Piece, and the interaction between Emotion and Piece factors for the six different cues. Linear mixed models (LMMs) were applied for each cue except for mode, with and without the factors in question, utilising Participant as the random factor in the models. A generalised mixed model (GLMM) with a binomial distribution was used for mode, due to its binary nature. A likelihood ratio test was then computed to assess whether the contribution of the factor (i.e., Emotion, Piece, or their interaction) offered statistically significant improvements to the model. The main effect of interest in Table 2 is between the different cues and the Emotion factor, which are all significant, suggesting that the cues were utilised in a specific way depending on the emotion to be portrayed. The Piece factor had a statistically significant effect on all cues except for brightness which suggests that certain structures of musical pieces

Table 2. LMM Estimates for Tempo, Articulation, Pitch, Dynamics, and Brightness Cues and GLMM Estimate for the Mode Cue for the Main Effect of Emotions, Musical Pieces, and Their Interactions, Using a Likelihood Ratio Test.

	Emotion	Piece	Piece $\times$ Emotion
Tempo	657.49***	42.46***	45.66
Articulation	645.56***	14.85*	88.12***
Pitch	695.79***	31.67***	38.60
Dynamics	280.96***	13.86*	45.15
Brightness	303.83***	9.96	27.61
Mode	613.48***	23.72***	50.71

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$, $df = 6$ for Emotion, $df = 6$ for Piece, $df = 36$ for Interaction for the likelihood ratio test.

also had an influence on how the cues were utilised by participants. This is understandable as the musical pieces had been originally composed to convey different emotions, and thus, might require the cues to be utilised slightly differently to portray the same emotion across the pieces. A further investigation of the Piece factor in relation to the different cues showed that the pieces composed and validated as conveying calmness and sadness were the ones that mostly affected the use of the cues, with tempo, pitch, and mode having a significant interaction with the calmness piece, while tempo and articulation had a significant effect with the sadness piece. A breakdown of the effect of the cues on each musical piece is presented in Table S2 in the Supplementary Material. Articulation was the only significant effect on the interaction between Piece and Emotion. The most relevant result from Table 2 for the purpose of this experiment is the fact that all cues had a significant effect on the conveyed emotions. The rest of this experiment's analysis focusses on cue usage and combinations used to communicate different emotions.

Figure 2 portrays the mean cue values utilised by participants to convey the different emotions across the musical pieces. A slow tempo was utilised to portray calmness and sadness whilst power and fear featured a moderately fast tempo. Joy and anger had a very similar fast tempo, and surprise had the fastest tempo. Nearly identical pitch values were utilised for fear, surprise, and joy. Power had the highest pitch, with anger being a close second. Participants utilised a lower pitch to convey calmness and sadness. Interestingly, participants opted for soft dynamics in general, with anger and sadness having the softest dynamics. Surprise and joy were the only two emotions conveyed via loud dynamics. The brightness parameter alters the amount of harmonic content outputted by having participants control the cut-off frequency value of a low-pass filter (Micallef Grimaud & Eerola, 2021). The smaller the value, the fewer high frequencies are passed through the filter, which makes the sound darker. A dark timbre was used to portray sadness, whilst surprise and joy featured the brightest timbre. Although the articulation and mode cues hold categorical data, the means of these two cues are visualised in the same manner as the other cues for the purpose of clarity. Nevertheless, these two cues were regarded as discrete categories in the analysis. The articulation cue consisted of three discrete levels: *legato*, *detaché*, and *staccato*, which were available to participants in sequence from the longest note-duration to shortest via a slider. Participants chose *legato* for sadness and calmness, *detaché* for fear and power, and *staccato* to portray anger, surprise, and joy. Mode was utilised as a binary parameter (major, minor), with participants opting for minor to express negative emotions: sadness, fear, and anger; and major for calmness, joy, surprise, and power. Although mode works with distinct values, allowing for a categorical violation in the visualisation (see Figure 2) helps identify the emotions which participants were indecisive about

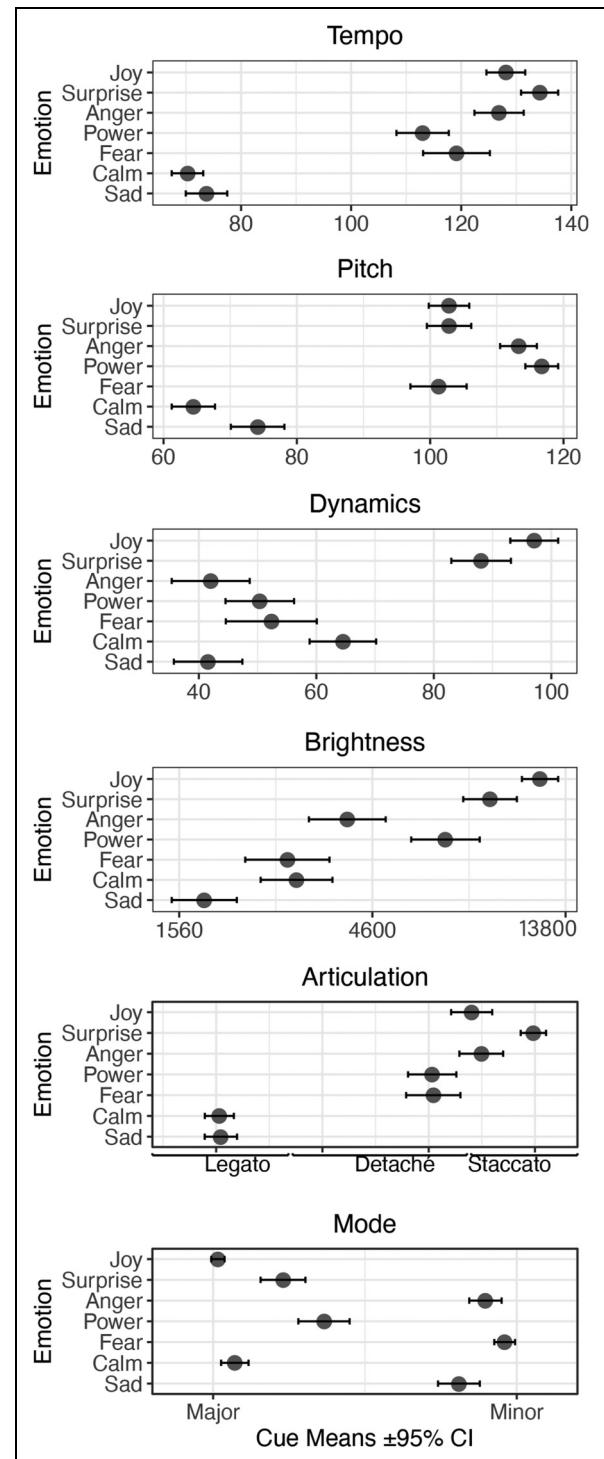


Figure 2. Means and 95% confidence intervals of cues utilised by participants to portray different emotions.

Note. The X-axes refer to actual values in the cue manipulation experiment. Tempo is denoted in beats per minute (BPM). Pitch and Dynamics are denoted in MIDI values. Brightness values represent the cut-off frequency in Hertz (Hz) of the low-pass filter. Articulation levels consist of *legato*, *detaché*, and *staccato*. Mode consists of major and minor levels.

when choosing between major and minor mode. The most prominent example in this respect is power, where although the overall cue mean indicates major mode was utilised for

Table 3. Normalised Cue Coefficients Across Emotions for Each Cue Utilising Linear Discriminant Analysis. The Overall Correct Prediction Rate Is 60.67%.

	Anger 59.6%	Calmness 91.0%	Fear 44.9%	Joy 67.4%	Power 49.4%	Sadness 75.3%	Surprise 37.1%	R^2	P
Tempo	0.50	-1.15	0.39	0.56	0.09	-1.07	0.68	0.52	< .001
Articulation	0.48	-1.04	0.21	0.46	0.22	-1.12	0.80	0.50	< .001
Pitch	0.66	-1.22	0.16	0.27	0.75	-0.80	0.19	0.47	< .001
Dynamics	-0.63	0.04	-0.22	0.91	-0.26	-0.48	0.63	0.28	< .001
Brightness	-0.14	-0.39	-0.34	0.78	0.36	-0.76	0.49	0.27	< .001
Mode	0.88	-0.79	0.99	-0.92	-0.31	0.70	-0.54	0.59	< .001

$n = 623$ cases used in estimation; null hypotheses: two-sided; multiple comparisons correction: False Discovery Rate correction applied simultaneously to the entire table.

the emotion, Figure 2 shows how the mode value for power also leans towards minor mode. A similar pattern can be seen for the surprise emotion.

Discrimination of emotions using the cues. To explore the efficacy of the cue combinations in characterising each emotion, we carried out a linear discriminant analysis with the cues to predict the emotions. This analysis, which we carried out with a training set (70% of observations stratified across the emotions) provided a set of transformations where the first two functions carry the majority of weight (93.47%) and could predict 60.67% of emotions in the test set (baseline being 14.3%). To understand how specific emotions and cues consistently operated in this mapping, Table 3 portrays the accuracy percentage of correctly predicting the emotions (anger, calmness, fear, joy, power, sadness, and surprise) and the normalised cue coefficients across emotions for each cue.

Table 3 outlines cue combinations and their values that have a good percentage of predicting the intended emotion. The first seven columns in Table 3 present the cues as discriminant functions of each of the seven emotions. The values in bold mark the cue values that have significant weight in predicting emotions. Values with a minus (-) sign represent low/negative values, whilst values with no sign represent high/positive values, excluding the values for mode, where a positive value points to minor, and a negative value points to major mode (0 denotes major and 1 denotes minor in the interface). For example, sadness can be accurately predicted 75.3% of the time utilising the cue combination presented in Table 3: slow tempo, *legato* articulation, low pitch, soft dynamics, low brightness, and minor mode. Calmness is the emotion that could be identified most correctly with 91% accuracy. Tempo, articulation, pitch, brightness, and mode are all significant parameters for characterising calmness; however, dynamics does not have a significant effect on the shaping of calmness. Following calmness, the sadness (75.3%), joy (67.4%), and anger (59.6%) emotion profiles are the ones with the highest correct identification rates. Power (49.4%) and fear (44.9%) have less than 50% accuracy rate, with the

least correctly predicted emotion being surprise, with 37.1% accuracy of recognition.

Table 3 also denotes the R^2 value of each cue across all emotions, which indicates the power of the individual cues in conveying the different emotions. Mode is the strongest discriminator when characterising different emotions ($R^2 = 0.59$), followed by tempo ($R^2 = 0.52$), articulation ($R^2 = 0.50$) and pitch ($R^2 = 0.47$). Dynamics ($R^2 = 0.28$) and brightness ($R^2 = 0.27$) hold the least weight in shaping different emotional expressions in music. It is important to note that this ranking of the different cues' communicative weight is done in respect to the other available cues investigated here. For example, tempo is overall, the second strongest discriminator in shaping different emotions, however, tempo was not significant in the conveying of power.

Correct prediction of emotions by cue selections. The confusion matrix in Table 4 presents the proportion of participants' cue selections in the linear discriminant analysis (LDA) which correctly predicted the intended emotion. The column headers display the intended emotions, whilst the rows show the emotions predicted, in proportions of the discriminant. Ratings along the diagonal in bold are expected to be higher than the other ratings in their relative column, following the hypothesis that the cue selections can correctly predict the intended emotion. This is confirmed in the confusion matrix, as the intended emotion is typically preferred over other emotions. The last row of the table denotes the balanced accuracy ($\frac{\text{sensitivity} + \text{specificity}}{2}$, where sensitivity is the true positive rate and specificity is the true negative rate) presented as proportions (Chen et al., 2004). Calmness is the emotion most correctly predicted (0.92) by the utilised cues, and also features the highest accuracy rating of 0.88. Fear is the least predicted emotion by the participants' cue selections in the LDA (0.35), with a correct emotion prediction accuracy rating of 0.62.

Feedback from participants. At the end of the experiment, participants were free to leave comments on any aspect of the experiment. Twenty-nine of the 49 participants (69%) gave us feedback. 48% of the feedback was about participants liking the experiment and commenting on how

Table 4. Confusion Matrix Displaying Prediction Proportion Rates of the Discriminant Model to Test Data.

		Intended						
		Joy	Surprise	Anger	Power	Fear	Calmness	Sadness
Predicted	Joy	0.46	0.41	0	0.03	0.03	0.08	0
	Surprise	0.16	0.51	0.14	0.05	0.14	0	0
	Anger	0	0	0.65	0	0.32	0	0.03
	Power	0.08	0.08	0.22	0.43	0.11	0	0.08
	Fear	0.03	0.08	0.38	0	0.35	0	0.16
	Calmness	0	0	0	0	0	0.92	0.08
	Sadness	0	0	0	0.05	0.05	0.19	0.70
Accuracy		0.77	0.70	0.70	0.84	0.62	0.88	0.81

“quick and easy” and user-friendly the interface was. 34.48% of comments mentioned that some musical pieces were harder to change to convey a specific emotion than others, and flagged power and/or surprise emotions as being the most difficult to portray in the pieces. Two participants commented that pitch was the trickiest cue, whilst another participant mentioned mode as being difficult. Individual participants mentioned mode, articulation, pitch, and brightness as being important cues in the conveying of emotions, whilst one participant commented that they thought dynamics was not of importance in expressing emotions.

Discussion

In this experiment, seven musical pieces previously validated as conveying a particular emotion were altered by participants via six cues to express the intended emotions. The main results identified cue values and combinations used to convey specific emotions across musical pieces. The overall success of the cues in predicting the emotions was estimated, and in general, the results suggested clear cue-emotion patterns.

Emotions expressed in cue combinations. Table 5 gives an overview of the cue combinations utilised by participants for each intended emotion across the different musical

pieces, which generally complement previous literature and other production studies (Bresin and Friberg, 2011; Kragness & Trainor, 2019; Saarikallio et al., 2019, 2014). The discrepancies between the current study’s results and four previous production studies by Bresin & Friberg (2011); Saarikallio et al. (2014), Kragness and Trainor (2019), and Saarikallio et al. (2019) are denoted in Table 5 by numerical values in subscript. Comparisons for power and surprise emotions could not be made as they were not investigated during the previous studies.

The cue combination expressing sadness featured a slow tempo, *legato* articulation, a low pitch, soft dynamics, a dark sound and minor mode, complementing previous literature (Akkermans et al., 2019; Hevner, 1936; Scherer & Oshinsky, 1977; Sievers et al., 2013; Thompson & Robitaille, 1992). Joy was communicated with a fast tempo, *staccato* articulation, high pitch, loud dynamics, bright sound, and major mode (Akkermans et al., 2019; Peretz et al., 1998; Quinto et al., 2014). The dynamics level for joy contrasted with one of the studies that registered low dynamics rather than high (Saarikallio et al., 2019). However, a low dynamics level for joy is not the norm, as most studies have reported a high dynamics level for joy (Akkermans et al., 2019; Bresin & Friberg, 2011; Gabrielsson & Lindström, 1995; Juslin & Laukka, 2003; Kragness & Trainor, 2019; Quinto et al., 2014; Saarikallio et al., 2014). Calmness was represented by a

Table 5. Cue Combinations Utilised by Participants for Each Emotion With Discrepancies to Past Production Studies Highlighted.

Emotion		Sadness	Joy	Calmness	Anger	Fear	Power	Surprise
Cues	Tempo	--	++	--	++	+	+	++
	Articulation	<i>leg.</i>	<i>stac.</i>	<i>leg.</i>	<i>stac.</i> ²	<i>det.</i> ¹	<i>det.</i>	<i>stac.</i>
	Pitch	-	+	- ¹	+ ²	+ ¹	++	+
	Dynamics	--	++ ⁴	/ ^{1, 3}	- ^{2, 3, 4}	- ¹	-	+
	Brightness	-	++	/	/ ²	/	+	++
	Mode	-	+	+	-	-	+	+

Notes. -- = very low/slow, - = low/slow, / = moderate, + = high/fast, ++ = very high/fast. For articulation, *leg.* = legato, *stac.* = staccato, *det.* = détaché. For mode, + = major, - = minor. The numeric values in subscript refer to the following studies: 1 = Bresin and Friberg (2011), 2 = Saarikallio et al. (2014), 3 = Kragness and Trainor (2019), 4 = Saarikallio et al. (2019). The differences in results between the current study and any of the aforementioned results are indicated by the corresponding numeric value of the previous study being written in subscript in the columns of the table.

slow tempo (Sievers et al., 2013), *legato* articulation, low pitch, moderate dynamics, a rich sound, and major mode (Eerola et al., 2013; Kragness & Trainor, 2019). The pitch level for calmness varied between the current study and a previous one, as participants opted for a low pitch in this study, and a high pitch in the previous study (Bresin & Friberg, 2011). However, looking beyond production studies, both low pitches (Gundlach, 1935) and high pitches (Eerola et al., 2013; Hevner, 1937) have been registered as conveying calmness. The dynamics level for calmness sits between low to moderate, which is slightly different from previous studies, where a low dynamics level was consistent across studies (Bresin & Friberg, 2011; Eerola et al., 2013; Watson, 1942). Anger was characterised by a fast tempo, *staccato* articulation, high pitch, a moderate level of harmonic content, minor mode (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Sievers et al., 2013), and most interestingly, very soft dynamics. Articulation, pitch, and brightness levels differ from the production study carried out by Saarikallio et al. (2014). However, the articulation, pitch, and brightness levels for the anger emotion resulting from this current study are in line with the other studies being compared in Table 5 as well as other previous studies not following a production approach (Gabrielsson & Juslin, 1996; Juslin & Lindström, 2010; Quinto et al., 2014). Saarikallio et al. (2014) had proposed that these differences may be due to the participant pools utilised, as their study focussed on adolescents, rather than adult participants. The authors had in fact noted that the discrepancies in results might be due to the variance between the socio-emotional abilities of adolescents and adults (Saarikallio et al., 2014). The starkest contrast lies between the low dynamics level achieved for anger in this current study, as against previous literature, where very loud dynamics have been associated with anger (Chau & Horner, 2015; Kragness & Trainor, 2019; Saarikallio et al., 2014). This discrepancy might be due to participants' differing views on what constitutes anger and what type of anger they were trying to portray (e.g., passive aggressiveness, open aggression, assertive anger). The variances between participants' definition of anger might stem from participants' different experiences and social interactions (Susino & Schubert, 2017). Perhaps providing definitions of the target emotions to the participants prior to the musical task, would have ensured that participants were aiming to convey the same type of emotion through their compositions. Alternatively, anger being represented by a soft dynamics level might be distinct to this particular musical piece due to the particular cue combination used. All cues except for brightness had a significant effect on the portrayal of anger in the musical piece, which suggests that participants specifically chose to use a soft dynamics level together with a fast tempo, *staccato* articulation, a high pitch level and minor mode. This finding provides further motivation to investigate the effect of cues on emotional expressions in music as combinations of multiple cues

rather than as individuals, as an individual cue might not give enough information to portray a specific emotion in music (Eerola et al., 2013; Gabrielsson, 2008).

Fast tempo, *detaché* articulation, high pitch, soft dynamics, a dark sound, and minor mode represented fear (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Juslin, 1997b, 2000; Scherer & Oshinsky, 1977; Sievers et al., 2013). Previous literature suggests that fear may be expressed by both low (Bresin & Friberg, 2011; Eerola et al., 2013) and high pitches (Scherer & Oshinsky, 1977). A fast tempo, *detaché* articulation, high pitch, soft dynamics, a bright sound, and major mode conveyed power (Rigg, 1940). Finally, surprise was expressed through a fast tempo, *staccato* articulation, high pitch, loud dynamics, a bright sound, and major mode (Scherer & Oshinsky, 1977). Fear and surprise have been characterised by *staccato* articulation in previous literature (Juslin, 1997b). However, this might be because mostly two levels (*legato*, *staccato*) of articulation have been investigated (Juslin, 1997b; Wedin, 1972).

Effectiveness of cue combinations to predict emotions. The discrepancies in certain cue values across previous literature and this current study might suggest why specific emotions might be more challenging to predict utilising certain cues. Table 3 gives a summary of which cues provide a significant weight and thus the most influence in characterising an emotion. Furthermore, it identifies the cues which are not adding flavour to the emotion recognition process. The cues' influence on the emotions conveyed may be more easily reconciled if they are considered through a modified version of Brunswik's lens model (Juslin, 1997b, 2000). The lens model theory suggests that successful emotion communication through music is determined by the layering of cues and their interaction (see also Argstatter, 2016; Eerola et al., 2013; Gabrielsson, 2008; Ramos et al., 2011).

Limitations of the experiment. A possible limitation of this experiment is the fact that participants were not given definitions of the emotion terms utilised. Although participants were asked whether the emotion terms were clear to them, which everyone agreed to be the case, it raises the question of whether participants were trying to convey the same type of emotion or not, as the different emotion terms might have different meanings to the participants. It has been reported that different sub-types of an emotion term account for variance in a musical piece's structure (Warrenburg, 2020b) and thus, might explain the inconsistencies present in interpretations of the same emotional expressions. Therefore, in future studies, asking the participants to provide definitions of their understanding of the different emotion terms might make understanding what sub-type of the emotion they were trying to express, clearer. A post-cue manipulation task question on whether the participants were satisfied with their musical creations might also have given a better understanding of the available musical pieces and

cues' roles in conveying the intended emotions. Additionally, although cue ranges of *EmoteControl* allow for a large combination of cue levels to be explored, there is always the possibility of making these ranges larger to increase the cue space being investigated. In particular, cues such as pitch and mode could be altered to include more semitones and modes, to investigate whether bigger cue ranges would influence how users utilise the cues to portray the different emotions.

Finally, to further explore how certain cue combinations affect the portrayal of different emotions in music, another music evaluation experiment was conducted (Experiment 3) where we took the musical pieces featuring the optimal cue combinations used by the participants to portray specific emotions in Experiment 2, as well as the original versions of the pieces which were rated in Experiment 1, to investigate how well they communicate their intended emotions to other listeners.

Experiment 3: Evaluation of More Music Stimuli

This final experiment investigated whether the musical creations produced through cue alterations by participants in Experiment 2 also conveyed the intended emotion to other listeners. Furthermore, we wanted the participants of Experiment 3 to rate the original musical pieces already validated in Experiment 1 to confirm whether the pieces successfully conveyed their intended emotion to other listeners. Therefore, Experiment 3 allowed us to examine how the variations of the same musical pieces were perceived and delve into a more detailed investigation of how well the cue combinations used conveyed the intended emotion, with the aim of gaining more insight into cue combinations and their role in emotion shaping in music.

Method

Participants. Participants were recruited via social media and email notices. Ninety-one participants (23 men, 67 women, one individual did not indicate their gender) between 18 and 71 years of age ($M = 34.99$, $SD = 15.86$) took part in the study. A one-question version of the OMSI (Ollen, 2006; Zhang & Schubert, 2019) was utilised to determine the participants' level of musical expertise. Fifty-seven of the participants were non-musicians, and 34 were musicians. Participation in the study was voluntary.

Material. Fourteen musical pieces were utilised as stimuli. These encompassed the original seven piano musical pieces, previously validated as conveying one of the following emotions: joy, sadness, calmness, anger, fear, power, or surprise, in Experiment 1, as well as the new variations of these pieces, created using the mean cue values participants

utilised in Experiment 2, which were rendered utilising the default virtual instrument used as output in the *EmoteControl* interface; a chamber strings virtual instrument from the Vienna Symphonic Library sound library. The musical stimuli can be found on OSF repository².

Procedure. The study was administered online and carried out in English with the exact instructions and scales as in Experiment 1 (for details, see Exp.1 procedure). Full instructions can be seen in the Supplementary Material.

Results

The consistency of how the participants used the individual rating scales was calculated using Cronbach's alpha (intra-class correlation coefficient) to examine the inter-rater reliability within each emotion rating scale across each participant and musical piece. High consistency was observed for all emotion rating scales across the participants, especially in the sadness emotion rating scale $\alpha = 0.995$, joy $\alpha = 0.995$, calmness $\alpha = 0.994$, followed by fear $\alpha = 0.990$, anger $\alpha = 0.990$, and surprise $\alpha = 0.978$ rating scales, with the power rating scale having the lowest consistency score $\alpha = 0.962$.

Table 6 presents a general summary of whether the different emotion scales rated had an overall significant effect on Piece, Source, and the interaction between Piece and Source fixed factors, with Participant being the random factor. Linear mixed models (LMMs) were run for each emotion scale rated, one with the factor in question and one without (e.g., one LMM run with Piece as fixed factor, and the other LMM without the fixed factor), with the significance of the difference depending on the factor (if any) being calculated using a likelihood ratio test. The Source factor indicates whether the musical pieces are the original seven from Experiment 1 or the participant-proposed versions from Experiment 2. Table 6 shows how all interactions between the different emotions rated across pieces are significant. The source of the pieces had a significant effect on two of the emotions (anger and joy), which indicates that participants utilised the emotion scales differently for those particular music pieces,

Table 6. LMM Results for Seven Rated Emotions for the Main Effect of Musical Pieces, Sources, and Their Interactions Using the Likelihood Ratio Test.

	Piece	Source	Piece $\times$ Source
Anger	667.15***	56.702***	105.12***
Calmness	1001.6***	0.358	180.34***
Fear	859.4***	2.04	48.45***
Joy	1162.5***	15.84***	215.06***
Power	231.96***	0.06	73.2***
Sadness	1281.7***	1.47	165.57***
Surprise	438.13***	1.06	40.278***

Notes. * $p < .05$, ** $p < 0.01$, *** $p < .001$. $df = 6$ for Emotion, $df = 1$ for Source, $df = 6$ for Interaction for the likelihood ratio test.

Table 7. Mean Ratings of Emotions Perceived in the Musical Pieces.

Intended emotion (musical piece)															
Source	Anger		Calmness		Fear		Joy		Power		Sadness		Surprise		
	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	
Rated Emotion	Anger	3.40*(1.29)	2.45*(1.02)	1.04*** (0.25)	1.05*** (0.23)	2.99*(1.24)	2.12*** (1.05)	1.03*** (0.18)	1.07*** (0.29)	2.14*** (1.23)	1.16*** (0.48)	1.27*** (0.52)	1.29*** (0.64)	1.70*** (0.84)	1.27*** (0.65)
	Calmness	1.05*** (0.31)	1.14*** (0.55)	4.14 (0.96)	3.27*(1.18)	1.05*** (0.27)	1.42*** (0.84)	1.66*** (0.79)	1.38*** (0.65)	1.05*** (0.23)	2.12 (1.03)	2.80*** (1.05)	2.59*** (1.05)	1.59*** (0.80)	1.14*** (0.55)
	Fear	3.18 (1.23)	3.45*** (1.38)	1.08*** (0.31)	1.20*** (0.50)	3.26*(1.31)	3.29*(1.38)	1.07*** (0.29)	1.02*** (0.15)	2.02*** (1.16)	1.35*** (0.86)	1.56*** (0.67)	1.71*** (0.91)	1.78*** (0.98)	1.21*** (0.59)
	Joy	1.11*** (0.31)	1.11*** (0.41)	2.44*** (0.95)	2.37*** (1.06)	1.31*** (0.64)	1.21*** (0.51)	4.42 (0.86)	4.18 (0.90)	2.22*** (1.32)	3.07*** (1.18)	1.03*** (0.18)	1.04*** (0.21)	1.79*** (0.96)	3.55*** (1.08)
	Power	3.20*(1.19)	2.77** (1.29)	1.38*** (0.73)	1.95*** (1.19)	2.84** (1.24)	2.33*** (1.20)	1.73*** (1.07)	2.35*** (1.33)	2.74*(1.32)	2.07 (1.25)	1.43*** (0.86)	1.70*** (1.06)	2.32*** (1.15)	2.57*** (1.26)
	Sadness	1.56*** (0.76)	1.53*** (0.77)	2.11*** (1.00)	1.93*** (1.00)	1.36*** (0.66)	2.16*** (1.08)	1.00*** (0.00)	1.01*** (0.10)	1.18*** (0.41)	1.36*** (0.68)	4.47 (0.79)	4.14 (1.04)	2.15*** (1.03)	1.08*** (0.31)
	Surprise	1.90*** (1.09)	2.00*** (1.15)	1.15*** (0.47)	1.24*** (0.54)	2.57*** (1.28)	1.84*** (1.13)	2.07*** (1.13)	2.36*** (1.28)	2.29** (1.18)	1.96 (1.12)	1.10*** (0.37)	1.09*** (0.41)	2.90 (1.34)	3.04 (1.32)

Notes. The values in brackets represent the standard deviation of the means. * $p < .05$, ** $p < .01$, *** $p < .001$. Bonferroni corrected values from the one-way repeated measures ANOVA. Df (1, 90). The asterisks note when the mean rating of the intended emotion was significantly different from the other rated emotions. Values without any asterisks (apart from the intended emotion) represent the emotion ratings which were not significantly higher than the intended emotion's rating.

depending on their source. This can also be seen in the Piece $\times$ Source factor where all emotions had a significant effect on that interaction.

Due to source having a significant impact on how emotions were rated, the rest of this experiment's analysis will regard the music pieces from the different sources (Source 1: Exp.1, Source 2: Exp.2) separately. This will help determine how well the emotions are efficiently recovered in the musical pieces with the cue combinations used by participants in Experiment 2, as well as the original pieces composed in Experiment 1.

The data, filtered by source, was subjected to a one-way repeated measures ANOVA with the intended emotion being the independent variable and the collapsed ratings of the remaining emotion scales being the dependent variable. The main effect of intended emotion on emotion scale rating was significantly different, suggesting that in general, participants rated the intended emotion higher than the other six emotions in the pieces, in both sources; Experiment 1, $F(1, 90) = 1098.00$, $p < .001$; and Experiment 2, $F(1, 90) = 875.50$, $p < .001$. These findings reaffirm the hypothesis that emotions can be effectively encoded in music and communicated to the listener (Juslin, 1997a).

Table 7 displays the mean emotion ratings given by participants for each of the musical pieces in their respective source and allows for a contrast of means between the original pieces (Exp.1) and the participant-proposed musical variations (Exp.2) conveying the same emotion. Columns in the table refer to the seven different types of intended emotions in the music excerpts. Rows in the table refer to the seven emotion scales that participants rated for each excerpt, to establish how much of each emotion was conveyed through the excerpts. The ratings along the diagonal in bold are expected to be higher than the other ratings in their relative column. This is true for all Exp.1 music pieces, where the intended emotion was always rated highest. However, this was not the case for all Exp.2 musical pieces. The participant-proposed variations aiming to convey calmness, fear, joy, and sadness were given the highest ratings for their intended emotion, whilst excerpts composed to convey anger, power, and surprise were rated highest for other emotions.

A series of one-way repeated measures ANOVA were then computed for each musical piece to establish whether the difference between the intended emotion's and the other emotions' mean ratings was significant or not. The asterisks following the mean ratings in Table 7 represent how significantly different the intended emotion's rating was to the other emotions' ratings. Exp.1 music pieces composed to express calmness, joy, power, surprise, and sadness were rated significantly higher for their intended emotion than other emotions. Although the anger piece was rated highest for its intended emotion, the difference between anger, power, and fear in the anger-conveying piece was not significant. Similarly, the

difference in mean ratings between fear and anger for the fear-intended piece was non-significant in Experiment 1. Exp.2 pieces aiming to convey calmness, fear, joy, and sadness were all rated significantly higher for their intended emotion than other emotions, whilst excerpts intending to convey anger, power, and surprise were rated higher for a different emotion.

Discussion

The results confirmed that: the seven pieces carried forward from Experiment 1 are strong representatives of their intended emotion as they were all rated significantly highest for their target emotion. The pieces conveying calmness, fear, joy, and sadness from Experiment 2 have also been rated highest for their intended emotions, whilst the remaining three pieces were rated highest for other emotions. These findings allow us to gather more information on whether emotions are efficiently recovered in two variations of the same musical pieces aiming to convey the same emotion.

Comparisons between musical variations conveying the same emotion. As the Exp.2 pieces were variations created from the Exp.1 pieces, the excerpts expressing the same emotion from the two sources had quite similar characteristics overall. Calmness, fear, joy, and sadness pieces from both sources were rated highest for their intended emotion. Calmness pieces consisted of major mode, *legato* articulation, and a slow tempo (Bresin & Friberg, 2011; Hevner, 1937), with the highest rated piece for calmness having a high pitch. Fear candidates were both moderately fast and in minor mode, with the Exp.2 piece having *detaché* articulation rather than *legato* and a higher pitch (Juslin, 2000; Scherer & Oshinsky, 1977). Joy pieces featured major mode, fast tempo, and *staccato* articulation (Kragness & Trainor, 2019), however, the strongest candidate had the lowest pitch level and slowest tempo (Ilie & Thompson, 2006; Juslin & Lindström, 2010). Both sadness pieces were in minor mode, had a slow tempo, *legato* articulation, and a similar low pitch (Eerola et al., 2013).

Exp.1 representatives of anger, power, and surprise were rated highest for their intended emotion, whilst Exp.2 pieces for the aforementioned emotions were rated highest for other emotions. Both anger candidates were in minor mode; however, the Exp.1 piece had a lower pitch, slower tempo, less detached articulation, and louder dynamics (Saarikallio et al., 2014). The excerpts for anger had the biggest difference in mean ratings (Exp.1 = 3.40, Exp.2 = 2.45). Surprise and power candidates had the most variations between their counterparts. The strongest candidate of surprise (Exp.1) featured a moderate tempo, minor mode, and semi-detached articulation, whilst its counterpart had a fast tempo, major mode, a higher pitch and *staccato* articulation (Scherer & Oshinsky, 1977). The strongest power candidate (Exp.1) had a faster

tempo, louder dynamics, and a lower pitch level than its corresponding Exp.2 piece and was composed in minor rather than major.

Overall, Exp.1 pieces were rated as better representatives of their intended emotion than their corresponding Exp.2 pieces, with the fear candidate being the exception. In general, the differences between the two sources' musical variations were relatively subtle, bar for a contrast in the mode cue for power and surprise pieces. Another variable to be considered as a potential influencer on emotion perception in music is the instrument timbre (Balkwill & Thompson, 1999; Hailstone et al., 2009). All Exp.1 excerpts utilised a piano sound, whilst Exp.2 excerpts used a chamber strings sound, which might have also had a role in the perception of emotion in the music. The difference in fear ratings for the two fear representatives was minimal (Exp.1 = 3.26, Exp.2 = 3.29). A possible explanation for the Exp.2 piece having a higher rating might be that the characteristics of fear music, such as its roughness, loud dynamics, and high pitch on violins mimic acoustic features of human screams and thus communicate a notion of fear more effectively (Trevor et al., 2020). It has also been suggested that the piano timbre is relatively emotionally neutral in comparison to other instruments such as violins, guitars, and marimbas (Chau et al., 2015). Having an emotionally neutral timbre might provide the piano with more versatility when representing multiple emotions, which might account for the listeners' preference for the piano excerpts. Due to the possibility of timbre playing a role in how emotions were perceived across the two music sources, in future studies, timbre could be included as another parameter to be investigated, in order to properly determine its role in emotion communication in music.

Furthermore, the instrument being composed for might influence how the music is constructed, as the instrument's idiomatic features would be utilised and maximised to portray the desired emotion, and this uniqueness might limit their ability to translate to other instruments (Huron & Berec, 2009). The original Exp.1 pieces were composed for piano, and thus, playing to the piano's features' strengths, whilst the Exp.2 pieces retained the original piano pieces' structures but were played with an instrument which was not originally intended - a strings instrument timbre. As these two instruments have different timbres and onset and envelope characteristics, it is likely that the compositions are better suited for the original instrument, which might explain why Exp.1 pieces worked better than the others. Furthermore, the composer in Experiment 1 had use of all available musical features used to create the different musical pieces and full control of the piano, whilst participants in Experiment 2 were limited to changing six cues of the music and using the chamber strings instrument. Nevertheless, the findings of Experiment 3 revealed a similar overall pattern emerging across the two versions of the pieces, which confirms that regardless of a change in the makeup of the piece, cue combinations are

utilised in the same manner to portray a specific emotion in music (Eerola et al., 2013; Juslin & Lindström, 2010).

Limitations of the experiment. Similar to Experiment 1, due to the online administration of this experiment, the environment in which participants carried out the survey could not be controlled by the researchers. Therefore, although instructions to test their sound prior to the experiment, wear headphones or use good quality speakers in a quiet environment were given, we cannot tell whether the participants adhered to these specifications, which might have hindered their attention and performance in the study. Another limitation is that the Exp.1 musical pieces were rendered with the original piano timbre, while the Exp.2 pieces were rendered with the chamber strings virtual instrument utilised in the *EmoteControl* interface in Experiment 2. The mismatch in timbre did not allow for a direct comparison of stimuli. Rendering the two sets of stimuli with both piano and strings timbres would have allowed for a comparison between the stimuli from the two sources, and also provide information on how the timbre affects the emotion perceived in the music.

General Discussion

In this paper, three studies were carried out with the main aim of exploring a big cue space utilising an interactive paradigm where participants themselves changed musical pieces in real-time through a combination of structural and expressive cues to communicate different emotional expressions through the music. Furthermore, this paper also had two secondary aims: the first was to create new polyphonic music validated as expressing a target emotion, which would allow for experimental flexibility whilst also retaining ecological validity and avoiding the use of commercial music and any familiarity bias. The second sub-aim was to explore how well the participants' own musical interpretations expressed their intended emotion to other listeners and collect more information on how the structure and expression of musical pieces affect the emotional expression conveyed. These aims were attained through an iterative process.

First, 28 new, unfamiliar, musical pieces were specifically created, each with the intent of expressing one particular emotion from a selection of nine emotions (calmness, sadness, joy, anger, fear, power, longing, love, and surprise) which may be expressed through music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Zentner et al., 2008), to be used as stimuli for research purposes, and be compatible with the apparatus used in our main experiment (Experiment 2). Furthermore, the stimuli were strategically composed to allow a flexible amount of manipulation whilst retaining a delicate balance between ecological validity and experimental control (Eerola et al., 2013; Gabrielsson & Lindström, 2010; Juslin & Lindström, 2010; Juslin & Västfjäll, 2008) that might not be achievable with

commercial music. Experiment 1 provided participants' ratings on nine emotion scales for all pieces to determine which emotion(s) were successfully conveyed, with sixteen pieces of the stimuli being successful representatives of their intended emotion (a selection of sadness, joy, calmness, anger, fear, power, and surprise pieces), adding knowledge to previous literature about emotions expressed through music other than the most common ones which are sadness, happiness, and anger (Warrenburg, 2020a). Thus, this newly composed musical stimuli set together with emotion ratings is in itself a new contribution to the field, being available and accessible to others in an online OSF repository for future use.

Experiment 2 utilised an interactive computer interface, *EmoteControl* (Micallef Grimaud & Eerola, 2021) to investigate the importance and efficacy of a selection of structural and expressive cues (tempo, articulation, pitch, dynamics, brightness, and mode) and their combinations, rather than focussing only on expressive cues, where participants, irrelevant of any prior musical knowledge, altered the cues themselves to convey different emotions in music in real-time (Bresin & Friberg, 2011; Friberg, 2006; Friberg et al., 2000; Friberg et al., 2006; Kragness & Trainor, 2019; Ramirez & Hazan, 2005). This production approach allowed for an extensive cue space to be explored, one which would have been limited if utilising a traditional systematic approach which requires cue levels and combinations to be predetermined and rendered. Thus, Experiment 2 contributed new findings on the combinations and interactions of six cues in relation to seven emotions to the field of music and emotions research. Furthermore, this interactive approach allowed participants to take on the roles of composers and performers, irrelevant of their musical knowledge, or lack of, and *show* their understanding of how different emotions are expressed through music. The findings are generally in line with previous literature (see: Gabrielsson and Lindström, 2010), with some interesting exceptions like anger being conveyed with soft dynamics, and *detaché* articulation being preferred for fear and power.

Experiment 3 offered another opportunity to investigate how well different emotions were encoded by the composer in the original musical pieces (Experiment 1) and by Experiment 2 participants in their musical variations and decoded by other listeners. This was attained by carrying out an online rating experiment. The findings confirm that calmness, fear, joy, and sadness emotions were rated highest when they each were the intended emotion and thus successfully conveyed in both Exp.1 and Exp.2 pieces. On the other hand, anger, power, and surprise were identified as the intended emotions in the Exp.1 pieces, but not correctly recognised as the intended emotions in Exp.2 pieces. The comparison of how these variations of the same musical pieces were ranked with regards to the different emotions provides further insight in cue values and combinations used to portray the aforementioned seven emotions in music. Furthermore, these

results highlight the potential influence of the instrument timbre on the emotional expression conveyed by the music (Balkwill & Thompson, 1999; Eerola et al., 2012; Hailstone et al., 2009; Huron & Berec, 2009).

The experiments in this study contribute to and expand knowledge on the notion that certain musical cues used by composers and performers are important contributors in shaping the emotional expression conveyed by music (Eerola et al., 2013; Gabrielsson & Lindström, 2010; Juslin, 2000; Kragness & Trainor, 2019). They also provide corroborating evidence that specific cue combinations are consistently used to map particular emotions (Juslin, 2000; Juslin & Lindström, 2010). This paper supports the claim across two methodologies and provides mean values of tempo, articulation, pitch, brightness, dynamics, and mode cues utilised to distinguish each of seven emotions: anger, calmness, fear, joy, power, sadness, and surprise. Calmness, sadness, and joy were the three emotions most reliably communicated while power, fear, and surprise had the least reliable cue combinations produced in Experiment 2. Previous literature has suggested that basic emotions (Ekman, 1992), in particular, joy and sadness are the most accurately recognised emotions (Bigand et al., 2005; Gabrielsson & Juslin, 1996; Kallinen, 2005), and can also be recognised cross-culturally (Balkwill et al., 2004; Balkwill & Thompson, 1999; Fritz et al., 2009; Laukka et al., 2013). Although sadness and joy were the second and third most accurately recognised emotions, it is interesting to note that the emotion with the highest accuracy rating, which is calmness (91%), is not a basic emotion. Furthermore, the two least predicted emotions (fear and surprise) are indeed basic emotions. These findings suggest that musical emotions other than basic emotions can also be highly recognised in music. A potential explanation to how different emotions which also include non-basic ones may be communicated through music and perceived by the listener is due to the emotions' organisation on dimensional planes such as valence and arousal, stemming from core affects (Cespedes-Guevara & Eerola, 2018). An alternative theory is that listeners better recognise emotions due to the emotions' frequency in music (Kallinen, 2005), such as calmness being listed as one of the highest ranking emotions that can be conveyed through music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Zentner et al., 2008).

The findings of this paper also denoted how certain emotions were confused by participants, such as fear and anger, perhaps due to their similar musical characteristics as well as being negative emotions. Joy and surprise emotions also tended to be confused. Similarly, these two emotions share musical characteristics, which might explain why participants sometimes wrongly rated joy and surprise musical pieces. A small number of participants from Experiment 2 commented that power and surprise emotions were the trickiest to try and convey through the musical pieces. This was reflected in the results of

Experiment 2 as surprise (37.1%) had the lowest cue-emotion reliability model from all the emotions. The difficulty in conveying and identifying surprise in music might be due to the fact that surprise may be regarded as both a positive and negative emotion (Kallinen, 2005). The musical variations created by participants in Experiment 2 for power and surprise emotions were also not successful in conveying their target emotions, as they were rated highest for other emotions than their intended ones in Experiment 3.

This paper also presented new information on the weight of the cues in characterising an emotion and the prediction rate of correctly identifying an emotion via the selected cue combinations, giving a better understanding of the impact of cues on the creation of emotional code in music. Overall, mode was the strongest contributor to the portrayal of different emotions across different pieces, followed by tempo, which complements previous findings (Eerola et al., 2013). Dynamics and brightness had the least communicative weight across emotions, in comparison to the other cues. It is interesting to note that albeit the ranking of cues with respect to their contribution to the intended emotion, in certain situations, cues which have been ranked as strong discriminators of emotions, did not significantly contribute to the conveying of a specific emotion (e.g., tempo did not have a significant effect on the portrayal of power). This suggests that the cues' combination and interaction affects how an emotion is shaped in the music (Argstatter, 2016; Eerola et al., 2013). Therefore, this paper also gives motivation for future research to further explore multiple combinations of cues in order to better understand their effect on the emotion communicated.

The series of empirical observations presented multiple, novel contributions to the current field of music emotion research. The systematic production approach used (Experiment 2) determined levels and combinations of six musical cues for the communication of seven emotional expressions, providing researchers with a form of presets which may be utilised to create the desired emotions across different musical pieces. The findings support and expand on previous literature (Bresin & Friberg, 2011; Kragness & Trainor, 2019; Saarikallio et al., 2019, 2014). Furthermore, this paper presented new, rated music stimuli (Experiment 1) and novel data leading towards a better understanding of perceived emotions in music (all experiments). Future studies adopting a production approach should explore different cues, such as the role of timbre and its interactions with other cues to portray different emotions. Other emotions that may be communicated in music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003) should also be investigated. Emotional states that are currently encapsulated under a single emotion term, such as melancholy and grief (Warrenburg, 2020b) should also be studied to identify distinctions between emotional states that music, as a highly expressive medium, has to offer.

Author Contributions

AMG and TE conceived and designed the experiments. AMG composed the music stimuli, carried out data collection, and participated in a small portion of the data analysis. TE carried out the experiments' analyses. Both AMG and TE contributed to the writing, reviewing, and editing of the manuscript and approved the final version of the manuscript.

Action Editor

Isabel Martínez, Universidad Nacional de La Plata, Facultad de Bellas Artes.

Peer Review

Two anonymous reviewers.

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Tertiary Education Scholarship Scheme (Malta).

ORCID iDs

Annaliese Micallef Grimaud  <https://orcid.org/0000-0002-4772-0983>

Tuomas Eerola  <https://orcid.org/0000-0002-2896-929X>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. The 28 musical pieces can be found on the OSF repository by following this link <https://osf.io/t9nhp/>
2. The 14 musical pieces can be found on the OSF repository by following this link <https://osf.io/m26bu/>

References

- Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., & Frieler, K. (2019). Decoding emotions in expressive music performances: A multi-lab replication and extension study. *Cognition and Emotion*, 33(6), 1099–1118. <https://doi.org/10.1080/02699931.2018.1541312>
- Argstatter, H. (2016). Perception of basic emotions in music: Culture-specific or multicultural? *Psychology of Music*, 44(4), 674–690. <https://doi.org/10.1177/0305735615589214>
- Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception: An Interdisciplinary Journal*, 17(1), 43–64. <https://doi.org/10.2307/40285811>
- Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. I. E. (2004). Recognition of emotion in Japanese, western, and hindustani music by Japanese listeners. *Japanese Psychological Research*, 46(4), 337–349.
- Behrens, G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music*, 21(1), 20–33. <https://doi.org/10.1177/030573569302100102>
- Bigand, E., Viellard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, 19(8), 1113–1139. <https://doi.org/10.1080/02699930500204250>
- Bresin, R., & Friberg, A. (2011). Emotion rendering in music: Range and characteristic values of seven musical variables. *CORTEX*, 47(9), 1068–1081. <https://doi.org/10.1016/j.cortex.2011.05.009>
- Brunswick, E. (1956). *Perception and the representative design of psychological experiments. Perception and the representative design of psychological experiments, 2nd ed.* University of California Press.
- 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, 1–19. <https://doi.org/10.3389/fpsyg.2018.00215>
- Chau, C. J., & Horner, A. (2015). The effects of pitch and dynamics on the emotional characteristics of piano sounds. ICMC 2015 - Proceedings of the 2015 International Computer Music Conference, (September), 372–375.
- Chau, C. J., Wu, B., & Horner, A. (2015). The emotional characteristics and timbre of nonsustaining instrument sounds. *AES: Journal of the Audio Engineering Society*, 63(4), 228–244. <https://doi.org/10.17743/jaes.2015.0016>
- Chen, C., Liaw, A., & Breiman, L. (2004). Using random forest to learn imbalanced data. *University of California, Berkeley*, 110(2), 1–12.
- Collier, G. L. (2007). Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, 35(1), 110–131. <https://doi.org/10.1177/0305735607068890>
- Costa, M., Fine, P., & Ricci Bitti, P. E. (2004). Interval distributions, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Perception: An Interdisciplinary Journal*, 22(1), 1–14. <https://doi.org/10.1525/mp.2004.22.1.1>
- Costa, M., Ricci Bitti, P. E., & Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music*, 28(1), 4–22. <https://doi.org/10.1177/0305735600281002>
- Cunningham, J. G., & Sterling, R. S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion*, 12(4), 399–413. <https://doi.org/10.1007/BF00992362>
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80, 1–10.
- Eerola, T., Ferrer Flores, R., Alluri, V., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated

- instrument sounds. *Music Perception*, 30(1), 49–70. <https://doi.org/10.1525/mp.2012.30.1.49>
- Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: Contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, 4(July), 1–12. <https://doi.org/10.3389/fpsyg.2013.00487>
- Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal*, 30(3), 307–340. <https://doi.org/10.1525/mp.2012.30.3.307>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169–200.
- Friberg, A. (2006). pDM: An expressive sequencer with real-time control of the KTH music-performance rules. *Computer Music Journal*, 30(1), 37–48. <https://doi.org/10.1162/014892606776021308>
- Friberg, A., & Battel, G. U. (2002). Structural communication. In R. Parncutt & G. E. McPherson (Eds.), *The science & psychology of music performance: Creative strategies for teaching and learning* (pp. 199–218). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195138108.003.0013>
- Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, 2(2–3), 145–161.
- Friberg, A., Bresin, R., & Sundberg, J. (2014). Analysis by synthesis. In W. F. Thompson (Ed.), *Music in the social and behavioral sciences: An encyclopedia* (Vol. 1, pp. 44–45). Sage Publications, Inc. <https://doi.org/10.4135/9781452283012>
- Friberg, A., Colombo, V., Frydén, L., & Sundberg, J. (2000). Generating musical performances with director musices. *Computer Music Journal*, 24(3), 23–29. <https://doi.org/10.1162/014892600559407>
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19(7), 573–576. <https://doi.org/10.1016/j.cub.2009.02.058>
- Gabrielsson, A. (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae VO - Special Issue 2001–2002*, 123–147. <https://doi.org/10.1177/10298649020050S105>
- Gabrielsson, A. (2008). The relationship between musical structure and perceived expression. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (1st ed, pp. 141–150). Oxford University Press. <https://doi.org/10.1093/oxfordhob/9780199298457.013.0013>
- Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24(1), 68–91. <https://doi.org/10.1177/0305735696241007>
- Gabrielsson, A., & Lindström, E. (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology*, 14(1), 94–116.
- Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In *Handbook of music and emotion: Theory, research, applications* (pp. 367–400). Oxford University Press.
- Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to “happy-sad” judgements in equitone melodies. *Cognition and Emotion*, 17(1), 25–40. <https://doi.org/10.1080/02699930302279>
- Goebel, W., & Bresin, R. (2001). Are computer-controlled pianos a reliable tool in music performance research? Recording and reproduction precision of a Yamaha Disklavier grand piano.
- Gundlach, R. H. (1935). Factors determining the characterization of musical phrases. *The American Journal of Psychology*, 47, 624–643. <https://doi.org/10.2307/1416007>
- Hailstone, J. C., Omar, R., Henley, S. M. D., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *Quarterly Journal of Experimental Psychology*, 62(11), 2141–2155. <https://doi.org/10.1080/17470210902765957>
- Hevner, K. (1935). The affective character of the Major and Minor modes in music. *The American Journal of Psychology*, 47(1), 103–118.
- Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, 48(2), 246–268. <https://doi.org/10.2307/1415746>
- Hevner, K. (1937). The affective value of pitch and tempo in music. *The American Journal of Psychology*, 49(4), 621–630.
- Huron, D., & Berec, J. (2009). Characterizing idiomatic organization in music: A theory and case study of musical affordances. *Empirical Musicology Review*, 4(3), 103–122. <https://doi.org/10.18061/1811/44531>
- Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, 23(4), 319–329.
- Juslin, P. N. (1997a). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception*, 14(4), 383–418. <https://doi.org/10.2307/40285731>
- Juslin, P. N. (1997b). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae*, 1(2), 225–256.
- Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6), 1797–1813. <https://doi.org/10.1037/0096-1523.26.6.1797>
- Juslin, P. N. (2013). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, 4(596), 1–14. <https://doi.org/10.3389/fpsyg.2013.00596>
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770–814. <https://doi.org/10.1037/0033-2909.129.5.770>
- Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, 33(3), 217–238. <https://doi.org/10.1080/0929821042000317813>
- Juslin, P. N., & Lindström, E. (2010). Musical expression of emotions: Modelling listeners' judgements of composed and performed features. *Music Analysis*, 29(1–3), 334–364. <https://doi.org/10.1111/j.1468-2249.2011.00323.x>

- Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In *Handbook of music and emotion: Theory, research, applications* (pp. 453–489). 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/S0140525X08005293>
- Kallinen, K. (2005). Emotional ratings of music excerpts in the western art music repertoire and their self-organization in the kohonen neural network. *Psychology of Music*, 33(4), 373–393.
- Kamenetsky, S. B., Hill, D. S., & Trehub, S. E. (1997). Effect of tempo and dynamics on the perception of emotion in music. *Psychology of Music*, 25(2), 149–160. <https://doi.org/10.1177/0305735697252005>
- Kastner, M. P., & Crowder, R. G. (1990). Perception of the Major/Minor distinction: IV. *Emotional Connotations in Young Children. Music Perception: An Interdisciplinary Journal*, 8(2), 189–201.
- Kragness, H. E., Eitel, M. J., Baksh, A. M., & Trainor, L. J. (2021). Evidence for early arousal-based differentiation of emotions in children's Musical production. *Developmental Science*, 24(1), 1–11. <https://doi.org/10.1111/desc.12982>
- Kragness, H. E., & Trainor, L. J. (2019). Nonmusicians express emotions in musical productions using conventional cues. *Music & Science*, 2, 205920431983494. <https://doi.org/10.1177/2059204319834943>
- Kreutz, G. (2000). Basic emotions in music. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the 6th international conference on music perception and Cognition*. Keele, Staffordshire, UK, (CD-rom), Department of Psychology, Keele University. <https://www.escom.org/proceedings/ICMPC2000/Tue/Kreutz.htm>
- Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. In *Canadian Journal of experimental psychology/revue canadienne de psychologie expérimentale*. 51(4), 336–353. Canadian Psychological Association. <https://doi.org/10.1037/1196-1961.51.4.336>
- Lahdelma, I., & Eerola, T. (2016). Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychology of Music*, 44(1), 37–54. <https://doi.org/10.1177/0305735614552006>
- Laukka, P., Eerola, T., Thingujam, N. S., Yamasaki, T., & Beller, G. (2013). Universal and culture-specific factors in the recognition and performance of musical affect expressions. *Emotion (Washington, D C)*, 13(3), 434–449. <https://doi.org/10.1037/a0031388>
- Lindström, E. (2006). Impact of melodic organization on perceived structure and emotional expression in music. *Musicae Scientiae*, 10(1), 85–117. <https://doi.org/10.1177/102986490601000105>
- Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). “Expressivity comes from within your soul”: A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, 20(1), 23–47.
- Livingstone, S. R., Muhlberger, R., Brown, A. R., & Thompson, W. F. (2010). Changing musical emotion: A computational rule system for modifying score and performance. *Computer Music Journal*, 34(1), 41–64. <https://doi.org/10.1162/comj.2010.34.1.41>
- Lyvers, M., Cotterell, S., & Thorberg, F. A. (2018). “Music is my drug”: Alexithymia, empathy, and emotional responding to music. *Psychology of Music*, 48(5) 626–641. <https://doi.org/10.1177/0305735618816166>
- Micallef Grimaud, A., & Eerola, T. (2021). Emotecontrol: An interactive system for real-time control of emotional expression in music. *Personal and Ubiquitous Computing*, 25(4), 677–689. <https://doi.org/10.1007/s00779-020-01390-7>
- Midya, V., Valla, J., Balasubramanian, H., Mathur, A., & Singh, N. C. (2019). Cultural differences in the use of acoustic cues for musical emotion experience. *PLOS ONE*, 14(9), e0222380. <https://doi.org/10.1371/journal.pone.0222380>
- Mohn, C., Argstatter, H., & Wilker, F. W. (2011). Perception of six basic emotions in music. *Psychology of Music*, 39(4), 503–517. <https://doi.org/10.1177/0305735610378183>
- Morreale, F., Masu, R., & Angeli, A. D. (2013). Robin: An algorithmic composer For interactive scenarios. *Sound and Music Computing Conference 2013, SMC, 2013*, 207–212.
- Morreale, F., Masu, R., Angeli, A. D., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*.
- Ollen, J. E. (2006). A criterion-related validity test of selected indicators of musical sophistication using expert ratings. *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC9)*, 690–696.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(3), 111–141.
- Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89(4), 344–350.
- Quinto, L., Thompson, W. F., & Keating, F. L. (2013). Emotional communication in speech and music: The role of melodic and rhythmic contrasts. *Frontiers in Psychology*, 4, 1–8. <https://doi.org/10.3389/fpsyg.2013.00184>
- Quinto, L., Thompson, W. F., & Taylor, A. (2014). The contributions of compositional structure and performance expression to the communication of emotion in music. *Psychology of Music*, 42(4), 503–524. <https://doi.org/10.1177/0305735613482023>
- Ramirez, R., & Hazan, A. (2005). Modeling Expressive Performance in Jazz. In *Proceedings of 18th International Florida Artificial Intelligence Research Society Conference* (pp. 86–91). Menlo Park, California: Association for the Advancement of Artificial Intelligence Press.
- Ramos, D., Bueno, J. L. O., & Bigand, E. (2011). Manipulating Greek musical modes and tempo affects perceived musical emotion in musicians and nonmusicians. *Brazilian Journal of Medical and Biological Research*, 44(2), 165–172. <https://doi.org/10.1590/S0100-879X2010007500148>
- Rigg, M. G. (1940). The effect of register and tonality upon musical mood. *Journal of Musicology*, 2(2), 49–61.

- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
- Saarikallio, S., Tervaniemi, M., Yrtti, A., & Huotilainen, M. (2019). Expression of emotion through musical parameters in 3- and 5-year-olds. *Music Education Research*, 21(5), 596–605. <https://doi.org/10.1080/14613808.2019.1670150>
- Saarikallio, S., Vuoskoski, J., & Luck, G. (2014). Adolescents' expression and perception of emotion in music reflects their broader abilities of emotional communication. *Psychology of Well-Being*, 4, 21. <https://doi.org/10.1186/s13612-014-0021-8>
- Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, 1(4), 331–346.
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, 21(4), 561–585. <https://doi.org/10.1525/mp.2004.21.4.561>
- Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1), 70–75. <https://doi.org/10.1073/pnas.1209023110>
- Silverman, M. J. (2008). Nonverbal communication, music therapy, and autism: A review of literature and case example. *Journal of Creativity in Mental Health*, 3(1), 3–19. <https://doi.org/10.1080/15401380801995068>
- Smith, L. D., & Williams, R. N. (1999). Children's artistic responses to musical intervals. *The American Journal of Psychology*, 112(3), 383–410.
- Susino, M., & Schubert, E. (2017). Cross-cultural anger communication in music: Towards a stereotype theory of emotion in music. *Musicae Scientiae*, 21(1), 60–74. <https://doi.org/10.1177/1029864916637641>
- Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, 10(1), 79–89.
- Trevor, C., Arnal, L. H., & Frühholz, S. (2020). Terrifying film music mimics alarming acoustic feature of human screams. *The Journal of the Acoustical Society of America*, 147(6), EL540–EL545. <https://doi.org/10.1121/10.0001459>
- Turnbull, D., Barrington, L., Torres, D., & Lanckriet, G. (2008). Semantic annotation and retrieval of music and sound effects. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), 467–476.
- Vidas, D., Dingle, G. A., & Nelson, N. L. (2018). Children's recognition of emotion in music and speech. *Music & Science*, 1, 205920431876265. <https://doi.org/10.1177/2059204318762650>
- Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, 22(4), 720–752. <https://doi.org/10.1080/02699930701503567>
- Warrenburg, L. A. (2020a). Choosing the right tune: A review of music stimuli used in emotion research. *Music Perception*, 37(3), 240–258. <https://doi.org/10.1525/MP.2020.37.3.240>
- Warrenburg, L. A. (2020b). Redefining sad music: Music's structure suggests at least two sad states. *Journal of New Music Research*, 49(4), 373–386. <https://doi.org/10.1080/09298215.2020.1784956>
- Watson, K. B. (1942). The nature and measurement of musical meanings. *Psychological Monographs*, 54(2), i–43. <https://doi.org/10.1037/h0093496>
- Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, 13, 241–257.
- Zentner, M., 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>
- Zhang, J., & Schubert, E. (2019). A single item measure for identifying musician and nonmusician categories based on measures of musical sophistication. *Music Perception*, 36(5), 457–467. <https://doi.org/10.1525/MP.2019.36.5.457>