Towards a Better Understanding of Emotion Communication in Music: An Interactive Production Approach.

MICALLEF-GRIMAUD, ANNALIESE

Towards a Better Understanding of Emotion Communication in Music: An Interactive Production Approach.

Annaliese Micallef Grimaud

Abstract

It has been well established that composers and performers are able to encode certain emotional expressions in music, which in turn are decoded by listeners, and in general, successfully recognised. There is still much to discover, however, as to how musical cues combine to shape different emotions in the music, since previous literature has tended to focus on a limited number of cues and emotional expressions. The work in this thesis aims to investigate how combinations of tempo, articulation, pitch, dynamics, brightness, mode, and later, instrumentation, are used to shape sadness, joy, calmness, anger, fear, power, and surprise in Western tonal music. In addition, new tools for music and emotion research are presented with the aim of providing an efficient production approach to explore a large cue-emotion space in a relatively short time. To this end, a new interactive interface called *EmoteControl* was created which allows users to alter musical pieces in real-time through the available cues. Moreover, musical pieces were specifically composed to be used as stimuli. Empirical experiments were then carried out with the interface to determine how participants shaped different emotions in the pieces using the available cues. Specific cue combinations for the different emotions were produced. Findings revealed that overall, mode and tempo were the strongest contributors to the

conveyed emotion whilst brightness was the least effective cue. However, the importance of the cues varied depending on the intended emotion. Finally, a comparative evaluation of production and traditional approaches was carried out which showed that similar results may be obtained with both. However, the production approach allowed for a larger cue-emotion space to be navigated in a shorter time. In sum, the production approach allowed participants to directly show us how they think emotional expressions should sound, and how they are shaped in music.

Towards a Better Understanding of Emotion Communication in Music: An Interactive Production Approach.

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

by

Annaliese Micallef Grimaud

Department of Music

Durham University

2022

Table of Contents

List of Tables

List of Figures

Declaration of Academic Achievement

This thesis consists of six chapters, with three of the chapters containing work published in peer-reviewed journals. The thesis comprises an introductory chapter which presents an overview of the current literature pertaining to this work. The second and third chapters give an account of the creation of the interactive interface *EmoteControl* and a set of newly composed musical pieces which were used to carry out empirical experiments in this thesis. Chapters 4 and 5 are empirical chapters. Lastly, Chapter 6 presents a general discussion of the overall work in this thesis.

Chapters 2 and 4 have been published in peer-reviewed scientific journals, whilst Chapter 5 has been submitted to a scientific journal and is currently under review. The material presented in Chapters 2, 4, and 5 is mostly the same as the content in the respective published papers. However, where appropriate, reference to other chapters in this thesis have been added – which would not be present in the published papers. The author of this thesis is the primary author of all six chapters. I developed the interactive interface called *EmoteControl* described in Chapter 2 under the guidance of my supervisors Tuomas Eerola and Nick Collins. I composed all the musical excerpts described in Chapter 3, from which subsets were used as stimuli in the experiments reported in this thesis. For each study, I was the primary individual responsible for the experiment design, preparation of the stimuli and other required material, data collection, data analysis, and the writing up. The experiments were carried out in consultation with Tuomas Eerola, who initially led the data analyses and contributed to the writing and editing of the published manuscripts, and thus is a co-author of Chapters 2, 4, and 5.

Statement of Copyright

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Tuomas Eerola, for believing in me since Day One, and providing continuous support throughout my doctoral studies, both academically and professionally. I am extremely lucky to have had such a wonderful mentor who shared his immense knowledge and expertise with me, gave me invaluable advice, and always encouraged me to think outside of the box and not shy away from a challenge – you are truly an inspiration. I am also deeply grateful to my secondary supervisor, Professor Nick Collins, for giving me constant support across multiple aspects of the PhD, from troubleshooting in Max/MSP to providing insightful feedback on my writing and encouraging me to pursue opportunities beneficial to my professional development. I am very fortunate to have had such a fantastic supervisory team.

I would also like to thank the Tertiary Education Scholarship Scheme (Malta) who funded my PhD and provided me with financial support. A huge thanks also goes to all the individuals who participated in my studies and were happy with being remunerated in chocolate – this project would not have been possible without you.

Thanks goes to all the wonderful people who I am grateful to call my colleagues and friends at the Music and Science Lab. Thank you for providing a friendly lab environment, giving me insightful comments and suggestions, and inspiring me to do my best. I am also grateful for all the people I met through the university, college, and conferences – thank you for all the fantastic experiences that enriched my doctoral studies. I am extremely grateful for all my friends all around the globe, old and new, who supported me on this journey, and checked in on me one way or another. A heartfelt thank you goes to Yvonne and Giselle for being wonderful office mates and keeping me caffeinated – sharing a workspace with you made the difficult times a little less stressful and a little more cheerful. Special thanks also go to Steve and Aby for being awesome flatmates, and for the memorable adventures we shared. I am deeply grateful for my fellow Maltese-ers who I met in Durham – Noelle, Jessica, Mark, Matthew, Andrew, Roberta, and Jeanette – I feel extremely lucky to have found a piece of home in Durham. A heartfelt thank you also goes to my life-long friend Althea Marie, who has been cheering me on in my ventures ever since we were kids.

My deepest gratitude goes to my family members for their unconditional love and support throughout the years. A huge thank you goes to my wonderful parents Aurelio and Giulia, who are my greatest inspiration and taught me that anything is possible with determination and plenty of coffee. Thank you for always encouraging me in everything that I do and for making it possible for me to pursue my dreams. I am deeply grateful for my brother André and his constant support with everything, from being my moving assistant, to buying me treats, coming to visit, and organising lovely trips whenever I was visiting him in London, allowing me to decompress. I would like to extend my sincere thanks to my grandparents, for their endless love and for always believing in me – my Nanna Dolores, as well as my late Nanna Lela and Nannu Manuel, who saw the beginning of this adventure, but sadly not the end. Finally, I could not have completed this journey without the unwavering support of my partner Joseph, who constantly pushed me to achieve my goals, patiently listened to all my venting, sent over numerous pick me ups in the form of coffee and cake, and magically always knew how to anticipate my needs and what to say – thank you from the bottom of my heart.

List of Publications

The work presented in Chapter 2 is published in:

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), 677-680. https://doi.org/10.1007/s00779-020-01390-7

The research reported in Chapter 4 is published in:

Micallef Grimaud, A., & Eerola, T. (2022). An interactive approach to emotional expression through musical cues. *Music and Science*, *5*, 1-23. https://doi.org/10.1177/20592043211061745

The research reported in Chapter 5 is under review in:

Micallef Grimaud, A., & Eerola, T. (in review). Emotional expression through musical cues: A comparison of production and evaluation approaches. *PLOS One.*

Additional publications and material related to the present work:

Micallef Grimaud, A., Eerola, T., & Collins, N. (2019). EmoteControl: A system for live-manipulation of emotional cues in music. In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound (AM'19)*. Association for Computing Machinery, New York, NY, USA, 111-115. https://doi.org/10.1145/3356590.3356608

The *EmoteControl* project can be found online at:

https://github.com/annaliesemg/EmoteControl2019

For the openly accessible collection of musical pieces described in Chapter 3 and used in Chapter 4, please see:

Micallef Grimaud, A., & Eerola, T. (2022, January 13). Emotion manipulation through music. osf.io/fy4h6

For the openly accessible collection of musical pieces used and data collected in Chapter 5, please see:

Micallef Grimaud, A., & Eerola, T. (2022, March 16). Emotional expression through musical cues: A comparison of production and evaluation approaches. Dataset and stimuli. osf.io/atxhk

Chapter 1. Current Perspectives in Music and Emotion Research

Abstract

The work in this thesis investigated how a combination of musical cues were used to change the emotional expression conveyed by the music. To do so, a new interactive interface which allows users to change the music themselves in real-time was created. This interface together with a newly composed set of Western tonal musical pieces were used to explore how individuals used the cue-emotion space to shape different emotions in the music. This first chapter provides an introduction to previous research on music and emotion, particularly looking at how emotional expressions are communicated through music and perceived by the listener in a Western musical context. An overview of previous studies on emotion perception in music is given, and the different methodologies utilised throughout the field are mentioned. Limitations in the existing literature are highlighted. The chapter concludes with an overview of the work that will be presented in the remainder of the current project and how it will tackle existing limitations in the current literature on emotional expression through cues in Western music.

1.1 Introduction

Music has been said to be a language of emotion (Cooke, 1959) due to its capacity as a vehicle for emotional communication (Juslin & Laukka, 2004), as well as its ability to evoke emotional responses from the music listeners (Saarikallio & Erkkilä, 2007). These aspects of music are put to use in different scenarios: music listening is used for affect regulation by music listeners in everyday life (Lonsdale & North, 2011; Lyvers, Cotterell, & Thorberg, 2018; Saarikallio, Maksimainen, & Randall, 2018), as therapeutic aids for people with dementia (Baird & Thompson, 2018; Sakamoto, Ando, & Tsutou, 2013) and non-verbal patients (Silverman, 2008), and as a marketing tool for brand advertising and retail (Brodsky, 2011; Lepa et al., 2020; North, Hargreaves, & McKendrick, 1999), to name a few. The numerous uses of music, which are not limited to the aforementioned ones, has motivated researchers to try gain a better understanding of the effect of music on emotion over the last century. The first theories on the relationship of music and emotion came from a philosophical perspective, in the works of Plato and Aristotle (Stamou, 2002), while music and emotion research from a music psychology perspective flourished during the last century (Eerola & Vuoskoski, 2013; Gabrielsson & Lindström, 2010; Juslin & Timmers, 2010; Schubert, 2013a; Swaminathan & Schellenberg, 2015). With emotion being investigated simultaneously in different fields, various theories and concepts on emotions and how they function have emerged, giving rise to inconsistencies in the terminology used. In the music psychology field, certain terms, such as *affect, mood, feeling,* and most significantly, *emotions* (Juslin & Sloboda, 2010; Lamont & Eerola, 2011), have been used interchangeably or used to represent different meanings across research due to there not being an agreement on specific definitions for the terminologies. To tackle this conundrum, Juslin and Sloboda presented a set of working definitions of key terms in the *Handbook of Music and Emotion* (2010), with the aim of having a consistent framework of definitions across the field.

> The term affect is used as a general term that addresses the different affective phenomena, such as moods, feelings, and emotions. The term mood refers to affective states that are longer in duration and have a lower intensity than that of an emotion. The term feeling denotes the subjective experience of moods or emotions. (Juslin & Sloboda, 2010, p. 10)

Due to emotions being multifaceted and examined in different ways, it has been proven to be rather difficult to identify one working definition of emotions which can be used for all aspects of emotion studies. However, although one distinct definition still hasn't been agreed upon, a growing consensus on the general characteristics of emotions has been established in the affective sciences (Izard, 2009), with the following definition being employed in the field:

> Emotions are relatively brief, intense and rapidly changing responses to potentially important events (subjective challenges or opportunities) in the external or internal environment, usually of a social nature, which involve a number of subcomponents (cognitive changes, subjective feelings, expressive behaviour, and action tendencies) that are more or less 'synchronized'. (Sloboda & Juslin, 2010, p. 74)

Therefore, *emotion* is seen as a scientific construct of a set of phenomena that occur together at different levels and produce different outcomes (Dennett, 1987): phenomenological (e.g., emotions as experienced), psychological (functions, appraisal, information processes, such as information recognition), and physiological (e.g., changes in hormones or brain function). In music emotion research, emotional processes that occur during a musical experience are also differentiated with regards to whether they are *perceived* or *felt* by the listener.

Perceived emotion refers to the listeners' perception and recognition of the emotional expression conveyed by the music, i.e., the listener recognises that the music is communicating a particular emotion, while *felt* or *induced* emotion refers to the listener's emotional response to the music. A relationship between emotion perception and emotion induction may exist (Evans & Schubert, 2008), and it has been suggested that the two processes may occur simultaneously, with evidence proposing that people perceive emotions more intensely than experienced (felt) ones (Schubert, 2013a). Gabrielsson (2002, pp. 131–136) describes how the relationship between felt and perceived emotions can either be positive, negative, non-systematic, or non-existent. If the relationship is positive, this suggests that the perceived emotion and the listener's emotional response to the music are the same, for example, listening to happy music induces happiness in the listener, or sad music inducing sadness (Garrido & Schubert, 2013). This effect is referred to as emotional contagion (Juslin & Västfjäll, 2008) where the listener's emotional state mirrors the emotion being expressed by the music (Davies, 2013) and might suggest how music is utilised as a tool to change or maintain one's emotional state, which is referred to as mood regulation (Knobloch & Zillmann, 2002; Saarikallio & Erkkilä, 2007). A negative relationship occurs when the experienced emotion is the opposite of the perceived emotion, such as music perceived as sad inducing a positive effect in the listener (Kawakami et al., 2013; Schubert, 2013b; Vuoskoski et al., 2012). A non-systematic relationship describes how a perceived emotion might evoke various emotions in different listeners or within an individual, or else have no effect on the listener who thus remains emotionally 'neutral'. Lastly, it must be said that the emotional expression of a musical composition does not necessarily affect the emotion experienced by the listener, and a relationship between felt and perceived emotions might not exist (Evans & Schubert, 2008; Gabrielsson, 2002). This might be due to the fact that certain emotions expressed by music are not necessarily emotions that might be experienced by listeners and vice-versa (Hunter & Schellenberg, 2010; Scherer, Zentner, & Schacht, 2001; Schubert, 2013a). Due to the various types of relationships that can take place between felt and perceived emotions, it is imperative to distinguish between these two types of emotional processes. Although perceived and felt emotions might be differentiated by a rather fine line, and the distinction between the two has not always been observed in empirical research, they are different modes of emotional responses, which may produce disparate results (Kallinen & Ravaja, 2006; Zentner, Grandjean, & Scherer, 2008) and thus, are treated as two distinct categories of emotional processes. Consequently, they are investigated using different emotion theory frameworks and methodologies.

The next section gives an overview of different current emotion models employed in the music and emotion field, particularly when investigating *perceived* emotions in music, since the work in the current project focusses on the communication of emotional expressions through Western tonal music. This is followed by an in-depth look at which emotions have been reported as capable of being expressed in music, and theories on how emotions are communicated through music are also discussed. A summary of previous studies on music and perceived emotion, with a focus on different methodologies and stimuli used within a Western context will be presented, together with limitations of the existing literature. Finally, the aims and significance of this project will be detailed, highlighting how current limitations and gaps in the literature will be addressed, for a better understanding of the communication of emotions through music within a Western framework.

1.2 Emotion Models in Music

Over the years, several different psychological emotion theories have been utilised in the affective sciences, which has made reaching an accord on what constitutes an emotion in music and emotion research difficult. Emotions have been modelled as discrete categories (Ekman, 1992; Izard, 1977), dimensions (Russell, 1980), prototypes (Shaver et al., 1987), and componential appraisal processes (Scherer, 2001), to name a few. Most of these models have been employed for both emotion perception and induction research. However, emotion perception and induction are two distinct emotional processes which may employ different underlying mechanisms, which may sometimes overlap in an emotional experience (Kallinen & Ravaja, 2006), or diverge (Taruffi & Koelsch, 2014). Furthermore, certain emotions have been reported to be more likely induced by music than perceived in music (Juslin, 2019c; Juslin & Laukka, 2004). The differences between the two emotional processes have thus served as motivation for models specific to induced emotions (e.g., GEMS) (Cowen et al., 2020; Zentner, Grandjean, & Scherer, 2008). For the purposes of the current project, the emotion model theories used to investigate emotion perception in music will be addressed. Although several psychological emotion model theories have been employed, the two dominant models used are the categorical and dimensional emotion models

1.2.1 Categorical Emotion Model (Basic Emotion)

This model, known as the categorical or discrete model of emotion, states that individuals experience distinct emotions one at a time. This model is commonly linked with the notion of 'basic emotion', which affirms that all emotions can be derived from a finite number of innate universal fundamental emotions. These typically include joy, sadness, anger, fear, and disgust, however, other emotions like tenderness, surprise, interest, and contempt can also be recognised as basic emotions (Ekman, 1992; Izard, 1977; Panksepp, 1998; Plutchik, 1994; Tomkins, 1962). Although commonly used, the categorical model has been criticised extensively. One of the main issues of this model is the ongoing debate on which emotions comprise the set of basic emotions, and which ones are relevant to the music and emotions sector. Different theorists have proposed different sets and numbers of basic emotions (Ortony & Turner, 1990), contributing to the lack of consistency across the field of music and emotions. Furthermore, with regards to music research, the basic emotion model tends to undergo modifications to eliminate certain basic emotions that rarely occur as perceived emotional expressions in music, such as disgust, to accommodate other emotions which are more prone to be represented by music, such as love (Balkwill & Thompson, 1999; Eerola & Vuoskoski, 2011; Vieillard et al., 2008). It has been suggested that certain emotions such as disgust and surprise are less recognised in music since music tends to be composed for pleasure and/or functional reasons, and thus, may not be expected to be present in music (Kallinen, 2005; Mohn, Argstatter, & Wilker, 2010). On the other hand, happiness and sadness have been reported as being easily identified in music, and this may be due to them having rather recognisable characteristics similar to ones used regularly in other emotion communication channels (e.g., vocal communication) which individuals would be exposed to (Juslin & Laukka, 2004). The need for modifications in the basic emotion model to be utilised in music and emotion research raises the question of its suitability (Cespedes-Guevara & Eerola, 2018; Scherer, 2004). Consequentially, a dispute on whether certain emotion models that intrinsically address everyday emotions can also be applied to emotions in a musical context, has emerged. This debate specifically targets the issue from the music-induced emotions perspective (Zentner, Grandjean, & Scherer, 2008). Another concern which makes the basic emotion model problematic, is that basic emotion theory builds on the notion that every discrete basic emotion has an independent neural system supporting it (Eerola & Vuoskoski, 2011). However, this theory has yet to be supported by hard evidence, as physiological and neuro-imagining studies have not yet been successful in delivering results that correspond to the basic emotions categories (Barrett & Wager, 2006).

1.2.2 Dimensional Emotion Model

The dimensional model of emotion states that all emotions can be conceptualised by mapping them out on affect dimensional continuums. Throughout the years, there have been different forms of dimensional models, emerging after Wilhelm Wundt's (1897) three-dimensional model of emotions, with pleasantness-unpleasantness, tension-relaxation, and arousal-calmness as the different dimensions. Following Wundt's model of emotion, models varied from one-dimensional *arousal* models (Duffy, 1941), to two-dimensional and three-dimensional models. Different dimensions of affect have been utilised together, such as valence and arousal (Russell, 1980), and tension and energy (Ilie & Thompson, 2006) in two-dimensional models. Three-dimensional models included valence, activation, and power (Osgood, Suci, & Tannenbaum, 1957), or valence, activity, and interest (Leman et al., 2005), or valence, energy, and tension (Eerola & Vuoskoski, 2011), to name a few. In general, two-dimensional models seem to be the ones most regularly used with respect to dimensional models. In particular, Russell's (1980) two-dimensional circumplex model of affect has been the most influential model of emotion, being utilised in more than 70% of empirical research taking a dimensional approach (Eerola & Vuoskoski, 2013).

Russell's circumplex model of affect is composed of a two-dimensional, circular structure, with *valence* (pleasantness-unpleasantness) and *arousal* (activation-deactivation) as its two continuums. Russell and Feldman Barrett (1999) refer to the two continuums as *core affect dimensions*, following the proposition that all affective states can be derived from two core bodily and neural affects; *valence* and *arousal* (Posner, Russell, & Peterson, 2005; Russell, 1980). On the other hand, Thayer (1989) suggested that the core affects arise from two distinct dimensions of activation; *energetic arousal* and *tension arousal*, in the process, giving rise to a variation of the circumplex model. Other variations of the circumplex model have also been attained, a prominent example being the Positive and Negative Affect Schedule Scale; PANAS (Watson, Clark, & Tellegen, 1988). However, the circumplex model of affect (Russell, 1980) was and still is the most extensively two-dimensional model used in music emotion research. A significant characteristic of the circumplex model of affect is that the model depicts emotional differences by minute degrees. Additionally, some semantic descriptors of emotions lying on opposite sides of the axes may be conceived as bipolar. Nevertheless, dimensional models have been criticised of not distinguishing between closely situated distinct emotions, such as anger and fear, on the valence-arousal space (Tellegen, Watson, & Clark, 1999). Another issue with dimensional models is that valence might not be a bipolar dimension, as experiencing simultaneous 'opposite', mixed emotions, such as happiness and sadness, while watching films or listening to music, might be possible. However, it is argued that this variance in emotions is not addressed by the two-dimensional model (Bigand et al., 2005; Collier, 2007; Ilie & Thompson, 2006; Sloboda & Juslin, 2010).

1.2.3 Complex and Aesthetic Emotions

A concern regarding both the categorical and dimensional emotion models is whether emotions perceived or induced by music can be explained by basic everyday emotions or dimensions. Scherer (2004) made a distinction between utilitarian emotions and aesthetic emotions; where utilitarian emotions are ones present and important in everyday life, such as joy, anger, fear, and sadness, and aesthetic emotions are ones that are experienced when engaging with forms of art, such as feeling moved and in awe. This issue primarily targets music-induced emotions, as emotional responses induced by music (and other forms of art) might not utilise the same biological mechanisms used when individuals experience emotions in everyday life situations (Eerola, 2018). Zentner, Grandjean, and Scherer (2008) addressed the issue of whether emotional experiences induced by music could be reduced to either a small number of basic emotions or dimensions by creating a new model for emotion induction in music. First, they compiled a list of comprehensive emotions terms that were relevant to music, and then determined the fundamental emotion structure by using exploratory factor analysis. The list of emotion terms was reduced to nine final emotion factors: sadness, tension, wonder, transcendence, nostalgia, peacefulness, power, joyful activation, and tenderness - known as the Geneva Emotional Music Scale (GEMS). Other models have also been developed, such as the Positive and Negative Affect Scale PANAS (Watson, Clark, & Tellegen, 1988) and the Aesthetic Emotions Scale AESTHEMOS (Schindler et al., 2017). More recently, Cowen et al. (2020) derived 13 dimensions of subjective experiences associated with music across two different cultures, with subjective experiences evoked by music including the likes of amusing, beautiful, triumphant, and sadness.

1.2.4 Constructionist Theory of Emotion

Another emotion model that attempts to ascertain how emotions are perceived and experienced draws on constructionist theories, which proposes that emotional expressions and experiences are not discrete events in themselves as suggested by the basic emotion theory, but rather, are psychologically constructed and perceived as categories by individuals. Barrett's conceptual-act model suggests that individuals experience or recognise emotions by using conceptual knowledge about emotion to categorise a state of affect, which relies on the individual's prior experience and knowledge about emotion, culture, and language, combined with information about the present event. This model complements appraisal theories with regards to the categorisation of emotions - emotions are characterised depending on the conditions evoking them, rather than by biological mechanisms (Barrett, 2006a, 2006b). Barrett states that certain emotion concepts may be common across different cultures, due to potentially being effective communication tools in a social climate. Cespedes-Guevara and Eerola's constructionist approach to emotion perception in music (2018) builds on Barret's conceptual-act model. They propose that contrary to basic emotion theory, the mechanisms utilised to perceive and feel musical emotions do not stem from a biological nature, but from psychological processes. Acoustical cues in music provide information which can be mapped to fluctuations of core affect (valence and arousal). These are then interpreted by individuals via an active psychological process that uses the individual's prior emotional and musical knowledge, together with information about the cultural and situational context of the musical event and supported by common linguistic characteristics to categorise and perceive these core affect changes as discrete emotions. While constructionist approaches theoretically address the social context and conditions surrounding the experience of perceiving and feeling musical emotions, the psychological mechanisms behind a constructionist approach might be challenging to test and assess how they operate.

Evidently, several different theories on how emotions are modelled have been proposed. Although commonalities in results have been found when using categorical and dimensional models (Eerola & Vuoskoski, 2011), the use of different emotion models has produced variances across music emotion studies. Eerola (2018) presented a synthesis of models that aims to reconcile emotion theories and explain both musically-induced emotional experiences and emotion recognition in music processes within one framework. The model comprises of three different affect concepts at three different levels. The lower two levels principally represent the perceived emotional expression process, while the top level represents the process of emotion induction. Eerola postulates that at the lowest level, the two-dimensional model of valence and arousal deals with core affects. The second level of affect, which is how emotions are consciously perceived, is attributed to the basic emotions model, where, similar to the conceptual-act model (Barrett, 2006b), individuals recognise different emotions as categories. The top level represents musically-induced emotions as complex ones, such as nostalgia or wonder. Eerola (2018) attributes the eight mechanisms of the BRECVEMA framework developed by Juslin and Västfjäll (Juslin, 2013a; Juslin et al., 2010; Juslin & Västfjäll, 2008) – brain stem reflexes, rhythmic entrainment, evaluative conditioning, contagion, visual imagery, episodic memory, musical expectancy, and aesthetic judgment – which propose how music might express or evoke emotional expressions, to four different processes: physiology, embodied, memory, and appraisal, and are utilised at the distinct affect levels depending on whether the emotional process is that of perception or induction. The synthesis of the three affect levels proposed by Eerola aims to describe how individuals perceive and also experience emotions at an intricate level (Eerola, 2018, p. 542), providing one dynamic model for all emotional processes in music.

In the current project, no strong position is taken with regards to following one specific emotion framework. The implications of this decision will be discussed in section 1.6.2. Instead, this project focusses on investigating how emotions that have been reported as possible to be expressed in music are communicated through the music and recognised by the listeners. This will be addressed in the next sections.

1.3 Emotions Expressed in Music

Each individual's musical interpretation is subjective and unique (MacDonald, Kreutz, & Mitchell, 2012). Listeners may interpret any emotion in a musical work and essentially, no emotion perceived may be deemed as incorrect. However, for the purposes of research, a more restrictive view on emotional expression is commonly employed, which states that for a musical work to be recognised as expressive of a particular emotion, there must be a minimum level of agreement among a pool of listeners. If an agreement on a specific emotion being perceived in a musical work is reached by a number of different listeners, it is indicative that an element in the musical work is producing a congruous impression across the listeners (Juslin, 2013b, 2019c). In fact, listener agreement has been long used as one way of investigating emotion expressivity in music (Campbell, 1942; Juslin, 2013b). In order to explore *which* emotions may be expressed (and perceived by listeners) in music, researchers have asked music listeners which emotions they think music can express and measured the degree of agreement in responses.

Lindström et al. (2003) carried out a questionnaire where 135 music students from Sweden, England, and Italy were first asked what they thought music can express, if anything, followed by a question on which emotions they thought music can communicate. For the first question, participants responded by selecting items from a given list of terms, based on a survey of the literature on expressivity (Gabrielsson & Juslin, 2003). They also had the option of writing down alternative responses where appropriate. The results revealed that "emotions" was the most selected item (99%) in response to what can be expressed through music, followed by "psychological tension/relaxation" (92%), "personality characteristics" (89%), "experiences that cannot be described in words" (86%), "physical aspects"

(86%), and “beauty” (85%) (Lindström et al., 2003, pp. 31–32). As music is a subjective form of art, each listener will evidently process and understand a musical work differently, depending on the individual’s thought processes. However, in this study, most participants (99%) came to an agreement that music undeniably expresses emotions. For the question regarding which emotions they thought music could express, Lindström et al. (2003) instructed participants to tick emotions from a given list of 38 randomised emotion labels, based on a review of the literature (Oatley & Jenkins, 1996; Plutchik, 1994). The participants also had the choice of writing down alternative emotion terms. The top ten emotion terms most frequently chosen were joy (98%), sadness (91%), anxiety (90%), love (89%), calm (89%), tension (89%), humour (87%), pain (86%), tenderness (86%), and anger (83%) (Lindström et al., 2003, p. 32). Juslin and Laukka (2004) ran a similar study with 141 music-listeners from Sweden, which included both musically-trained participants (N = 72) and non-musicians (N = 69). Juslin and Laukka utilised the same survey questions and material as Lindström et al. Similar to the Lindström et al. (2003) study, when asked about what can be expressed through music, all participants (100% of responses) selected the “emotions” item. This was followed by “psychological tension/relaxation” (89%), “physical aspects” (88%), “beauty” (82%), and “sound patterns” (80%) (Juslin & Laukka, 2004, p. 228). Although certain items varied in response frequency, the participants across the two studies responded similarly, with the term “emotions” being the most commonly selected item. With regards to which emotions can be expressed in music, Juslin and Laukka also got similar results to Lindström et al. Joy had the highest response frequency (99%), followed by sadness (91%), love (90%), calm (87%), anger (82%), tenderness (82%), longing (77%), solemnity (76%), anxiety (75%), and hate (74%) (Juslin & Laukka, 2004, p. 229).

Juslin (2013b) compared the results from the two aforementioned studies and an additional study by Kreutz (2000), who had 50 participants from a university student population rate their agreement/disagreement on five-point Likert scales on which emotions from a 32-item list could be represented in music. Kreutz's findings support the responses gathered in Lindström et al. and Juslin and Laukka's studies. Happiness/joy and sadness were the two most selected items across the three studies. It is worth noting that apart from the basic emotions joy and sadness, the emotions anger, love, fear, and tenderness also made the top ten rankings – which, as Juslin (2013b) points out, are all classified as basic emotions (Plutchik, 1994). Zentner, Grandjean, and Scherer (2008) carried out a set of experiments to produce a list of affect terms that were relevant to a musical context. Two hundred sixty-two undergraduate psychology students rated on four-point Likert scales how often they perceived and felt the given 146 affect terms during a musical experience. Participants also had to rate the 146 affect terms with regards to how often they experienced the same affect terms in non-musical everyday events. The list of affect terms was first reduced to 89 terms, by eliminating items such as "guilty" and "jealous" that had a low mean frequency rating of being perceived or felt by the listeners during a musical experience. Utilising exploratory factor analysis, the terms were grouped into ten extracted factors. Similar to other musical expressivity studies (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003), affect terms indicating "joy" (e.g., joyful, happy, content) were the most frequently selected as being perceived in music. Affect terms representing dysphoria (e.g., anger, fear, anxiety, and tension), tender and longing (e.g., affectionate, nostalgic, melancholic), tranquillity (e.g., calm, serene, soothed), activation (e.g., excited, active, energetic), power (proud, strong, heroic), and sadness (sorrowful, sad, depressed) were also commonly selected as emotional states perceived in music. Affect terms indicating amazement, transcendence, and sensuality were also picked, however, at a lower frequency rate.

The previous research on musical expressivity has thus provided substantial evidence that music listeners do in fact perceive emotions in music, and there is also some agreement across studies on which emotional states can be expressed by music. Certain emotions such as joy and sadness – ones that are categorised as basic emotions, are more frequently perceived in music. However, other non-basic emotions have also been rated as frequently perceived, such as calmness (Cespedes-Guevara & Eerola, 2018; Juslin & Laukka, 2004; Lindström et al., 2003; Zentner, Grandjean, & Scherer, 2008), which supports the notion that emotional states perceived in a musical context should be explored depending on whether they may be perceived in music, rather than limiting the investigation to a specific selection, determined by the use of one restrictive distinct emotion theoretical framework. Furthermore, evidence on perceived emotional states in musical works have led to questions about the processes of *how* an emotion is expressed in music. Therefore, widening the scope to simultaneously investigate a diverse range of perceived emotional expressions might help us understand better how different emotions are conveyed through music and which aspects of a musical work play a role in the communication of emotions.

1.4 How Does Music Express Emotions?

Several theories concerning what can be expressed through music have been presented by philosophers, music theorists, and others. Meyer (1956) highlighted two dichotomies in relation to the meaning of music: *absolutist* versus *referentialist*. The absolutist view focusses on intra-musical relations, where musical meaning is found within the music, without any consideration of external influences. On the other hand, the referentialist view suggests that musical meaning of a work refers to events, emotions, or other phenomena in the extra-musical world. Although these are two distinct views, the difference between absolutism and referentialism is rather ambiguous, and it has been proposed that "absolute meanings and referential meanings can and do coexist in one and the same piece" (Meyer, 1956, p. 1).

The musicologist Cooke (1959) described three different aspects of how music is an expression of emotion: 1) an *architectural* aspect, where the musical work is aesthetically pleasing to the listener due to its contrapuntal form; 2) a *pictorial* aspect in a selected number of works, which comprise imitation of nature sounds; 3) a *literary* aspect, where music expresses emotion in a similar way as speech. On the other hand, Langer (1957) argued that form, tones, and other elements of a musical work do not convey meaning like a language does. Following her theory of symbols in reason, rite, and art (1953, 1957), Langer claimed that music and language are two entities with different symbolic meanings. Language portrays discursive symbolism; where symbols have a specific meaning, whereas music is of presentational symbolism, where the meaning of elements in a musical work can only be understood in relation to the musical work's whole structure. As a result, Langer argued that music can only represent a *general* form of feelings, due to commonalities between the structure of feelings and that of a musical work, such as "patterns of motion and rest" and "of tension and release" (Langer, 1957, p. 228). Similar to Langer, Hanslick (1854) had affirmed that instrumental music is non-referential and pure; it is self-contained absolute music that does not represent anything outside of the music itself, such as feelings. However, Hanslick posits that certain features of feelings might also exist in a musical work, which motivates listeners to make associations between music and extra-musical items, such as emotional meaning. Following on the notions of philosopher Charles S. Peirce, Dowling and Harwood (1986) adapted a semiotic theory to music, whereby music could be perceived as a source of emotional expression due to it being viewed as a symbol, an index, or an iconic representation of another event (Juslin & Sloboda, 2012, p. 599). Music may be coded as a *symbol* of emotional expression, where listeners respond to the music and perceive emotions in it based on the internal content of the music itself and the relationships existing between properties within the musical structure context such as the creation of tension due to harmonic motion (Krumhansl, 1996), or variance in emotion activity due to ornamentation (Timmers & Ashley, 2007). Listeners may relate a musical piece to a specific emotion if a particular feature of the music (e.g., such as a distinct melody) may have been regularly associated with a specific event or object in the past. Consequentially, the emotional expression linked to the specific event/object being reminisced would then be associated with the musical piece, and in this situation, music may be viewed as an *index*. Music may also be categorised as an *icon* where emotions are perceived in music due to similarities between the code used in music and signals stemming from physiological behaviour that may portray emotional expression (Juslin, 2019e), such as the tone of voice (Juslin, 1997a) and movement (Davies, 1994). In particular, Juslin and Laukka (2003) postulate that musicians may communicate specific emotions in music by imitating acoustic patterns occurring in involuntary responses to an emotional reaction, such as a change in the tone of voice when expressing anger, referred to as Spencer's Law.

A common theme across different theories of how emotional expression is communicated in music seems to be the concept that elements within a musical work are responsible for the communication of emotional expression (or characteristics of emotion) in music. Juslin (2000) suggested that a modified version of Brunswik's (1956) *Lens Model*, which was originally intended for visual perception, could be used to illustrate the emotional communication process in music. Juslin (2000) indicated that there are two factors responsible for the origin of the code. The first factor builds on the idea that the brain is programmed to understand cues in vocal expression which translate into emotions, and that the shared acoustic similarities between speech prosody and music when expressing emotional expressions are how listeners are able to perceive emotions through music (Ilie & Thompson, 2006; Juslin & Laukka, 2004; Kivy, 1980; Scherer et al., 2015; Scherer, Clark-Polner, & Mortillaro, 2011). This concept reflects the idea of music being coded as an icon. Juslin explained how a performer encodes a particular emotion by combining several uncertain and partly redundant acoustic cues and utilises them to express said emotion. In turn, the listener absorbs the intended perceived emotional expression and utilises the same array of cues to decode and recognise the intended emotion. The second factor responsible for the origin of the code is the individual's social learning experiences and idiosyncratic memories gained throughout their life. As society is an important factor that influences all individuals, Juslin stipulated that social situations outside of music can influence music performers and give them a deeper understanding of how acoustic cues are used to communicate emotional expressions. An example of an early on social learning experience would be an infant recognising the different emotional tones of their mother (Juslin & Timmers, 2010). The collation of this ever growing knowledge on expressive cues gained from multiple social experiences throughout their life would consequently assist in the performers' ability to portray and communicate emotions in their music to listeners.

It is worth noting that the communication of emotion through music is not comprised only from the performer's interpretation of the music, but also from the composer's intention. Juslin and Laukka (2004) later expanded the lens model to incorporate both performer-related (known as *expressive* cues) as well as composer-related (known as *structural* cues) features (Figure 1.1[1]). Structural cues refer to features of the notated score, which include tempo, pitch, and mode, while expressive cues refer to features utilised by performers when playing a musical work, such as variations in tempo, articulation, and timbre (Gabrielsson, 2003).

[1] Permission to reuse the figure was granted through a License Agreement (Order License ID: 1202780-1) with the Copyright Clearance Center, Inc. on behalf of the rightsholder. The copyright of the figure is held by John Wiley & Sons.

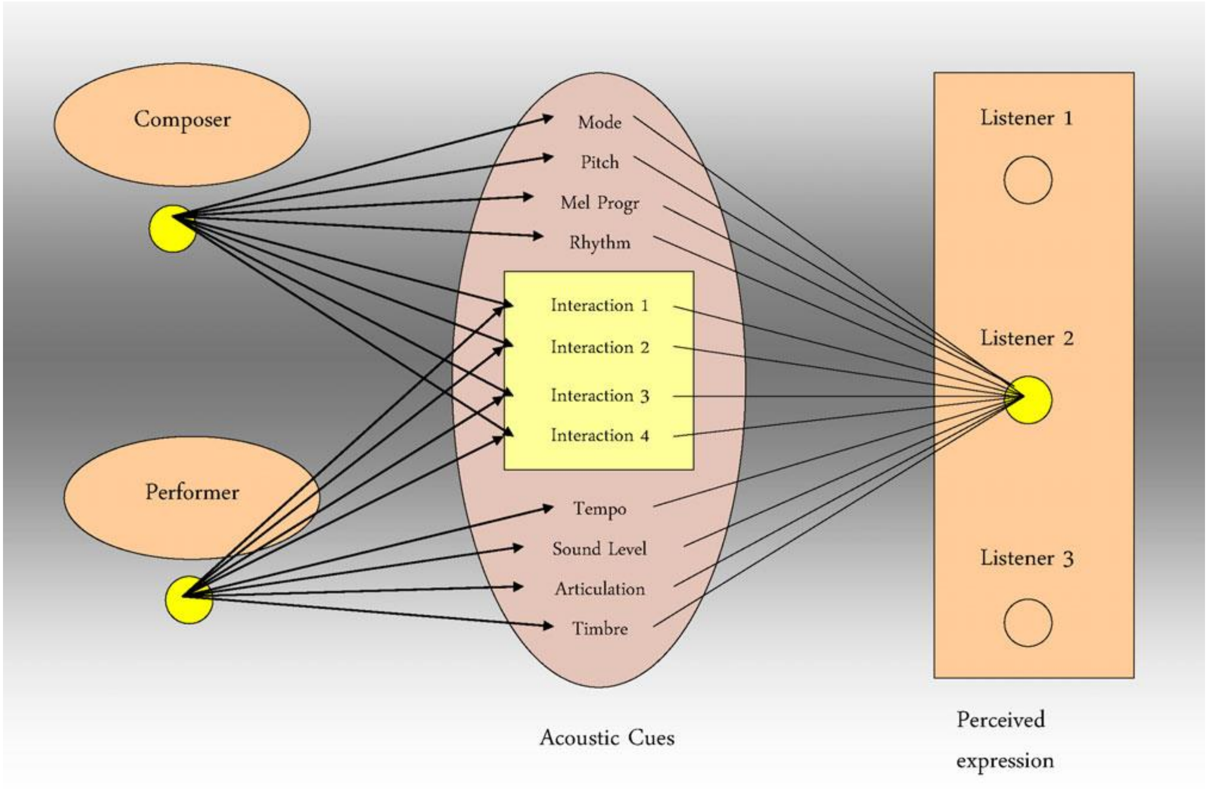


Figure 1.1. The expanded lens model of communication of emotions in music (taken from Juslin & Lindström, 2010)

The expanded lens model considers both the composer's and performer's expressive intentions, as well as the potential interactions between the two different types of cues. The cues are said to be probabilistic and partly redundant, as the individual cues on their own are not definitive in the way that they will point to one specific emotion. For example, a fast tempo may have a tendency of pointing to both joy and anger emotions (Gagnon & Peretz, 2003; Hevner, 1937; Juslin, 1997b; Thompson & Robitaille, 1992). However, cues may be used differently to portray the same emotion. Additionally, the fact that the cues are said to be partially redundant suggests that multiple cues may communicate similar information. Since the cues are said to be probabilistic in nature, the probability of successful emotion communication relies on the addition of multiple cues, which may help pinpoint to a specific emotion (Juslin, 2019c). Juslin and others (Eerola, Friberg, & Bresin, 2013; Gabrielsson & Lindström, 2001; Juslin & Laukka, 2004; Juslin & Lindström, 2010; Scherer & Oshinsky, 1977) have postulated that the interactions between the individual musical cues have slight importance in increasing the accuracy of emotion communication in music, while other researchers have emphasised the significance of the interactions between the cues (Argstatter, 2016; Hevner, 1936; Lindström, 2006). Nevertheless, the bigger the number of cues utilised together, the better chance of the communication of the intended emotional expression being reliable since it is the additivity of cues that helps shape different emotions in the music (Juslin, 2001).

Apart from the musicians' use of a combination of cues to attempt to convey a particular emotion in the music, the success of emotion communication also depends on whether the listener can correctly decode and recognise said emotion. Therefore, the cues used by the musician have to be ones which are familiar to the listener. The listener utilises their knowledge and personal experiences to understand the cues and attempt to interpret the emotional content in the music. However, the listener's ability to recognise the intended emotion also depends on the context in which the musical piece exists. Juslin postulates that a baseline or a "reference level" is needed in order to interpret how the cues are being used and what they are potentially communicating, since cue levels are dependent on the context (Juslin, 2019a, p. 124). Thus, identifying the particular style or genre of the music helps listeners make sense of the emotional information being communicated through the music, that is, if the listener has the required knowledge to understand the music in question (already mentioned by Davies, 2001). This is where knowledge gained from the social situations and culture of the listener may come into play. Some cues are said to be universal, which suggests that they are likely to be used similarly across cultures, due to being based on core psychophysical elements, such as tempo (Athanasopoulos et al., 2021; Balkwill, Thompson, & Matsunaga, 2004; Balkwill & Thompson, 1999; Fritz et al., 2009). On the other hand, the way other cues are used may be culture-specific (e.g., mode, melody), which suggests that in order for the cue information to be correctly interpreted, the listener may need prior knowledge on and exposure of the particular culture and tradition (Balkwill & Thompson, 1999). Since culture-specific cues convey emotional meaning due to their association with a particular event or object existing in a specific tradition or culture, they are viewed as an index (Juslin & Sloboda, 2001). Albeit having different origins (i.e., whether emotional meaning is conveyed through music due cues being seen as an icon, index, or symbol), universal and culture-specific cues may exist concurrently in a musical work, and it is suggested that their layering may help in the emotion communication process (Juslin, 2013b).

To determine how different musical cues contribute to encoding emotional expressions in music, which are subsequently decoded and identified by listeners, numerous studies were carried out, investigating multiple musical cues and employing different approaches.

1.5 Overview of Musical Cues and Perceived Emotional Expression Research

Empirical research on emotion perception in music has been carried out since the late nineteenth century, with the earliest studies tracing back to the 1890s by Gilman (1891, 1892) and Downey (1897) who investigated listener agreement on perceived emotions in music. Participants were asked to listen to selected classical musical works performed live on piano or piano and violin, and report which emotional expression they perceived in the music by freely writing down descriptive terms. Huber (1923) ran a similar study where musically-trained participants listened to short tone sequences and provided descriptions of the sequences using free choice of terms. Allowing participants free choice of terms resulted in too many descriptive terms being used, some of which described minute variations of emotions, making classification of terms difficult. To address this issue and limit ambiguities and inconsistencies in results, later researchers put together a list of descriptive emotion terms for participants to choose from (Gundlach, 1935; Hevner, 1935). Other studies asked participants to rate on Likert scales how well selected descriptive emotion terms were relevant to the music (Dalla Bella et al., 2001; Gabrielsson, 1973; Juslin, 2000; Laukka & Gabrielsson, 2000; Quinto, Thompson, & Taylor, 2014) or rate stimuli on dimensional models of affect (Ilie & Thompson, 2006).

Research on the effect of musical cues on the perceived emotional expression surged in the 1930s, with Hevner's work (1935, 1936, 1937) still being one of the most note-worthy. Hevner introduced the use of a systematic manipulation design in music and emotion research, where she created different versions of the same musical pieces that varied in compositional cues to study the effect each cue had on the perceived emotion. Hevner investigated two levels of mode (major and minor) by having 205 students listen to ten pairs of short musical compositions performed live on a piano. Each pair of musical pieces differed in mode – one version was played in major mode and the second version was played in minor mode (Hevner, 1935). Hevner arranged a large number of emotion terms in eight different clusters in a circular fashion, referred to as an 'adjective circle', foreshadowing Russell's (1980) circumplex model of affect. Participants were then instructed to choose as many of these terms as appropriate for each musical stimulus. Getting participants to listen to the same musical piece which varied in one feature (in this case, the mode) allowed Hevner to attribute any differences in the participants' responses between the two variations of the same stimulus to the specific type of mode (major or minor) utilised. Hevner investigated other compositional cues such as tempo, pitch level, complexity of harmony, rhythmic quality, and melodic line direction utilising similar systematic manipulation designs (1936, 1937). From there onwards, more than a hundred studies were carried out on musical cues and emotion research, with different methodologies being implemented over the years. Some of the favoured approaches are discussed in the next section.

1.5.1 Methodological Approaches

1.5.1.1 Systematic Manipulation

Several researchers followed Hevner's lead in testing different structural cues and cue levels in relation to various emotion terms using a systematic manipulation method. Rigg (1939, 1940b, 1940a) utilised five four-bar phrases played on the piano, which were systematically varied in pitch level, tonality, and tempo, and asked participants to select a term from the available list that best described the variations. Results showed that in general, a fast tempo and a high pitch level increased the perception of happiness, while minor mode and a slow tempo increased the perception of sadness in the musical phrases. Although earlier experiments (Hevner, 1935, 1936, 1937; Rigg, 1940b) utilising this methodology focussed on *structural* cues (properties pertaining to the notation of a musical work, e.g., mode, tempo), a systematic manipulation approach was later also adopted for *expressive* cues (cues utilised by the performer, e.g., articulation, timbre). Researchers utilised this method to investigate individual cues such as mode (Kastner & Crowder, 1990) and timbre (Eerola, Ferrer, & Alluri, 2012), or different combinations of cues, such as tempo and dynamics (Kamenetsky, Hill, & Trehub, 1997), tempo and mode (Gagnon & Peretz, 2003; Morreale et al., 2013;

Peretz, Gagnon, & Bouchard, 1998), rhythm, melodic contour, and melodic direction (Lindström, 1997), pitch level and rhythm (Schellenberg, Krysciak, & Campbell, 2000), and pitch, brightness, and loudness (Cousineau et al., 2014), to name a few. Although there have been numerous studies that have used a systematic approach to look at different cues, only a couple of studies tested a bigger number of cues, with six to eight cues being investigated simultaneously (Eerola, Friberg, & Bresin, 2013; Juslin, 1997b; Juslin & Lindström, 2010; Scherer & Oshinsky, 1977).

Scherer and Oshinsky (1977) investigated two levels of each of seven different musical cues (a 2^7 factorial design resulting in 128 different stimuli) consisting of both acoustical and structural parameters: amplitude variation (small/large), pitch variation (small/large), pitch level (low/high), pitch contour (up/down), tempo (slow/fast), envelope (low attack-decay ratio/equal attack-decay ratio), and filtration cut-off level (intermediate/high). Synthesised sawtooth tone sequences were manipulated on a MOOG synthesizer for all 128 cue combinations and rated by participants on different polar scales. Juslin (1997b) also attempted a sizeable factorial design (108 total cue combinations) to investigate two or three different levels of five musical cues: tempo (slow/medium/fast), sound level (low, medium, high), frequency spectrum (soft/bright/sharp), articulation (*legato/staccato*), and attack (slow/fast). Juslin (1997b) utilised a synthesised version of the *Nobody Knows* melody by Stephen Foster as the initial stimulus. The different cue combinations were implemented to the melody which resulted in 108 different variations. Due to the substantially large number of musical variations, each participant rated a subset of the stimuli on six different emotion scales (happy, sad, fear, tender, and expressive). A more recent study by Eerola, Friberg, and Bresin (2013) utilised a systematic manipulation design to explore how 2-6 levels of six different cues impacted the emotion being communicated in four different musical excerpts. However, as the full factorial design produced 14,400 different musical excerpts, the authors utilised an optimal design and reduced the amount of cue combinations tested whilst retaining a balance between different cue levels and combinations explored. The resulting 200 musical stimulus set was rated by participants on four different emotion scales (happy, sad, peaceful, and scary).

An advantage of a systematic manipulation methodology is that it allows the researcher to have complete experimental control on the cues and cue levels being investigated. However, this approach is restrictive with regards to how many cues and/or cue levels can be investigated simultaneously. As each cue combination results in a new musical variation that has to be listened to and rated by participants, a factorial design with a large number of iterations becomes unfeasible, due to potential lack of engagement and fatigue from the participants (Lee & Müllensiefen, 2020).

1.5.1.2 Correlation Studies

A different approach to investigating the relationship between the emotional content and properties of a musical work is a correlation study. Gundlach (1935) carried out a study where participants first listened to single phrases taken from 40 musical works by Chopin, Haydn, and Tchaikovsky, to name a few, and selected descriptive terms from a given list (or wrote their own term if none from the given list were satisfying) which best described the music. Musical excerpts were grouped depending on the descriptive terms used by participants to characterise them. Excerpts in the same categories were then compared with regards to different musical attributes, such as tempo, intervals range, loudness, mean pitch value, and melodic range. This methodology allowed for correlations to be made between underlying features that were common across musical pieces and the terms used to describe the pieces. Wedin (1972) ran a similar study in which 100 students first described 40 musical excerpts with terms available from a list of 125 adjectives. Fifteen musicians then judged the excerpts on different features such as intensity, pitch, rhythm, and harmony, and correlations were made between the properties of the music and the different adjectives used to describe the excerpts. Other studies have used a similar experiment design to further analyse different features in the musical score with regards to emotional content (Battcock & Schutz, 2019; Imberty, 1979; Kleinen, 1968; Krumhansl, 1996; K. Watson, 1942; Wedin, 1972). Schubert (2004) had participants listen to one of four pieces and continuously rate the perceived emotion on a two-dimensional space presented on a computer. Some researchers have asked musicians to specifically create music representing different emotions to analyse similarities between pieces across composers (Quinto, Thompson, & Taylor, 2014; Ramos & Mello, 2021; Thompson & Robitaille, 1992). Thompson and Robitaille (1992) asked musicians to compose short melodies to communicate six different emotions: joy, sorrow, excitement, dullness, anger, and peace, and the structural cues of the melodies in relation to the intended emotional expression. Fourteen participants then rated on six individual emotion scales (joy, sorrow, excitement, dullness, anger, and peace) how much of each emotion they thought the musical excerpts were conveying.

A correlation approach has also been used to measure the expressive (performer) cues of a musical work (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Juslin, 1997a; Laukka & Gabrielsson, 2000). These types of correlation studies asked musicians to perform a set piece multiple times; each time trying to convey a different emotional expression. Gabrielsson and Juslin (1996) asked nine musicians to perform short melodies with the aim of communicating happiness, sadness, anger, fear, tenderness, and solemnity through the music. The musicians also had to perform the melodies under a 'no expression' condition. The different performances were subjected to a listening study where participants rated each performance on its expression. The performance cues used by the musicians were then analysed in relation to the intended emotion.

Some researchers investigated both structural and expressive cues in the same study (Quinto, Thompson, & Taylor, 2014; Ramos & Mello, 2021). Quinto, Thompson, and Taylor (2014) analysed the role of structural and expressive cues in conveying different emotions by carrying out a study with three conditions. First, they asked musicians to compose short, monophonic melodies intending to communicate tenderness, anger, fear, happiness, sadness, and a neutral expression, to analyse how structural cues were employed. They also explored how musicians used expressive cues to convey the six emotions through ambiguous melodies composed by one of the authors. Their final condition combined the two types of cues by having the musicians perform their own melodies to convey the different intended emotions. Forty-two listeners listened to a subset of the musical stimuli and rated which emotions they thought each musical trial was conveying. A selection of expressive and structural cues was then analysed in relation to the different emotions communicated. Results revealed that the efficacy of a conveyed emotion varied depending on the condition. For example, happiness was best communicated in the condition with both structural and expressive cues, while fear was best expressed through the structural cues. These findings indicated that emotional expressions are communicated differently through the two types of cues. Furthermore, the combination of structural and expressive cues seemed to result in more accurate emotion recognition.

Studies that use ‘real music’ performances as their stimuli allow researchers to investigate how emotional expression is communicated in an organic setting, since the music retains its ecological validity. However, although using this approach can lead to associations between cues and specific emotions, the cues are analysed as the overall combination, rather than individuals. Thus, it is difficult to recognise the effect of the individual cues on the emotion communicated. Furthermore, having musicians play the stimuli live with different conditions, such as ‘no expression’, might not be the most reliable approach. It is highly unrealistic that a musician will be able to perform a musical work without adding any ‘humanising’ qualities to it, such as slurring, or articulation, which may contribute to the communication of emotion, and thus, might skew results (Ramos & Mello, 2021; Shoda & Adachi, 2012).

1.5.1.3 Production Approach

A newer type of approach that has been used to examine the relationship between musical cues and emotion perception studies is a production approach. This concept uses an analysis-by-synthesis methodology (Friberg, Bresin, & Sundberg, 2014; Gabrielsson, 1985) where the different emotional expressions are reconstructed in musical pieces utilising a computer simulation to vary selected musical cues. The novelty of this approach is that rather than having a two-part experiment which entails having an encoding task (composers create new music or musicians perform different musical works) and then a decoding task (listeners judge the stimuli or performances in relation to emotional expressions), participants are presented with musical stimuli and asked to alter them themselves via a number of available musical cues in real-time, utilising an interactive paradigm. Bresin and Friberg (2011) devised an apparatus that let users alter seven cues (tempo, sound level, articulation, phrasing, register, timbre, and attack speed) of the melodic part of four short musical excerpts. Participants carried out the changes using a physical slider for each individual cue. This approach allowed users to explore a substantially vast cue space in real-time, which would not be possible to carry out if all cue combinations had to be exported as individual musical stimuli. Although Bresin and Friberg recruited musicians to take part in their study, this interactive paradigm may also be used to explore how non-musicians would use the cues to express different emotional expressions, as no prior musical knowledge is required to utilise the interface. Saarikallio, Vuoskoski, and Luck (2014) used a similar interface called *MusicBox* to investigate how 61 adolescents (32 participants had received music lessons while the rest had not) would use tempo, loudness, pitch, articulation, and timbre cues to convey happiness, sadness, and anger through three instrumental melodies of children's songs. Results showed that in general, the adolescent participants used the cues in a similar manner to adults when portraying the different emotions (Juslin & Laukka, 2003, 2004; Juslin & Timmers, 2010). Saarikallio et al. (2019) later ran a study where 3- and 5-year-olds used a modified version of the *MusicBox* interface to control tempo, loudness, and pitch cues of instrumental melodies taken from three children's songs, to convey happiness, sadness, and anger through the music. The findings suggested that children as young as 3-year-olds successfully used the three cues to differentiate between the different emotional expressions. Kragness et al. (2021) also used an interactive paradigm to investigate how 3-, 5-, and 7-year-old children used tempo, loudness, and articulation to convey joy, sadness, anger, and peace in chord sequences taken from Bach chorales. Participants controlled the three cues via the Middle C key on a MIDI piano keyboard. Results showed that 5- and 7-year-olds used a relatively faster tempo and higher loudness level to distinguish high-arousal emotions from low-arousal ones. Furthermore, the 7-year-olds also used articulation to distinguish between emotions, while the 3- and 5-year-old children did not. Kragness and Trainor (2019) had previously carried out a similar study utilising the same self-pacing paradigm to explore how adults with no expert musical training altered the Bach chorales chords to express different emotions. Sievers et al. (2013) employed an interactive interface to compare how anger, happiness, peacefulness, sadness, and fear emotions are expressed in music and movement, using two participant pools from the US and Cambodia. Participants were presented with a musical excerpt and an animation of a ball. They could control the pitch direction/ball movement, beats per minute/ball bouncing rate, interonset interval/jitter, consonance/smoothness of ball texture, and pitch interval size/ball bounce height to portray the different emotional expressions. Findings suggested that similar dynamic contours were used in both music and movement across the two different participant pools.

The use of a production approach has notable advantages. Firstly, a bigger number of cue combinations can be explored in a relatively shorter time. Cues can have substantially wide ranges (e.g., a tempo range of 100 beats per minute, a dynamics range of 20 decibels) which produce a large number of possible cue combinations, as the different cue combinations will not need to be rendered as individual musical excerpts. Secondly, these types of experiments can be carried out by individuals who do not have any prior musical knowledge, and also seem to be successful with a younger participant pool (Kragness et al., 2021; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014). The current literature on these types of production studies is rather limited in comparison to the other methodologies mentioned. Existing studies that have adopted a production approach have focussed on utilising either simple monophonic melodies or a sequence of static chords as stimuli. To further advance studies using this methodology, it would be interesting to give participants control of polyphonic musical works, rather than just the melodic line. Furthermore, most studies employing this approach have investigated a small number of emotions (3-5 emotions, mostly basic ones) and musical cues (3-4 as an average), which leaves an extensive cue-emotion space yet to be explored.

1.5.2 Musical Cues

Across the different methodologies, various structural and expressive musical cues have been studied in relation to emotional expression in music (for a review see: Gabrielsson & Lindström, 2010). Musical cues investigated include tempo (Behne, 1972; Dalla Bella et al., 2001; Gabrielsson & Lindström, 1995; Gagnon & Peretz, 2003; Gundlach, 1935; Hevner, 1937; Ilie & Thompson, 2006; Imberty, 1979; Peretz, Gagnon, & Bouchard, 1998; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Thompson & Robitaille, 1992; K. Watson, 1942), mode (Battcock & Schutz, 2019; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Hevner, 1935; Kleinen, 1968; Lahdelma, Athanasopoulos, & Eerola, 2021; Lindström, 2006; Morreale et al., 2013; Quinto, Thompson, & Taylor, 2014; Rigg, 1939; Scherer & Oshinsky, 1977), pitch level (Gundlach, 1935; Rigg, 1939; Schellenberg, Krysciak, & Campbell, 2000; K. Watson, 1942), rhythm (Imberty, 1979; Juslin & Lindström, 2010; Lindström, 2006; Wedin, 1972), timing (Gabrielsson & Juslin, 1996; Gabrielsson & Lindström, 1995; Laukka & Gabrielsson, 2000), sound level (dynamics) (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Gundlach, 1935; Juslin, 1997a, 1997b; Kamenetsky, Hill, & Trehub, 1997; Kragness & Trainor, 2019; Rigg, 1939; Scherer & Oshinsky, 1977; K. Watson, 1942), melody (Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014), harmony (Athanasopoulos et al., 2021; Behne, 1972; Hevner, 1936; Imberty, 1979; Lindström, 2006), consonance (Lahdelma & Eerola, 2016a, 2020), articulation (Eerola, Friberg, & Bresin, 2013; Juslin, 1997a, 1997b, 2000), intervals (Costa, Fine, & Ricci Bitti, 2004; Costa & Nese, 2020), tonality (Scherer & Oshinsky, 1977; Thompson & Robitaille, 1992; Wedin, 1972), and timbre (Behrens & Green, 1993; Eerola, Ferrer, & Alluri, 2012; Hailstone et al., 2009; Schutz et al., 2008), to name a few.

Research on different cues varies in frequency. Certain cues, such as tempo, dynamics, and mode have been investigated more often than, for example, rhythm and articulation (Gabrielsson & Juslin, 2003). This is potentially due to particular cues being deemed as having more influence on the emotional expression conveyed in the music. Hevner (1937) proposed that tempo, mode, pitch level, rhythm, and harmony had significant weight on how emotions were perceived in music. Juslin (2001) also suggested that tempo, together with articulation, dynamics, and timbre were the cues that mostly affected the perceived emotion (Juslin & Lindström, 2010). From a review of the literature, Gabrielsson and Lindström (2010) observed that tempo and dynamics seemed to have the biggest role in emotion portrayal to the other cues, while findings from a systematic manipulation design by Eerola, Friberg, and Bresin (2013) ranked mode (followed by tempo, pitch, and dynamics) as the most important in emotional expression amongst six cues in total. In summary, it seems that across different studies, tempo, mode, and dynamics have been reported as being three significantly salient cues in conveying different emotional expressions in music (Dalla Bella et al., 2001; Kamenetsky, Hill, & Trehub, 1997; Scherer & Oshinsky, 1977). However, it is good to note that the contribution of cues may be relative to the other cues investigated and dependent on the particular musical pieces used in the study (Hevner, 1937). Furthermore, most musical cues and emotion studies have been carried out in a Western context, with only the minority of studies having explored the effect of musical cues on the perceived emotion in non-Western music and/or non-Western cultures (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). Therefore, the usage and significance of cues in shaping different emotions in music may vary across different cultures. In addition, whether cues are culture-specific or universal may also affect the contributory weight of the cue in question. For example, Gagnon and Peretz (2003) confirmed that tempo and mode contributed to conveying happiness and sadness in music. However, tempo had a bigger effect on the perceived emotion than mode. Gagnon and Peretz suggested that this may be due to tempo having a similar function in other communication channels such as speech, whilst mode is a music-specific cue, and particular to Western tonal music (Gagnon & Peretz, 2003, p. 37).

When considered as individuals outside of a musical context, cues can contribute to multiple, different emotional expressions. For example, a fast tempo is usually associated with expressions that contain high arousal, such as happiness, anger, surprise, and fear (Hevner, 1937; Juslin & Laukka, 2004; Kragness & Trainor, 2019; Scherer & Oshinsky, 1977). A slow tempo may be associated with low arousal emotions, such as sadness, tenderness, peacefulness, calmness, and longing (Gundlach, 1935; Juslin, 1997b; Quinto, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 2014; Scherer & Oshinsky, 1977). A major mode tends to be correlated with happiness, serenity, and gracefulness in Western music (Gagnon & Peretz, 2003; Hevner, 1936), while a minor mode is affiliated with sadness, agitation, and anger (Eerola, Friberg, & Bresin, 2013; Hevner, 1936; Peretz, Gagnon, & Bouchard, 1998; Rigg, 1937; Scherer & Oshinsky, 1977). In the dimensional models, the difference between major and minor mode is associated with positive or negative valence (Costa, Fine, & Ricci Bitti, 2004). A high pitch level is usually seen as corresponding to happiness, but also surprise, anger, and fear (Hevner, 1937; Juslin, Friberg, & Bresin, 2002; Juslin & Lindström, 2010; Rigg, 1940b; Scherer & Oshinsky, 1977). A low pitch level is typically associated with sadness (Hevner, 1937; Watson, 1942), tranquillity (Gundlach, 1935), solemnity, and agitation (Rigg, 1940b). Loud dynamics are usually indicative of excitement, tension, power, anger, and joy (Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014), while soft dynamics are related to sadness, fear, tenderness, and peacefulness (Eerola, Friberg, & Bresin, 2013; Juslin, 1997b; K. Watson, 1942).

Each musical cue on its own may refer to several different emotional expressions. Furthermore, different cue levels may contribute to the same emotion. For example, both high and low pitch levels have been attributed to fear in music with a similar style (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010). The fact that the same cue can contribute to multiple emotions implies that the perceived emotional expression in a musical piece is not determined by one single cue, but rather, by a function of multiple cues, working together either in an additive and/or interactive way (Gabrielsson, 2008; Gabrielsson & Juslin, 2003). Therefore, the bigger the number of cues used simultaneously, the more chance there is for the intended perceived emotional expression to be accurately conveyed and identified (Juslin, 2001). A number of studies have thus explored a combination of cues, such as tempo and mode (Dalla Bella et al., 2001; Gagnon et al., 2012; Gagnon & Peretz, 2003; Peretz, Gagnon, & Bouchard, 1998) or tempo and dynamics (Kamenetsky, Hill, & Trehub, 1997). Only the minority of studies have investigated more ambitious cue combinations such as pitch level, mode, melody, rhythm, tempo, sound level, articulation, and timbre (Juslin & Lindström, 2010). The fact that only a handful of studies have attempted to investigate numerous cue combinations simultaneously might be attributed to constraints of current methodological approaches discussed in section 1.5.1.

1.5.3 Musical Cues and Emotional Expression

Musical cues have been investigated in relation to different emotional expressions or affect dimensions. Common emotions investigated usually tend to be selected from the group known as discrete or basic emotions (Ekman, 1992; Plutchik, 1994), which include happiness, anger, fear, sadness, and tenderness (Dalla Bella et al., 2001; Eerola & Vuoskoski, 2013; Gosselin et al., 2011; Juslin, 1997a; Juslin & Lindström, 2010; Peretz, Gagnon, & Bouchard, 1998; Vieillard et al., 2008). Happiness and sadness seem to be the two emotions most frequently studied in conjunction to emotional expression. In a review on music stimuli and emotion studies, Warrenburg (2020a) reported that sadness was the mostly frequently studied emotion, followed by happiness. This might be due to happiness and sadness tending to be the two emotions most easily recognised in childhood (Dalla Bella et al., 2001; Kastner & Crowder, 1990). Anger and fear have also been consistently studied in relation to musical cues (Kragness & Trainor, 2019; Warrenburg, 2020a), potentially due to also tending to have a good recognition accuracy rate across cultures (Balkwill, Thompson, & Matsunaga, 2004; Balkwill & Thompson, 1999; Fritz et al., 2009). Furthermore, basic emotion theory postulates that discrete emotions are easier to identify, due to being experienced in everyday life (Akkermans et al., 2019; Gabrielsson & Juslin, 1996). Other emotions that are not deemed as basic emotions, such as humour, solemnity, peacefulness, longing, and awe have also been explored, but to a significantly lesser extent (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Laukka et al., 2013; Vieillard et al., 2008), which leaves a need for further research (Juslin & Laukka, 2004; Saarikallio, Vuoskoski, & Luck, 2014). The other typical approach in music emotion studies is mapping musical cues on dimensions (G. L. Collier, 2007; Costa, Fine, & Ricci Bitti, 2004; Gagnon & Peretz, 2000; Ilie & Thompson, 2006; Quinto, Thompson, & Taylor, 2014; Schubert, 2004), with valence and arousal tending to be the two most commonly used dimensions (Eerola & Vuoskoski, 2013). Although findings show that different points on the dimensions are related to particular cue levels, such as, high arousal is related to fast tempo and low arousal is related to slow tempo (G. L. Collier, 2007; Wedin, 1972), the valence-arousal model does not allow for proper differentiation between emotions with similar valence-arousal properties (e.g., fear and anger are both mapped on low valence and high arousal dimensions), and thus, might produce ambiguous results (Bigand et al., 2005; G. L. Collier, 2007; Eerola & Vuoskoski, 2013; Leman et al., 2005).

1.5.4 Music Stimuli

Various musical pieces have been utilised as stimuli across music and emotion research, with Western tonal music being the dominant type of music used (Warrenburg, 2020a). The earliest experiments entailed having the selected musical excerpt, typically from the Classical music repertoire, presented to the participants either played live by one or more musicians or as recordings of the performances (Behne, 1972; Downey, 1897; Gilman, 1891, 1892; Hevner, 1935, 1936, 1937; Rigg, 1937, 1940b). Later, technological advances introduced the use of synthesised tone sequences as musical stimuli, rather than live performances or recordings of live music (Gagnon & Peretz, 2003; Juslin, 1997a; Kamenetsky, Hill, & Trehub, 1997; Lindström, 2006; Schellenberg, Krysciak, & Campbell, 2000; Scherer & Oshinsky, 1977). Stimuli used also varied from simple melodic monophonic progressions (Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014; Thompson & Robitaille, 1992), to polyphonic synthesised musical examples (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Peretz, Gagnon, & Bouchard, 1998). Musical stimuli differed in style, with excerpts chosen from classical works such as that of Beethoven, Bach, and Handel (Battcock & Schutz, 2019; Hevner, 1935, 1936; Hunter, Schellenberg, & Schimmack, 2010; Ilie & Thompson, 2006; Kallinen, 2005; Kawakami et al., 2013) or variations of well-known tunes, such as *Frère Jacques* and *Nobody Knows* (Gabrielsson & Juslin, 1996; Juslin, 1997a, 1997b, 2000; Lindström, 2006). Extensive reviews of the literature observed that classical music was the preferred choice of stimuli amongst researchers (Eerola & Vuoskoski, 2013; Juslin & Laukka, 2003), most specifically between the 1920s and late 2000s (Warrenburg, 2020a). Although popular music seems to be more frequently opted for since the 2010s, the most favoured composers were still from the Classical era, with Mozart, Beethoven, and Bach being the top three composers for chosen musical stimuli in music emotion research (Warrenburg, 2020a). Either way, it seems that the majority of studies utilise pre-existing, commercial recordings as stimuli (Eerola & Vuoskoski, 2013).[2]

Although using commercial music enhances the ecological validity of the music stimuli, as it represents 'real music' that exists in a real-life context (Battcock & Schutz, 2019; Schubert, 2004; W. Zhang et al., 2019), it is difficult to dissect the musical work and manipulate cue levels to analyse the effect of the individual cues without tampering with the work's ecological validity. To ensure better control of the investigated cues, the minority of researchers specifically composed musical excerpts to be used as stimuli (Hailstone et al., 2009; Paquette, Peretz, & Belin,

[2] From Eerola and Vuoskoski's 2013 review, findings showed that 76% of stimuli used in 170 studies were commercial recordings.

2013; Quinto, Thompson, & Taylor, 2014; Rigg, 1937, 1939; Thompson & Robitaille, 1992; Vieillard et al., 2008). Hailstone et al. (2009) created 80 new single-line melodies in total, with the aim of conveying happiness, sadness, anger, and fear (20 melodies were composed for each emotional expression). These melodies were tested in a pilot study with regards to whether they conveyed their intended emotion. Subsequently, the best conveyors of the intended emotions were carried forward as stimuli to be used in an investigation of timbre, where participants listened to the four variations of the stimuli – played on a piano, violin, trumpet, and synthesizer. Paquette, Peretz, and Belin (2013) asked 10 violinists and 10 clarinettists to improvise ‘musical bursts’ to communicate fear, happiness, sadness, and neutrality to the listener, with the best candidates being accurately identified with an emotion recognition score of 80.4% (2013, p. 4), albeit being new, unfamiliar stimuli. Unlike most research-specific stimuli, Vieillard et al. (2008) composed *polyphonic* musical excerpts exported with a piano timbre to portray happiness, sadness, threat (scary), and peaceful emotions in music. This stimulus set was also used in other studies to investigate whether anteromedial temporal excisions and amygdala resections impaired recognition of emotions in music (Gosselin et al., 2005, 2011). Although the stimulus set was not altered in the subsequent studies, the creators of the set (Vieillard et al., 2008) declared that the available stimuli could also withstand cue alterations, for potential research on specific musical cues, such as tempo. Specifically-composed musical excerpts which allow for systematic manipulation of cues whilst also retaining a form of ecological validity would permit a better investigation of cues in a ‘real’ musical context (Eerola & Vuoskoski, 2013), and are significantly needed.

1.6 Limitations of Current Research

From this overview of current research, one can discern that music emotion research has come a long way. However, the different emotion frameworks, methodological approaches, and material used give rise to multiple limitations in the research area, which hinder future progress in the field. This section will highlight limitations identified across the different sections of this chapter, which have already been touched on in the relevant sections.

1.6.1 Methodological Constraints

The first set of limitations regard the *methodological approaches* (systematic manipulation, correlation studies, and production approach) mostly employed in musical cues and emotion research. Although a systematic manipulation approach provides the most experimental control over the different cues and cue levels in order to identify how the individual cues and their combinations affect the perceived emotional expression, there is a *finite number of iterations* which can be tested simultaneously. Huge factorial designs would realistically be time-consuming and also risk participant fatigue and lack of engagement (Juslin, 1997b; Lee & Müllensiefen, 2020). On the other hand, a correlation study would allow for a number of cues of 'real' musical compositions and/or performances to be investigated. However, this methodology *does not offer experimental control* over the cues and does not allow for the cues to be explored as individuals, thus making it difficult to determine how the different cues are contributing to the perceived emotional expression. The most promising methodology for an extensive exploration of musical cues is the production approach, as this method bypasses the need for fixed iterations as per a systematic manipulation design, whilst allowing for the individual cues to be analysed, unlike correlation studies featuring performance and score analyses. Although a production approach allows for substantial ranges of cue levels and cue combinations to be tested, only the minority of studies have employed this methodology (Bresin & Friberg, 2011; Kragness et al., 2021; Kragness & Trainor, 2019; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Sievers et al., 2013). Furthermore, current research utilising this approach have mostly focussed on altering a *small number of musical cues* (Kragness et al., 2021; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014) of monophonic melodies or a sequence of chords, albeit having potential to provide control of more cues simultaneously (Bresin & Friberg, 2011; Sievers et al., 2013). A need for an interactive paradigm that withstands bigger cue combinations and involves a mix of *both* expressive and structural cues is evident. A final point which seems pertinent to a better understanding of current knowledge in the field is whether the use of different methodological approaches produces similar results, or whether findings must be interpreted with respect to the methodology used.

1.6.2 Emotions Investigated

Although music expressivity studies (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003) have reported that music listeners deem numerous emotions to be perceived in music (e.g., joy, sadness, love, surprise, pride, hope, humour, loneliness), the majority of studies have mostly investigated discrete emotions that follow basic emotion theory (Ekman, 1992; Plutchik, 1994), with sadness and joy being the two mostly studied emotions (Kragness & Trainor, 2019; Warrenburg, 2020a), or followed a dimensional model of affect (Russell, 1980; Russell & Barrett, 1999) which might not be effective in distinguishing between emotions similarly mapped on the valence-arousal dimensions. The use of one specific emotion framework restricts researchers to *a small selection of emotions* and thus, limits research on musical cues and emotion studies to those particular emotions present in the chosen emotion framework. Therefore, when researchers opt for a specific emotion framework, they are constrained to investigate only the emotions pertaining to the framework, which results in the same emotional expressions being tested repeatedly across different studies, which might not necessarily provide us with new information. Other emotional expressions that have been reported as being communicated in music (see: Juslin & Laukka, 2004), such as peacefulness, longing, and humour (Akkermans et al., 2019; Laukka et al., 2013) should also be investigated. Laukka et al. (2013) conducted a cross-cultural study where 11 different emotions were investigated, which included basic emotions (sadness, happiness, anger, and fear), as well as other emotions that have been reported as being accurately identified in music (affection, solemnity, and peacefulness). Most importantly, the authors also included three emotional expressions that had not been previously explored (humour, longing, and spirituality). Findings indicated that most tested emotions were identified above chance by participants. Although the basic emotions had a higher identification accuracy rate, the non-basic emotions humour and peacefulness were also successfully recognised across cultures. The rationale behind Laukka et al.'s study did not focus on keeping to one specific framework. Although this meant that they could not use a specific emotion framework to test their theories, not restricting their study to one particular framework resulted in new knowledge on a wide array of musical emotions. Future work should focus on unveiling how other musical emotions, irrespective of which theoretical framework they belong to, are encoded in a musical work, and communicated to listeners.

1.6.3 Choice of Musical Stimuli

An ongoing issue pertaining to musical excerpts used as stimuli is the use of pre-existing commercial music (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). Although when opting to use ‘real music’ as stimuli, the ecological validity of the stimuli is held, explicit control of the different musical cues is limited, and hence, recognising the effects of the different cues is rather difficult. On the other hand, when systematic manipulation of musical cues is carried out, the stimuli tend to lose their ‘real music’ properties and become artificially sounding, thus forfeiting their ecological validity. Other studies who utilised synthesised excerpts tended to manipulate monophonic melodies (Bresin & Friberg, 2011; Juslin, 1997b; Kamenetsky, Hill, & Trehub, 1997; Lindström, 2006; Saarikallio et al., 2019; Scherer & Oshinsky, 1977) rather than polyphonic excerpts. Current studies have grappled to retain a balance between ecological validity of the stimuli and experimental control (Gabrielsson & Lindström, 2010). Furthermore, when utilising pre-existing commercial music (with the most common music genres used in music emotion research being classical and popular music), it is possible that participants are already familiar with the stimuli due to prior exposure through mediums such as films and adverts, and might result in an unconscious familiarity bias, which cannot be controlled (Juslin & Västfjäll, 2008).

Specifically composing music for research would avoid any probability of familiarity bias, and it helps researchers move away from the tendency to overuse commercial recordings. At present, only a minor percentage (less than 1%) of stimuli sets used in perceived emotional expression research have been specifically composed (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). Creating new musical excerpts specifically for research would also allow the researcher to attend to the balance between ecological validity and experimental control during the creation process of the stimuli by composing music which is flexible and allows cue manipulations. Lastly, a review on stimuli used in music emotion research carried out by Warrenburg (2020a) that covers stimuli used from 1928 up till 2018, revealed that more than half of the stimuli used in music emotion studies represented either sadness, happiness, or anger, with other emotions that have been considered as being expressed in music, such as peace and tenderness, appearing only 1% of the time across studies (2020a, p. 243). These findings demonstrate that other emotional expressions deemed to be perceived in music need to start being considered more. Thus, future research should aim to compose musical stimuli that represent said emotions, for a better understanding of how different musical emotions are encoded and communicated to the listener.

1.7 Summary of Thesis Contributions

The sections described above review how research on perceived musical expression has unfolded over the years, advancing our understanding of how and which emotional expressions are perceived in music, mostly within a Western musical context, and which factors in a musical work contribute to communicating these emotions. Although large strides have been made in this field, more work is needed to better understand the role of specific musical cues in conveying intended emotions, probing the large cue space on which emotions are mapped.

This current thesis project is based on the following assumptions: i) emotional meaning can be perceived in music (Dowling & Harwood, 1986; Juslin & Sloboda, 2012) ii) particular emotional expressions may be encoded in music through the additivity of probabilistic and partially redundant universal and culture-specific cues (Balkwill & Thompson, 1999; Gabrielsson & Juslin, 1996; Juslin, 1997a, 2019d)

iii) emotional expressions in music may be successfully decoded and recognised by listeners who use their knowledge and personal experiences to understand the music in question (Davies, 2001; Juslin, 2019c) iv) providing a context in which the music under investigation exists helps set a baseline for understanding the emotional meaning (Balkwill & Thompson, 1999; Fritz et al., 2009; Juslin, 2019c). This project investigates how six and then later, seven musical cues, are used together to try and communicate a selection of emotional expressions which have been reported as being highly relevant and frequently perceived in music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003). The current project employs an interactive process to investigate the effect of various musical cues in communicating different emotions through music, whilst also addressing significant limitations highlighted in this chapter that have been detrimental to emotion perception research. This research was carried out in a Western context, utilising Western tonal musical pieces and participants mainly from a Western culture.

This thesis contains six chapters in total, including the current introductory chapter. Chapters 2 and 4 contain work published in journals, while Chapter 5 contains work that has been submitted to a journal and is currently under review. To ensure the work in this thesis is cohesive, prefaces have been added at the beginning of Chapters 2 to 5 to contextualise the work done in each individual chapter and its role as part of the whole project. All findings and contributions of this thesis are summed up and discussed in the final chapter of this thesis - Chapter 6.

The first limitation addressed in this project concerns how previous studies only investigated a *limited number of cues* simultaneously and/or only a small number of levels per cue. This was mostly a feasibility issue due to methodological constraints. Utilising an interactive paradigm bypasses the taxing methodology of a listening experiment and its constraints on the number of cues and combinations tested, by providing users with substantially bigger cue ranges to explore in a relatively shorter time. As a resolution to this methodological issue, **Chapter 2** introduces a new, interactive computer interface called *EmoteControl* which allows users to change six cues (tempo, pitch, articulation, dynamics, brightness, and mode) of a tonal instrumental musical piece in real-time. *EmoteControl* was inspired by interfaces such as *pDM* (Friberg, 2006) and the gesture controller built by Bresin and Friberg (2011). *EmoteControl* was specifically created for the research carried out in this thesis and for future prospective music and emotion research.

Chapter 2 describes how *EmoteControl* was designed, utilising the visual programming language Max/MSP, the music production software Logic Pro X, and the extensive sound library Vienna Symphonic Library (VSL) as its three main tools. The details of the cues and the rationale behind it are explained, together with its system architecture. Unlike previous interfaces, *EmoteControl* provides users control over a selection of *both* structural and expressive cues of a musical piece which they can personally alter. The interface uses digital sliders and buttons to manipulate the *whole polyphonic* musical piece, rather than just the monophonic melodic line or individual chords. Furthermore, no musical expertise is required to utilise the interface, which allows for researchers to investigate how a general population (both musicians and non-musicians) thinks different emotions should be conveyed in music. The system was also subjected to a formal evaluation via a Music-HCI (Human Computer Interaction) study to gather feedback from users on the learnability and usability of the interface for its target use. Multiple use cases of *EmoteControl* are highlighted; first and foremost, as a new interactive system for musical cues and emotion research, but also as an educational tool.

To ensure maximum compatibility with *EmoteControl* whilst also addressing *choice of stimuli* and *familiarity bias* issues, 28 new, tonal polyphonic musical pieces were composed to be used as stimuli in this work and for future research, detailed in **Chapter 3**. Each piece was composed within a Western tonal framework and with the intention of conveying one specific emotion from a selection of nine emotional expressions: sadness, joy, calmness, anger, fear, power, surprise, love, and longing. The creation of new stimuli is a contribution to the field in itself, as it furthers the current repository of stimuli available for music emotion research, providing the field with stimuli representing other emotions apart from the most common ones (happy, sad, and angry) and moving away from the *overuse of commercial music* (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). It also eliminates any possibility of *familiarity bias* by the participants which might be encountered when utilising commercial music as stimuli, due to previous exposure. Most importantly, when composing new music with prior knowledge that the musical pieces will be modified for experimental use, it allows for a *balance* between composing *ecologically valid music* whilst ensuring that the pieces permit the *necessary flexibility for experiment control*.

Chapter 4 features three empirical experiments: Experiment 1 is a validation study where participants listened to and rated the newly composed 28 musical pieces with regards to their emotional content. This was done to determine whether the intended emotional expression of each piece was successfully conveyed to the listener. Collecting emotion ratings for the different musical pieces also provides future researchers utilising the stimuli with data on how a sample of listeners have already perceived the material. Experiment 2 detailed in Chapter 4 utilised the *EmoteControl* interface together with a subset of the composed and rated musical pieces for a cue manipulation study, testing the production approach that the tools were created for. Participants used the interactive paradigm to change the musical pieces via six cues in real-time, with the aim of reaching different emotional expressions through the music. The results of Experiment 2 showed how the combination of tempo, pitch, dynamics, brightness, articulation, and mode affects the conveying of different emotional expressions in music – with mostly similar results to previous literature, but also some novel findings. Experiment 2 also produced a new set of stimuli comprised of the variations of the original music created by the mean cue values used by participants. The variations created by the participants as well as the original musical pieces were subjected to a rating study in Experiment 3. This last rating study explored whether the musical variations created in Experiment 2 successfully conveyed their intended emotion to a new set of listeners. The original pieces were also rated by the new set of listeners, allowing for an examination of whether the different cue combinations used to portray the same emotion were successful in doing so or not. Findings showed that the original seven pieces were all rated as strong representatives of their intended emotion, whilst four of the variations created by participants were successful in conveying their intended emotion.

The aim of the last two empirical experiments described in **Chapter 5** was two-fold. First, the two experiments aimed to further probe a larger cue-emotion space, following the work done in the preceding chapter. Experiment 1 utilised a second version of the *EmoteControl* interface, which was modified to allow for *seven* musical cues to be altered: tempo, pitch, articulation, brightness, dynamics, mode, and instrumentation. The same musical stimuli used in the cue manipulations study reported in Chapter 4 were also used in this experiment. Apart from providing new information on how the seven cues were simultaneously used to shape different emotions in instrumental Western tonal music, the second aim addressed in this chapter was to compare the results produced by the *EmoteControl* interactive paradigm with the traditional systematic manipulation approach mostly utilised in musical cues and emotion perception research. As discussed in this current chapter, several methodologies have been utilised in music and emotion studies. However, there are *no assurances that these different methodologies produce compatible results.* In light of this, a second experiment was run in Chapter 5. Experiment 2 consisted of a listening experiment where participants rated musical excerpts varying in cue levels of the same seven cues explored in Experiment 1. The comparison provides the novel contribution of addressing pros and cons of the two different approaches and sheds light on the compatibility of results across the two methodologies, offering information on the best approach to a wider exploration of the cue-emotion space for future reference.

All findings and contributions of this thesis are then summarised in **Chapter 6**. The implications of this thesis are discussed in the context of music emotion research as well as applied directions such as marketing, music therapy, and education settings. Limitations of the project are also discussed, together with recommendations for future studies based on the work stemming from this thesis.

In summary, the work accomplished in this thesis presents a new interactive tool and an emotion-rated musical stimulus set to explore an extensive cue space in which different emotional expressions are mapped out within a Western musical context. The empirical chapters in this project provide new data on how a combination of structural and expressive cues are used by individuals to shape different emotions in the same unfamiliar tonal instrumental polyphonic musical pieces. An evaluation of the new interactive paradigm and the traditional systematic manipulation approach is also carried out, presenting new information on best approaches to musical cues and emotion perception studies.

1.8 References

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., Busch, V., Lothwesen, K., Elvers, P., Fischinger, T., Schlemmer, K., & 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

Athanasopoulos, G., Eerola, T., Lahdelma, I., & Kaliakatsos-Papakostas, M. (2021). Harmonic organisation conveys both universal and culture-specific cues for emotional expression in music. *PLoS ONE*, *16*(1), 1–17. https://doi.org/10.1371/journal.pone.0244964

Baird, A., & Thompson, W. F. (2018). The Impact of Music on the Self in Dementia. *Journal of Alzheimer's Disease*, *61*, 827–841. https://doi.org/10.3233/JAD-170737

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.

Barrett, L. F. (2006a). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Barrett, L. F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20–46.

https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F., & Wager, T. D. (2006). The Structure of Emotion: Evidence From Neuroimaging Studies. *Current Directions in Psychological Science*, *15*(2), 79–83. https://doi.org/10.1111/j.0963-7214.2006.00411.x

Battcock, A., & Schutz, M. (2019). Acoustically expressing affect. *Music Perception*, *37*(1), 66–91. https://doi.org/10.1525/MP.2019.37.1.66

Behne, K. E. (1972). *Der Einfluss des Tempos auf die Beurteilung von Musik [The influence of tempo on judgement of music]*. Arno Volk Verlag.

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., Vieillard, 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

Brodsky, W. (2011). Developing a functional method to apply music in branding: Design language-generated music. *Psychology of Music*, *39*(2), 261–283. https://doi.org/10.1177/0305735610387778

Brunswik, E. (1956). Perception and the representative design of psychological experiments. In *Perception and the representative design of psychological experiments, 2nd ed.* University of California Press.

Campbell, I. G. (1942). Basal emotional patterns expressible in music. *The American Journal of Psychology*, *55*, 1–17. https://doi.org/10.2307/1417020

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

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

Cooke, D. (1959). *The Language of Music* (First). Oxford University Press.

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., & Nese, M. (2020). Perceived Tension, Movement, and Pleasantness in Harmonic Musical Intervals and Noises. *Music Perception*, *37*(4), 298–322.

Cousineau, M., Carcagno, S., Demany, L., & Pressnitzer, D. (2014). What is a melody? On the relationship between pitch and brightness of timbre. In *Frontiers in Systems Neuroscience* (Vol. 7, p. 127).

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(4), 1924–1934. https://doi.org/10.1073/pnas.1910704117

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.

Davies, S. (1994). *Musical Meaning and Expression*. Cornell University Press.

Davies, S. (2001). Philosophical perspectives on music's expressiveness. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 23–44). Oxford University Press.

Davies, S. (2013). Music-to-Listener Emotional Contagion. In T. Cochrane (Ed.), *The Emotional Power of Music: Multidisciplinary perspectives on musical arousal, expression, and social control*. Oxford University Press. https://doi.org/10.1093/acprof

Dennett, D. C. (1987). *The Intentional Stance* (Issue 1). MIT Press.

Dowling, W. J., & Harwood, D. L. (1986). *Music Cognition*. NY: Academic Press.

Downey, J. E. (1897). A musical experiment. *The American Journal of Psychology*, *9*(1), 63–69. https://doi.org/10.2307/1411476

Duffy, E. (1941). An Explanation of "Emotional" Phenomena without the use of the Concept "Emotion." *The Journal of General Psychology*, *25*(2), 283–293. https://doi.org/10.1080/00221309.1941.10544400

Eerola, T. (2018). Music and emotion. In *Springer Handbook of Systematic Musicology* (1st ed.). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-662-55004-5

Eerola, T., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Perception*, *30*(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. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49. https://doi.org/10.1177/0305735610362821

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.

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75–99. https://doi.org/10.1177/102986490801200105

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., Bresin, R., & Sundberg, J. (2014). Analysis by Synthesis. In W. F. Thompson (Ed.), *Music in the Social and Behavioral Sciences: An Encyclopedia* (p. Vol.1, 44-45). SAGE Publications, Inc. https://doi.org/http://dx.doi.org/10.4135/9781452283012

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & 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. (1973). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology*, *14*(1), 244–260. https://doi.org/10.1111/j.1467-9450.1973.tb00118.x

Gabrielsson, A. (1985). Interplay between Analysis and Synthesis in Studies of Music Performance and Music Experience. *Music Perception: An Interdisciplinary Journal*, *3*(1), 59–86. https://doi.org/10.2307/40285322

Gabrielsson, A. (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae VO - Special Issue 2001-2002*, 123–147. https://doi.org/https://doi.org/10.1177/10298649020050S105

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology* *of Music*, *31*(3), 221–272. https://doi.org/10.1177/03057356030313002

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/oxfordhb/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., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503–534). Oxford University Press.

Gabrielsson, A., & Lindström, E. (1995). Emotional Expression in Synthesizer and Sentograph Performance. *Psychomusicology*, *14*(1), 94–116.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In *Music and emotion: Theory and research.* (pp. 223–248). Oxford University Press.

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., Gosselin, N., Provencher, V., & Bier, N. (2012). Perception and Emotional Judgments of Music in Dementia of the Alzheimer Type: A Short Case Study. *Music Perception: An Interdisciplinary Journal*, *29*(5), 509–519.

Gagnon, L., & Peretz, I. (2000). Laterality effects in processing tonal and atonal melodies with affective and nonaffective task instructions. *Brain and Cognition*, *43*(1–3), 206–210. https://doi.org/10.1006/brcg.1999.1135

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

Garrido, S., & Schubert, E. (2013). Adaptive and maladaptive attraction to negative emotions in music. *Musicae Scientiae*, *17*(2), 147–166. https://doi.org/10.1177/1029864913478305

Gilman, B. I. (1891). Report on an experimental test of musical expressiveness. *American Journal of Psychology*, *4*, 558–576.

Gilman, B. I. (1892). Report of an experimental test of musical expressiveness (continued). *American Journal of Psychology*, *5*, 42–73.

Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., & Samson, S. (2011). Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *47*(9), 1116–1125. https://doi.org/10.1016/j.cortex.2011.05.012

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., & Samson, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*(3), 628–640. https://doi.org/10.1093/brain/awh420

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

Hanslick, E. (1854). *Vom musikalisch Schönen* (21st ed.). Breitkopf and Härtel.

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.

Huber, K. (1923). *Der Ausdruck musikalischer Elementarmotive: Eine experimentalpsychologische Untersuchung*. Verlag von Johann Ambrosius Barth.

Hunter, P. G., & Schellenberg, E. G. (2010). Music and Emotion. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception, Springer Handbook of Auditory Research* (Vol. 36, pp. 129–164). Springer Science+Business Media. https://doi.org/10.1007/978-1-4419-6114-3

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and Perceptions of Happiness and Sadness Induced by Music: Similarities, Differences, and Mixed Emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(1), 47–56. https://doi.org/10.1037/a0016873

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.

Imberty, M. (1979). *Entendre la musique: sémantique psychologique de la musique*. Dunod.

Izard, C. E. (1977). *Human Emotions*. Plenum Publishing Corporation. https://doi.org/10.1007/978-1-4899-2209-0

Izard, C. E. (2009). Emotion Theory and Research: Highlights, Unanswered Questions, and Emerging Issues. *Annual Review of Psychology*, *60*(1), 1–25. https://doi.org/10.1146/annurev.psych.60.110707.163539

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*, *I*(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. (2001). A Brunswikian approach to emotional communication in music performance. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik: Beginnings, Explications, Applications.* (pp. 426–430). Oxford University Press.

Juslin, P. N. (2013a). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N. (2013b). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, *4*(596), 1–14. https://doi.org/10.3389/fpsyg.2013.00596

Juslin, P. N. (2019a). Breaking the code. In *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (pp. 122–138). Oxford University Press.

Juslin, P. N. (2019b). *Musical emotions explained: Unlocking the secrets of musical affect.* Oxford University Press.

Juslin, P. N. (2019c). Viewing Music Through a Brunswikian Lens. In P. N. Juslin (Ed.), *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (pp. 147-C10.N1). Oxford University Press. https://doi.org/10.1093/oso/9780198753421.003.0010

Juslin, P. N. (2019d). The Voice of Angels? Iconic Expression. In P. N. Juslin (Ed.), *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (pp. 156–

C11.N5). Oxford University Press. https://doi.org/10.1093/oso/9780198753421.003.0011

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae*, *5*(Special Issue 2001-2002), 63–122. https://doi.org/10.1177/10298649020050s104

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., Liljeström, S., Västfjäll, D., & Lundqvist, L. O. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230143.003.0022

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., & Sloboda, J. A. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 71–104). Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (2010). Introduction: aims, organization, and terminology. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion* (pp. 3–11). Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (2012). Music and Emotion. In D. Deutsch (Ed.), *The Psychology of Music* (Third Edit, pp. 583–632). Academic Press.

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 452–489). 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.

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191–213. https://doi.org/10.1177/102986490601000203

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.

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, *4*(June), 1–15. https://doi.org/10.3389/fpsyg.2013.00311

Kivy, P. (1980). *The Corded Shell: Reflections on Musical Expression*. Princeton University Press.

Kleinen, G. (1968). *Experimentelle Studien zum musikalischen Ausdruck [Experimental* *studies on musical expression]*. Universität Hamburg.

Knobloch, S., & Zillmann, D. (2002). Mood management via the digital jukebox. *Journal of Communication*, *52*(2), 351–366. https://doi.org/10.1093/joc/52.2.351

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*. Department of Pscyhology, Keele (CD-rom).

Krumhansl, C. L. (1996). A Perceptual Analysis of Mozart's Piano Sonata K. 282: Segmentation, Tension, and Musical Ideas. *Music Perception*, *13*(3), 401–432. https://doi.org/10.2307/40286177

Lahdelma, I., Athanasopoulos, G., & Eerola, T. (2021). Sweetness is in the ear of the beholder: chord preference across United Kingdom and Pakistani listeners. *Annals of the New York Academy of Sciences*, *1502*(1), 72–84. https://doi.org/10.1111/nyas.14655

Lahdelma, I., & Eerola, T. (2016). Mild dissonance preferred over consonance in single chord perception. *I-Perception*, *7*(3). https://doi.org/10.1177/2041669516655812

Lahdelma, I., & Eerola, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-65615-8

Lamont, A., & Eerola, T. (2011). Music and emotion. *Musicae Scientiae*, *15*(2), 139–

145. https://doi.org/10.1177/1029864911403366

Langer, S. K. (1953). Feeling and Form; a Theory of Art Developed from Philosphy in a New Key. *Journal of Philosophy*, *52*(11).

Langer, S. K. (1957). *Problems of art; ten philosophical lectures.* Scribner.

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*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Laukka, P., & Gabrielsson, A. (2000). Emotional Expression in Drumming Performance. *Psychology of Music*, *28*(2), 181–189. https://doi.org/10.1177/0305735600282007

Lee, H., & Müllensiefen, D. (2020). The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability. *Attention, Perception, and Psychophysics*, *82*(7), 3658–3675. https://doi.org/10.3758/s13414-020-02058-3

Leman, M., Vermeulen, V., De Voogdt, L., Moelants, D., & Lesaffre, M. (2005). Prediction of musical affect using a combination of acoustic structural cues. *Journal of New Music Research*, *34*(1), 39–67. https://doi.org/10.1080/09298210500123978

Lepa, S., Herzog, M., Steffens, J., Schoenrock, A., & Egermann, H. (2020). A computational model for predicting perceived musical expression in branding scenarios. *Journal of New Music Research*, *49*(4), 387–402. https://doi.org/10.1080/09298215.2020.1778041

Lindström, E. (1997). Impact of melodic structure on emotional expression. *Proceedings of the Third Triennial ESCOM Conference, Uppsala*, 292–297.

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.

Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. *British Journal of Psychology*, *102*(1), 108–134. https://doi.org/10.1348/000712610X506831

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

MacDonald, R., Kreutz, G., & Mitchell, L. (2012). What is Music, Health, and Wellbeing and Why is it Important? In *Music, Health, and Wellbeing* (pp. 3–11). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586974.003.0001

Meyer, L. B. (1956). *Emotion and Meaning in Music* (Vol. 1, Issue 1). The University of Chicago Press. https://doi.org/10.2307/843099

Mohn, C., Argstatter, H., & Wilker, F. W. (2010). Perception of six basic emotions in music. *Psychology of Music*, *39*(4), 503–517. https://doi.org/10.1177/0305735610378183

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*. University of Jyväskylä, Department of Music.

North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology*, *84*(2), 271–276. https://doi.org/10.1037/0021-9010.84.2.271

Oatley, K., & Jenkins, J. M. (1996). Understanding emotions. In *Understanding* *emotions*. Blackwell Publishing.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315–331.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.

Panksepp, J. (1998). *Affective Neuroscience: The foundations of human and animal emotions* (1st ed.). Oxford University Press.

Paquette, S., Peretz, I., & Belin, P. (2013). The "Musical Emotional Bursts": A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*(AUG), 1–7. https://doi.org/10.3389/fpsyg.2013.00509

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. (1994). *The Psychology and Biology of Emotion*. Harper Collins College Publishers.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*, 715–734.

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

Ramos, D., & Mello, E. K. S. (2021). Communication of emotions in music between Brazilian composer, performers, and listeners. *Psychomusicology: Music, Mind, and Brain*, *31*(1), 1–17. https://doi.org/10.1037/pmu0000269

Rigg, M. G. (1937). Musical expression: an investigation of the theories of Erich

Sorantin. *Journal of Experimental Psychology*, *21*(4), 442–455. https://doi.org/10.1037/h0056388

Rigg, M. G. (1939). What features of a musical phrase have emotional suggestiveness? *Bulletin. Oklahoma A & M College*, *36*, *13*, 38.

Rigg, M. G. (1940a). Speed as a determiner of musical mood. *Journal of Experimental Psychology*, *27*(5), 566–571. https://doi.org/10.1037/h0058652

Rigg, M. G. (1940b). 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

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. In *Journal of Personality and Social Psychology* (Vol. 76, Issue 5, pp. 805–819). American Psychological Association. https://doi.org/10.1037/0022-3514.76.5.805

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109. https://doi.org/10.1177/0305735607068889

Saarikallio, S., Maksimainen, J. P., & Randall, W. M. (2018). Relaxed and connected: Insights into the emotional–motivational constituents of musical pleasure. *Psychology of Music*, *47*(5), 644–662. https://doi.org/10.1177/0305735618778768

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. K., & 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

Sakamoto, M., Ando, H., & Tsutou, A. (2013). Comparing the effects of different individualized music interventions for elderly individuals with severe dementia. *International Psychogeriatrics*, *25*(5), 775–784. https://doi.org/10.1017/S1041610212002256

Schellenberg, G., Krysciak, A. M., & Campbell, J. (2000). Perceiving emotion in melody: interactive effects of pitch and rhythm. *Music Perception*, *18*(2), 155–171. https://doi.org/10.2307/40285907

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In *Appraisal processes in emotion: Theory, methods, research.* (pp. 92–120). Oxford University Press.

Scherer, K. R. (2004). Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them? *Journal of New Music Research*, *33*(3), 239–251. https://doi.org/10.1080/0929821042000317822

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. In *International Journal of Psychology* (Vol. 46, Issue 6, pp. 401–435). Taylor & Francis. https://doi.org/10.1080/00207594.2011.626049

Scherer, K. R., & Oshinsky, J. S. (1977). Cue Utilization in Emotion Attribution from Auditory Stimuli. *Motivation and Emotion*, *1*(4), 331–346.

Scherer, K. R., Sundberg, J., Tamarit, L., & Salomão, G. L. (2015). Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech and Language*, *29*(1), 218–235. https://doi.org/10.1016/j.csl.2013.10.002

Scherer, K. R., Zentner, M. R., & Schacht, A. (2001). Emotional states generated by music: An exploratory study of music experts. *Musicae Scientiae*, *5*(Special issue 2001-2002), 149–171. https://doi.org/10.1177/10298649020050s106

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions : A review of the literature and a new assessment tool. *PloS ONE*, *12*(6), 1–45. https://doi.org/10.17605/OSF.IO/Q8ZV5

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

Schubert, E. (2013a). Emotion felt by the listener and expressed by the music: Literature review and theoretical perspectives. *Frontiers in Psychology*, *4*(DEC), 1–18. https://doi.org/10.3389/fpsyg.2013.00837

Schubert, E. (2013b). Loved music can make a listener feel negative emotions. *Musicae Scientiae*, *17*(1), 11–26. https://doi.org/10.1177/1029864912461321

Schutz, M., Huron, D., Keeton, K., & Loewer, G. (2008). The Happy Xylophone: Acoustics Affordances Restrict An Emotional Palate. *Empirical Musicology Review*, *3*(3), 126–135. https://doi.org/10.18061/1811/34103

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion Knowledge: Further Exploration of a Prototype Approach. *Journal of Personality and Social Psychology*, *52*(6), 1061–1086. https://doi.org/10.1037/0022-3514.52.6.1061

Shoda, H., & Adachi, M. (2012). The Role of a Pianist's Affective and Structural Interpretations in his Expressive Body Movement: A Single Case Study. *Music Perception: An Interdisciplinary Journal*, *29*(3), 237–254. https://doi.org/10.1525/mp.2012.29.3.237

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

Sloboda, J. A., & Juslin, P. N. (2010). At the interface between the inner and outer world: Psychological perspectives. In *Handbook of music and emotion: Theory, research, applications.* (pp. 73–97). Oxford University Press.

Stamou, L. (2002). Plato and Aristotle On Music and Music Education: Lessons From Ancient Greece. *International Journal of Music Education*, *39*, 3–16. https://doi.org/10.1177/025576140203900102

Swaminathan, S., & Schellenberg, E. G. (2015). Current emotion research in music psychology. *Emotion Review*, *7*(2), 189–197. https://doi.org/10.1177/1754073914558282

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS ONE*, *9*(10), 1–17. https://doi.org/10.1371/journal.pone.0110490

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*(4), 297–303. https://doi.org/10.1111/1467-9280.00157

Thayer, R. E. (1989). The biopsychology of mood and arousal. In *The biopsychology of mood and arousal.* Oxford University Press.

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, *10*(1), 79–89.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, *25*(2), 117–134. https://doi.org/10.1525/mp.2007.25.2.117

Tomkins, S. S. (1962). Affect, imagery, consciousness, Vol. 1: The positive affects. In *Affect, imagery, consciousness, Vol. 1: The positive affects.* Springer Publishing Co. https://doi.org/10.1037/14351-000

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

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*(3), 311–317.

Warrenburg, L. A. (2020). 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

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Watson, K. (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.

Wundt, W. M., & Judd, C. H. (1897). *Outlines of psychology*. W. Engelmann; G.E. Stechert.

Zentner, M. R., Grandjean, D., & Scherer, K. R. (2008). Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Zhang, W., Liu, F., Zhou, L., Wang, W., Jiang, H., & Jiang, C. (2019). The effects of timbre on neural responses to musical emotion. *Music Perception*, *37*(2), 134–146.

Chapter 2. *EmoteControl:* An Interactive Computer Interface for Real-Time Control of Emotional Expression in Music

Micallef Grimaud, A., Eerola, T. (2021). *EmoteControl:* an interactive system for real-time control of emotional expression in music. *Personal Ubiquitous Computing, 25*(4), 677-689. https://doi.org/10.1007/s00779-020-01390-7

Preface

Chapter 2 introduces the first element of the project: an interactive computer interface called *EmoteControl*, which allows users to change six cues of a tonal, instrumental musical piece in real-time. The interface was presented to a focus group of researchers and subjected to a formal Human-Computer Interaction (HCI) evaluation study to determine its suitability for its intended purpose: a tool for musical cues and emotion research which allows for a substantially large cue space to be explored in real-time. *EmoteControl* has been used as an empirical tool in the lab. A portable version of the interface has been used for field work in

Pakistan, as well as an educational tool in a Schools' Science Festival, which showcases *EmoteControl* 's multiple uses. The interface has an element of versatility, as it can be re-programmed to add or edit the musical cues available to the user, making it a promising tool for musical cues and emotion research. In this chapter, the system architecture of *EmoteControl* is detailed, and the cues' settings are reported. Feedback from the focus group and Schools' Science Festival, together with results from the evaluation study, are discussed in relation to the usability and further development of the interface.

Abstract

Several computer systems have been designed for music emotion research that aim to identify how different structural or expressive cues of music influence the emotions conveyed by the music. However, most systems either operate offline by pre-rendering different variations of the music or operate in real-time but focus mostly on structural cues. We present a new interactive system called *EmoteControl*, which allows users to make changes to both structural and expressive cues (tempo, pitch, dynamics, articulation, brightness, and mode) of music in real-time. The purpose is to allow scholars to probe a variety of cues of emotional expression from non-expert participants who are unable to articulate or perform their expression of music in other ways. The benefits of the interactive system are particularly important in this topic as it offers a massive parameter space of emotion cues and levels for each emotion which is challenging to exhaustively explore without a dynamic system. A brief overview of previous work is given, followed by a detailed explanation of *Emote Control* 's interface design and structure. A portable version of the system is also described, and specifications for the music inputted in the system are outlined. Several use-cases of the interface are discussed, and a formal interface evaluation study is reported.

Results suggested that the elements controlling the cues were easy to use and understood by the users. The majority of users were satisfied with the way the system allowed them to express different emotions in music and found it a useful tool for research.

Keywords

Music, emotion, cues, interactive, interface, design

2.1 Introduction

This work will focus on music and emotion research, in particular, musical cues and their effect on emotional expression research. Given the large possible feature range of cues related to emotional expression, understanding how these cues operate through traditional experiments with feature manipulations is severely limited. However, in an interactive paradigm where the user is allowed to control the cue levels in real-time, this constraint is largely absent. Such interface is challenging to design and implement in a way that would satisfy users. In this chapter, we propose a new real-time interactive system called *EmoteControl* which taps into a direct user experience and allows users to control emotional expression in music via a selection of musical cues. The main contribution of *EmoteControl* is that it is a tool that can be utilised by non-experts, allowing users without any musical skill to be able to change cues of music to convey different emotions in real-time.

This chapter will be structured as follows. First, an overview of previous research in musical cues and emotions will be offered, ranging from the traditional experiment methodologies to the computer systems designed for this research field, and highlighting some shortcomings in previous research that we aim to redress with our system. *EmoteControl* is then introduced in the third section of this chapter, giving an account of the system design and how the interface works. Specifications of the source material that can be utilised with the interface will then be covered, and possible limitations of the system will be described. The fourth section will describe two interactive use-cases of the system, as well as a formal evaluation to test the system's usability with non-musicians, drawing on Human-Computer Interaction principles (Friberg & Battel, 2002; Livingstone et al., 2010). Finally, the chapter will conclude with a discussion about the implications of *EmoteControl* as well as potential future research.

2.2 Related Work

2.2.1 Musical Cues for Expressions in Emotions in Music

Previous literature suggests that listeners can recognise emotions expressed by music due to the composer's and performer's use of musical cues in encoding emotions in music. In return, the listener uses these same cues to successfully decode and recognise the emotion being expressed by the music (Juslin, 1997a, 2000; Juslin & Lindström, 2010). Musical cues can be loosely divided into two categories: structural and expressive cues. Structural cues are the aspects of music which relate to the composed score of the music, such as tempo and mode, while expressive cues refer to the features used by performers when playing a piece of music, such as timbre and articulation (Gabrielsson, 2002; Livingstone & Brown, 2005). Although this distinction has been made between the two categories, which features belong in the structural and expressive categories is still open to interpretation, as some features such as dynamics can be modified by both composers and performers. Thus, it is not always possible to completely separate the performance features from the composition features (Livingstone et al., 2010). Either way, both structural and expressive cues are responsible for conveying emotion (Friberg & Battel, 2002) and are known to operate together in an additive and sometimes also in an interactive fashion (Gabrielsson, 2008; Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014); thus, a combination of both structural and expressive cues should be investigated together. In this work, dynamics, mode, tempo, and pitch will be referred to as structural cues, whilst brightness and articulation will be interpreted as expressive cues.

Throughout the years, empirical studies on emotional expression in relation to musical cues have been approached with different methodologies. Hevner introduced systematic manipulation of musical cues, where structural cues such as mode (1935), tempo, or pitch (1937) were varied on different levels (e.g., rising or falling melodic contour, or key from major to minor) in short pieces of tonal music, resulting in the creation of distinct versions of the same musical samples. The emotion conveyed in each variation was assessed by participants listening to the different music versions and rating which emotion(s) they perceived, which assisted in identifying how the different cues affected the communicated emotion. Thompson and Robitaille (1992) also inspected the structural cues of the music, by identifying similarities and differences in the structure of different melodies. In contrast to this, other studies investigated the effect of expressive cues of the music by instructing musicians to perform musical excerpts in different ways to convey various emotions and analyse how listeners use the expressive cues to decode the communicated emotions (Juslin, 1997a, 2000; Schubert, 2004).

Technological advancements have enabled new approaches to music emotion research, as computer systems have been specifically created for modelling musical cues to allow control over the emotions expressed by music (Collins, 2010;

Quinto, Thompson, & Keating, 2013). Such systems include KTH's *Director Musices* (Friberg et al., 2000; Friberg, Bresin, & Sundberg, 2006), a rule-based system that takes in a musical score as input and outputs a musical performance by applying a selection of expressive cues; CMERS (Livingstone et al., 2010) which is a similar rule-based system that provides both expressive and structural additions to a score; and linear basis function framework systems (Cancino Chacón, 2018; Grachten & Widmer, 2012) capable of modelling musical expression from information present in a musical score. Inductive machine-learning systems have also been used to model expressive performance in jazz music (Ramirez & Hazan, 2005).

Although the aforementioned systems allow multiple variations of musical cues, particularly, expressive cues, to be generated, this can only be achieved in non-real-time due to the need to pre-render cues. Systems that would carry more applied potential for musical interaction would be real-time systems which do not require pre-defined and pre-rendered discrete levels of cues but allow for control over a continuous range of each musical cue, as well as multiple cues to be explored together. A field of research that concerns real-time computer systems and the user's interaction with them is Music Human-Computer Interaction (HCI).

2.2.2 Music Interaction and Expression in Music

Music HCI (Human-Computer Interaction) looks at the various new approaches, methodologies, and interactive computer systems that are utilised for musical activities and music research. Music HCI focusses on multiple aspects of these methodologies and computer systems, such as their design, evaluation, and the user's interaction with the systems (Holland et al., 2014).

Interactive systems enable a versatile approach to music performance and research, as they allow for real-time manipulation of music. The advantage of the real-time manipulation of musical cues is that it affords the exploration of the large parameter space associated with altering expression of music, such as the production of different harmonies through a whole body interaction with the system (Bouwer, Holland, & Dalgleish, 2013; Holland et al., 2011), creation of rhythmic improvisation (Gifford, 2013), and pitch (Wong, Yuen, & Choy, 2008). These systems are shifting the focus to a direct user approach, where users can directly interact with the system in real-time to manipulate music and emotional expression. For instance, Robin (Morreale, Masu, & Angeli, 2013; Morreale & De Angeli, 2016) is an algorithmic composer utilising a rule-based approach to produce new music by changing certain structural musical cues depending on the user's choice of emotion terms or inputted values on a valence-arousal model; EDME (Lopez, Oliveira, & Cardoso, 2010), an emotion-driven music engine, generates music sequences from existing music fragments varying in pitch, rhythm, silence, loudness, and instrumentation features, depending on emotion terms or valence-arousal values selected by the user; the Emotional Music Synthesis (EMS) system uses a rule-based algorithm which controls seven structural parameters to produce music depending on valence-arousal values chosen by the user (Wallis et al., 2011). Hoeberechts and Shantz (2009) created an Algorithmic Music Evolution Engine (AMEE) which is capable of changing six musical parameters in real-time to convey ten selected emotions, while Legaspi et al. (2007) programmed a Constructive Adaptive User Interface (CAUI) which utilises specific music theory concepts such as chord inversions and tonality as well as cadences and other structural parameters to create personalised new music or alter existing ones depending on the user's preferences.

Although the aforementioned systems allow for real-time user interactions using MIDI representation, important shortcomings with respect to gaining insights into the actual musical cues utilised are identified. First, although these past real-time systems work with multiple musical cues, the cues utilised are mainly structural cues and expressive ones are generally disregarded. This issue arises from the fact that most systems work with a real-time algorithmic composer which would require expressive cues to be pre-defined beforehand. Secondly, although the user is interacting with the systems, the user can only partially control the music created or altered, as the user only has access to emotion terms and valence-arousal values. Therefore, the user might not agree with how the music is being created or altered by the system to convey the desired emotion. Thirdly, not all systems have carried out validation studies that confirm that the music generated by the systems is indeed conveying the desired emotion (Legaspi et al., 2007; Morreale & De Angeli, 2016; Wallis et al., 2011). Finally, these systems are unfortunately not accessible in any form.

By drawing from both music emotion and HCI research, we can summarise that the current offline rule-based systems are capable of modelling several cues simultaneously, but they do not allow for real-time manipulations of the cues and direct interaction with the user. Also, the potential cue space is enormous since there are seven to fifteen cues linked to musical expression which each may be represented on a continuum (e.g., dynamics, pitch, timbre) or at different levels (e.g., articulation, harmony). Mapping the potential cue space to emotions is too large to be exhaustively studied with static factorial experiments, and such experiments would not have any guarantee of having the appropriate cue levels in the first place. On the other hand, current real-time systems allow for partial user interaction; however, the users do not directly change the musical cues themselves. In order to make progress in emotional expression in music research, focus should be directed to real-time systems which allow for the continuous adjustment of musical cues by users without requiring any musical expertise. In this approach, it is productive to allow for a direct user experience, where participants have direct interactions with the system and can personally change cues of a musical piece to convey different perceived emotions as they desire, in real-time (Bresin & Friberg, 2011; Friberg, 2006; Kragness & Trainor, 2019). This would allow one to explore the potentially large parameter space consisting of cues and cue levels (e.g., consider a simple example for four cues, tempo, articulation, pitch, and dynamics, each having potentially five cue levels, amounting to cue 1,024 cue-level combinations) to emotions in an efficient and natural way. This latter property refers to the dilemma of traditional experiments where the scholars need to decide the appropriate cue levels (e.g., three levels of tempi, 60, 100, and 140 BPM) which might not be the optimal values for specific emotions and emotion-cue combinations. A real-time system allows to offer a full range of cues to the user which is by far a more ecological valid and effective way of discovering the plausible cue-emotion combinations. Additionally, these interactive music systems should be formally evaluated to confirm their usability and usefulness within their paradigm (Poepel, 2005; Wanderley & Orio, 2002). Following this line of argument and with the aim of expanding the research on music and emotions from a non-expert perspective, *EmoteControl,* which is a new real-time interactive music system was designed.

2.3 Interface Design

2.3.1 System Architecture

For the creation of the system, two computer programs were used: Max/MSP and Logic Pro X. Max/MSP is used as the main platform while Logic Pro X[4] works in the background as the rendering engine. Additionally, Vienna Symphonic Library[5] (VSL) is used as a high-quality sound synthesizer operated through the Vienna Instruments sample player in Logic[6]. Functions in a Max/MSP patch allow the alterations of different cues of a musical track as it plays in real-time. Musical tracks to be utilised in the system are presented in MIDI format, as this allows for better manipulation of the file information than an audio file would.

A MIDI file is read by a sequencer in Max/MSP, which in turn sends the data through to Logic. The output of the MIDI file is played with a chosen virtual instrument in Logic. A chamber strings sound synthesizer from VSL is utilised as the default virtual instrument in the *EmoteControl* interface, since instruments from the strings family seem to have rather versatile features, such as the ability of playing quite slow or soft (Huron, Anderson, & Shanahan, 2014), and provide a substantially large pitch range. During the playback of the MIDI track, structural and expressive cues of the music can be manipulated in a continuous or discrete manner, and changes can be instantly heard as the MIDI track plays in real-time. The system architecture design can be seen in Figure 2.1.

---

[4] Logic Pro X will be denoted in the text as Logic for short.
[5] https://www.vls.co.at/en
[6] The *EmoteControl* project is available at github.com/annaliesemg/EmoteControl2019

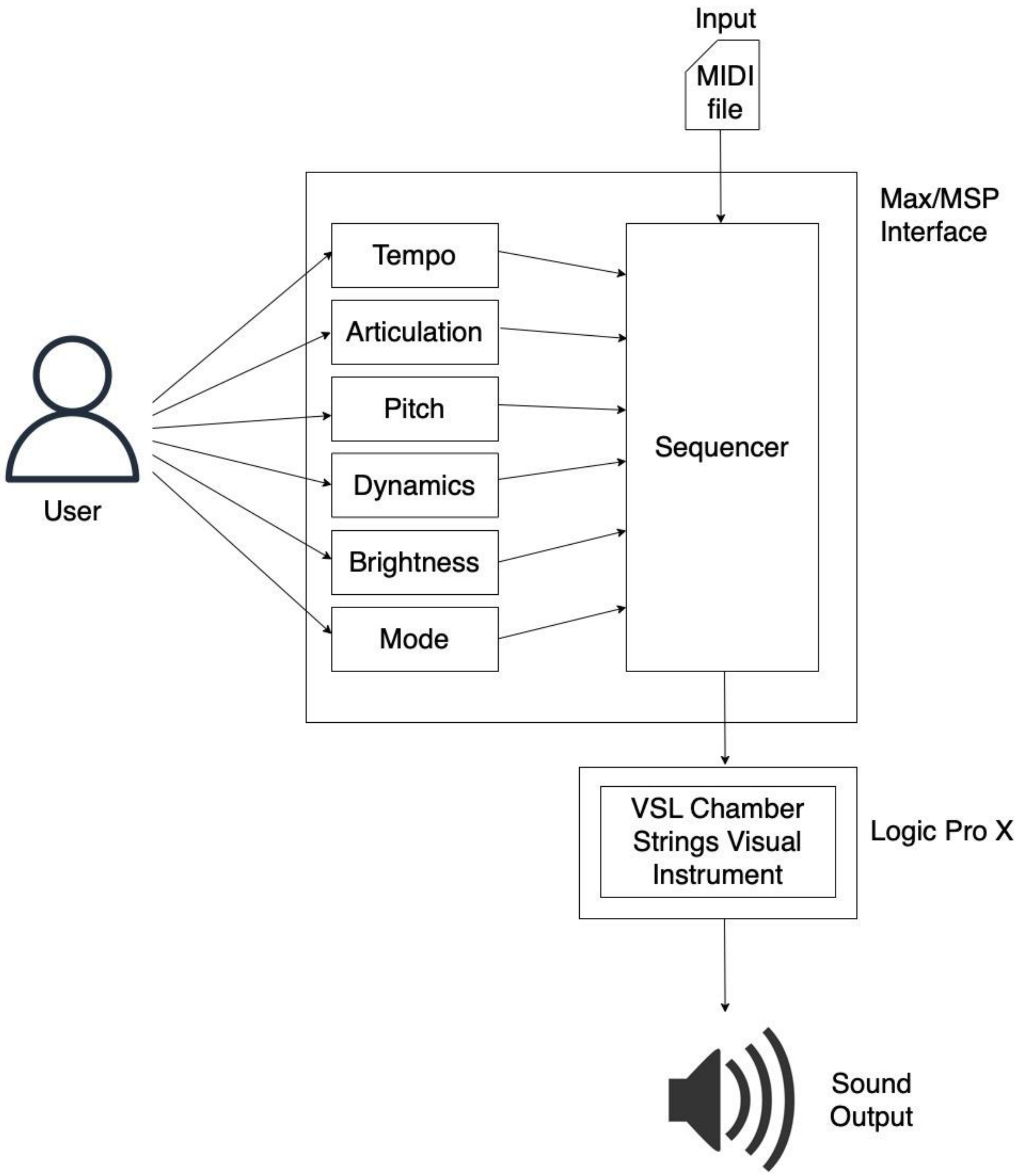


Figure 2.1. System architecture of *EmoteControl.*

2.3.2 User Interface

EmoteControl (Figure 2.2) is targeted at a general audience and aims to allow users to utilise the interface and manipulate cues of a musical piece with no prior music knowledge required. All functions present in the interface are accessed via visual elements manipulated by a computer mouse. The following six cues can be altered via the interface: tempo, articulation, pitch, dynamics, and brightness settings can be changed using the sliders provided, while mode can be switched from major to minor mode and vice-versa with a toggle button.

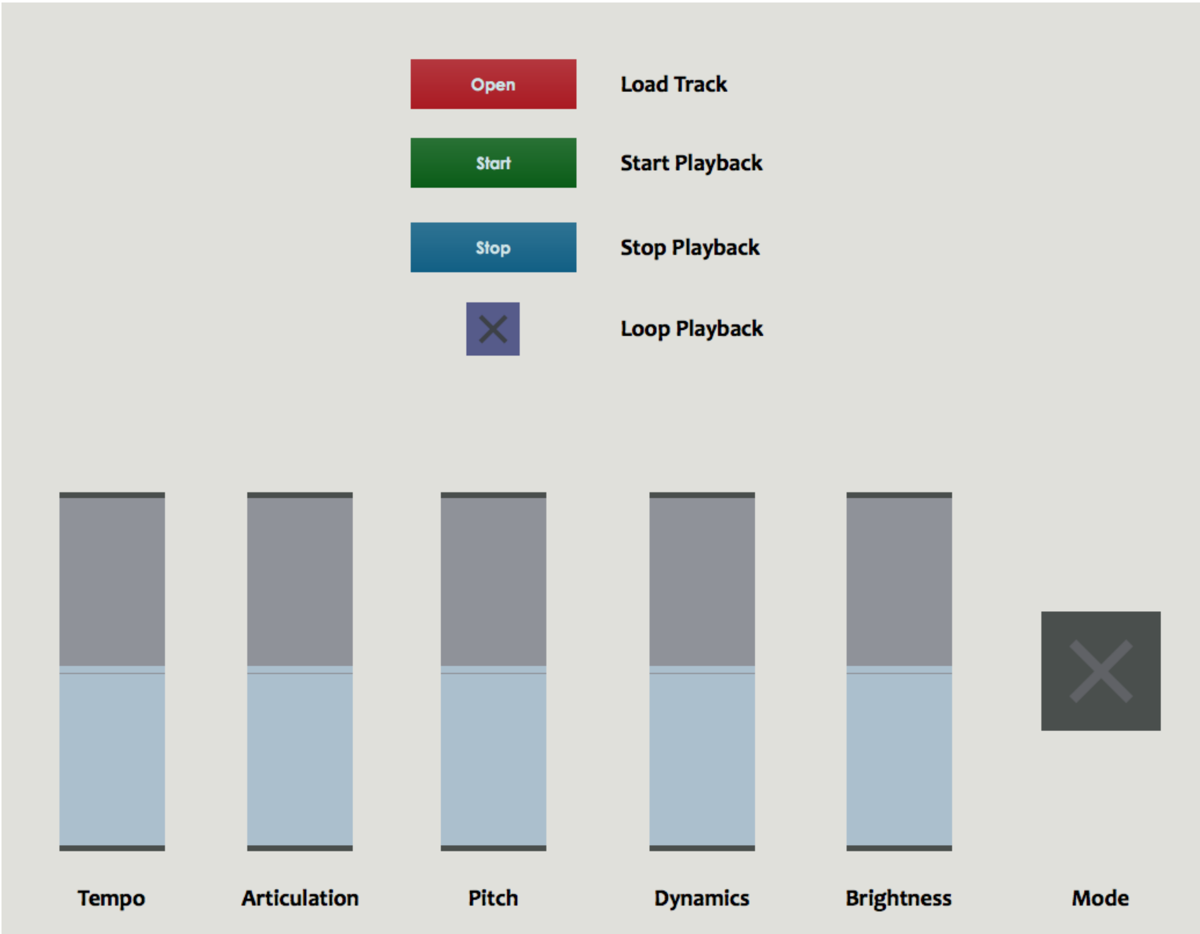


Figure 2.2. *EmoteControl* user interface.

2.4 Cues

As *EmoteControl* was specifically designed for music and emotion research, the cues available for manipulation were based on previous research, with a selection of frequently studied and effective cues in past research, as well as other less studied cues. Furthermore, the selection of cues features structural and expressive cues since both contribute to the conveyed emotional expression.

Tempo, mode, and dynamics are considered as the most influential structural cues in conveying emotion (Dalla Bella et al., 2001; Kamenetsky, Hill, & Trehub, 1997; Morreale et al., 2013); thus, we have implemented them as possible cues to manipulate to re-affirm and refine the results of past studies, and simultaneously provide a baseline for research carried out with *the EmoteControl* interface. The structural parameter pitch, and expressive cues articulation and brightness, have been periodically documented to contribute to emotions expressed by music (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 2014), albeit to a lesser degree. Although previous studies have investigated multiple cue combinations (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014), most of them were limited to a small number of levels per cue, as different levels had to be pre-rendered. Only a handful of studies utilised live cue manipulation systems to assess the effect of cues on perceived emotions, and these studies were mainly focussed on expressive cues (Bresin & Friberg, 2011; Kragness & Trainor, 2019; Saarikallio, Vuoskoski, & Luck, 2014).

Thus, the current cue selection (tempo, mode, dynamics, pitch, articulation, and brightness) will allow for a better understanding of how the cue combinations and full range of cue levels affect emotion communication. The six cues available in *EmoteControl* have been implemented in the following fashion.

2.4.1 Tempo

A slider controls the tempo function in Max/MSP which works as a metronome for incoming messages. The function reads the incoming data and sets the speed at which the data is outputted, in beats per minute (bpm). The slider was set with a minimum value of 50 bpm and a maximum value of 160 bpm, to cover a broad tempo range from *lento* to *allegrissimo*. When the slider is moved, the tempo function overrides the initial set tempo, and the MIDI file is played with the new tempo.

2.4.2 Articulation

Three levels of articulation (*legato*, *detaché,* and *staccato)* in the Vienna chamber strings instrument plug-in are controlled via a slider in the Max/MSP patch. The selected articulation methods were put in a sequence from longest note-duration to shortest, and changes are made as the slider glides from one articulation method to the next. Although the output of the articulation slider is continuous, the current implementation of the cue is based on real instrument articulation types and therefore offers limited and discrete choices ranging from *legato* to *detaché* to *staccato* levels.

2.4.3 Pitch

A slider in the interface controls the pitch of the virtual instrument, with a total range of ±2 semitones. Pitch shifts are made in terms of semitones, but the implementation of the shift utilises a pitch bend function that allows the pitch to be shifted in a smooth progression from one semitone to the other. This produces the perception of a continuous pitch shift of the discrete MIDI values of the slider.

2.4.4 Dynamics

The dynamic slider controls the volume feature within VSL. Dynamic changes in a real instrument also affect other acoustical attributes of an instrument, such as the timbre and envelope. In sound libraries such as VSL, sounds are sampled at different levels of loudness to emulate the distinct sound changes produced by the real instruments. On the other hand, changing the decibel (dB) level of the virtual instrument track will amplify or reduce the overall sound, not taking into consideration the other acoustical properties that are also affected. Thus, controlling the volume within the VSL plug-in produces a more realistic output, although some of the other acoustic parameters (e.g., brightness) will also change. The dynamic slider has a range of 99 MIDI volume, with the minimum being 30 and the maximum possible dynamic level being 129. It is worth noting that accenting patterns in the MIDI file (i.e., different velocity values) are retained when dynamic changes are made. The dynamics cue changes the overall level of the whole instrument and therefore retains the difference between the note velocities.

2.4.5 Mode

The mode feature uses the Transposer plug-in in Logic to shift the music from its original mode to a new chosen mode and vice-versa. Two types of mode (major and harmonic minor) are utilised at the moment, and changes between the two options are made via a toggle on/off button.

2.4.6 Brightness

Brightness refers to the harmonic content of a sound; the more harmonics present in a sound, the brighter it is, while a sound with fewer harmonics will produce a duller sound (Cousineau et al., 2014). A Logic Channel EQ plug-in is utilised for a low-pass filter with a steep slope gradient of 48dB/Oct, and the Q factor was set to 0.43 to diminish frequency resonance. A slider in the interface controls the cut-off frequency value of the low-pass filter which affects the overall sound output. The low-pass filter has a cut-off frequency range of 305Hz to 20,000Hz; the slider in the Max/MSP patch increases/decreases the cut-off frequency depending on the amount of harmonic content desired in the sound output.

2.5 Portable Version of *EmoteControl*

In an effort to make the system portable and accessible for field work outside of the lab, a small-scale version of the *EmoteControl* system was configured. This small-scale version of *EmoteControl* also works with Max/MSP and Logic. However, the system does not make use of the extensive VSL library as its sound synthesizer. Instead, one of Logic's own sound synthesizers is used, the EXS24 sampler, making the system Logic-based, and thus, more lightweight than a full system with VSL. Substituting VSL with a Logic sound synthesizer makes the system easier to replicate on multiple devices and more cost-effective. Furthermore, it can also be configured on less powerful devices such as small laptops. To limit the hindrance on the ecological validity of the music, a good-quality grand piano was chosen as the default virtual instrument of the portable version of the *EmoteControl* system.

The portable version of *EmoteControl* has a similar user interface (Figure 2.3) to the original version and functions the same way, allowing for five cues to be altered: tempo, brightness, mode, pitch, and dynamics. As most of the cue alterations possible in the full version were controlled through VSL, some of the parameters had to be reconfigured. The tempo, brightness, and mode features are external to the sound synthesizer plug-in; hence, they are not affected by the change in sound synthesizer used. On the other hand, pitch and dynamics settings had to be modified. As articulation changes in the full version of the system were made in VSL, the articulation parameter was omitted from the portable version.

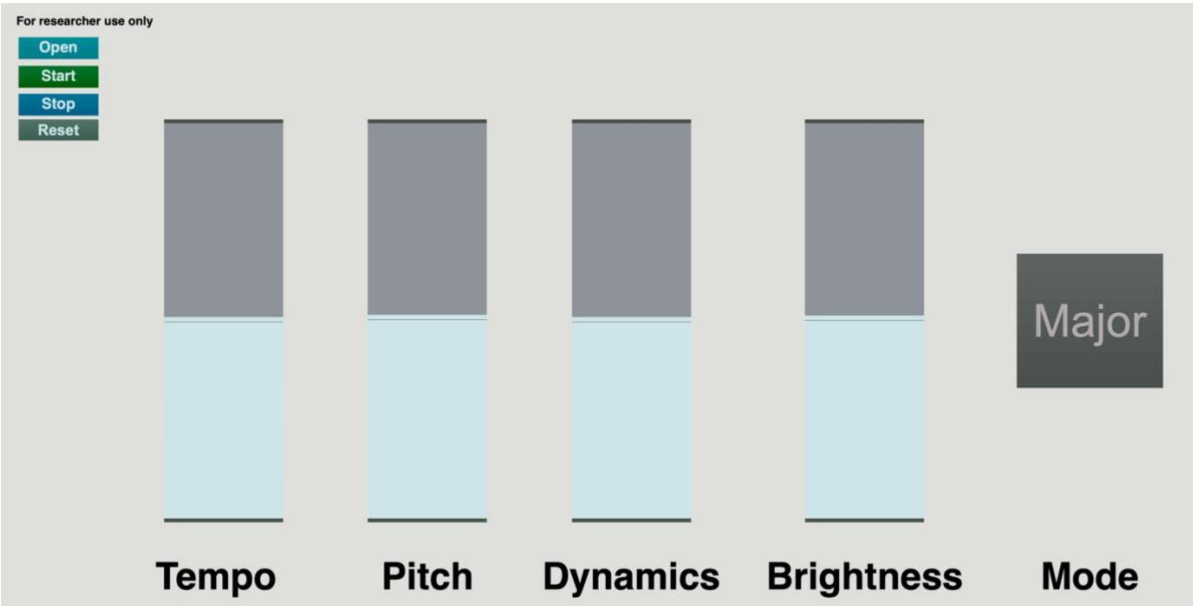


Figure 2.3. The user interface of the portable version of *EmoteControl*.

2.5.1 Pitch

The pitch feature in the portable version of the interface allows for incoming MIDI notes to be shifted up or down in semitone steps. Transposition of the music is made stepwise, and the feature allows for a pitch shift up and down 24 semitones from the mid-point of the slider used as the starting point, which totals to a range of 48 semitones.

2.5.2 Dynamics

The dynamics cue is controlled by changing the volume (in decibels) of the virtual instrument track in Logic.

2.6 Source Material Specifications

There is an amount of flexibility with regard to the MIDI music files which can be inputted in the *EmoteControl* system, although restrictions on the structure of the musical compositions also apply, as certain assumptions of the music are held. For the purpose of the research experiment that the system was originally designed for, new musical pieces were composed specifically for the experiment, taking into consideration the necessary assumptions needed for the music to be inputted into the system.

The main assumptions of the music are as follows:

1. Although the interface accepts both MIDI types 0 and 1, music composed should be targeted for one instrument, as the music output will be played using one virtual instrument; thus, multiple parts in the MIDI file will all receive the same manipulations.
2. Music notes should have note durations that allow for different articulation changes of the instrument to be possible (e.g., if a chamber strings synthesizer is utilised as a virtual instrument, then *legato*, *detaché*, and *staccato* articulations should be possible).
3. The pitch range of the music should be compatible with the chosen virtual instrument's register range. As the default virtual instrument of the interface is VSL's chamber strings, the pitch range of B0 to C7 has to be considered.

4. Music should allow for the mode changes (major/minor); therefore, a musical piece should be written in a specific key with no notes outside of the chosen scale, in order for the mode changes to be possible.

These assumptions should be taken into consideration and ideally followed to achieve reasonable and natural sounding variation in the cues across a musical piece.

2.7 Interface Limitations

As with any system, *EmoteControl* has a number of limitations. A notable constraint is that the interface is based on MIDI, and hence does not support other file formats such as audio files or direct inputs during live performance recordings. The required specifications of the pre-composed music may also be seen as a limitation, as the musical compositions to be inputted in the interface have to be purposely chosen or created prior to using the system. Also, the scope of the musical choices does put restrictions in terms of how many MIDI channels are processed and the non-separation of accompaniment and melodic instruments means that cue manipulations affect the whole piece, which may be seen as a constraint. Although the interface allows for the live manipulation of musical cues in real-time by the user, the music files have to be pre-made in order to adhere with the required specifications.

With regards to specific cue, the current pitch range limits the pitch shifts to ±2 semitones across all voices, thus only allowing for minute changes to the register, which may not be too perceptible. Furthermore, although some of the interface's musical cues may be common to most music around the world, making them universal, such as tempo, pitch, and timbre (Argstatter, 2016; Balkwill, Thompson,

& Matsunaga, 2004; Balkwill & Thompson, 1999; Fritz, 2013; Laukka et al., 2013), it is highly unlikely that *EmoteControl* will fully support non-Western music due to other musical parameters available for manipulation (such as the mode parameter and the exact tuning of the pitches) which vary across cultures. Also, the existing palette of cues may not be optimal for specific genres of Western music, but this interactive framework lends itself to easier prototyping of potential cue-emotion mappings in genres that have so far eluded music and emotion research (e.g., minimalism).

Although the *EmoteControl* system makes use of commercial software, the *EmoteControl* project is readily available and free to the public[7]. Furthermore, the system can be compiled as a standalone application via Max/MSP, which makes the system more accessible.

2.8 Interface User Experience Evaluations

As *EmoteControl* is aimed to be utilised by users who do not have any prior music knowledge and investigate their use of musical cues to convey different emotions for music emotion research, the interface was tested and evaluated by multiple user groups in different contexts, showcasing its use-cases ranging from a science fair to a data collection setting as well as a formal evaluation study utilising HCI methodology. In this section, the various user experiences of both versions of *EmoteControl* will be described, showing how the interface was utilised in various contexts by different users.

[7] github.com/annaliesemg/EmoteControl2019

2.8.1 Focus Group

First, the interface was presented to a group of seven music emotion researchers in order to gain feedback from researchers who would potentially utilise *EmoteControl* for their experiments, as well as identify any preliminary issues in the interface design prior to its completion. The group session was held at the Music and Science Lab at Durham University which houses the full version of *EmoteControl.* The researchers tried out the interface in a casual group session where they had to alter instrumental musical pieces utilising the cues provided and give feedback on potential adjustments and improvements on the interface design, feature labels, parameter ranges, and the system in general. This resulted in alterations such as the redesign of sliders, change in labelling terms, an increase in the tempo parameter range, and utilising a different plug-in to control the brightness feature. Participation in the focus group was on an unpaid voluntary basis.

2.8.2 *EmoteControl* as a Data Collection Tool

The revised full version of *EmoteControl* was then used as the main tool for data collection in a music emotion research experiment, where both musicians and non-musicians were studied to investigate how the two groups (musicians and non-musicians) altered musical cues to convey different emotions. The study was administered at the Music and Science Lab at Durham University. Full ethical consent was sought and approved prior to the study. Forty-two participants recruited from social media and university communications were given the musical task of manipulating musical pieces via the available cues in the *EmoteControl* interface, in order to convey a particular emotion specified by the researcher. The mean age of participants was 26.17 years (*SD* = 8.17) and 29 of the participants were female and 12 were male; one participant did not indicate their gender. Twenty-two of the participants were musicians whilst 20 were non-musicians as defined by the Ollen Music Sophistication Index (OMSI) (Ollen, 2006). The participants were first supplied with an information sheet on how the interface works, and a short demonstration was also presented by the researcher. Participants were offered a practice trial before they began the experiment in order to get them accustomed to the interface and to allow them to ask clarifying questions about the interface.

An open-ended question at the end of the experiment asked participants for any feedback on their experience utilising the interface during the experiment. Feedback revealed that the majority of participants thought the interface design satisfied the requirements for the musical task at hand, with 65.52% of comments mentioning liking the interface and the musical task; participants also commented that the interface was clearly labelled and had all the necessary instructions needed for users to utilise it. Six participants added they would have preferred for certain cues to have bigger ranges, such as the tempo range, when they were adjusting the tempo of a slow musical piece. Overall, participants felt that they had no issues with understanding the interface, and therefore quickly became accustomed to the interface during the practice trial. All participants were able to complete the task successfully. The whole experiment took approximately 30 minutes to complete. Participants were remunerated with chocolate.

2.8.3 Portable *EmoteControl* in an Educational Context

The portable version of *EmoteControl* was utilised as part of a presentation on music and emotions research at a local science fair known as the Schools' Science Festival, where students interacted with the interface in an educational setting. Nineteen groups consisting of approximately ten school children each, between the ages of 14 and 16 years, attended the University-approved demonstration at individual times. The students were first subjected to a short presentation on how different emotions can be communicated through music. They were then divided into teams of four to five students per team and given the musical task of randomly selecting an emotion term from a choice provided. The aim of the game was to utilise the cues available to alter the pre-composed music supplied and successfully convey the selected emotion term to the rest of the team. A point was given to the team if the other team members correctly guessed which emotion was trying to be conveyed. The game consisted of two or three rounds per team. The team with the most correct guesses in all the rounds would win. Prior to starting the musical task, the students were given instructions and a brief demonstration of how the interface works. The demonstration and musical task had a total duration of approximately 25 minutes. At the end of the activity, the students were asked if they had any comments about the interface, which led to informal anonymised feedback which was given as a collective. Although students did not necessarily have prior knowledge of the musical cues, the general consensus was that the students found the interface intuitive, and successfully understood what the musical cues do and how to operate it, and for the most part managed to convey the intended emotion correctly to their team members.

2.8.4 HCI Evaluation Study of *EmoteControl*

To properly assess and gather in-depth feedback on the design and usability of the interface as a data collection tool for music emotion research by the target users (non-experts), a formal evaluation study was carried out. The field of HCI provides methodologies and software to evaluate devices as well as the user's interaction with the device (Carroll, 2002). Wanderley and Orio (2002) drew on methodologies from the field of HCI to provide an evaluation framework for musical interfaces. They propose that the usability of an interface can be assessed by its learnability, explorability, feature controllability, and timing controllability, in a defined context. This can be attained by having target users carry out a simplistic musical task utilising the interface, and afterwards provide feedback. The current evaluation study followed their proposed evaluation method and aimed to assess the usability and controllability of the interface and to provide an overview of how well the interface could be utilised by a general audience.

2.8.4.1 Method

The study followed a task-based method, where participants were given the task of changing musical pieces via the possible features in order to make the music convey a specific emotion. Prior to the musical task, the participants were given 2-3 minutes to get accustomed to the interface. The musical task consisted of three trials featuring different musical pieces. For each piece, the participants altered the music via the six musical features available to make it convey one of the following three emotions: sadness, joy, or anger. The musical task had no time constraint. After the musical task, participants answered both close-ended and open-ended questions pertaining to their experience utilising the interface, usability of the interface, controllability, and range of the features, as well as questions regarding the aesthetic design of the interface. Close-ended questions consisted of 5-point Likert scales, ranging from ‘extremely well’ to ‘not well at all’, ‘a great deal’ to ‘none at all’, and so on, depending on the question posed. The study took approximately 15 minutes to complete. Participants were remunerated with chocolate. Full ethical consent was sought and approved prior to testing.

2.8.4.2 Participants

Twelve participants were recruited via social media and emails. Nine of the participants were female, and three were male. Participants ranged in age from 24 years to 62 years, with the mean age being 33.67 years (*SD* = 11.18). A one-question version of the Ollen Music Sophistication Index (OMSI) (Ollen, 2006; J. Zhang & Schubert, 2019) was utilised to distinguish between the level of musical expertise. The participants were an equal number of musicians and non-musicians.

2.9 Results

2.9.1 Summary of Quantitative Results

Most of the participants felt they were highly successful in conveying the desired emotion via the interface, with six participants selecting the maximum 'a great deal', five participants selecting the second highest 'a lot', and one participant selecting 'a moderate amount' on the rating scale. All participants indicated that they were satisfied with the usability of the interface to complete the musical task, with nine of the twelve participants selecting the maximum 'extremely satisfied', and the remaining three participants, all of whom being musicians, selecting the second highest 'somewhat satisfied' on the rating scale. Nine participants thought that the sliders and toggle button worked 'extremely well' to alter the features, and the remaining three thought they worked 'very well'.

All participants were satisfied with the ranges of the features available, with eight participants selecting the maximum 'extremely satisfied' and the remaining four participants selecting the second highest 'somewhat satisfied' on the rating scale. All participants agreed that the interface is extremely easy to get accustomed to, and the terms utilised as labels for the different features were clear, with nine participants selecting the highest 'extremely clear', and the remaining three participants, who were all non-musicians, selecting the second highest 'somewhat clear' on the rating scale. Figure 2.4 gives an overview of the participants' ratings for the aforementioned results.

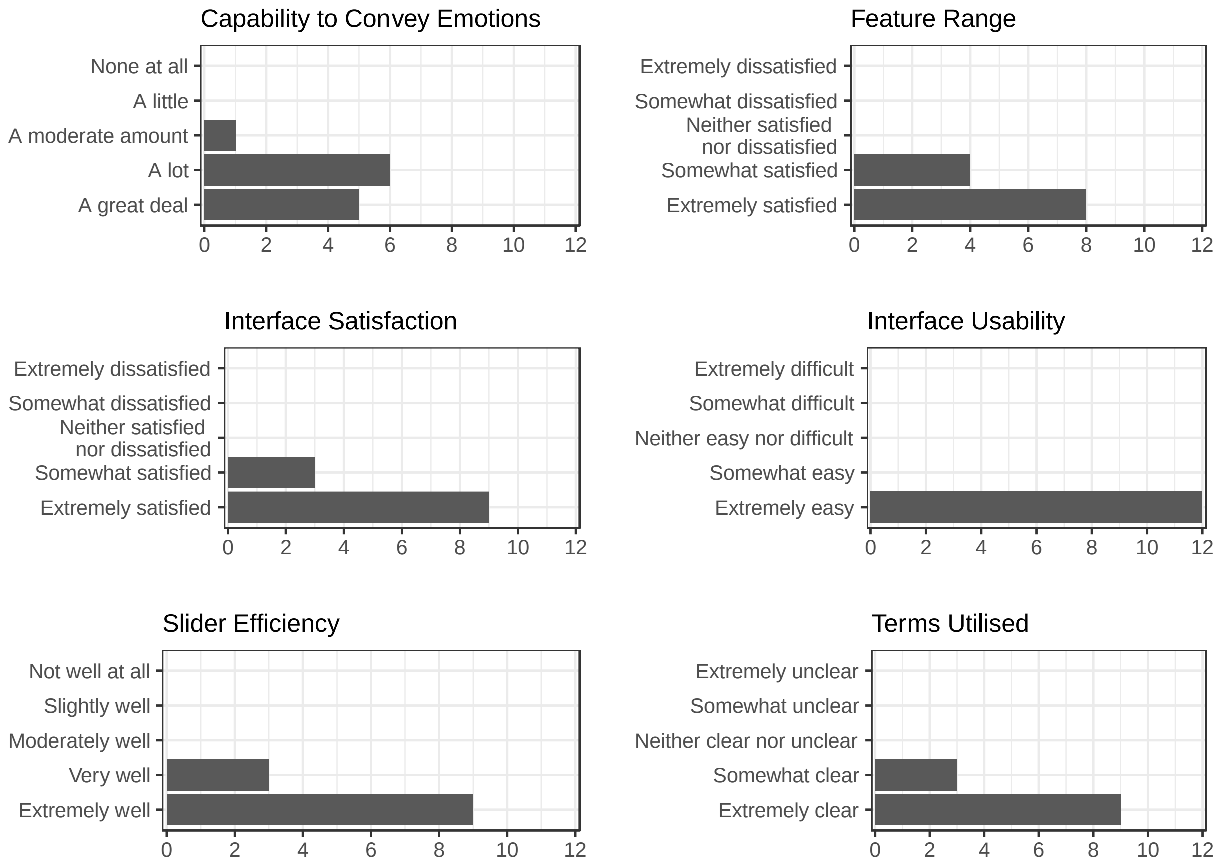


Figure 2.4. Participants' responses to the quantitative part of the evaluation study.

2.9.2 Summary of Qualitative Results

Participants were asked open-ended questions on what they liked about the interface, what they disliked, what changes they would make to the interface, and what additional features they would like to see in the interface.

Seven participants mentioned that they appreciated the usability of the interface; five participants directly stated how 'easy to use' the interface is, and two participants said that the interface is 'very intuitive'. Half of the participants stated that they liked the variety of musical cues available, as well as the clear design of the interface. Other comments mentioned they liked the good sound quality of the system, and the ability to instantly hear cue changes made to the music as it plays. When asked about what they did not like in the interface, half of the participants commented they had no dislikes, while two participants mentioned the continuous pitch-frequency shift emitted by the pitch manipulation until it stabilises on a new pitch. Individuals noted that the sliders have 'slightly bleak colours', the mode feature is the least intuitive feature for a non-musician, 'the range of the highest and the lowest can be more extreme', and that there should be clearer visual distinctions between each feature's range, potentially by having colour-changing sliders.

When asked about possible changes the participants would make to the interface, the two most prominent answers mentioned adding terms for the features with definitions to aid non-musicians (25% of comments) and making the interface more colourful (25% of comments). Two participants mentioned expanding the mode feature to include more than the two current levels (major/minor), such as adding an atonal option. Participants were also asked about the interface design, by providing them with mock-ups of the original layout and another possible interface layout featuring dials rather than sliders. The alternative design layout presented to the participants was chosen 16.7% of the time, suggesting that the original interface layout was the preferred design, being chosen 83.3% of the time.

The additional features suggested were the option of choosing different instruments (25% of comments), using colours for the musical cues (25% of comments), and adding a rhythm parameter (16.7% of comments). Four participants did not suggest an additional feature. All participants agreed that they would utilise the interface again (100% of participants selecting the maximum of 'extremely likely' on the rating scale) to complete a similar musical task.

Participants were also asked if they thought the interface could be adapted for non-research contexts. All participants agreed that the interface could be utilised outside of a research context, with the two main situations being an educational game for children (mentioned in 58.3% of comments) which would help 'children to get enthusiastic about music', and music therapy, in particular when working with special needs or non-verbal individuals (being mentioned 41.7% of the time). Participants suggested that the interface can be adapted for non-research contexts by installing it on mobile phones and tablets, making it more colourful, and providing definitions for the musical cues and a user guide.

This study was set up to evaluate the interface on its design and usability. The participants' responses indicated that the interface allowed them to successfully carry out the musical task and that they would utilise the interface again for a similar task. This affirms that the aim of *EmoteControl*, which is allowing users to change features of the music to convey different emotions without the need of any musical skills, is possible to achieve. Therefore, the *EmoteControl* interface has potential to be utilised for musical cues and emotion research, which was the primary aim of the project. Furthermore, this study got valuable information from users on what could be the next feature in the interface, such as the possibility of changing the instruments playing, or adding colour to the sliders for a better aesthetic. An interesting outcome was the suggestions for potential adaptation of the interface, where participants strongly agreed on *Emote Control* 's ability to be utilised in an education setting, as well as an aid for therapeutic methods.

2.10 Discussion and Conclusions

This chapter presented a new interactive system for real-time manipulation of both structural and expressive cues in music. Most importantly, the *EmoteControl* system allows for a direct user experience approach, where users are able to directly interact with the music and change its properties, having potential uses as a tool in musical cues and emotion research, in music composition, and also performance. Furthermore, *EmoteControl* is designed to be utilised by a general population who do not need significant music knowledge to be able to interact with the interface. The interface gives researchers the opportunity to investigate study the topic in a cost effective way. The fact the current cues available in the interface may be updated, changed, and/or new cues may be added, makes *EmoteControl* flexible and potentially extends its longevity. Furthermore, the interface allows researchers to investigate how individuals would convey different emotions in music, if given the chance, without requiring explicit competence about the cues. Thus, possible research may involve investigating how musicians create emotional expression in music as against non-musicians. Other potential interesting research questions would be exploring whether cues are used differently cross-culturally, and also looking at how individuals would convey emotional expression in different genres of music, both within the Western context (e.g., pop, rock, jazz music) as well as non-Western music (e.g., Hindustani music). To increase accessibility of the interface, a portable version of *EmoteControl* has also been designed to be utilised outside of lab settings.

Similar to other systems, certain specifications have to be considered for the system to work at its optimal condition. In *EmoteControl* 's case, the MIDI files to be utilised in the interface have to be compatible with the requirements of the system, mainly concerning the range and the duration of the pitches. However, in the unlikely circumstance that this is not possible, certain settings of the interface can be re-mapped to accommodate the MIDI file used as input. Potential changes to the interface include utilising a different virtual instrument as output, rather than the default chamber strings virtual instrument. Any virtual instrument from the sound library used in Logic can be selected as the sound output. However, most of the cues available in the *EmoteControl* interface are dependent on the selected virtual instrument; thus, a change in virtual instrument will also affect which cues will be available for manipulation. Therefore, although the aforementioned alterations are possible, the ideal approach would be to utilise music that complements the system, rather than the other way around.

Possible interface changes would be increasing or decreasing the number of cue levels, such as having more articulation methods to choose from (e.g., adding a *pizzicato* option as well, if utilising a string virtual instrument). The range of parameters, such as of tempo and brightness, can be amended as needed. The mode feature can be altered by choosing different mode scales or expanded by adding multiple mode scales to choose from. For instance, the portable *EmoteControl* system has been used for cross-cultural research in Pakistan, where three different scales were being investigated: major, minor, and whole-tone scales. To make this possible, the mode parameter was adapted to incorporate all three aforementioned scales. Another possible alteration would be to add other cues into the system, which would allow for non-Western music to be more compatible with the system, such as rhythm (e.g., event density, expressive microtiming) or timbre (e.g., the shape of the envelope, spectral distribution, or onset type) (Balkwill & Thompson, 1999; Midya et al., 2019), allowing users to vary the tuning system used, and potentially approaching cues such as mode in an adaptive fashion to reflect the organisation of scales in non-Western cultures. It might also be feasible to allow users to deselect the cues that they feel are irrelevant for the expression of emotions.

Following participants' feedback from the evaluation study, potential changes to the interface design could be the use of different terms as parameter labels to provide a clearer definition of what the parameter does. Given the response of participants about the mode parameter being the least intuitive feature for non-musicians, an example of this could be changing the 'mode' label term to something simpler such as 'major/minor swap' or 'change the pitch alphabet in the music'. Introducing more colour to the interface could also be a possibility, especially as it was frequently mentioned from participants during the formal evaluation study.

In this chapter, *EmoteControl* has been utilised in different contexts, which include utilising the interface as an interactive game for students in an educational setting and a data collection tool for music emotion research. However, the system has other potential uses. Following participants' responses from the formal evaluation study as well as the uses of similar systems (Wallis et al., 2011), *EmoteControl* could be utilised as a therapeutic tool, for instance, as an emotional, expressive tool for non-verbal patients (Silverman, 2008). As *EmoteControl* is targeted for non-experts and a general population, it could also be utilised to investigate the recognition of emotions in a developmental context. Saarikallio et al. (2019) report how children who were either three years old or five years old utilised three musical cues (tempo, loudness, and pitch) to convey three emotions (happy, sad, and angry) in music, but large individual variation exists at these age stages. *EmoteControl* could be utilised in a similar context, with the opportunity to explore more musical cues which are already programmed in the system. Finally, it could be possible to implement another feature that integrates user feedback about the appropriate cues utilised by others to convey different emotions back into the system. By presenting *EmoteControl* as an online multiplayer game, users have to try and guess each other's cue usage and cue values to convey the same emotion, with the gathered cue information fed back into the system. This method is known as a game with a purpose (von Ahn, 2006), which might aid in exploring how users utilise the large cue space linked to emotional expression. Utilising this approach could have potential due to its suitability for non-experts and the simplicity of the task.

Future research could potentially focus on making the system more accessible, such as reprogramming it as a web-based application or making it available for tablets and mobile devices (Fabiani, Dubus, & Bresin, 2011). Adding more parameters to the interface, such as the ability to switch instruments while the music plays, or changing the rhythm of the music, would be interesting ventures and allow for more cue combinations to be explored. Another possible amendment would be the ability to input both audio files (Bresin, Friberg, & Dahl, 2001) and MIDI files into the interface, potentially broadening the system's use-cases. Utilising audio files would allow for more flexibility and a greater amount of source material to be available, as it would be possible to use commercial music and other ready-made audio files rather than having to make the MIDI source files specifically. Furthermore, an audio format would allow for a more detailed analysis of how the cues were used and manipulated to convey specific emotions utilising audio feature extraction tools such as the *MIRToolbox* (Lartillot, Toiviainen, & Eerola, 2008) and *CUEX* (Friberg, Schoonderwaldt, & Juslin, 2007). Subsequently, the musical cues available for manipulation would have to be re-evaluated as the data in audio format is represented differently than MIDI files (Cataltepe, Yaslan, & Sonmez, 2007; Friberg, 2004). Cues such as articulation apply changes to individual notes in the music, which would not be possible in audio file formats, as they work with acoustic signals rather than music notation. Therefore, a system that allows for both MIDI and audio files would have to differentiate between the cues that can be utilised depending on the input file format (Oliveira & Cardoso, 2007).

EmoteControl provides a new tool for direct user interaction that can be utilised as a vehicle for continuous musical changes in different contexts, such as in music performances (Soydan, 2018), live interactive installations (Rinman et al., 2003; van 't Klooster & Collins, 2014), computer games (Livingstone & Brown, 2005), as well as a tool in a therapeutic context. Most importantly, *EmoteControl* can be used as a medium through which music psychology research on cues and emotion can be undertaken, with the capability of adapting it to suit different experiment requirements, allowing for flexibility and further development in the understanding of the communication of perceived emotions through the structural and expressive cues of a musical composition.

2.11 References

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.

Bouwer, A., Holland, S., & Dalgleish, M. (2013). Song walker harmony space:

embodied interaction design for complex musical skills. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 207–221). Springer.

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

Bresin, R., Friberg, A., & Dahl, S. (2001). Toward a new model for sound control. *COST G-6 Conference on Digital Audio Effects*, 1–5.

Cancino Chacón, C. E. (2018). *Computational Modeling of Expressive Music Performance with Linear and Non-linear Basis Function Models*. Johannes Kepler University, Linz.

Carroll, J. M. (2002). Introduction: Human-Computer Interaction, the Past and the Present. In J. M. Carroll (Ed.), *Human-Computer Interaction in the New Millenium* (pp. xxvii–xxxvii). ACM Press and Addison Wesley.

Cataltepe, Z., Yaslan, Y., & Sonmez, A. (2007). Music genre classification using MIDI and audio features. *Eurasip Journal on Advances in Signal Processing*, *2007*. https://doi.org/10.1155/2007/36409

Collins, N. (2010). *Introduction to computer music*. John Wiley & Sons.

Cousineau, M., Carcagno, S., Demany, L., & Pressnitzer, D. (2014). What is a melody? On the relationship between pitch and brightness of timbre. In *Frontiers in Systems Neuroscience* (Vol. 7, p. 127).

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., 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

Fabiani, M., Dubus, G., & Bresin, R. (2011). MoodifierLive: Interactive and collaborative expressive music performance on mobile devices. *Proceedings of NIME '11*, 116–119.

Friberg, A. (2004). A fuzzy analyzer of emotional expression in music performance and body motion. In J. Sundberg & B. Brunson (Eds.), *Proceedings of Music and Music Science Stockholm*.

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., 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

Friberg, A., Schoonderwaldt, E., & Juslin, P. N. (2007). CUEX: An algorithm for automatic extraction of expressive tone parameters in music performance from acoustic signals. *Acta Acustica United with Acustica*, *93*(3), 411–420.

Fritz, T. (2013). The Dock-in Model of Music Culture and Cross-cultural Perception. *Music Perception: An Interdisciplinary Journal*, *30*(5), 511–516.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae VO - Special Issue 2001-2002*, 123–147. https://doi.org/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/oxfordhb/9780199298457.013.0013

Gifford, T. (2013). Appropriate and Complementary Rhythmic Improvisation in an Interactive Music System. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 271–286). Springer.

Grachten, M., & Widmer, G. (2012). Linear Basis Models for Prediction and Analysis of Musical Expression. *Journal of New Music Research*, *41*(4), 311–322. https://doi.org/10.1080/09298215.2012.731071

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. (1937). The Affective Value of Pitch and Tempo in Music. *The American Journal of Psychology*, *49*(4), 621–630.

Hoeberechts, M., & Shantz, J. (2009). Realtime Emotional Adaptation in Automated Composition. *Audio Mostly*, 1–8.

Holland, S., Wilkie, K., Bouwer, A., Dalgleish, M., & Mulholland, P. (2011). Whole Body Interaction in Abstract Domains. In D. England (Ed.), *Whole Body Interaction. Human-Computer Interaction Series.* (pp. 19–34). Springer. https://doi.org/https://doi.org/10.1007/978-0-85729-433-3_3

Holland, S., Wilkie, K., Mulholland, P., & Seago, A. (2014). Open Research Online. *Choice Reviews Online*, *51*(06), 51-2973-51–2973.

Huron, D., Anderson, N., & Shanahan, D. (2014). "You Can't Play a Sad Song on the Banjo:" Acoustic Factors in the Judgment of Instrument Capacity to Convey Sadness. *Empirical Musicology Review*, *9*(1), 29. https://doi.org/10.18061/emr.v9i1.4085

Juslin, P. N. (1997). 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. (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., & 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

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

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

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab Toolbox for Music Information Retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data Analysis, Machine Learning and Applications* (pp. 261–268). Springer Berlin Heidelberg.

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*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Legaspi, R., Hashimoto, Y., Moriyama, K., Kurihara, S., & Numao, M. (2007). Music compositional intelligence with an affective flavor. *International Conference on Intelligent User Interfaces, Proceedings IUI*, 216–224. https://doi.org/10.1145/1216295.1216335

Livingstone, S. R., & Brown, A. R. (2005). Dynamic Response: Real-Time Adaptation for Music Emotion. *2nd Australasian Conference on Interactive Entertainment*, 105–111.

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

Lopez, A. R., Oliveira, A. P., & Cardoso, A. (2010). Real-time emotion-driven music engine. *Proceedings of the International Conference on Computational Creativity, ICCC-10*, 150–154.

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/https://doi.org/10.1371/journal.pone.0222380

Morreale, F., & De Angeli, A. (2016). Collaborating with an Autonomous Agent to Generate Affective Music. *Comput. Entertain.*, *14*(3). https://doi.org/10.1145/2967508

Morreale, F., Masu, R., & Angeli, A. De. (2013). Robin: An Algorithmic Composer For Interactive Scenarios. *Sound and Music Computing Conference 2013, SMC 2013*, 207–212.

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*. University of Jyväskylä, Department of Music.

Oliveira, A., & Cardoso, A. (2007). *Control of Affective Content in Music Production*.

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.

Poepel, C. (2005). On Interface Expressivity: A Player-Based Study. *Proceedings of* *the International Conference on New Interfaces for Musical Expression*, 228–231.

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. *Proceedings of 18th International Florida Artificial Intelligence Research Society Conference*, 86–91.

Rinman, M.-L., Friberg, A., Kjellmo, I., Camurri, A., Cirotteau, D., Dahl, S., Mazzarino, B., Bendiksen, B., & McCarthy, H. (2003). EPS - an interactive collaborative game using non-verbal communication. *Proceedings of SMAC 03, of the Stockholm Music Acoustics Conference*, *2*(Smac 03), 561–563.

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. K., & 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

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

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

Soydan, A. S. (2018). *Sonic Matter: A Real-time Interactive Music Performance System*.

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, *10*(1), 79–89.

van 't Klooster, A., & Collins, N. (2014). In A State: Live Emotion Detection and Visualisation for Music Performance. *Proceedings of the International Conference on New Interfaces for Musical Expression*, 545–548.

von Ahn, L. (2006). Games with a Purpose. *Computer*, *39*(6), 92–94.

Wallis, I., Ingalls, T., Campana, E., & Goodman, J. (2011). A Rule-Based Generative Music System Controlled by Desired Valence and Arousal. *Proceedings of the 8th Sound and Music Computing Conference, SMC 2011*.

Wanderley, M., & Orio, N. (2002). Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*, *26*(3), 62–76. https://doi.org/10.1162/014892602320582981

Wong, E. L., Yuen, W. Y. F., & Choy, C. S. T. (2008). Designing Wii controller: a powerful musical instrument in an interactive music performance system. *MoMM*. https://doi.org/10.1145/1497185.1497205

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

Chapter 3. A Set of New Musical Pieces for Emotion Perception in Music Research

Preface

Chapter 3 presents a new set of 28 musical excerpts specifically composed to be used as stimuli in the current project and more broadly, for music and emotion research. The musical pieces were each composed to portray one of a possible nine emotional expressions which have been reported as being expressed through music: sadness, joy, calmness, anger, fear, power, surprise, love, and longing. The novel excerpts were created with the aim of representing 'real music'. Thus, the pieces featured polyphonic music and complete musical ideas. Furthermore, the pieces were recorded as performances on a piano, to encapsulate a proper musical experience. The pieces were created to complement the *EmoteControl* interface, which means that they can withstand real-time cue manipulations without heavily compromising their ecological validity. For convenience, the excerpts are represented as traditional notation in this chapter, but the original audio files are also accessible via the OSF repository. Furthermore, this chapter also provides a self-reflexive account on the compositional task by me, as the composer of the pieces, and fundamentally, the first individual to assess the pieces' perceived emotional content, where I consider the challenges of conveying certain emotions and distinguishing between similar emotional expressions in the compositional task.

Abstract

Researchers have utilised numerous types of musical excerpts as stimuli in musical cues and perceived emotion research. Most excerpts selected were from pre-existing commercial music, with pieces by renowned composers from the

Classical musical genre being generally favoured. However, using pre-existing commercial music might restrict the researcher's experimental control on the music, and brings about the possibility of involuntary familiarity bias from the participants. To circumvent these potential limitations, a small number of researchers specifically composed their own music. However, most stimuli created mainly represented sadness and happiness emotions, and comprised simple, monophonic sequences which do not emulate real musical contexts. Moreover, it is unclear whether the stimuli could withstand cue manipulations. To this end, 28 new polyphonic musical pieces were created, each representing one of nine emotions (sadness, joy, calmness, anger, fear, power, surprise, love, and longing). In this chapter, the compositional process behind the 28 novel pieces is described, detailing both general properties of the excerpts and specific characteristics of pieces based on the intended target emotion. The newly composed pieces were recorded as performances played on a piano rather than playing the score using MIDI playback, with the aim of emulating a real musical context and preserving the ecological integrity of the music. A first-hand account of the composer's experience creating the pieces is reported, also noting which emotions were challenging to depict, such as surprise, and which emotions were more straightforward to portray, such as sadness and joy.

3.1 Introduction

The communication of different emotional expressions through music has been investigated in numerous ways, using several research approaches, frameworks of emotion, methods, and explorations of different musical cues (for a review, see Chapter 1). In addition, different types of music material varying in such aspects as style of music, duration, and familiarity (Warrenburg, 2020a), have been used in investigations as to whether musical cues affect the perceived emotional expression in music.

3.1.1 Use of Pre-Existing Commercial Music

Early studies used musical excerpts from the Classical music repertoire as stimuli, played to participants either by live musicians (Behne, 1972; Downey, 1897; Gilman, 1891; Rigg, 1940b) or from recordings of the musical works (Bigand et al., 2005; Gagnon & Peretz, 2003; Ilie & Thompson, 2006; Juslin, 1997a). Hevner utilised relatively short, complete musical phrases from classical works by Schumann, Rameau, Arensky, Bach, Beethoven, and others, performed live by a pianist to investigate the effect of mode (1935), and tempo and pitch (1937) on the pieces' emotional content. Bigand et al. (2005) used 27 musical excerpts taken from real performances from baroque, classical, romantic, and modern periods of Western classical music in an investigation of whether musical expertise affected emotional response. Parke, Chew, and Kyriakakis (2007) similarly selected excerpts from CD recordings of works from classical music to accompany short film clips in a study on the impact of music on the emotion perception of film, including pieces by Mahler, Debussy, and Grieg, to name a few, while Kawakami et al. (2013) sampled around 30 seconds from piano works by Glinka, Blumenfeld, and Granado when investigating how emotional content is perceived by listeners in music regarded as sad.

Later, with the aid of technological progress, researchers also started creating and using synthesised musical excerpts as stimuli. Scherer and Oshinsky (1977) manipulated synthesised sawtooth tone sequences and a Beethoven melody using a MOOG synthesiser to investigate the effect of amplitude, pitch variation, pitch level and contour, tempo, envelope, and filtration on the emotional expression of the sequences. Kamenetsky, Hill, and Trehub (1997) altered MIDI versions of 30-second excerpts from works by Bach, Chopin, and Liszt to examine the effect of tempo and dynamics on the perceived emotional expression. Peretz, Gagnon, and Bouchard (1998) utilised computer-generated versions of sequences from piano works such as Chopin's Nocturne Opus 48 no. 1 and Mozart's Piano Concerto no. 23, as well as orchestral works including Beethoven's Symphony no. 3 and Mahler's Symphony no. 5. Although utilising synthetic music material helped the researchers have more control over the musical cues and their manipulations, the technology at the time was limited with regards to synthesizing stimuli that sounded like 'real' music. Therefore, the stimuli used may not have been the best representatives of ecologically valid music material due to the artificial synthesised sounds available at the time. Ilie and Thompson (2006) tried to find a way around the use of artificial sounds by sampling from CD recordings. Complete musical phrases from works by Vivaldi, Handel, Mozart, Stradella, Haydn, and Alberti were captured from CD recordings and then edited using computer software to create different versions varying in loudness, tempo, and pitch height. However, this method limits the number of cue manipulations possible as superimposing cue changes on existing audio files can easily make the audio files lose their ecological properties. In fact, only two levels of loudness, tempo, and pitch were used for each musical excerpt. The cue changes were limited to mild variations such as a 26% faster tempo and a 21% slower tempo than the original, to try and preserve a natural sound.

In early studies up until the late 2000s, Western classical music remained the most prominent choice of stimuli used, with other genres such as popular music, jazz, and film music utilised notably less (Eerola & Vuoskoski, 2013; Juslin & Laukka, 2003; Warrenburg, 2020a). Later, the 2010s saw a shift in genres utilised, with popular music increasing in use as stimuli choice (Warrenburg, 2020a). Juslin (1997a) used instrumental melodies from popular songs including *Nobody Knows* and *When the Saints* as stimuli to be performed by multiple musicians in a study assessing how emotions are communicated through expressive cues. Lindström (2006) generated 72 variations of the well-known *Frère Jacques* tune in an investigation on melodic organisation. Saarikallio et al. (2019; 2014) also utilised instrumental versions of popular melodies taken from children's songs such as *Mary had a little lamb* and *Chim Chim Cher-ee* (from the film *Mary Poppins*) as stimuli for studies with children and adolescents on emotional expression communication through musical cues. Burger et al. (2013) selected excerpts from popular music which included *Bla bla bla* by Gigi D'Agostino and *Closer* by Kings of Leon in an investigation on the relationship between emotions perceived in music and music-induced movement. Omar et al. (2011) utilised a mix of short, 11-second excerpts from Western classical music and film scores to investigate emotion recognition in patients with frontotemporal lobar degeneration (FTLD). Eerola and Vuoskoski (2011) curated a set of 110 excerpts from film music and collected ratings on their expressed emotion as well as ratings on valence-arousal dimensions, with the aim of producing a validated set of non-synthetic stimuli for music and emotion research, which subsequently, was frequently used by other researchers in later studies (Punkanen, Eerola, & Erkkilä, 2011; Saarikallio, Vuoskoski, & Luck, 2014; Vuoskoski & Eerola, 2017).

3.1.1.1 Limitations of Using Pre-Existing Commercial Music as Stimuli

Across decades and styles of music, most music materials used as stimuli over the last century were from pre-existing commercial recordings, which whilst preserving ecological validity, brings about certain limitations (Eerola & Vuoskoski, 2013; Warrenburg, 2020a). Firstly, altering musical properties of already existing commercial recordings may result in the musical material becoming artificially sounding, as the original material would not have been created for manipulation purposes. Therefore, using music samples from commercial recordings might restrict the amount of experimental control held over the stimuli if ecological validity of the music material is to be retained (Bigand et al., 2005; Gabrielsson & Lindström, 2010; Morreale et al., 2013). Secondly, the use of commercial recordings and pre-existing music grants the possibility that participants may have already been familiar with the material due to prior uncontrolled exposure such as in commercials, television shows, and radio shows, for example. Therefore, participants may have already formed opinions that spark an involuntary bias about the music, influenced by their previous experience listening to it (Eerola & Vuoskoski, 2013; Juslin & Västfjäll, 2008).

3.1.2 Specifically-Composed Musical Stimuli

To avoid the potential hindrances brought about by the use of pre-existing commercial recordings, a minority of studies in music and emotion research had music material specifically composed to be used as stimuli. A recent review of 306 studies on music and emotion by Warrenburg (2020a) found that only 1% of musical stimuli used over the last 90 years were specifically created for research purposes. Specifically created musical excerpts varied from simple, monophonic melodies (Dolgin & Adelson, 1990; Gagnon & Peretz, 2003; Hailstone et al., 2009) to melodies with chords as accompaniments (Kleinsmith, Friedman, & Neill, 2016). Other studies had musical excerpts composed in a specific style such as film music (Gosselin et al., 2005; Vieillard et al., 2008) or popular music (Parke, Chew, & Kyriakakis, 2007), with the aim of presenting ecologically valid music.

Gagnon and Peretz (2003) composed short, monophonic, eight tone sequences in C major, using different note patterns. They also created two other variants of the sequences (a total of 24 sequences), differing in mode (minor and whole tone) with the aim of systematically investigating the effect of mode and tempo on the perception of happiness and sadness in the music. Ramos, Bueno, and Bigand (2011) also created novel musical pieces to investigate the effect of tempo and mode on the perception of emotional expression in music. They composed three 36-second pieces in Ionian mode played with a piano timbre. The mode and tempo of the pieces were systematically varied on different levels. Three levels of tempo (slow, moderate, and fast) were investigated. The original three pieces in Ionian mode were also transposed to Dorian, Phrygian, Lydian, Mixolydian, Aeolian, and Locrian modes. The different tempo and mode combinations of pieces resulted in a total of 63 variations of the musical pieces which were then assessed by participants on their emotional content, whether the pieces were conveying happiness, sadness, serenity, or fear/anger.

Dolgin and Adelson (1990) composed monophonic melodies intending to convey happiness, sadness, anger, and fear for a study investigating children's emotion recognition abilities. Four, 15-20 seconds long melodies were composed for each emotional expression, amounting to 16 melodies in total, which were presented to participants in two different conditions; performed on a viola as well as sung by a soprano. Hailstone et al. (2009) also investigated how timbre affected the perception of emotions in music. They created 20 single-line, 8-bar melodies portraying each of the following four basic emotions: happiness, sadness, anger, and fear, amounting to 80 melodies in total. All melodies were subjected to a pilot study to assess how successful they were in conveying their intended emotion. The ten best representatives of each emotion which ranged between 9-31 seconds long, were carried forward and recorded with four different timbres – piano, trumpet, violin, and electronic synthesizer, and then assessed by participants with regards to their perceived emotional expression. Kleinsmith, Friedman, and Neill (2016) composed their own stimuli to investigate whether variations in cadences would affect the participants' judgment of the pieces. They composed eight stem melodies in C major featuring solely crotchet notes within a pitch range of one octave in 4/4 time. They combined the eight melodies with eight final chords in either a conventional sequence ending with a plagal cadence or an unresolved cadence. In total, they produced 64 unique stimuli which were presented to participants with a synthesised grand piano timbre. Gosselin et al. (2005) also composed polyphonic excerpts consisting of a melody with an accompaniment generated with a synthesised piano timbre. A set of fourteen excerpts were composed to convey each of the following emotions: happiness, sadness, fear, and peacefulness, totalling 56 musical pieces. The excerpts were composed in the style of film music utilising the Western tonal system. These pieces were used to investigate how patients with amygdala resection identify emotional expression in music. The music material was later validated by Vieillard et al. (2008) across three experiments to investigate whether the stimuli were good representatives of their intended emotion and also collect ratings of the stimuli on valence-arousal dimensions. Sutcliffe et al. (2017) also composed and validated musical excerpts for their research on whether age affects emotion recognition in music. They composed multiple music clips to communicate happiness, peacefulness, sadness, anger, or fear, with tempo and mode being the two cues predominantly manipulated to convey the different emotional expressions. Musical clips also varied with regards to timbre. Thirty pieces were recorded using an acoustic guitar and another 30 pieces were recorded with an electric guitar. The other musical pieces were generated with different keyboard timbres: an electric keyboard, a grand piano sound, an electric piano, and an organ sound.

Other studies had musical excerpts specifically created to accompany visual stimuli. Parke, Chew, and Kyriakakis (2007) devised a study investigating whether adding music to film clips affects the emotional expression perceived by the participants. The researchers paired a set of film clips with four familiar classical musical pieces conveying content, depressed, exuberant, and anxious emotions. They also had a composer create 20 instrumental popular-style pieces to accompany the film clips. The composer was informed that the excerpts should convey a specific emotion, however, it was up to the composer to decide on the emotional expression to be communicated through the music. The newly-composed pieces featured emotions such as sadness, anxiety, courage, and dreaminess. Participants were then presented with the film clips in three different conditions: visual-only (film clip), audio-only (musical excerpt), and audio-visual (film clip with musical excerpt) conditions, which they rated on valence, activation, and dominance dimensions. Järvinen et al. (2010) created 5-second long music segments intending to convey fear, happiness, and sadness to accompany visual images in an investigation of emotional reactivity of individuals with Williams Syndrome (WS). The same musical segments were later used in a study on the sensitivity of the autonomic nervous system to auditory affective stimuli in individuals with WS (Järvinen et al., 2012).

A small number of researchers recruited music composers as participants and instructed them to compose melodies with the intention of portraying specific emotional expressions. Thompson and Robitaille (1992) tasked five highly trained musicians to create short, monophonic melodies conveying joy, sorrow, excitement, dullness, anger, and peace. These melodies were later assessed by other musicians with regards to their emotional content, with the results suggesting that overall, the composers were successful in communicating all intended specific emotional expressions through their melodies, except for peace. Quinto, Thompson, and Taylor (2014) investigated the role of compositional features and performance expression to the conveying of emotions in music by asking four violinists and four vocalists to create brief, monophonic melodies containing a maximum of nine notes intending to convey each of the following emotions: anger, fear, happiness, sadness, and tenderness, as well as a neutral emotion portrayal condition. The musicians were also asked to perform brief, emotionally ambiguous musical phrases composed by one of the researchers, to investigate how musicians also utilise expressive cues to portray the same six emotion states. A third condition consisted of the musicians performing their own melodies to further emphasise the intended emotion. The performances of all stimuli across the three conditions were then assessed on their emotional content by undergraduate students. Behrens and Green (1993) had two violinists, two vocalists, two trumpet players, and two timpanists perform 30-second improvisations to convey sadness, anger, or fear. The performers' improvisational efforts were rated by undergraduate students on 4-point Likert emotion scales, with findings supporting the notion that individuals can correctly identify emotional expressions in short musical pieces. Paquette, Peretz, and Belin (2013) ran a similar study where 20 musicians (10 violinists and 10 clarinettists) performed 10 short improvisations between one to four seconds in duration, to portray happiness, sadness, fear, and neutrality. These improvisations, dubbed as 'musical bursts' were evaluated by listeners in a rating task to assess whether the short stimuli were successfully conveying their intended emotion, with an average recognition score of 80.4% from listeners.

3.1.2.1 Limitations of Existing Specifically-Composed Stimuli

Although the aforementioned studies specifically composed musical stimuli and thus, ensured there was no possibility of familiarity bias present, most studies utilised simple, monophonic melodies (Dolgin & Adelson, 1990; Gagnon & Peretz, 2003; Hailstone et al., 2009) which might not be the clearest representation of a 'real' musical context. Furthermore, although the studies created unique musical excerpts, this does not necessarily imply that the stimuli, in particular the polyphonic ones, were composed in a way that they could withstand cue manipulations. Gagnon and Peretz (2003), and Ramos, Bueno, and Bigand (2011) systematically altered three levels of tempo and three to seven levels of mode to create different musical variations, while Vieillard et al. (2008) mentioned that their 56 validated musical excerpts could potentially be altered with respect to certain cues such as tempo, however, they did not attempt it. Some studies (Parke, Chew, & Kyriakakis, 2007; Sutcliffe et al., 2017) did not share the stimuli or any details pertaining to how the musical structure of the stimuli was composed, which makes it difficult to determine the cues used. Therefore, it is unclear whether the stimuli sets mentioned could support further manipulations subjected to them such as in systematic cue manipulation studies (Eerola, Friberg, & Bresin, 2013; Gabrielsson & Lindström, 2010; Hevner, 1935) or in analysis-by-synthesis approaches where musical stimuli are manipulated in real-time (Bresin & Friberg, 2011; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Sievers et al., 2013), without compromising their ecological validity. In addition, the musical excerpts specifically composed in previous studies represent a limited number of emotions (Behrens & Green, 1993; Dolgin & Adelson, 1990; Hailstone et al., 2009; Paquette, Peretz, & Belin, 2013), mainly sadness and happiness, followed by fear and anger, which is a fraction of the possible emotions reported as being expressed in music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003).

In order to address the highlighted limitations of existing stimuli sets, I composed a new collection of polyphonic musical pieces to be used as stimuli in the current project. This ensures that there is no familiarity bias with the use of the musical excerpts in the current project. The new musical creations were produced with prior knowledge that they were going to undergo a validation test with regards to their emotional content (reported in **Chapter 4**) and also be utilised as stimuli in the interactive interface *EmoteControl* (detailed in **Chapters 2, 4,** and **5**) which enables the user to alter the stimuli in real-time via a selection of available musical cues (Micallef Grimaud & Eerola, 2021). Therefore, the balance between experimental control and ecological validity required musical creations to have a flexible structure to support the cue manipulations subjected to them in *EmoteControl* whilst not compromising their ecological validity. Furthermore, any source material requirements specified by the interface could also be addressed in the creation process to ensure maximum compatibility. Lastly, another aim of creating new musical pieces was to broaden the repertoire of emotional expressions represented in the existing stimuli sets, composing musical pieces that convey emotions other than those most commonly found in prior literature: sadness, happiness, and anger (Warrenburg, 2020a).

To this end, I composed 28 musical excerpts, each portraying one of nine emotional expressions: happiness, sadness, calmness, anger, fear, power, surprise, longing, or love. The rationale behind the chosen emotional expressions and the compositional process of the musical pieces are detailed in the next section. Finally, I will discuss the novel contribution of the new musical excerpts to the field together with a self-reflection on the compositional task.

3.2 Compositional Task

I composed novel musical excerpts with the intention of conveying one of the following emotional expressions: sadness, joy, calmness, anger, fear, power, surprise, love, and longing. Rather than focussing on one emotion framework, the list of emotions incorporates a combination of emotional expressions present in different frameworks that have been reported in music expressivity surveys as being communicated through music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Zentner, Grandjean, & Scherer, 2008). The basic emotions sadness, joy, fear, and anger (Ekman, 1992; Izard, 1977) are the most common emotions presented in previous stimuli sets, being represented in more than half of musical stimuli used between 1928 and 2018 (Warrenburg, 2020a). Although other emotions such as calmness, love, and longing have ranked high in emotions expressed by music, with calmness also being one of the most frequently perceived emotional expression in music (Cespedes-Guevara & Eerola, 2018; Juslin & Laukka, 2004; Lindström et al., 2003), these emotions are severely underrepresented in music emotion studies stimuli (Warrenburg, 2020a). Laukka et al. (2013) postulated how focussing on one particular emotion framework potentially restricts the investigation to a limited number of basic emotions, complex emotions, or affect states. In light of this, the selection process of emotional expressions investigated in the current project looked past distinct emotion frameworks and incorporated both basic and complex emotions that have been reported as being expressed in music (Juslin & Laukka, 2004; Zentner, Grandjean, & Scherer, 2008), with the aim of expanding existing knowledge on musical expressions.

I decided to compose three musical pieces for each target emotional expression (sadness, joy, calmness, love, longing, anger, fear, power, and surprise). The rationale behind creating multiple pieces for each emotion was influenced by previous knowledge that the excerpts were then going to be rated on their emotional content by a group of participants to assess whether I was successful in portraying the target emotion through the music. The strongest representative of each emotional expression would then be selected and used in the cue manipulation studies with *EmoteControl.* Therefore, creating multiple candidates for each emotional expression would provide me with more possibilities in the following screening process.

The musical pieces were composed mostly by following intuition and previous knowledge about how certain musical features tend to be used to convey different emotions, due to experience gained as a composer and awareness of existing literature (for a review on musical cues associated with different emotional expressions see: Gabrielsson & Lindström, 2010). I tried to compose the excerpts in a naturalistic way, similar to how I would compose a musical work which was not meant for research purposes. This entailed working on a piano to develop distinct musical phrases that encapsulate the different emotional expressions in question. Since one of the goals was to create music which emulates 'real' musical pieces that individuals might listen to in everyday life, I decided to record the pieces as piano performances. A MIDI controller keyboard was used to record the pieces in Logic Pro X using a grand piano from the Vienna Symphonic Library (VSL) as the virtual instrument. Performing the pieces captures additional information encoded as expressive cues in the performance, such as minute variations in velocities, and perhaps subtle rhythmic and timing variations (Juslin & Timmers, 2010), which happen naturally whilst performing, even when performers attempt to play music in a deadpan condition, and adds a humanising feel to the music (Ramos & Mello, 2021; Shoda & Adachi, 2012; Timmers & Ashley, 2007).

In the next two sub-sections, general properties of the musical excerpts such as duration and overall pitch range are described, followed by an account of features specific to musical excerpts aiming to convey the same emotional expression.

3.2.1 General Properties of the Musical Excerpts

The musical pieces were composed as complete musical phrases, with the aim of portraying 'real' music with high ecological validity. Following previous studies which reported that most musical excerpts used in emotion perception studies were less than one minute (Warrenburg, 2020a), with a mean duration of around 30 seconds (Eerola & Vuoskoski, 2013), and that emotion recognition in music is possible in short tone sequences (Vieillard et al., 2008), the musical pieces composed were kept relatively short, and varied in duration, ranging from 14 to 33 seconds. Pieces aiming to convey joy and surprise had the shortest durations, averaging around 18 seconds, whilst pieces aiming to convey longing and sadness had the longest durations, averaging around 30 seconds and 28 seconds long, respectively. Furthermore, the pieces mostly did not contain large musical variations, with the aim of providing excerpts that portrayed a single, specific emotional expression. The instrumental musical excerpts were composed using a Western, tonal framework, and unlike most musical stimuli used in previous studies (Bresin & Friberg, 2011; Dolgin & Adelson, 1990; Gagnon & Peretz, 2003; Hailstone et al., 2009; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014), they are polyphonic works. In addition, since the musical pieces were also going to be used in conjunction with *EmoteControl* (Micallef Grimaud & Eerola, 2021), certain interface compatibility requirements were attended to during the compositional process. As previously mentioned in **Chapter 2**, the *EmoteControl* interface plays the inputted musical piece with one chosen virtual instrument, with its current default being a chamber strings virtual instrument from the Vienna Symphonic Library (VSL). The pitch range of the virtual instrument, albeit a large one (B0 – C7), was also taken into consideration when composing the musical pieces, to ensure that all pieces could be played with the default chamber strings instrument in *EmoteControl.* Musical pieces were composed with a variety of note durations, ranging from semiquavers to semibreves, which would allow for articulation changes from *legato* to *staccato* and vice-versa in the interface. Furthermore, as one of the functions of *EmoteControl* is the ability to change the mode of a piece, by playing the pieces with either a major mode or harmonic minor mode, the musical pieces had to be composed with a specific tonic centre, and have no modulations or notes outside of the designated scale.

3.2.2 Emotion-Specific Musical Features

A total of 28 musical pieces were created. As previously mentioned, more than one excerpt was composed for each emotion category to provide multiple exemplars from which the best representative of each emotion would then be selected after an emotional content assessment task (reported in **Chapter 4**). Three musical pieces were composed with the intention of conveying each of sadness, joy, anger, calmness, fear, love, power, and surprise, while four pieces were composed with the intention of portraying longing. The reason why four pieces were composed for the longing category instead of three, as originally planned, was simply due to inspiration during the actual compositional task, where I inadvertently created four distinct musical ideas that I thought could represent longing. Since the overall aim was to create new material to be used as stimuli and expand the current repertoire of available excerpts, I decided to make use of all four pieces rather than choosing to omit one.

In this section, the commonalities between musical excerpts composed with the intention of conveying the same emotion are detailed, with some discrepancies between excerpts within an emotion category also highlighted. For convenience, the musical excerpts are provided as conventional notation in the Supplementary Material (<u>Section 3.5</u>) at the end of this chapter and referenced throughout this section. However, the portrayal of excerpts as traditional scores does not allow for all the information, such as MIDI velocities, to be conveyed. Furthermore, certain elements which were present during the actual performance of the pieces which contribute to the pieces' ecological property, such as potential variations in the expressive timing had to be quantised to fit the traditional notation. To this end, the audio files of each musical excerpt are also provided via OSF repository[8], as they hold the most information pertaining to the source material.

3.2.2.1 Anger Pieces

The musical excerpts aiming to convey anger (<u>ANG_001</u>, <u>ANG_002</u>, <u>ANG_003</u>) were written in minor mode with tempi ranging from 100 beats per minute (bpm) to 180 bpm (*M* = 136 bpm). All excerpts featured intervallic leaps in the melodic line as well as diatonic dissonance. Melodic lines had a range of two octaves, except for one of the pieces (<u>ANG_001</u>) which had a narrow melodic range of one octave. In two of the excerpts (<u>ANG_001</u>, <u>ANG_002</u>), the melodies were accompanied by a repetition of short duration notes that changed with small interval steps. ANG_001 and ANG_002 mostly had a regular rhythmic pattern, while the third excerpt

[8] The audio files of the musical pieces can be accessed here: https://bit.ly/28musicpieces

(ANG_003) featured mostly chords with an irregular rhythmic pattern as accompaniment.

3.2.2.2 Sadness Pieces

Musical pieces composed to portray sadness (SAD_001, SAD_002, SAD_003) were also written in minor mode, however, with slower tempi ranging from 70 to 80 bpm (M = 75 bpm). All excerpts had a relatively simple musical structure, featuring a melodic pitch range of less than one octave. Melodies followed a descending stepwise motion and had repetitive patterns. Two of the excerpts had regular melodic motions (SAD_001, SAD_003), with the melodies consisting of all quavers apart from the final note, while the other excerpt (SAD_002) comprised less melodic movement. The accompaniments were made up of static, long notes in octaves that moved in a stepwise motion or with small intervallic leaps. All pieces held a simple rhythmic pattern and were mostly consonant, apart from some minor hints of dissonance.

3.2.2.3 Joy Pieces

The musical excerpts intending to portray joy (JOY_001, JOY_002, JOY_003) were composed in major mode and had relatively fast tempi ranging from 110 to 120 bpm (M = 113 bpm). The excerpts featured consonance and a relatively simple harmonic structure. Constant pitch leaps with short note durations were present in the melodic line, and pitches across the melodic line ranged over one to two octaves. The accompaniments varied across pieces from an accompaniment pattern consisting of fifths to an ascending chord pattern which mostly remained consistent within a musical piece.

3.2.2.4 Fear Pieces

Musical pieces composed with the intention of conveying fear (FEA_001, FEA_002, FEA_003) were all written in minor mode, with an average tempo of 103 bpm (range from 100 to 110 bpm). Excerpts featured dissonance and retained a regular rhythmic pattern throughout the pieces. The three musical pieces varied with regards to melodic lines which either moved mostly upwards in a stepwise fashion or with small intervallic leaps, varying in range between two to four octaves. Two of the pieces' accompaniments consisted of repetitive patterns of running sextuplets or semiquaver octaves. The third musical piece (FEA_003) differed from the other two as the melodic line ran across four octaves and did not have an accompaniment.

3.2.2.5 Love Pieces

The musical excerpts intending to portray the emotion love (LOV_001, LOV_002, LOV_003) were all written in major mode with tempi ranging from 60 to 110 bpm (M = 77 bpm). All excerpts featured consonance and had a regular rhythmic pattern. The melodic lines were relatively simple, mostly moving in a stepwise fashion or with small intervallic leaps within a narrow pitch range of an octave. A steady pattern of arpeggios accompanied the melodies for most part of the excerpts. Two of the pieces (LOV_002, LOV_003) had accompaniments at a relatively high pitch level (D4 to A5).

3.2.2.6 Power Pieces

Power may be considered as both a negative or a positive valenced emotion. For example, Kawakami et al. (2013) found that music in minor mode was rated higher for 'heightened' emotions (strong, fiery, heroic, to name a few) than the same music played in major mode. On the other hand, Zentner et al. (2008) grouped high arousal but positive valenced emotions such as heroic, strong, and triumphant under ‘Power’ and subsequently categorised them under the factor ‘Vitality’. When investigating the relative weight of six musical factors on the different emotion clusters in her adjective circle, Hevner (1936, 1937) reported that mode did not have a significant weight in the representation of the high activity emotion cluster containing emotion terms similar to power (e.g., triumphant). Thus, it seems that power may be interpreted as both negative and positive valenced. The musical pieces intending to convey power (POW_001, POW_002, POW_003) in this thesis were all written in minor mode as I felt that I could potentially be more successful in conveying power using the chosen key signatures, in those particular musical contexts. The pieces had fast tempi, ranging from 125 to 175 bpm (*M* = 153 bpm). All excerpts had a narrow melodic range consisting mostly of short duration repetitive notes, differing in small intervals. The pieces also featured dissonance and for the most part, had relatively complex structures. Two excerpts (POW_002, POW_003) had accompaniments consisting of repetitive semiquaver notes that held regular rhythmic patterns with little to no variations, while the other musical piece (POW_001) had an accompaniment which varied in both harmonic and rhythmic structures.

3.2.2.7 Calmness Pieces

All musical excerpts portraying calmness (CAL_001, CAL_002, CAL_003) were written in major mode, with tempi ranging from 60 to 100 bpm (*M* = 83 bpm). All musical pieces were consonant and retained a simple, consistent rhythmic pattern throughout. Two of the excerpts (CAL_001, CAL_002) had melodic lines moving in stepwise motion or with small intervallic leaps, while the other piece (CAL_003) presented the melody in an arpeggiated form. All musical excerpts had relatively static accompaniments consisting of minims or dotted minims. Melodic lines also had narrow pitch ranges, mostly spanning one octave in total.

3.2.2.8 Longing Pieces

Four musical pieces were composed with the aim of conveying longing (LON_001, LON_002, LON_003, LON_004). Three of the longing excerpts (LON_001, LON_003, LON_004) were written in different minor modes (E*b* minor, D minor, and A minor) while the other piece was written in C major (LON_002). The pieces also varied in tempo, ranging from slow tempi (65 bpm) to relatively fast (120 bpm), with an average moderate tempo of 103 bpm across pieces. Two of the excerpts (LON_001, LON_002) featured stepwise movement in the melodic lines accompanied by arpeggios, while the other two pieces (LON_003, LON_004) consisted of melodic movement in an arpeggiated manner paired with mostly sustained notes in the accompaniment. The pitch range of melodies varied from less than an octave to nearly two octaves. The pieces were composed with a simple harmony and featured a constant rhythmic pattern.

3.2.2.9 Surprise Pieces

The three excerpts composed with the aim of conveying surprise (SUR_001, SUR_002, SUR_003) varied in certain compositional features. The pieces were written in different modes: E major, D harmonic minor, or A natural minor. All pieces were composed with a tempo of 100 bpm and mostly had complex harmonic structures. Both the melodic line and accompaniment of pieces featured mostly upward pitch leaps and had irregular rhythmic patterns and variations. Rests were used throughout the musical pieces to potentially instil more elements of surprise. An element of dissonance was also present in some instances with the use of cluster notes.

3.3 Discussion

This chapter introduces 28 new, instrumental musical excerpts which I composed with the intention of conveying either sadness, happiness, calmness, anger, fear, power, love, longing, or surprise emotional expressions through the music, with the aim of creating unfamiliar, complete musical phrases representative of real music which can withstand cue manipulations whilst retaining ecological validity. Additionally, these new musical pieces expand the existing repertoire of specifically-composed music for research to include music portraying emotions other than sadness, happiness, fear, and anger, which are the most common ones portrayed in musical stimuli used in current literature (Warrenburg, 2020a).

Unlike most previous specifically composed stimuli, the musical excerpts composed were polyphonic pieces rather than monophonic sequences (Dolgin & Adelson, 1990; Gagnon & Peretz, 2003; Hailstone et al., 2009; Quinto, Thompson, & Taylor, 2014; Thompson & Robitaille, 1992). The pieces were created as complete musical ideas, rather as incomplete snippets of a musical phrase (Morreale et al., 2013), to emulate 'real music' that exists in a real-life context which is more intricate than simple tone sequences that are usually used as stimuli for research (Battcock & Schutz, 2019; Juslin & Lindström, 2010; Schubert, 2004). The musical excerpts were provided as real performances of myself playing them on the piano, rather than presenting the pieces solely as notation, which allowed me to include other elements of 'real music' which are not portrayed in a score, such as expressive timing and velocities, thus moving away from previous stimuli sets created, by making these new musical excerpts as realistic as possible. Apart from creating ecologically valid musical pieces that depict specific emotional content, I also composed the pieces with the premise that they will undergo experimental manipulations of tempo, pitch, dynamics, brightness, articulation, and mode cues using *EmoteControl* [9] to change the conveyed emotion (detailed in **Chapter 4**). With this in mind, the excerpts were composed using certain musical characteristics which have been linked to specific emotional expressions, which I gained knowledge of through music composition experience, but also due to evidence in existing literature (Dalla Bella et al., 2001; Gabrielsson & Juslin, 1996; Gabrielsson & Lindström, 2010; Juslin & Laukka, 2003), whilst allowing for flexibility in the musical pieces for experimental manipulations and the shaping of different emotion profiles.

In some instances, musical features were utilised in the same manner when trying to portray different emotion profiles during the compositional task, due to the features being associated as potential characteristics of multiple emotional expressions. For example, musical excerpts representing negative-valenced emotions were written in minor mode while pieces portraying emotions with a positive valence were written in major mode (Thompson & Robitaille, 1992; Vieillard et al., 2008). Different mode types were used for musical excerpts intending to convey emotional expressions which can be interpreted as both positive and negative on the valence dimension, such as surprise. Fast tempi were utilised in compositions portraying happiness, anger, fear, surprise, and power, while slower tempi were used for excerpts intending to convey sadness, calmness, love, and longing (Dolgin & Adelson, 1990; Hailstone et al., 2009; Quinto, Thompson, & Taylor, 2014). Anger, power, and fear pieces featured repetition of quick notes and dissonance (Lindström, 2006; Scherer & Oshinsky, 1977), while sadness, calmness, love, and longing excerpts were consonant and had steady

[9] A selection of these musical pieces was also manipulated with regards to tempo, pitch, dynamics, brightness, articulation, mode, and instrument timbre, in a following experiment utilising a second version of *EmoteControl*, detailed in **Chapter 5**.

accompaniments made up of arpeggios or static notes moving mostly in stepwise motion. Anger, fear, joy, and surprise piece featured intervallic leaps in the melody, while stepwise movement in the melodic line was present more in calmness, love, sadness, and longing excerpts (Dolgin & Adelson, 1990; Thompson & Robitaille, 1992).

The overlap of musical properties in the conveying of different emotions has been consistently reported (Gabrielsson & Lindström, 2010; Hailstone et al., 2009) and suggests that even subtle variations in musical features affect the emotional expression communicated in a musical excerpt. The use of similar musical features across emotion profiles might also explain why certain emotional expressions, such as fear and anger have been reported as being commonly confused by listeners in musical pieces (Kragness et al., 2021; Lindström, 2006; Vidas, Dingle, & Nelson, 2018).

The modified lens model (Brunswik, 1956; Juslin, 1995, 1997a, 2000) suggests that the individual musical cues used to encode emotions in the music are probabilistic and partially redundant. This means that a certain cue level may point to a particular emotional expression (e.g., a slow tempo may suggest sadness), however, the cue may not always be used in the same manner to portray the same emotion (e.g., fast and slow tempi both being used for fear). Furthermore, the cues' redundancy suggests that similar emotional information may be communicated through multiple cues. Due to the possible utilisation of the same cues in portraying different emotions, previous literature suggests that the accuracy of emotion communication in music is increased when these cues are added together, rather than the influence of an individual cue, that helps shape different emotional expressions in musical excerpts (Argstatter, 2016; Eerola, Friberg, & Bresin, 2013; Lindström et al., 2003).

For example, a fast tempo can help shape multiple emotional expressions, such as joy, anger, and fear (Balkwill & Thompson, 1999; Juslin, 1997a, 1997b, 2000; Scherer & Oshinsky, 1977), however, the combination of a fast tempo and a major mode might further pinpoint to a specific emotion, such as joy, in this case (Gagnon & Peretz, 2003; Hevner, 1936; Peretz, Gagnon, & Bouchard, 1998; Thompson & Robitaille, 1992). This simple example only mentioned two cues, however, investigating a bigger number of cues at the same time would allow for a clearer picture of how the cues operate together to shape specific emotional expressions (Juslin, 2000).

The fact that cues are said to be partially redundant and able to share similar information also potentially explains why individual cues may be used in different ways when shaping the same emotional expression. For example, multiple previous studies have reported that a fast tempo contributes to the conveying of fear (Bunt & Pavlicevic, 2001; Juslin & Madison, 1999; Laukka & Gabrielsson, 2000). However, a slow tempo has also been reported as being used to communicate fear in music (Bresin & Friberg, 2000; Juslin, 1997b). Similarly, the emotion calmness has sometimes been portrayed by a low pitch level (Gundlach, 1935), while other studies have suggested that a high pitch level contributes to the shaping of calmness in music (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013). A different use of individual cues may also be observed in the new set of musical excerpts detailed in this chapter. For example, the musical excerpts aiming to convey surprise were all written in different modes and with varying musical structures. The first surprise excerpt (SUR_001) was composed in A natural minor (Aeolian mode with the note A being its tonic centre), while the second (SUR_002) and third (SUR_003) excerpts were composed in D harmonic minor and E major respectively. The excerpts also vary with regards to the amount of dissonance present and the regularity of rhythmic patterns, where SUR_001 and SUR_002 pieces have more regular rhythmic patterns and accompaniments than the third excerpt. The variance in cues and combinations between the three surprise excerpts resulted from the fact that surprise may be understood in multiple ways. While I was creating the surprise pieces, I kept in mind that surprise could be interpreted as either a positive or a negative expression. From a composer's perspective, I was also not certain which degree and type of surprise would be best conveyed in the musical pieces. Therefore, I decided to try and portray surprise in different ways, which is why there is notable variation between the three surprise excerpts. Due to this multidimensional property, as well as surprise being usually attributed to a brief moment created due to a violation of expectancy, I think it was personally one of the most challenging emotional expressions to try and portray within the musical pieces during the compositional task.

On the other hand, I found that composing pieces aiming to convey joy, sadness, and calmness was more straightforward. Perhaps this might be because I have a rather definite idea of how joy, sadness, and calmness should stereotypically sound like in music, potentially due to these three emotions frequently occurring in music (Juslin & Laukka, 2004; Lindström et al., 2003), thus making me more familiar with them. My experience as a composer might have also played a role, as I am more used to creating musical works with sadness, joy, and calmness elements to them, as against creating pieces conveying surprise.

Distinguishing between certain emotional expressions such as anger, fear, and power was also rather challenging as I found myself utilising similar musical features, such as dissonance, minor mode, and high note density to portray these different target emotions in the musical excerpts. This might be due to anger, fear, and power all being regarded as high arousal emotional expressions. Furthermore, anger and fear are considered as negative valenced emotions whilst power may be seen as both a negative and positive valenced emotion, thus, it is possible that these emotions may be understood or interpreted in multiple ways. The same issues applied when trying to differentiate between love and calmness, and sadness and longing emotional expressions, which seem to exist on very similar planes in the valence-arousal dimensions. In my opinion, this is where the use of nuances of cues and cue combinations comes into play. For example, I portrayed both love and calmness using a major mode and consonance. However, with the aim of differentiating between the two, I composed excerpts meaning to convey love with more movement and at a faster tempo than the calmness pieces. Similarly, I made a distinction between sadness and longing excerpts by composing the longing pieces mostly with flowing accompaniments, at a faster tempo, and with more movement than the sadness pieces. I used narrow melody ranges and relatively static melodic movements when trying to convey power in musical excerpts, in comparison to pieces aiming to convey fear, which were represented with fleeting notes moving mostly in an upward melodic direction. Musical excerpts representing anger and power mainly featured a repeated note pattern in the accompaniment, whilst more dynamic changes were present in the pieces portraying fear. From the emotions that had similar valence and arousal properties, I found distinguishing between joy and surprise excerpts the least challenging. Although joy and surprise excerpts comprised common characteristics, such as fast tempi and similar pitch ranges, joy pieces had a consistent, regular rhythmic pattern while surprise pieces featured irregular rhythmic patterns with variations and instances of dissonance, with the aim of instilling the surprise element in the pieces.

My experience composing these 28 musical excerpts and essentially being the first individual to assess the emotional content conveyed by the musical pieces, amplified my awareness that the different emotional expressions in question (sadness, joy, calmness, anger, fear, love, longing, power, and surprise) may be interpreted in music in multiple ways. From a composer's perspective, the creation process of these new musical excerpts highlighted the multidimensional properties of the target emotions and the fact that emotions may not be as clear-cut as occasionally suggested. Therefore, although I composed the excerpts with the aim of conveying specific emotional expressions, these musical pieces should be regarded as open materials subjected to the communication of emotions. However, to properly assess the emotional content of these new musical excerpts and how successful I was in portraying particular emotions in the music, the pieces will undergo a validation experiment where a group of participants will listen to the excerpts and rate them on separate Likert scale representing each of the nine original emotional expressions (sadness, joy, calmness, anger, fear, longing, love, power, and surprise). This will give us a clear picture of which emotional expression(s) other individuals perceive in the musical excerpts. The empirical experiment will thus provide additional critical evaluation of the pieces. Due to the similarities between excerpts conveying highly associated emotions, I anticipate that participants might also find it challenging from a listener's perspective to distinguish between certain emotion categories, such as anger and fear, or love and calmness, similar to my experience from a composer's perspective.

In conclusion, this chapter introduced 28 novel Western tonal musical excerpts specifically composed for use in the current project. The fact that the musical pieces were explicitly created for research ensures that participants were not exposed to them through previous experiences, such as, for example, in concerts, adverts, or films, thus ensuring that potential participants do not have any previous attributions and an unconscious bias linked to these specific musical pieces during future studies, which may occur when participants are presented with commercial music in research studies. The musical pieces differ from most specifically-created musical stimuli used in previous research (Dolgin & Adelson, 1990; Gagnon & Peretz, 2003; Hailstone et al., 2009; Järvinen et al., 2010, 2012; Paquette, Peretz, & Belin, 2013; Quinto, Thompson, & Taylor, 2014; Thompson & Robitaille, 1992), as they represent complete, polyphonic musical phrases which were captured through actual performances, with the aim of representing a real-life musical context with *ecologically valid music* within a Western tonal framework. Each excerpt in this collection of musical pieces was created with the intention of conveying one of *nine emotional expressions* (sadness, joy, calmness, anger, fear, power, love, longing, or surprise) which may be expressed through music (Juslin & Laukka, 2004; Lindström et al., 2003; Zentner, Grandjean, & Scherer, 2008), contributing new material to the existing stimuli repertoire which until now had mostly featured excerpts portraying sadness, happiness, fear, and anger emotional expressions (Warrenburg, 2020a). Furthermore, the musical pieces were composed in a way which allows for features of the music to be manipulated (detailed in **Chapter 4** and **Chapter 5**) without crucially compromising the ecological validity of the musical pieces, enhancing the *versatility of the excerpts* and their potential uses. This chapter also provided a first-hand reflective account of my compositional process and experience as the composer and essentially, as the first participant carrying out a perceived emotion assessment task on these novel musical pieces. Moreover, apart from being new contributions to existing stimuli in music emotion research, these new musical excerpts will also be assessed by a group of participants with regards to their emotional content through an evaluation task, with the aim of providing a new, ecologically-valid, polyphonic musical stimuli set with perceived emotional expression ratings on nine different emotion scales, which is reported next in **Chapter 4**.

3.4 References

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

Battcock, A., & Schutz, M. (2019). Acoustically expressing affect. *Music Perception*, *37*(1), 66–91. https://doi.org/10.1525/MP.2019.37.1.66

Behne, K. E. (1972). *Der Einfluss des Tempos auf die Beurteilung von Musik [The influence of tempo on judgement of music]*. Arno Volk Verlag.

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., Vieillard, 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. (2000). Emotional Coloring of Computer-Controlled Music Performances. *Computer Music Journal*, *24*(4), 44–63.

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

Brunswik, E. (1956). Perception and the representative design of psychological experiments. In *Perception and the representative design of psychological experiments, 2nd ed.* University of California Press.

Bunt, L., & Pavlicevic, M. (2001). Music and emotion: Perspectives from music therapy. In *Music and emotion: Theory and research.* (pp. 181–201). Oxford University Press.

Burger, B., Saarikallio, S., Luck, G., Thompson, M. R., & Toiviainen, P. (2013). Relationships Between Perceived Emotions in Music and Music-induced Movement. *Music Perception: An Interdisciplinary Journal*, *30*(5), 517–533.

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

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.

Dolgin, K. G., & Adelson, E. H. (1990). Age Changes in the Ability to Interpret Affect in Sung and Instrumentally-Presented Melodies. *Psychology of Music*, *18*(1), 87–98. https://doi.org/10.1177/0305735690181007

Downey, J. E. (1897). A musical experiment. *The American Journal of Psychology*, *9*(1), 63–69. https://doi.org/10.2307/1411476

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. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49. https://doi.org/10.1177/0305735610362821

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.

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. (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

Gilman, B. I. (1891). Report on an experimental test of musical expressiveness. *American Journal of Psychology*, *4*, 558–576.

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., & Samson, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*(3), 628–640. https://doi.org/10.1093/brain/awh420

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.

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.

Izard, C. E. (1977). *Human Emotions*. Plenum Publishing Corporation. https://doi.org/10.1007/978-1-4899-2209-0

Järvinen, A., Dering, B., Neumann, D., Ng, R., Crivelli, D., Grichanik, M., Korenberg, J. R., & Bellugi, U. (2012). Sensitivity of the autonomic nervous system to visual and auditory affect across social and non-social domains in Williams syndrome. *Frontiers in Psychology*, *3*(September). https://doi.org/10.3389/fpsyg.2012.00343

Järvinen, A., Vines, B. W., Hill, K. J., Yam, A., Grichanik, M., Mills, D., Reiss, A. L., Korenberg, J. R., & Bellugi, U. (2010). Cross-modal influences of affect across social and non-social domains in individuals with Williams syndrome. *Neuropsychologia*, *48*(2), 456–466. https://doi.org/10.1016/j.neuropsychologia.2009.10.003

Juslin, P. N. (1995). Emotional communication in music viewed through a Brunswikian lens. In G. Kleinen (Ed.), *Music and expression: Proceedings of the conference of DGM and ESCOM* (pp. 21–25). University of Bremen.

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*, *I*(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., & 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., & Madison, G. (1999). The Role of Timing Patterns in Recognition of Emotional Expression from Musical Performance. *Music Perception*, *17*(2), 197–221. https://doi.org/10.2307/40285891

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 452–489). 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

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

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, *4*(June), 1–15. https://doi.org/10.3389/fpsyg.2013.00311

Kleinsmith, A. L., Friedman, R. S., & Neill, W. T. (2016). Exploring the impact of final ritardandi on evaluative responses to cadential closure. *Psychomusicology: Music, Mind, and Brain*, *26*(4), 346–357. https://doi.org/10.1037/pmu0000159

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

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*. Department of Pscyhology, Keele (CD-rom).

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*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Laukka, P., & Gabrielsson, A. (2000). Emotional Expression in Drumming Performance. *Psychology of Music*, *28*(2), 181–189.

https://doi.org/10.1177/0305735600282007

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.

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

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*. University of Jyväskylä, Department of Music.

Omar, R., Henley, S. M. D., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., Frost, C., Scott, S. K., & Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *NeuroImage*, *56*(3), 1814–1821. https://doi.org/10.1016/j.neuroimage.2011.03.002

Paquette, S., Peretz, I., & Belin, P. (2013). The “Musical Emotional Bursts”: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*(AUG), 1–7. https://doi.org/10.3389/fpsyg.2013.00509

Parke, R., Chew, E., & Kyriakakis, C. (2007). Quantitative and visual analysis of the impact of music on perceived emotion of film. *Computers in Entertainment*, *5*(3). https://doi.org/10.1145/1316511.1316516

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(3), 111–141.

Punkanen, M., Eerola, T., & Erkkilä, J. (2011). Biased emotional recognition in depression: Perception of emotions in music by depressed patients. *Journal of Affective Disorders*, *130*(1–2), 118–126. https://doi.org/10.1016/j.jad.2010.10.034

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

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

Ramos, D., & Mello, E. K. S. (2021). Communication of emotions in music between Brazilian composer, performers, and listeners. *Psychomusicology: Music, Mind, and Brain*, *31*(1), 1–17. https://doi.org/10.1037/pmu0000269

Rigg, M. G. (1940). The effect of register and tonality upon musical mood. *Journal of Musicology*, *2*(2), 49–61.

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. K., & 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

Shoda, H., & Adachi, M. (2012). The Role of a Pianist's Affective and Structural Interpretations in his Expressive Body Movement: A Single Case Study. *Music Perception: An Interdisciplinary Journal*, *29*(3), 237–254. https://doi.org/10.1525/mp.2012.29.3.237

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

Sutcliffe, R., Rendell, P., Henry, J., Bailey, P., & Ruffman, T. (2017). Music to my ears: Age-related decline in musical and facial emotion recognition. *Psychology and Aging*, *32*(8), 698–709. https://doi.org/10.1037/pag0000203

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, *10*(1), 79–89.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, *25*(2), 117–134. https://doi.org/10.1525/mp.2007.25.2.117

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

Vuoskoski, J. K., & Eerola, T. (2017). The pleasure evoked by sad music is mediated by feelings of being moved. *Frontiers in Psychology*, *8*(MAR), 1–11. https://doi.org/10.3389/fpsyg.2017.00439

Warrenburg, L. A. (2020). 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

Zentner, M. R., Grandjean, D., & Scherer, K. R. (2008). Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

3.5 Supplementary Material

The following scores denote the 28 musical excerpts created and detailed in this chapter. For convenience, the excerpts are presented in traditional notation form. The notations do not give all the information of the performance recordings of the excerpts, such as velocities and expressive timings. Furthermore, due to traditional notation conventions, the durations of the notes have been quantised for the notations, therefore, the score representations might vary from the actual performances. Similarly, an approximation of the tempo has been denoted in the notation representations of the musical excerpts, which gives an indication of how the quantised notes depicted in the score should be played to mirror the actual performances of the pieces. The notations have been provided for convenience, however, the performances which hold all compositional and expressive information may be heard in the audio files provided in the associated OSF repository. The 28 musical excerpts are grouped by the emotional expression they intend to convey. Furthermore, each excerpt was given a code, depending on the emotion category and the number of the excerpt in that category. For example, there are three excerpts in the 'Anger' emotion group. They were named excerpts ANG_001, ANG_002, and ANG_003. The same coding procedure was administered for all 28 musical pieces.

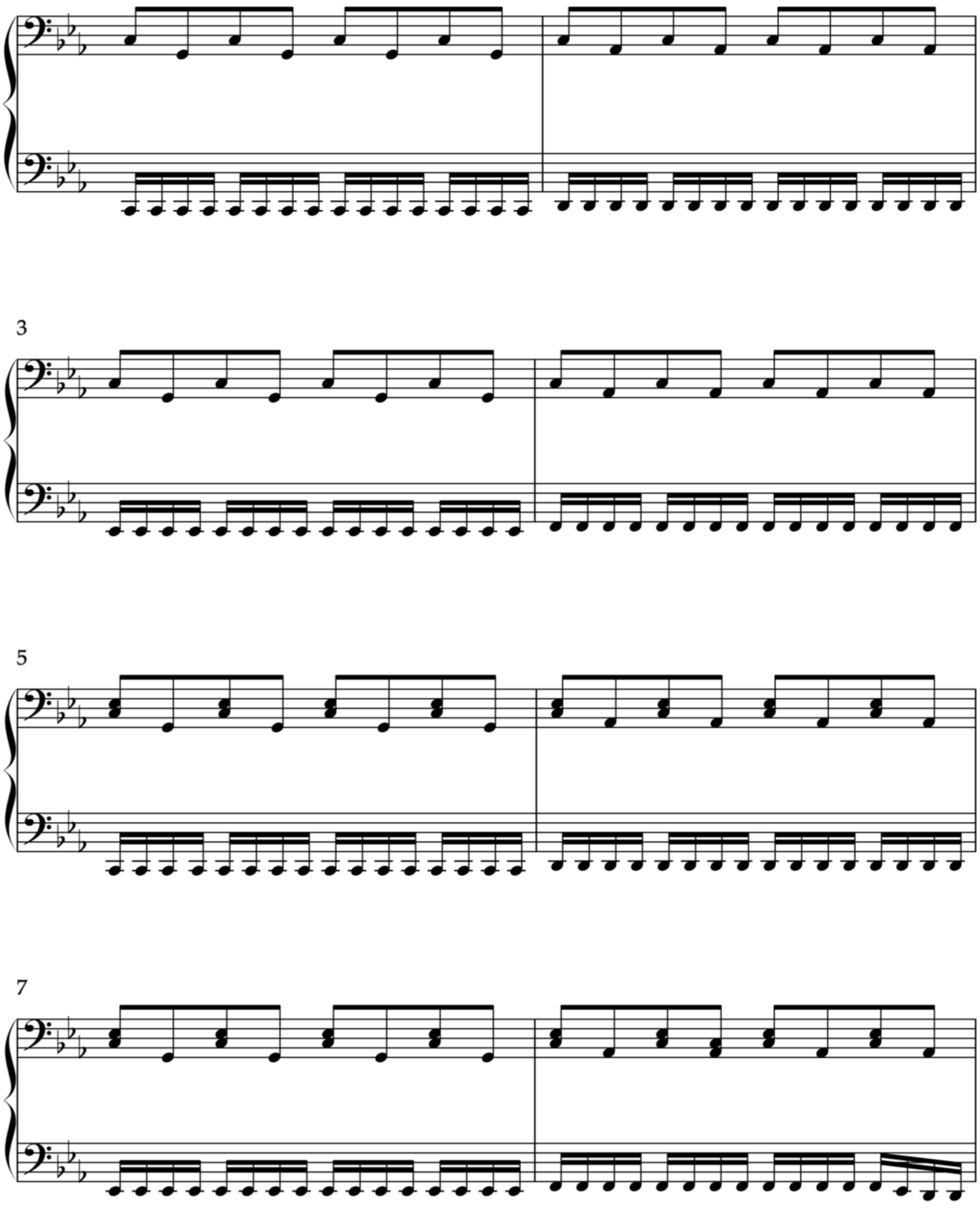
3.5.1 Anger Excerpts

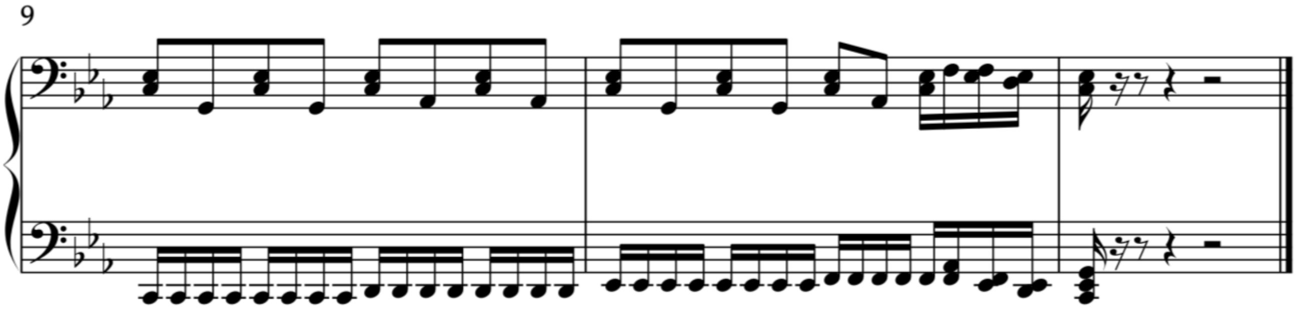
Musical pieces aiming to convey anger are denoted as ANG_001, ANG_002, and ANG_003.

3.5.1.1 ANG_001

The first excerpt in the anger emotion group was written in C minor and is represented in the score (on the next page) in 4/4 time, with a tempo of 100 bpm.

ANG_001





3.5.1.2 ANG_002

The second excerpt in the anger emotion group was also written in C minor and is represented in the score (on the next page) in 6/8 time, with a tempo of 130 bpm.

ANG_002

3.5.1.3 ANG_003

The third excerpt in the anger emotion group was composed in G minor, and is represented in the score in 4/4 time, with a tempo of 180 bpm.

3.5.2 Sadness Excerpts

Musical pieces aiming to convey sadness are denoted as SAD_001, SAD_002, and SAD_003.

3.5.2.1 SAD_001

The first excerpt in the sadness emotion group was written in A minor, and is represented in the score in 6/8 time, with a tempo of 65 bpm.

3.5.2.2 SAD_002

The second excerpt in the sadness emotion group was written in C minor and is represented in the score in 4/4 time, with a tempo of 70 bpm.

SAD_002

3.5.2.3 SAD_003

The third excerpt in the sadness emotion group was written in D minor, and is represented in the score in 6/8 time, with a tempo of 60 bpm.

SAD_003

3.5.3 Joy Excerpts

Musical pieces aiming to convey joy are denoted as JOY_001, JOY_002, and JOY_003.

3.5.3.1 JOY_001

The first excerpt in the joy emotion group was written in A major, and is represented in the score in 4/4 time, with a tempo of 110 bpm.

3.5.3.2 JOY_002

The second excerpt in the joy emotion group was written in C major, and is represented in the score in 6/8 time, with a tempo of 120 bpm.

3.5.3.3 JOY_003

The third excerpt in the joy emotion group was written in G major, and is represented in the score in 4/4 time, with a tempo of 110 bpm.

JOY_003

3.5.4 Fear Excerpts

Musical pieces aiming to convey fear are denoted as FEA_001, FEA_002, and FEA _003. The notation representation of the pieces can be found in the following pages.

3.5.4.1 FEA_001

The first excerpt in the fear emotion group was written in C minor, and is represented in the score in 2/4 time, with a tempo of 80 bpm.

3.5.4.2 FEA_002

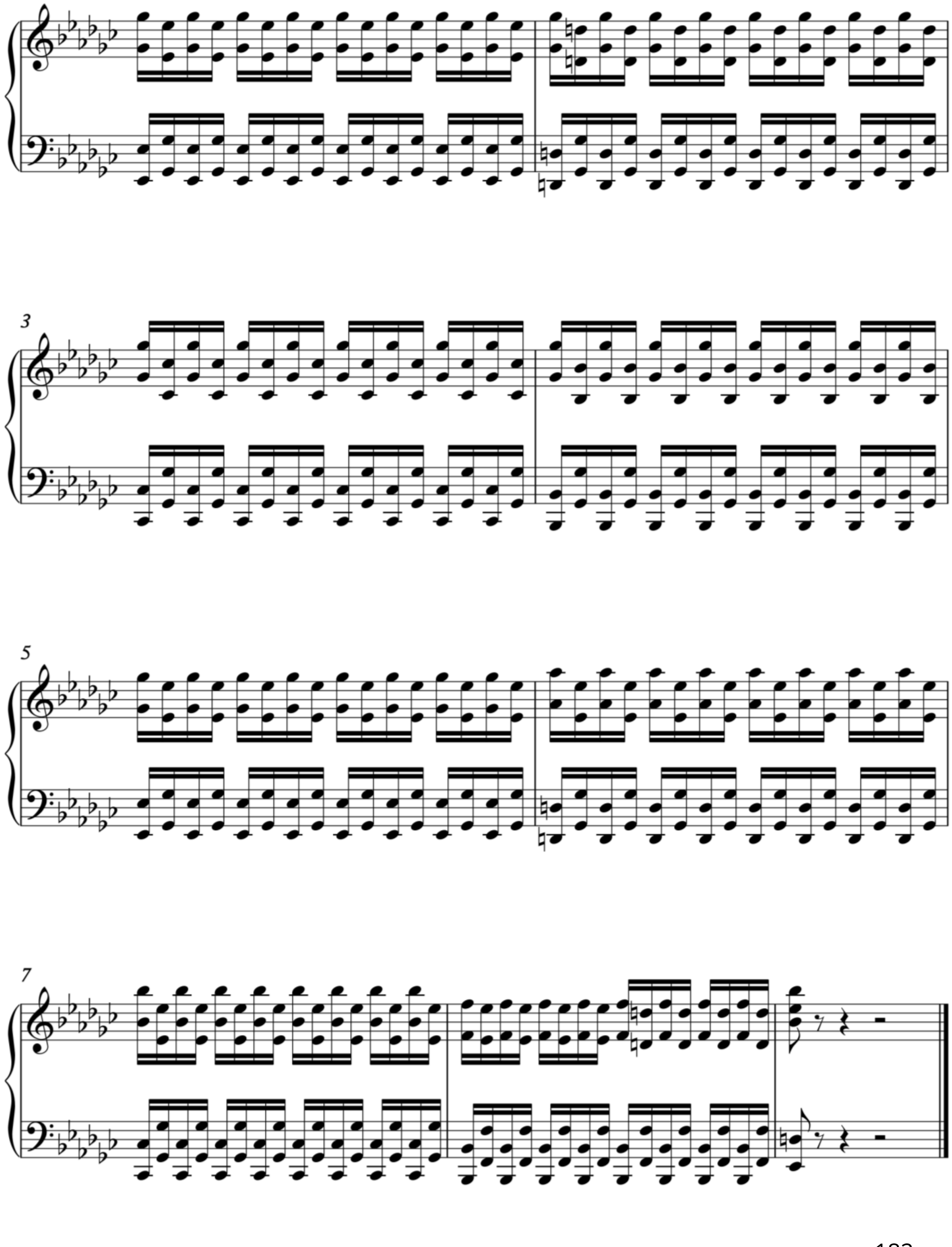
The second excerpt in the fear emotion group was written in E*b* minor, and is represented in the score in 4/4 time, with a tempo of 110 bpm.

3.5.4.3 FEA_003

The third excerpt in the fear emotion group was composed in G minor, and is represented in the score in 2/4 time, with a tempo of 100 bpm.

FEA_001

FEA_002



FEA_003

3.5.5 Love Excerpts

Musical pieces aiming to convey love are denoted as LOV_001, LOV_002, and LOV_003.

3.5.5.1 LOV_001

The first excerpt in the love emotion group was written in C major, and is represented in the score in 4/4 time, with a tempo of 110 bpm.

LOV_001

3.5.5.2 LOV_002

The second excerpt in the love emotion group was written in A major, and is represented in the score in 6/8 time, with a tempo of 90 bpm.

LOV_002

3.5.5.3 LOV_003

The third excerpt in the love emotion group was written in E major, and is represented in the score in 4/4 time, with a tempo of 100 bpm.

3.5.6 Power Excerpts

Musical pieces aiming to convey power are denoted as POW_001, POW_002, and POW_003.

3.5.6.1 POW_001

The first excerpt in the power emotion category was composed in C minor and is represented in the score in 4/4 time, with a tempo of 130 bpm.

POW_001

3.5.6.2 POW_002

The second excerpt in the power emotion category was composed in D minor and is represented in the score in 2/4 time, with a tempo of 80 bpm.

3.5.6.3 POW_003

The third excerpt in the power emotion category was also written in D minor and is represented in the notation form in 4/4 time, with a tempo of 90 bpm.

POW_003

3.5.7 Calmness Excerpts

Musical pieces aiming to convey calmness are denoted as CAL_001, CAL_002, and CAL_003.

3.5.7.1 CAL_001

The first excerpt in the calmness emotion group was written in A major and is represented in the score in 4/4 time, with a tempo of 90 bpm.

3.5.7.2 CAL_002

The second excerpt in the calmness emotion group was written in C*b* major and is represented in the score in 6/8 time, with a tempo of 90 bpm.

3.5.7.3 CAL_003

The third excerpt in the calmness emotion group was written in C major and is notated in 6/8 time, with a tempo of 90 bpm.

3.5.8 Longing Excerpts

The longing emotion group is the only category with four musical excerpts composed for the target emotion. Musical pieces aiming to convey longing are denoted as LON_001, LON_002, LON_003, and LON_004.

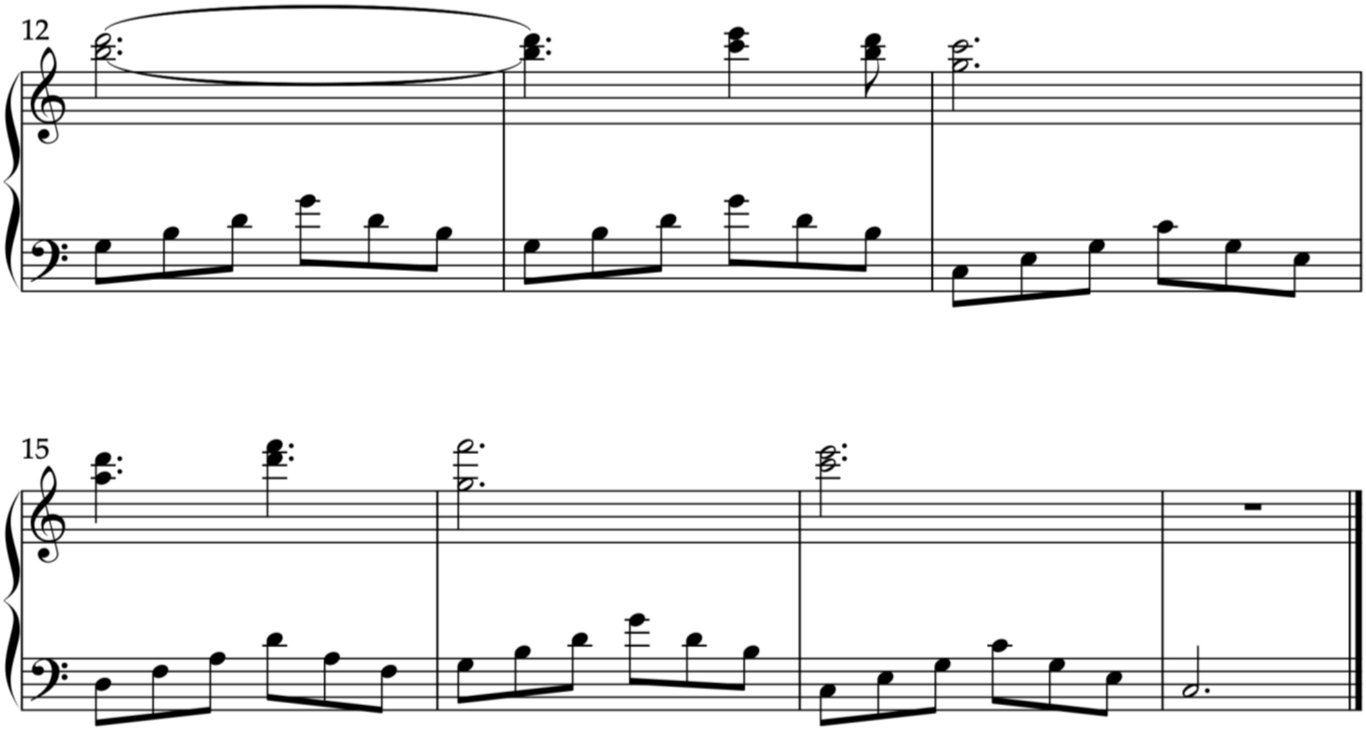
3.5.8.1 LON_001

The first excerpt in the longing emotion group was composed in A minor and is represented in the score in 6/8 time, with a tempo of 120 bpm.

3.5.8.2 LON_002

The second excerpt in the longing emotion group was written in C major and is represented in the score in 6/8 time, with a tempo of 100 bpm.

LON_002



3.5.8.3 LON_003

The third excerpt in the longing emotion group was composed in D minor and is notated in 4/4 time, with a tempo of 50 bpm.

LON_003

3.5.8.4 LON_004

The fourth and final excerpt in the longing emotion group was written in E*b* minor and is represented in the score in 6/8 time, with a tempo of 100 bpm.

LON_004

3.5.9 Surprise Excerpts

Musical pieces aiming to portay are denoted as SUR_001, SUR_002, and SUR_003.

3.5.9.1 SUR_001

The first excerpt in the surprise emotion group was composed in A natural minor (Aeolian mode). The musical piece is notated in 6/8 time, with a tempo of 100 bpm.

SUR_001

3.5.9.2 SUR_002

The second excerpt composed to convey surprise was written in D minor and is notated in 4/4 time, with a tempo of 100 bpm.

3.5.9.3 SUR_003

The third excerpt aiming to convey surprise was composed in E major and is notated in 4/4 time, with a tempo of 100 bpm.

Chapter 4. An Interactive Approach to Emotional Expression Through Musical Cues

Micallef Grimaud, A., & Eerola, T. (2022). An interactive approach to emotional expression through musical cues. *Music & Science*. https://doi.org/10.1177/20592043211061745

Preface

Chapter 4 is a natural continuation of the previous chapters, and details three empirical experiments. First, the newly composed 28 musical pieces detailed in Chapter 3 are subjected to an evaluation experiment, where a group of participants listens to the pieces and rates them on nine separate emotion Likert scales (the initial nine target emotions composed for: sadness, joy, calmness, anger, fear, power, longing, love, and surprise) to assess whether the target emotional content was successfully communicated through the music. The second experiment in this chapter puts to use both elements specifically created for the current project: the *EmoteControl* interface and the novel musical excerpts. Experiment 2 utilises a production approach where a subset of the musical pieces is altered in real-time by participants in *EmoteControl* to communicate different target emotions, providing us with distinct cue patterns used by participants in their efforts to convey the different emotions, and consequentially, new variations of the original excerpts. The third experiment comprises an evaluation study where participants assess the emotional content of these new versions of the excerpts together with the original versions, providing more information on how the same emotional expressions presented in varying excerpts were perceived by the listeners.

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 chapter 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.

4.1 Introduction

Previous literature suggests that emotions can be successfully conveyed through music and recognised by the listeners (Juslin, 1997a, 2013b). 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., 2019; Saarikallio, Vuoskoski, & Luck, 2014), and a tool for mood regulation (Lyvers, Cotterell, & Thorberg, 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 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 (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, 2000; Juslin & Laukka, 2004). 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 work, 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 & 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; K. Watson, 1942). Many of these cues – such as loudness, timbre, tonality, and rhythm – seem to be operating similarly in different cultures (Balkwill, Thompson, & Matsunaga, 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, Friberg, and Bresin (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, Ferrer, & Alluri, 2012), harmony (Lahdelma & Eerola, 2016b), mode (Kastner & Crowder, 1990), and harmonic intervals (Costa, Ricci Bitti, & Bonfiglioli, 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, Friberg, & Bresin, 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, Bresin, & Sundberg, 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; Saarikallio, Vuoskoski, & Luck, 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, 2020c), 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, Argstatter, & Wilker, 2010; Saarikallio et al., 2019). Other studies ask participants to rate musical pieces on valence and arousal dimensions (Costa, Fine, & Ricci Bitti, 2004; Morreale et al., 2013; Quinto, Thompson, & Taylor, 2014). However, utilising valence and arousal dimensions may be somewhat ubiquitous and limiting, as some emotional expressions might not be captured by these dimensions (G. L. 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 & Laukka, 2004; Kreutz, 2000; Lindström, Juslin, Bresin, & Williamon, 2003; Zentner, Grandjean, & Scherer, 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 work: 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, Argstatter, & Wilker, 2010), calmness (Juslin & Laukka, 2004; Laukka et al., 2013; Lindström et al., 2003; Thompson & Robitaille, 1992; Vieillard et al., 2008; Zentner, Grandjean, & Scherer, 2008), anger (Akkermans et al., 2019; Behrens & Green, 1993; Juslin & Laukka, 2004; Laukka et al., 2013; Mohn, Argstatter, & Wilker, 2010), fear (Behrens & Green, 1993; Kreutz, 2000; Mohn, Argstatter, & Wilker, 2010; Vieillard et al., 2008; Zentner, Grandjean, & Scherer, 2008), surprise (Juslin & Laukka, 2004; Lindström et al., 2003; Mohn, Argstatter, & Wilker, 2010; 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, Grandjean, & Scherer, 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, Friberg, & Bresin, 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 chapter, 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 chapter:

1) Our first sub-goal was to create a new set of musical stimuli (described in **Chapter 3**) 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, 2013b; 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; Saarikallio, Vuoskoski, & Luck, 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 this chapter, 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; Saarikallio, Vuoskoski, & Luck, 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.

4.2 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, detailed in **Chapter 3**. 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, 2013b; Juslin & Laukka, 2004; Lindström et al., 2003; Turnbull et al., 2008; Zentner, Grandjean, & Scherer, 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.

4.2.1 Method

4.2.1.1 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 (J. 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.

4.2.1.2 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 & Eerola, 2018; Juslin & 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 33 seconds. Further details on the musical pieces can be found in the supplementary material (Table 4.8). All musical pieces are available via the OSF repository[11].

4.2.1.3 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 will 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 at the end of this chapter, 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

[11] The 28 musical pieces can be found on the OSF repository by following this link https://osf.io/f9nhp/

emotions utilised to compose the pieces (joy, sadness, calmness, power, anger, fear, surprise, love, and 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 minutes to complete.

4.2.2 Results

The consistency among participants in using the emotion rating scales was calculated using Cronbach's alpha (intraclass 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 scale differently to the other emotion rating scales. To run this ANOVA, the nine emotion scales used where classified as 'Target' or 'Non-Target' emotions for each of the musical pieces, with the 'target' emotion being the independent variable and the dependent variable being the collapsed ratings across all other 'non-target' emotion scales. The ANOVA produced a significant result, which suggests that in general, participants rated the intended emotion significantly different to the other available eight emotion scales, depending on the target emotion of the pieces, $F(1, 95) = 1173.00$, $p < .001$, $\eta^2 = 0.14$.

The mean ratings given by participants for all nine emotion scales were calculated for the 28 musical excerpts. Table 4.1 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.

One-way repeated measures ANOVAs were 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 intended emotion was compared to each of the other eight possible emotions by running individual ANOVAs for each, totalling eight iterations for each emotion group. The results were corrected with the Bonferroni method, and the degrees of freedom for each ANOVA are denoted in the Notes section beneath Table 4.1. The asterisks following the mean emotion ratings in the rows in Table 4.1 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 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. For each musical piece, one-way repeated measures ANOVAs were run to assess whether the intended emotion of the piece was rated significantly higher than the other possible emotion rating scales. The pieces confirmed with having their intended emotion rated significantly higher than the other emotion rating scales may be considered as successfully representing their target emotion and are denoted in the 'Excerpts' row in Table 4.1. The pieces are listed in a ranked order starting with the strongest representative of the intended emotion. The mean emotion rating of the individual pieces given by participants is denoted in brackets. A '-' in the table denotes when none of the pieces in the emotion group were rated significantly higher for their intended emotion, and thus, may not be seen as representatives of their 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.

Table 4.1. 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.

		Rated Emotion								
	Rating:	Anger	Calmness	Fear	Joy	Longing	Love	Power	Sadness	Surprise
Intended Emotion in the Pieces	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 ANOVAs. Df (1, 95) for each one-way repeated measures ANOVA. 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.

4.2.3 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, 2013b; 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. Furthermore, love, longing, and calmness are all low-activity emotions that exist in a similar space on the arousal plane, which may be why they were not successfully recognised by listeners. In fact, in a study investigating whether mild and intense sadness, happiness, love, and anger may be conveyed through ornamentation in instrumental music, Timmers and Ashley (2004, 2007) reported that the low-activity emotions love and sadness were not successfully communicated by performers and distinguished by listeners.

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, 2013b; Peretz, Gagnon, & Bouchard, 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, Grandjean, & Scherer, 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, Ricci Bitti, & Bonfiglioli, 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, Dingle, & Nelson, 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, Argstatter, & Wilker, 2010). 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 beats per minute (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, Gagnon, & Bouchard, 1998) whilst surprise excerpts varied in modes. The strongest candidate for surprise was in a harmonic 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, Ricci Bitti, & Bonfiglioli, 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, Thompson, & Keating, 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; K. 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, 2013b). Furthermore, it may be that other musical features not investigated here have a bigger role in shaping love and longing in music. For example, Lahdelma and Eerola (2014, 2015) discovered that single chords in a minor triad root position and major 7th chords in a third inversion position were rated as expressing nostalgia/longing and rated particularly higher as conveying nostalgia/longing when played with a strings timbre as against a piano timbre. Perhaps the utilisation of these types of chords may have helped communicate longing effectively. Another musical feature not present in this work which may have a significant role in conveying love in music is lyrics. When analysing the emotional content of lyrics in a database of English and German songs, Kreutz (2000) found out that the most frequently represented emotion in the lyrics was love. Longing was also reported as being frequently portrayed in lyrics (Juslin, 2019c). Therefore, perhaps adding lyrics may have helped distinctly communicate love and longing in the pieces.

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, Friberg, & Bresin, 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.

4.2.3.1 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.

4.3 Experiment 2: Cue Manipulation Task

Experiment 2 addressed the main aim of this chapter, 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.

4.3.1 Method

4.3.1.1 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; J. 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.

4.3.1.2 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.

4.3.1.3 Apparatus

EmoteControl, a graphical user interface created for music emotion research, was utilised for the study (Micallef Grimaud & Eerola, 2021). Figure 4.1 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 10Hz.

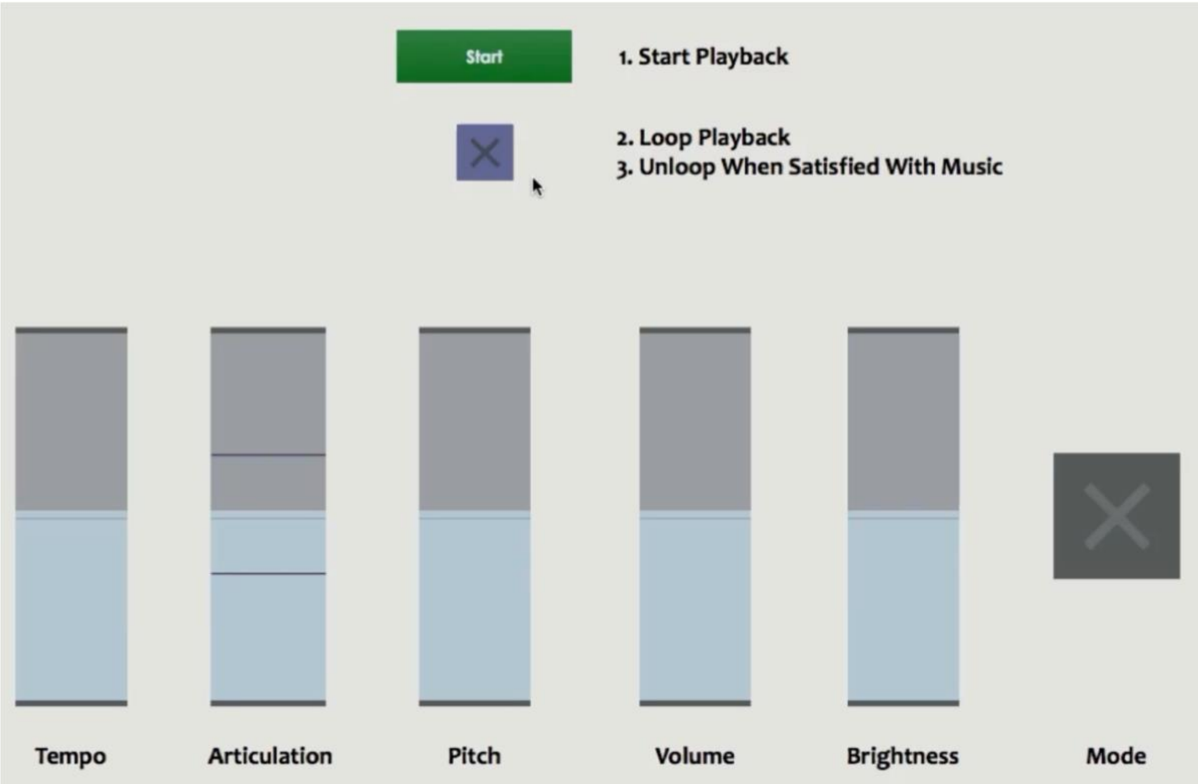


Figure 4.1. The *EmoteControl* user interface.

4.3.1.4 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, Friberg, & Bresin, 2013; Kamenetsky, Hill, & Trehub, 1997; Morreale, Masu, & Angeli, 2013), while the remaining three cues, pitch, articulation, and brightness, have been investigated to a lesser extent (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 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 combination 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/or interactive nature (Gabrielsson, 2008), which were two additional motivations taken into consideration when choosing the cues to be investigated in this experiment. The following sub-sections describe the ranges and levels used for each cue in the *EmoteControl* 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 305Hz to 20,000Hz.

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.

4.3.1.5 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 at the end of this chapter. 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 × 7 pieces). The experiment took approximately 30 minutes to complete.

4.3.2 Results

The consistency and reliability of the participants' use of the six cues were determined by using Cronbach's alpha to calculate the inter-rater agreement (intraclass correlation coefficient) of participants for each of the six cue sliders or buttons across the different combinations of music stimuli and target emotions. Since participants were split into three different groups, the inter-rater agreement for the six cues were calculated separately within each group, resulting in three alpha values for each cue (one for each cue in the three individual groups). High consistency among participants within each group was observed, particularly in Tempo (with α values ranging from 0.943 to 0.964 across the groups), Mode (α = 0.950-0.960), Articulation (α = 0.943-0.957), Pitch (α = 0.939-0.956), Brightness (α = 0.833-0.869), and Dynamics (α = 0.832-0.865).

Table 4.2 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 4.2 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 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 4.9 in the Supplementary Material (section 4.7.2). Articulation was the only significant effect on the interaction between Piece and Emotion. The most relevant result from Table 4.2 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.

Table 4.2. 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 x 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.

Figure 4.2 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 4.2) helps identify the emotions which participants were indecisive about 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 the emotion, Figure 4.2 shows how the mode value for power also leans towards minor mode. A similar pattern can be seen for the surprise emotion.

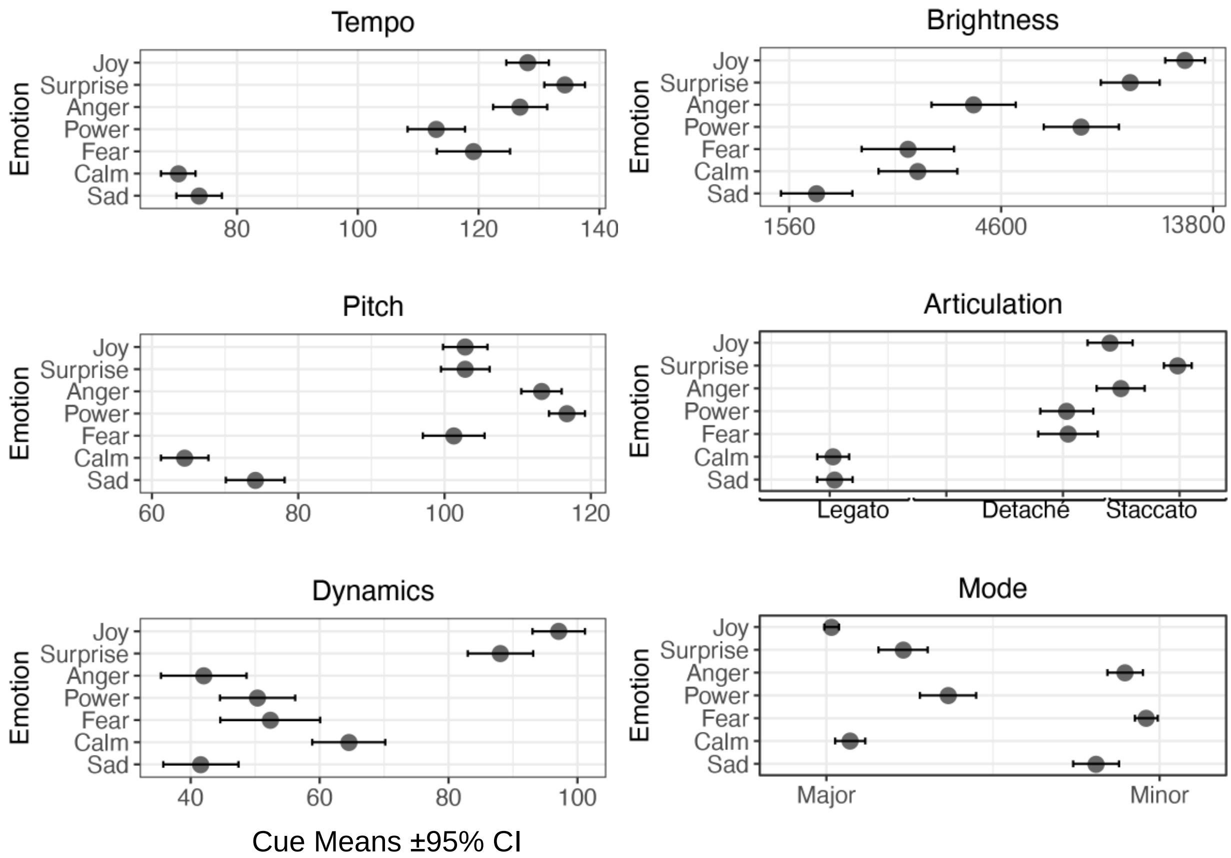


Figure 4.2. Means and 95% confidence intervals of cues utilised by participants to portray different emotions.

Notes. 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 harmonic minor levels.

4.3.2.1 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 (LDA) with the cues to predict the emotions. Linear discriminant analysis (LDA) is a statistical technique that allows to characterise objects into groups based on linear combinations of the features.

It is a variant of regression but LDA uses continuous independent variables to predict categorical dependent variable. In this case, the LDA algorithm uses the data collected from the six cues to predict a linear combination of features for the different emotion categories. 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 4.3 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 4.3 outlines cue combinations and their values that have a good percentage of predicting the intended emotion. The first seven columns in Table 4.3 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 4.3: 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 4.3 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 lowest R^2 values which indicate that they have low relevance in shaping the different emotional expressions in music. Furthermore, these small values may be a result of a greater proportion of variance in the settings. Further research where the variance is controlled may attempt to untangle this issue. 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.

Table 4.3. 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**

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

4.3.2.2 Correct Prediction of Emotions by Cue Selections

The confusion matrix in Table 4.4 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 ($\frac{sensitivity+specificity}{2}$, where sensitivity is the true positive rate and specificity is the true negative rate) presented as proportions (Chen, Liaw, & Breiman, 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.

Table 4.4. Confusion matrix displaying prediction proportion rates of the discriminant model to test data.

		Predicted						
		Joy	Surprise	Anger	Power	Fear	Calmness	Sadness
Intended	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

4.3.2.3 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 "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.

4.3.3 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.

4.3.3.1 Emotions Expressed in Cue Combinations

Table 4.5 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 & Friberg, 2011; Kragness & Trainor, 2019; Saarikallio et al., 2019; Saarikallio, Vuoskoski, &

Luck, 2014). The discrepancies between the current study's results and four previous production studies by Bresin and Friberg (2011), Saarikallio, Vuoskoski, and Luck (2014), Kragness and Trainor (2019), and Saarikallio et al. (2019) are denoted in Table 4.5 by numerical values in subscript. Since cue ranges varied across studies, the comparisons are relative ones rather than absolute values. Comparisons for power and surprise emotions could not be made as they were not investigated during the previous studies.

Table 4.5. 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	*leg.*	*stac.*	*leg.*	*stac.* [2]	*det.* [1]	*det.*	*stac.*
	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.* = detaché. For mode, + = major, - = minor. The numeric values in subscript refer to the following studies: 1 = Bresin & Friberg (2011), 2 = Saarikallio, Vuoskoski, & Luck (2014), 3 = Kragness & 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.

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, Gagnon, & Bouchard, 1998; Quinto, Thompson, & Taylor, 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, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 2014). Calmness was represented by a slow tempo (Sievers et al., 2013), *legato* articulation, low pitch, moderate dynamics, a rich sound, and major mode (Eerola, Friberg, & Bresin, 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, Friberg, & Bresin, 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, Friberg, & Bresin, 2013; K. 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, Vuoskoski, and Luck (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 4.5 as well as other previous studies not following a production approach (Gabrielsson & Juslin, 1996; Juslin &

Lindström, 2010; Quinto, Thompson, & Taylor, 2014). Saarikallio, Vuoskoski, and Luck (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, Vuoskoski, & Luck, 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, Vuoskoski, & Luck, 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. An alternative explanation may be that the musical pieces altered were composed in a way with the intention of conveying a particular emotion. Therefore, the low dynamics level may have been a result of changing musical pieces which were already structured to represent a different emotion. A closer look at the mean cue combinations used by participants to convey anger through the piece composed to express anger saw that a low dynamics level was also utilised by participants in their effort to communicate anger. This may suggest that the representation of anger by a soft dynamics level may be 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, Friberg, & Bresin, 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, Friberg, & Bresin, 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, 1940b). 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).

4.3.3.2 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 4.3 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) which proposes that the layering of cues determines the successful of emotion communication through music (see also Argstatter, 2016; Eerola et al., 2013; Gabrielsson, 2008; Ramos, Bueno, & Bigand, 2011). It is good to note that the lens model also mentions the potential interaction of cues as having an effect on the emotion communicated, however, previous studies have found little to no effect of cue interactivity on the communication of emotions, which is why this work has focussed on the additivity rather than interactivity of cues (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010; Scherer & Oshinsky, 1977).

4.3.3.3 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, 2020c) 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. Although deemed to be difficult to achieve (Quinto, Thompson, & Taylor, 2014; Ramos & Mello, 2021; Shoda & Adachi, 2012; Timmers & Ashley, 2007), providing participants with an emotionally neutral musical piece may have been beneficial to investigate whether certain cue combinations used by participants were influenced by the pieces, since they were already structured to portray different specific 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 (which currently has the limiting range of ±2 semitones) and mode (which currently gives the option to change from major to minor and vice-versa) 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.

4.4 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.

4.4.1 Method

4.4.1.1 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; J. 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.

4.4.1.2 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 repository[12].

4.4.1.3 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 at the end of this chapter.

4.4.2 Results

The consistency of how the participants used the individual rating scales was calculated using Cronbach's alpha (intraclass 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 4.6 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

[12] The 14 musical pieces can be found on the OSF repository by following this link https://osf.io/m26bu/

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 4.6 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 musical pieces, depending on their source. This can also be seen in the Piece × Source factor where all emotions had a significant effect on that interaction.

Table 4.6. 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 × 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. * $p < .05$, ** $p < .01$, *** $p < .001$. df=6 for Emotion, df=1 for Source, df=6 for Interaction for the likelihood ratio test.

Due to source having a significant impact on how emotions were rated, the rest of this experiment's analysis will regard the musical 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 was then filtered by source and subjected to a one-way repeated measures ANOVA to explore whether, in general, across the different pieces within each source, the participants rated the intended emotion of the pieces significantly different to the other available emotions. The intended emotion of each piece represented the independent variable whilst the collapsed ratings of the remaining emotion scales represented the dependent variable. The main effect of the intended emotion on the other emotion ratings 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 4.7 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. The standard deviation values of the mean emotion ratings are denoted in brackets. 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 musical 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 4.7 represent how significantly different the intended emotion's rating was to the other emotions' ratings. Exp.1 musical 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.

Table 4.7. Mean ratings of emotions perceived in the musical pieces.

		Intended Emotion (Musical Piece)													
		Anger		Calmness		Fear		Joy		Power		Sadness		Surprise	
	Source	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	2.00	1.15	1.24	2.57	1.84	2.07	2.36	2.29	1.96	1.10	1.09	**2.90**	**3.04**

	***	**	***	***	***	***	***	***	*	(1.12)	***	***	(1.34)	(1.32)
	(1.09)	(1.15)	(0.47)	(0.54)	(1.28)	(1.13)	(1.13)	(1.28)	(1.18)		(0.37)	(0.41)		

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.

4.4.3 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.

4.4.3.1 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, Friberg, & Bresin, 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, Vuoskoski, & Luck, 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, since the different instruments also contribute different kinds of articulation (e.g., there is more attack in piano than in strings). 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, Arnal, & Frühholz, 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, Wu, & Horner, 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 the 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, Friberg, & Bresin, 2013; Juslin & Lindström, 2010).

4.4.3.2 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.

4.5 General Discussion

In this chapter, three experiments 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 chapter also had two secondary aims: the first was to evaluate the newly-created polyphonic musical pieces presented in **Chapter 3** with regards to their expressed emotional content. These new pieces 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 in **Chapter 3**, 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, Grandjean, & Scherer, 2008), to be used as stimuli for research purposes, and be compatible with the apparatus used in our main experiment (Experiment 2). The excerpts created attempted to represent 'real' music that individuals may listen to in everyday life and were recorded as performances to also include human-like expression. On the other hand, previous stimuli tended to involve monophonic tone sequences, and only the minority of stimuli sets were created as polyphonic pieces. Secondly, this set of stimuli attempted to represent a bigger number of emotional expressions (nine emotions), rather than focussing on creating stimuli conveying the usual 3 basic emotions (happy, sad, and anger). Thirdly, the stimuli were strategically composed to allow a flexible amount of manipulation whilst retaining a delicate balance between ecological validity and experimental control (Eerola, Friberg, & Bresin, 2013; Gabrielsson & Lindström, 2010; Juslin & Lindström, 2010; Juslin & Västfjäll, 2008) that might not be achievable with commercial music. Most importantly, the pieces were subjected to a validation study, to explore whether the pieces were actual representative of their intended emotions, and it is good to note that only the minority of previous studies ran pilot tests to determine whether the stimuli created were conveying their intended emotion – only 3% of stimuli were pilot tested before using them in music and perceived emotions studies (Warrenburg, 2020).

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, adding potential new pieces to use as stimuli conveying sadness, happiness, anger, fear, calmness, as well as power and surprise. When looking at specifically-composed music pieces for music and perceived emotions studies in the PUMS database, which is a collection of stimuli used in 306 studies between 1928 and 2018, there are no excerpts which were composed to represent power and surprise. Therefore, these new stimuli have added 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.

If the stimuli set is consistently used in its original form, it may become familiar, if by any chance, the same individuals are participating in numerous studies using the stimuli. This can be easily controlled for by adding a familiarity question in future studies. As mentioned in the limitations of Experiment 1, the majority of participants who rated the stimuli set came from a Western culture (94.79%), therefore it would be beneficial to also test this stimuli set using participants from a non-Western culture.

Apart from the advantages of the stimuli set already mentioned above, the fact that the pieces allow for certain manipulations to be made to them, as shown in Experiment 2, is a useful property, and in fact, also a way to create variations of the original stimuli set. Although the purpose of this stimuli set was to provide new organic material with specific emotional content attributed to it which include emotions other than the usual three, future studies may find it useful to run an in-depth analysis of the musical properties, for example, using feature extraction with the MIRToolbox (Lartillot, Toiviainen, & Eerola, 2008). However, there are nevertheless certain limitations to the stimuli, such as the potential uneven spread of cues embedded into them, and overall musical style and the choice of textures and timbres is conventional and limited.

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, Bresin, & Sundberg, 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 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 & 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, Ferrer, & Alluri, 2012; Hailstone et al., 2009; Huron & Berec, 2009).

The experiments in this chapter 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, Friberg, & Bresin, 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 chapter 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, Thompson, & Matsunaga, 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, Grandjean, & Scherer, 2008).

The findings presented in this chapter 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. It is also interesting to note that the musical variation for anger created by participants in Experiment 2, which features an uncharacteristically low level of dynamics, was misidentified as fear in Experiment 3, which suggests that the dynamics level used by participants in Experiment 2 may have been result distinct to the particular musical pieces used. 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 work 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, Friberg, & Bresin, 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 affects how an emotion is shaped in the music (Argstatter, 2016; Eerola, Friberg, & Bresin, 2013). Therefore, this work 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 and 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; Saarikallio, Vuoskoski, & Luck, 2014). Furthermore, this chapter 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, 2020c) should also be studied to identify distinctions between emotional states that music, as a highly expressive medium, has to offer.

4.6 References

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., Busch, V., Lothwesen, K., Elvers, P., Fischinger, T., Schlemmer, K., & 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., Vieillard, 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

Brunswik, E. (1956). Perception and the representative design of psychological experiments. In *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, E., & Bonfiglioli, L. (2000). Psychological Connotations of Harmonic Musical Intervals. *Psychology of Music*, *28*(1), 4–22. https://doi.org/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, 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* (p. Vol.1, 44-45). SAGE Publications, Inc.

https://doi.org/http://dx.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., Friederici, A. D., & 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/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/oxfordhb/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

Goebl, W., & Bresin, R. (2001). *Are computer-controlled pianos a reliable tool in music* *performance research? Recording and reproduction precision of a Yamaha Diskavier 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*, *I*(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. (2019). *Musical emotions explained: Unlocking the secrets of musical affect.* Oxford University Press.

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 P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 452–489). 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*. Department of Pscyhology, Keele (CD-rom).

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. In *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* (Vol. 51, Issue 4, pp. 336–353). Canadian Psychological Association. https://doi.org/10.1037/1196-1961.51.4.336

Lahdelma, I., & Eerola, T. (2014). 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

Lahdelma, I., & Eerola, T. (2015). Theoretical Proposals on How Vertical Harmony May Convey Nostalgia and Longing in Music. *Empirical Musicology Review*, *10*(3), 245–263. https://doi.org/10.18061/emr.v10i3.4534

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*, *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/https://doi.org/10.1371/journal.pone.0222380

Mohn, C., Argstatter, H., & Wilker, F. W. (2010). 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. De. (2013). Robin: An Algorithmic Composer For Interactive Scenarios. *Sound and Music Computing Conference 2013, SMC 2013*, 207–212.

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*. University of Jyväskylä, Department of Music.

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. *Proceedings of 18th International Florida Artificial Intelligence Research Society*

Conference, 86–91.

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

Ramos, D., & Mello, E. K. S. (2021). Communication of emotions in music between Brazilian composer, performers, and listeners. *Psychomusicology: Music, Mind, and Brain*, *31*(1), 1–17. https://doi.org/10.1037/pmu0000269

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. K., & 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

Shoda, H., & Adachi, M. (2012). The Role of a Pianist's Affective and Structural Interpretations in his Expressive Body Movement: A Single Case Study. *Music Perception: An Interdisciplinary Journal*, *29*(3), 237–254.

https://doi.org/10.1525/mp.2012.29.3.237

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.

Timmers, R., & Ashley, R. (2004). *Communicating emotions through ornamentation*. *January*.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, *25*(2), 117–134. https://doi.org/10.1525/mp.2007.25.2.117

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. (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. R., Grandjean, D., & Scherer, K. R. (2008). Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, *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

4.7 Supplementary Material

4.7.1 Details of Musical Compositions for Experiment 1

The musical pieces were composed to be characteristic of their intended emotion which involved having partially similar as well as distinct features across the different emotions' excerpts. Sadness, calmness, and love excerpts featured a slow tempo, smooth note progressions, and stepwise motion in the melodic line, however, sad excerpts were written in minor mode whilst calm and love excerpts were in major mode. Excerpts composed to convey joy, anger, and power featured a fast tempo and short note durations. Joyful excerpts were composed in major mode and featured consonance and pitch leaps in the melodic line, whilst anger and power excerpts were composed in minor mode and featured note repetition. In general, fear, longing, and surprise excerpts had a moderate tempo, however, fear and surprise pieces featured dissonance whilst longing excerpts did not. Fear and longing excerpts had a constant rhythm and stepwise movement or small interval leaps in the melodic line, whilst surprise excerpts featured rhythmic variation and large pitch leaps. Table 4.8 provides more details of the compositional characteristics of the excerpts grouped by their intended emotion.

Table 4.8. Compositional characteristics of the musical excerpts.

		Compositional Characteristics				
		Tempo average in bpm (range)	Mode	Pitch range	Melody	Rhythm and Harmony
Excerpts	Sadness	75 (70 - 80)	Minor	D0 – F4	Stepwise or descending small intervallic leaps movement.	Constant rhythm and accompaniment were kept.

Joy	113 (110 - 120)	Major	F1 – A4	Melodies features pitch leaps and short note durations.	Constant rhythm was retained throughout. Excerpts were consonant.
Calmness	83 (60 – 100)	Major	B1 – E5	Stepwise movement, small intervallic leaps or arpeggiated melodic line.	Constant accompaniment and consonance kept throughout.
Anger	136 (100 – 180)	Minor	C1 – G5	Stepwise movement or small intervallic leaps in melodic line.	Excerpts featured dissonance. Repetitive notes with stepwise movement in accompaniment. No rhythmic variations.
Fear	103 (100 – 110)	Minor	A#0 – D5	Stepwise movement or small intervallic leaps in melodic line.	Excerpts features dissonance, a constant rhythm, repetitive notes in accompaniment.
Longing	103 (65 – 120)	Minor (3) & Major (1)	D1 – F5	Stepwise or arpeggiated movement in the melodic line.	Excerpts featured a constant rhythm and a sustained or arpeggiated accompaniment.

Power	153 (125 – 175)	Minor	C1 – G4	Stepwise movement or small intervallic leaps.	Quick repetitive short duration notes. No rhythmic variation.
Surprise	100 (all)	Major (1), Minor (2)	F1 – C5	Pitch leaps in melodic line.	Rhythmic variation, pitch leaps in both melody and accompaniment, and dissonance.
Love	77 (60 – 110)	Major	F1 – B4	Stepwise movement or small intervallic leaps in melodic line.	Excerpts featured consonance and a constant rhythm. Arpeggios were present in accompaniment.

All musical excerpts were rendered in Logic Pro X utilising a grand piano virtual instrument from the Vienna Symphonic Library (VSL) sound library. The musical pieces are available on OSF via this link: https://osf.io/m26bu/

4.7.2 Additional Analysis for Experiment 2

4.7.2.1 Cue Usage of Participants for the Different Musical Pieces

A linear mixed model (LMM) was run for each of tempo, articulation, brightness, pitch, and dynamics cues with Piece as fixed factor and Participant as random factor. A general mixed model (GLMM) with a binomial distribution was run for the mode cue due to its binary nature, with Piece as fixed factor and Participant as random factor. Table 4.9 presents a breakdown of how the cues were utilised with the different musical pieces. Mostly, a non-significant effect was obtained between the cues in relation to the musical pieces, with some exceptions. The piece composed and validated as conveying calmness had a significant effect on the utilisation of the tempo, pitch, and mode cues when portraying different emotions. The piece composed and validated as conveying joy had a significant effect on the mode cue. The piece composed to convey power affected the usage of the dynamics cue. Tempo and articulation cues had significant effects in relation to the piece composed to convey sadness. Finally, the piece originally composed to convey surprise had a significant effect on the tempo cue.

Table 4.9. LMM estimates and 2.5% and 97.5% confidence intervals of the seven musical pieces in relation to the six cues.

The different musical pieces, each composed to convey one specific emotion		Cues					
		Tempo	Articulation	Pitch	Dynamics	Brightness	Mode
	Anger Piece	-1.71 (-8.19, 4.75)	5.01 (-2.31, 12.29)	1.23 (-3.55, 6.01)	3.35 (-4.11, 10.86)	-0.62 (-4.81, 3.55)	0.33 (-0.05, 0.71)
	Calmness Piece	-13.85** (-22.11, -5.57)	-8.46 (-17.93, 1.06)	-11.30*** (-17.43, -5.15)	0.26 (-9.50, 9.97)	-2.94 (-8.37, 2.52)	-0.68** (-1.19, -0.18)
	Fear Piece	3.34 (-4.88, 11.57)	-1.34 (-10.79, 8.16)	-0.57 (-6.65, 5.51)	-6.12 (-15.85, 3.57)	-2.36 (-7.74, 3.03)	0.23 (-0.27, 0.73)
	Joy Piece	-0.18 (-8.40, 8.06)	-3.61 (-13.07, 5.88)	-0.79 (-6.86, 5.29)	3.31 (-6.42, 13.00)	3.41 (-1.97, 8.80)	-0.75** (-1.26, -0.25)
	Power Piece	1.58 (-6.62, 9.79)	-1.67 (-11.13, 7.80)	5.59 (-0.45, 11.63)	-11.54* (-21.25, -1.86)	2.55 (-2.79, 7.91)	-0.32 (-0.82, 0.18)
	Sadness Piece	8.49* (0.23, 16.77)	-14.74** (-24.21, -5.22)	-2.77 (-8.90, 3.38)	-6.59 (-16.35, 3.12)	0.06 (-5.37, 5.52)	-0.07 (-0.57, 0.43)
	Surprise Piece	10.83* (2.57, 19.11)	-0.51 (-9.98, 9.01)	0.54 (-5.60, 6.69)	0.98 (-8.78, 10.68)	2.52 (-2.91, 7.98)	-0.39 (-0.89, 0.11)

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$. Since ANOVA models calculate k-1 predictor levels, the LMM was rendering one missing factor (the Anger piece in this case). In order to calculate LMMs for the missing factor and ensure consistency in results across factors, LMMs were also run individually for each of the factors.

4.7.3 Experiment 1 Information Sheet, Demographic Questions, Format of Rating Task, and Additional Method Details

The information regarding Experiment 1 given to the participants is detailed below, together with additional method details not described in the main chapter. First, additional information on the procedure of the experiment is given, followed by the information sheet and instructions which were available to participants. The demographic questions participants were asked to answer are then listed. The instructions and format of the practice trial and rating task is also presented. The format of the rating task (and practice trial) was always the same, which is why one example is given here.

4.7.3.1 Additional Procedure Details

The online survey tool Qualtrics was utilised to carry out this experiment. Written instructions explained the difference between the emotion being expressed by the music (perceived emotion) and the emotion that the listener feels/experiences while listening to a piece of music (felt emotion) to ensure that participants understood which emotional process they will be assessing. Participants were subjected to a practice trial, which entailed listening to a musical piece outside of the stimuli set and rating on the nine separate five-point Likert emotion scales which emotions they thought the music was expressing.

4.7.3.2 Information Sheet

This study aims to investigate the perceived emotional expression in music, which emotions are expressed through music. It should take approximately 20 minutes to complete. Audio will be played in this survey; please ensure your device has sound, speakers or headphones are switched on and volume is not muted.

Test your sound now. (Sound test)

What is the purpose of this study?

This study aims to assess whether musical tracks can convey the intended emotion(s) to the listener. This study will contribute to a research project about recognising what features of a musical composition affect which emotions are expressed through it.

What does this study entail?

Throughout this experiment, you will be asked to listen to several musical tracks. All tracks are short piano pieces. For each track, you will be asked to identify which emotion(s) from a provided list you think the music is expressing.

There are two types of emotion that can be studied: *felt* and *perceived* emotion. *Felt* emotion is what the listener feels (experiences) while listening to a piece of music. *Perceived* emotion is the emotion being expressed by the music. In this experiment, you will be asked to assess what *perceived emotion* you think the music is conveying, i.e., not what the music is making you feel, but what you think the music in itself is expressing.

Why have I been approached?

Anyone over 18 is eligible to take part.

Participation

Your participation in this survey is voluntary. You may refuse to take part in the research or stop answering questions and leave the survey at any time.

Technical Requirements

Music tracks will be played throughout the survey; please ensure your device has sound, speakers or headphones are switched on and volume is not muted.

If using headphones, please ensure that the volume is not too loud, to prevent any hearing damage.

What are the benefits?

This study will aid in better understanding what features in the internal structure of a musical piece affect the piece's emotional expressivity. The findings will contribute to current research on how music is perceived to express specific emotions.

What are the risks?

There are no foreseeable risks involved in participating in this study.

What I something goes wrong?

It is highly unlikely that anything will go wrong. However, if you feel something has gone wrong or have any question, please contact the researcher Annaliese Micallef Grimaud at annaliese.micallef-grimaud@durham.ac.uk.

Confidentiality

The study is completely anonymous: participants will be given an automatically generate code and their identity will not be requested. Completely anonymised data may be shared with other researchers to support future research.

Contact information

If you have questions at any time about the study or the procedures, you may contact the researcher at annaliese.micallef-grimaud@durham.ac.uk.

Data Protection Information

When you start, this survey will store your answers and browser information on a secure Qualtrics data server. The responsibility for this survey rests entirely with the researcher(s) listed above. The audio played in this survey is stored in SoundCloud. The SoundCloud player uses cookies in accordance with their Cookie Policy which can be found here: http://soundcloud.com/pages/cookies
Usage data may be collected by SoundCloud for analytics purposes.

Confirm you want to do this survey

I confirm that:

- I have read the above information
- I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason
- I agree to take part in this study
- I consent to any data I submit being stored and used for academic research
- I am 18 years of age or older

o **I understand all the conditions of this study**

4.7.3.3 Demographic Questions

How old are you?

______

What is your gender?

- o Male
- o Female
- o Other ________
- o Prefer not to say

Which title best describes you?

- o Non-musician
- o Music-loving non-musician
- o Amateur musician
- o Serious amateur musician
- o Semi-professional musician
- o Professional musician

What is your country of origin?

<drop down menu with countries>

In which country do you currently reside?

<drop down menu with countries>

How proficient are you in the English language?

- o Good/fluent user
- o Competent user
- o Modest user
- o Limited user
- o Extremely limited user

What are your 3 preferred music genres?

________, _________, _________

4.7.3.4 Instructions and Format of Practice Trial and Rating Task

Practice Trial Instructions

You will now be presented with a musical track, similar to the ones used for this survey, so you can become familiar with the style of the music that will be played (instrumental piano pieces), and the emotion rating scales. There will be 9 emotion terms. For each track you listen to, you have to assess which emotion(s) you think the music is trying to convey.

Rating the Music

For each track, all emotion categories have to be rated on the scale (from 1 to 5; 1 being none and 5 being the most). If for example, you think a piece of music is expressing sadness, the sadness emotion is rated as '5 ' (i.e., the track is expressing mostly sadness). However, the other emotions have to be rated as well. So, if a track is only conveying sadness for example, all the other emotion terms are marked as '1', which means that none of those other emotions are being conveyed by the music.

If you perceive more than one emotion in a musical track, rate on the scale to which extent you think the music is expressing that particular emotion, with a rating of '5' being the most dominant emotion. It is important to rate **how much of each emotion** you think the music is conveying.

Practice Trial and Rating Task Format

Listen to the musical track and assess which emotion(s) you think the music is intending to express. For each emotion term, mark on the scale how much of that particular emotion (if any), you think the music is expressing.

<u>Scale: 1 being no emotion at all to 5 being the most emotion it is expressing.</u>

[embedded music link] Press the play button to listen to the track. You can replay the track as much as needed.

How much of each emotion do you think the music is expressing?

	1 (none at all)	2 (a little)	3 (a moderate amount)	4 (a lot)	5 (the most)
Calmness	○	○	○	○	○
Sadness	○	○	○	○	○
Power	○	○	○	○	○
Love	○	○	○	○	○
Fear	○	○	○	○	○
Anger	○	○	○	○	○
Longing	○	○	○	○	○
Joy	○	○	○	○	○
Surprise	○	○	○	○	○

Now that you are familiar with the format, you can start the experiment. You will listen to 28 pieces of music presented to you in a random order. For each musical piece, you will have to rate how much of each emotion listed is being expressed (if at all) by the music.

4.7.4 Experiment 2 Information Sheet, Instructions, Additional Procedure Details, and Demographic Questions

The information sheet and written instructions given to participants prior to Experiment 2 are detailed below as well as additional procedure details, the verbatim explanation given during the brief demonstration, and the demographic questions.

4.7.4.1 Information Sheet

This study aims to investigate how altering different parameters in the structure of a musical composition affects the emotional expression conveyed by the music. It should take approximately 30 minutes to complete.

What is the purpose of this study?

This study aims to investigate how people manipulate different features (tempo, timbre, articulation, pitch, dynamics, and mode) of a musical piece to convey a specific emotional expression. This study forms part of a research project, which will contribute to identifying how emotions are expressed through music.

Why have I been approached?

Anyone over 18 is eligible to take part.

Participation

Your participation in this study is voluntary. You may refuse to take part in the research or stop your participation in this study at any time.

What are the benefits?

This study will aid in better understanding what features in the internal structure of a musical piece affect the piece's emotional expressivity. The findings will contribute to current research on how music is perceived to express specific emotions.

What are the risks?

There are no foreseeable risks involved in participating in this study.

What if something goes wrong?

It is highly unlikely that anything will go wrong. However, if you feel something has gone wrong or if have any questions at any time about the study or the procedures, please contact the researcher Annaliese Micallef Grimaud on annaliese.micallef-grimaud@durham.ac.uk.

Confidentiality

The study is completely anonymous: participants will be given an automatically generated code and their real identity will not be requested. Completely anonymised data may be shared with other researchers to support future research.

Data Protection Information

This study will record all your answers and store them on a secure hard drive attached to the computer used. The responsibility for this study rests entirely with the researcher(s) listed above.

4.7.4.2 Instructions

In this experiment you will be asked to listen to different musical tracks. For each musical track, you will be presented with a word that describes a particular emotion. Your task is to alter how the music sounds in a way that you think makes it best convey the emotion word given. You will do this by changing the values of the 5 sliders and toggling on/off the button shown on the computer screen. Each slider (and the button) represents a musical feature of the track. As you move the sliders, the music will change. There is no time limit for this experiment. Each short musical piece can be looped by pressing a button on the screen, so that you listen to it and change it as much as you like. When you are satisfied with how the music sounds and convinced that the music is expressing the emotion word given, unloop the track and the researcher will stop the music. The researcher will load the next musical track, and you will be asked to do the same procedure. It will not be possible to go back to a previous piece of music. Please feel free to ask any questions at any time. Thank you in advance for your participation.

4.7.4.3 Consent Form

Study: Emotion Manipulation through Music

Researcher: Annaliese Micallef Grimaud

Confirm you want to do this study

I confirm that:

• I have read the above information	
• I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason	
• I agree to take part in this study	
• I consent to any data I submit being stored and used for academic research	
• I am 18 years of age or older	
• I understand the conditions of this study	

Participant

Name **Signature** **Date**

…………………………………………………… ……………………………………………… ……………………

Researcher

Name **Signature** **Date**

…………………………………………………… ……………………………………………… ……………………

4.7.4.4 Details of the Researcher's Demonstration and Participants' Practice Trial

The researcher gave a short demonstration of how the *EmoteControl* interface works prior to participants starting the musical task part of the experiment. The researcher explained in layperson terms what each of the six cues did to the music (e.g., the tempo cue speeds up or slows down the music, the dynamics cue makes the music sound louder or softer) but no information was given on how the cues affect the musical expressions, as we wanted participants to carry out the task using their intuition. The participants then had a couple of minutes to try out the interface to familiarise themselves with it. Participants were also subjected to a practice trial where they altered a musical piece outside of the actual pool of musical stimuli via the six available cues to try and convey the emotion sadness. Participants could ask questions pertaining to the use of the interface at any time.

4.7.4.5 Verbatim Explanation of Cues Accompanying Interface Demonstration

A verbatim explanation of what the cues did was also given to the participants while the research briefly demonstrated the changes that the individual cues made on a musical piece:

"The tempo cue is changed via this slider. The higher the slider level, the faster the speed of the music. The lower the slider value, the slows the speed of the music. Articulation controls the playing method of the instrument. The lowest level is *legato* which means that notes are played smoothly from one to the other. The next level is *detaché* where notes are played separated, and the top level is *staccato* where notes are played with a short duration. These are three different levels and there is no in-between. The pitch cue controls how low or high the instrument plays. The lower the slider value, the lower the pitch. The higher the slider value, the higher the pitch. The dynamics cue controls the loudness of the instrument. The higher the slider value, the louder the sound, and the lower the slider value, the softer the sound. The brightness slider controls how dull or bright the sound is. The lower the slider value, the duller the sound is as fewer high frequencies are played, while the higher the slider, the brighter the sound as more high frequencies are played. The mode cue works with a toggle button as it switches between two mode settings. The difference between the two settings is that when the toggle button is on and shows an 'x', two notes of the musical scale being used are lowered, and therefore, you may hear a difference in the music being played. You can change the cues as much as you'd like."

4.7.4.6 Demographic Questions

How old are you?

____

What is your gender?

- Male
- Female
- Other ________
- Prefer not to say

What is your home country?

<drop down menu with countries>

Which title best describes you?

- Non-musician
- Music-loving non-musician
- Amateur musician
- Serious amateur musician
- Semi-professional musician
- Professional musician

Have you played/do you play a music instrument?

- Yes (if yes, please write down instrument:) ______________
- No

If you are a musician/singer, which category comes nearest to the amount of time you spend practicing an instrument (or voice)?

- I rarely or never practice singing or playing an instrument
- About 1 hour per month
- About 1 hour per week
- About 15 minutes per day
- About 1 hour per day
- More than 2 hours per day

What is your preferred music of choice?

____

On average, how many hours do you purposely listen to music in a day?

____

4.7.5 Experiment 3 Information Sheet, Demographic Questions, and Format of Rating Task

The information given to participants in Experiment 3 is detailed below. First, the information sheet and instructions are reported. The demographic questions are then listed. Finally, the instructions and format of the practice trial and rating task is presented. The rating task and practice trial had the same format, which is why one example is given here.

4.7.5.1 Information About This Survey

This study aims to investigate the identification of emotions in music; which emotion(s) are conveyed to the listener through music. It should take approximately 15 minutes to complete.

Audio will be played in this survey; please ensure your device has sound, speakers or headphones are switched on and volume is not muted.

Test you sound now. <audio clip here>

What is the purpose of this study?

This study aims to assess whether musical pieces can convey the intended emotion to the listener. This study forms part of research project which will contribute to recognising what features of a musical composition affect which emotions are expressed through it.

What does this study entail?

In this online experiment, you will be asked to listen to 14 different, short instrumental musical pieces. For each musical piece, you will be asked to identify which emotion(s) from a provided list you think the music is expressing.

There are two types of emotion that can be studied: *felt* and *perceived* emotion. *Felt* emotion is what the listener feels (experiences) while listening to a piece of music. *Perceived* emotion is the emotion being expressed by the music. In this experiment, you will be asked to assess what *perceived emotion* you think the music is conveying, i.e., not what the music is making you feel, but what you think the music in itself is expressing.

Why have I been approached?

Anyone over 18 is eligible to take part.

Participation

Your participation in this survey is voluntary. You may refuse to take part in the research or stop answering questions and leave the survey at any time.

What are the benefits?

This study will aid in better understanding what features in the internal structure of a musical piece affect the piece's emotional expressivity. The findings will contribute to current research on how music is perceived to express specific emotions.

What are the risks?

There are no foreseeable risks involved in participating in this study.

What I something goes wrong?

It is highly unlikely that anything will go wrong. However, if you feel something has gone wrong or have any question, please contact the researcher Annaliese Micallef Grimaud at annaliese.micallef-grimaud@durham.ac.uk.

Confidentiality

The study is completely anonymous: at no point will the participant's identity be requested. After data collection, participants will only be known by an automatically generated code for data analysis purposes. Completely anonymised data may be shared with other researchers to support future research.

Technical Requirements

Music tracks will be played throughout the survey; please ensure your device has sound, speakers or headphones are switched on and volume is not muted.

If using headphones, please ensure that the volume is not too loud, to prevent any hearing damage.

Data Protection Information

When you start, this survey will store your answers and browser information on a secure Qualtrics data server. The responsibility for this survey rests entirely with the researcher(s) listed above. The audio played in this survey is stored in SoundCloud. The SoundCloud player uses cookies in accordance with their Cookie Policy which can be found here: http://soundcloud.com/pages/cookies Usage data may be collected by SoundCloud for analytics purposes.

Confirm you want to do this survey

I confirm that:

- I have read the above information
- I understand that my participation is voluntary and that I am free to withdraw at any time during the survey without giving any reason
- I agree to take part in this study
- I consent to any data I submit being stored and used for academic research
- I am 18 years of age or older

○ **I understand all the conditions of this study**

4.7.5.2 Demographic Questions

How old are you?

____

What is your gender?

- Male
- Female
- Other ________
- Prefer not to say

What is your home country?

<drop down menu with countries>

Which title best describes you?

- Non-musician
- Music-loving non-musician
- Amateur musician
- Serious amateur musician
- Semi-professional musician
- Professional musician

What is your preferred music genre?

____

Have you played/do you play a music instrument?

- Yes (if yes, please write down your instrument, which includes voice:)

  ____
- No

How many years of musical training have you had?

____

4.7.5.3 Instructions and Format of Practice Trial and Rating Task

Practice Trial Instructions

You will now be presented with a musical track, similar to the ones used for this survey, so you can become familiar with the style of the music that will be played (instrumental piano pieces), and the emotion rating scales. For each track you listen to, you have to assess which emotion(s) from the 7 listed emotions you think the music is trying to convey.

Rating the Music

For each track, all emotion categories have to be rated on the scale (from 1 to 5; 1 being 'none at all' and 5 being 'the most'). If for example, you think a piece of music is expressing sadness, the emotion term 'sadness' is rated as '5 ' (i.e., the emotion most expressed). However, the other emotion terms have to be rated as well. So, if a track is only conveying sadness for example, all the other emotion terms are marked as '1' (none at all), which means that none of those other emotions are being conveyed by the music.

If you perceive more than one emotion in a musical track, rate on the scale to which extent you think the music is expressing that particular emotion, with a rating of '5' being the most dominant emotion. It is important to rate **how much of each emotion** you think the music is conveying.

Practice Trial and Rating Task Format

Listen to the musical track and assess which emotion(s) you think the music is intending to express. For each emotion term, mark on the scale how much of that particular emotion (if any), you think the music is expressing.

Scale: 1 being no emotion at all to 5 being the most emotion it is expressing.

[embedded music link] Press the play button to listen to the track. You can replay the track as much as needed.

How much of each emotion do you think the music is expressing?

	1 (none at all)	2 (a little)	3 (a moderate amount)	4 (a lot)	5 (the most)
Sadness	○	○	○	○	○
Calmness	○	○	○	○	○
Joy	○	○	○	○	○
Anger	○	○	○	○	○
Fear	○	○	○	○	○
Power	○	○	○	○	○
Surprise	○	○	○	○	○

Chapter 5. Emotional Expression Through Musical Cues: A Comparison of Production and Evaluation Approaches

Micallef Grimaud, A., & Eerola, T. (under review). Emotional expression through musical cues: A comparison of production and evaluation approaches. *PLOS One.*

Preface

Chapter 5 details the last two empirical experiments carried out in the current project. The first experiment in this chapter employs a production approach (similar to Experiment 2 reported in Chapter 4) and uses a second version of the *EmoteControl* (V2.0) interface to alter *seven* cues (tempo, pitch, articulation, brightness, mode, dynamics, and instrumentation; with instrumentation being the new addition) of the musical excerpts utilised in Chapter 4, with the aim of simultaneously exploring *more* of the cue-emotion mapping space. Apart from the addition of the instrumentation cue, the second version of *EmoteControl* also features changes in terms used for the different cues and some alterations with respect to cue ranges/levels, following feedback gathered from users in the interface evaluation study (described in **Chapter 2**). All cue details of *EmoteControl* V2.0 are denoted in this chapter. A second aim of this chapter was to investigate whether similar results would be attained if a traditional systematic manipulation approach is used as against a production approach. Therefore, the second experiment detailed in this chapter uses a systematic manipulation approach, where different cue levels and combinations are computed through a fractional factorial design to render variations of the musical excerpts being used as stimuli in the production study. The emotional content of these variations is then rated on separate emotion Likert scales by participants. The results of the two experiments are compared, highlighting similarities and differences, together with pros and cons of the two approaches used, with the aim of providing insight on best approaches to the exploration of the cue-emotion space.

Abstract

We report two experiments to investigate the role of seven musical cues (tempo, pitch, dynamics, brightness, articulation, mode, and instrumentation) in communicating seven emotional expressions (sadness, joy, calmness, anger, fear, power, and surprise) in music. The first experiment utilised a production paradigm where participants adjusted the cues in real-time to convey each target emotion. The second experiment used an evaluation approach where participants rated pre-rendered systematic variations of the stimuli for all emotions. Both methods provided consistent patterns of cue-emotion mappings. However, the production approach offered more information about the cues. Future approaches to emotional expression in music are discussed.

5.1 Introduction

An important aspect of music is that it can communicate different emotional expressions (Juslin & Laukka, 2004; Lindström et al., 2003). A substantial amount of previous literature suggests that composers and performers can successfully encode a specific emotional expression in the music using particular musical cues (i.e., properties of the music) to communicate it to the listeners. In turn, listeners use these same cues to decode and, in general, can recognise the intended emotion conveyed (Juslin, 1997a, 2013b; Juslin & Lindström, 2010). Tempo, mode, pitch level, dynamics, timbre, rhythm, melodic range and direction, and harmony have all been identified as emotion cues (for an overview, see Juslin & Lindström, 2010). Understanding how musical cues affect the different emotions communicated through real music to the listeners has important applications, such as investigating emotion development and regulation in children and teenagers (Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014), utilising music as a medium for non-verbal patients (Silverman, 2008), and encapsulating specific branding identities in music for marketing purposes (Lepa et al., 2020).

Although musical cues and emotional expression have been investigated over the last century, research has only scratched the surface of how musical cues operate and shape the different emotion profiles. A number of studies have suggested that it is the additivity of musical cues that helps convey different emotions in the music, rather than the effect of an individual cue (Argstatter, 2016; Eerola, Friberg, & Bresin, 2013; Hevner, 1936; Lindström et al., 2003). However, the role of multiple musical cues as a combination has not been investigated as much (Gabrielsson & Lindström, 2001; Juslin & Lindström, 2010; Lindström, 2006). Previous research tended to focus on one musical cue, such as mode (Hevner, 1935; Kastner & Crowder, 1990), timbre (Behrens & Green, 1993; Chau, Wu, & Horner, 2015; Eerola, Ferrer, & Alluri, 2012; Elliott, Hamilton, & Theunissen, 2013; Hailstone et al., 2009), melody (Lindström, 2006), harmony (Lahdelma & Eerola, 2016a), or harmonic intervals (Costa & Nese, 2020). Other studies investigated two to three cues simultaneously, each with a limited number of variations/levels (e.g., tempo fast/slow) (Dalla Bella et al., 2001; Fritz et al., 2009; Hevner, 1937; Lindström, 2006; Scherer & Oshinsky, 1977), and only a few studies have tried to explore a bigger cue parameter space with seven or eight cues and multiple cue levels simultaneously and their interactions (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010).

The most prominent methodology used to investigate how musical cues affect the emotional expression communicated through music is the systematic manipulation approach, where similar musical excerpts are created by slightly varying the levels of different cues. These musical variations would then be assessed by participants evaluating the excerpts in terms of emotional expression (Eerola, Friberg, & Bresin, 2013; Hevner, 1935, 1936, 1937; Juslin, Friberg, & Bresin, 2002; Juslin & Lindström, 2010). This systematic manipulation and evaluation approach allows for minute changes in musical cues to be investigated, with complete experimental control. However, each systematic variation produces another musical stimulus that participants would need to listen to and evaluate. Therefore, the number of cues and cue level combinations that can be investigated simultaneously utilising a systematic manipulation design is limited, as a design with a large number of cue combinations becomes quickly unfeasible (Juslin, 1997b). Furthermore, running numerous systematic variations on a musical stimulus might tamper with the ecological validity of the music (Gabrielsson & Lindström, 2010).

An alternative method used in musical cues and emotion research is the production approach, where participants are in charge of changing a selection of musical cues in real-time to express different emotions through music. This methodology is referred to as analysis-by-synthesis (Friberg, Bresin, & Sundberg, 2014), and this interactive paradigm allows for a larger parameter space to be explored, as cue levels and combinations do not need to be pre-defined and rendered. Only a few studies have employed this methodology, with participants using either physical or digital sliders (Bresin & Friberg, 2011; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Sievers et al., 2013), or a one-key apparatus (Kragness & Trainor, 2016, 2019) to express three to five emotions by controlling three to seven cues. A downside to these studies was that cues only controlled either the melodic part of the musical stimuli or Bach chorales chord sequences, which perhaps are not the best representatives of real music.

Other researchers have used correlation studies to assess which cues help communicate certain emotions in music. This is usually attained by first asking composers to create music expressing different emotions or using already existing music from a repertoire, then asking listeners to assess which emotion or valence/arousal state is being portrayed by the music and finally, analysing the score to identify which cue combinations correlate to different emotions (Battcock & Schutz, 2019; Quinto, Thompson, & Taylor, 2014; Thompson & Robitaille, 1992). However, this methodology does not allow for the dissociation of the cues used. Thus, findings can only describe the effect of specific cue combinations, which cannot be freely changed and do not tell us the causal effect of the individual cues.

The use of different methodologies and their specific limitations begs the question of which methodology should be used to explore better the large cue space that exists and the cue combinations that help shape different emotions in real music. This leads to other questions of how reliable the methodologies are and whether their results converge.

This chapter aims to provide new information on how a selection of musical cues affects the emotion expressed in music using two approaches. Firstly, a production experiment will be carried out (Experiment 1), where participants will change musical expression by manipulating selected cues to shape different emotions in the music. This will allow exploring a substantial cue space at once. Secondly, an evaluation experiment will be carried out (Experiment 2). A pre-defined number of cue combinations and levels of the same cues used in Experiment 1 will be systematically manipulated and the emotions expressed will be evaluated by participants. This chapter will thus present a critical evaluation of two methodologies used in music and emotion research.

A combination of seven musical cues will be investigated in the present study: tempo, pitch, dynamics, brightness, articulation, mode, and instrumentation. Previous literature suggests that tempo, mode, and dynamics are three of the strongest contributing factors in emotional expression in music (Dalla Bella et al., 2001; Kamenetsky, Hill, & Trehub, 1997; Morreale, Masu, & Angeli, 2013). Pitch, articulation, brightness, and instrumentation have also been linked to affecting emotion perception in music (Behrens & Green, 1993; Eerola, Friberg, & Bresin, 2013; Hailstone et al., 2009; Juslin & Lindström, 2010; Quinto, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 2014); however, they have not been studied as much as the former cues. Therefore, this chapter aims to provide new data on these cues and their combinations and identify their role in shaping seven different emotional expressions in the music. The emotional expressions investigated in this chapter are sadness, joy, calmness, anger, fear, power, and surprise. Previous literature suggests that these seven emotions may be expressed through music (Juslin, 2013b; Juslin & Laukka, 2004; Lindström et al., 2003; Turnbull et al., 2008), with joy, sadness, anger, and fear being the most accurately recognised emotions, also cross-culturally (Balkwill & Thompson, 1999; Fritz et al., 2009; Kragness & Trainor, 2019; Laukka et al., 2013). Furthermore, these emotions cover a broad range on the emotion spectrum (Plutchik, 2001) and the valence-arousal circumplex model (Russell, 1980).

In summary, in this study, we investigate seven musical cues (tempo, pitch, dynamics, brightness, articulation, mode, and instrumentation) in relation to seven emotional expressions (sadness, joy, calmness, anger, fear, power, and surprise) with the intent of exploring the rich and complex cue space that underlies expression in music. Most importantly, we compare two approaches designed for mapping the cues and emotions. Experiment 1 is a production study and follows an analysis-by-synthesis methodology. Experiment 2 entails a systematic manipulation design to create the different cue combinations and an evaluation task to determine how the change in cues affects the emotion expressed in the music. Our research questions are:

(1) How do the musical cues and their combinations contribute to the expression of different emotions in music?
(2) To what extent do the results from the two experiments converge?

The first section details the production experiment (Experiment 1). The second section reports the systematic manipulation and evaluation experiment (Experiment 2). The following section compares the findings of Experiment 1 and 2, highlighting similarities and differences between the two experiments and the existing literature. Finally, the last section outlines the pros and cons of the two approaches utilised and gives insight on methodological considerations for future studies.

5.2 Experiment 1: Production Approach

This experiment investigated the influence of the seven musical cues (tempo, articulation, pitch, dynamics, brightness, mode, and instrumentation) on the perceived emotional expression in music using a production approach. In this experiment, participants actively engaged with a computer interface called *EmoteControl* (Micallef Grimaud & Eerola, 2021), allowing users to alter musical pieces via tempo, articulation, pitch, dynamics, brightness, mode, and instrumentation cues. Furthermore, it allows us to gain a deeper insight into tempo, pitch, dynamics, and brightness as they are not confined to pre-determined distinct cue levels, whilst also exploring different levels of articulation, mode, timbre, and the cue combinations created by these seven cues.

5.2.1 Method

5.2.1.1 Participants

Participants were recruited via social media and word-of-mouth. Forty-two participants (19 males, 23 females) between the ages of 20 and 68 years (M = 34.45, SD = 13.64) took part in the study. A one-question version of the Ollen Music Sophistication Index (OMSI) (Ollen, 2006; Zhang & Schubert, 2019) was utilised to distinguish between the participants' levels of musical expertise. Eight of the participants were musicians, whilst the remaining 34 were non-musicians. Participation in the study was voluntary.

5.2.1.2 Material

Seven tonal, instrumental musical pieces were utilised as material[13]. These pieces were derived from a previously validated musical stimulus set specifically composed to be utilised with the *EmoteControl* interface[14]. The seven musical excerpts were validated via an online listening study as conveying one of the

[13] The stimuli can be found in the Experiment 1: Production Approach, MIDI files folder following this link: https://osf.io/atxhk/?view_only=6d0ef66b819e476dbfcc40b54ae6c986

[14] The creation of stimuli is detailed in Chapter 3, and the validation study is described in Chapter 4 Experiment 1.

following seven emotions: joy, sadness, calmness, power, anger, fear, or surprise by having participants rate on Likert scales how much of each emotion was portrayed in the excerpts (Micallef Grimaud & Eerola, 2022). In this current study, participants were presented with these seven musical excerpts and asked to convey each of the seven emotions attributed to the stimulus set (joy, sadness, calmness, power, anger, fear, surprise) in all the seven excerpts.

5.2.1.3 Apparatus

A second version of the computer interface *EmoteControl* (V2.0) was utilised for this experiment (Micallef Grimaud & Eerola, 2021). The second version of the interface allows users to alter seven cues[15] (tempo, articulation, brightness, pitch, dynamics, mode, and the addition of the instrumentation cue) of instrumental musical pieces in MIDI format. A representation of the interface can be seen in Figure 5.1.

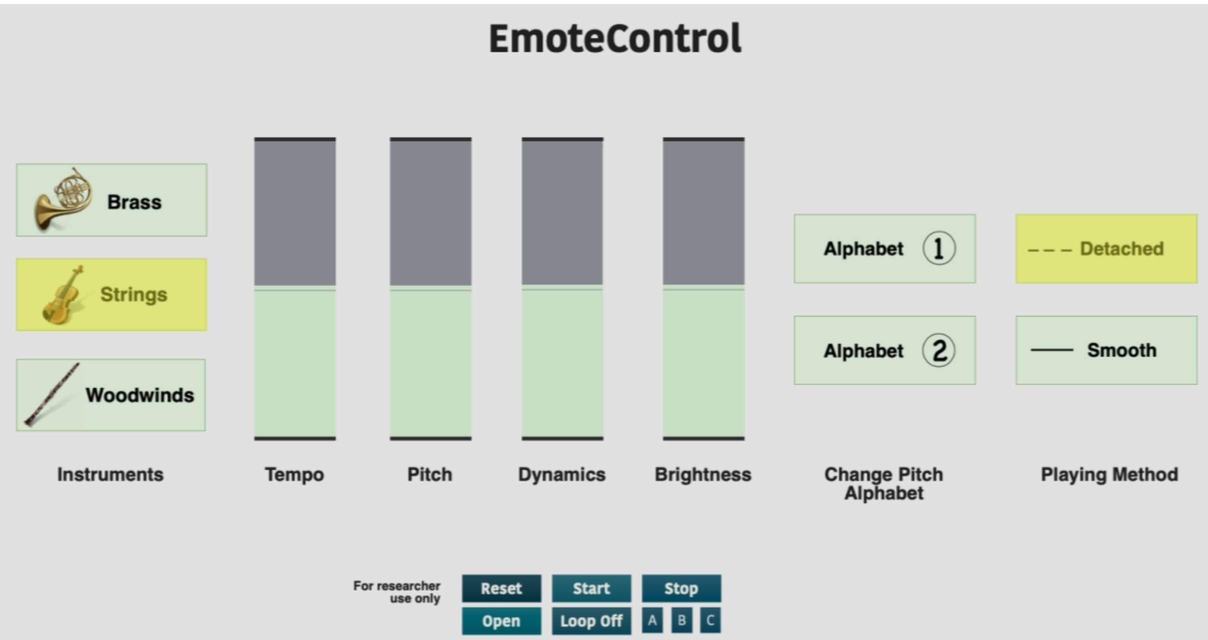


Figure 5.1. The *EmoteControl* V2.0 user interface.

[15] The first version of the interface which allows manipulations of 6 cues is detailed in **Chapters 2** and **4**.

The interface is aimed at a general population, and no prior musical skills are required to utilise the interface. For this reason, and following feedback gained from non-musicians who tested out the interface in **Chapter 2**, certain terms that are music-specific, such as mode and articulation, were renamed into more general terms. As can be seen in Figure 5.1, Mode is referred to as 'Change Pitch Alphabet', with Alphabet 1 being Major mode and Alphabet 2 being Minor. Articulation is referred to as 'Playing Method' in the interface, with detached indicating *staccato* and smooth referring to *legato*.

Cue changes for tempo, pitch, dynamics, and brightness are made via digital sliders in the interface. Digital buttons are used to switch between discrete levels of mode (pitch alphabet), articulation (playing method), and instrumentation. Changes to the music through the cues are instantly heard in real-time. When a MIDI file is inputted in *EmoteControl*, the properties of the musical piece are altered depending on the initial values of the cue sliders. Therefore, the users would not be exposed to the original version of the piece as it initially portrayed its intended emotion. The cue values were recorded at 10Hz.

5.2.2 Cue Details

5.2.2.1 Tempo

The tempo cue is measured in beats per minute (bpm). The slider is set with a minimum value of 40 bpm and a maximum value of 210 bpm to cover a wide tempo range.

5.2.2.2 Articulation

The articulation cue in this second version of the interface has two levels rather than three: *legato* and *staccato*. Since in this version of the interface users can change the instruments playing the music; from strings to brass and to woodwinds, the *detaché* articulation option was removed due to being specific to string instruments, and thus, would not be compatible with the brass and woodwinds instrument ensemble options. In addition, post-feedback gathered during the formal evaluation study in **Chapter 2**, the 'articulation' term was changed to 'playing method' in the interface, to provide users with a term which is not music specific. Similarly, *legato* is denoted in the interface as 'smooth' and *staccato* as 'detached'.

5.2.2.3 Pitch

The pitch slider controls a pitch shift range of ±2 semitones from the starting point.

5.2.2.4 Dynamics

The dynamics slider alters the MIDI volume of the virtual instrument used as sound output, rather than the overall volume via the dB level. The dynamics slider has a minimum MIDI volume value of 30 and a maximum value of 129.

5.2.2.5 Brightness

The brightness cue changes how bright or dull the musical piece sounds by altering the number of harmonics present in the sound. This is attained by changing the cut-off frequency value of a low-pass filter, with an available cut-off range of 305 Hz to 20,000 Hz. The low-pass filter has a steep slope gradient of 48dB/Oct and a Q factor of 0.43 to diminish frequency resonance.

5.2.2.6 Mode

Mode (labelled as pitch alphabet) gives participants the option to select a major mode denoted as 'pitch alphabet 1' or a harmonic minor mode (flattening the third and sixth degree of the scale to switch from major to minor) denoted as 'pitch alphabet 2'.

5.2.2.7 Instrumentation

Participants can also choose which group of instruments play the music: brass, strings, or woodwinds. Previous findings have suggested that difference in sound attributes such as brightness and spectral entropy may impact the emotional quality of the music (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013). Different instruments have been investigated with respect to emotional qualities, however, mostly as individual instruments rather than ensembles (Huron, Anderson, & Shanahan, 2014; Saitis & Siedenburg, 2020; Schutz et al., 2008). We wanted to test groups of instruments rather than individual instruments as this allowed us to test polyphonic music and a bigger register range simultaneously. Based on a pilot experiment that investigated the emotional expressivity range of a number of instruments (detailed in section 5.7.1 in the Supplementary Material at the end of this chapter) and using instruments with register ranges that could support the pitch ranges of the musical stimuli, the following instruments were chosen for the instrument family ensembles:

- Vienna horn (#3 in emotional expressivity range rank, Table 5.5 in the Supplementary Material), piccolo trumpet, euphonium (#6), and trombone for the brass ensemble
- violin (#4), viola (#13), cello (#2), and double bass for the strings ensemble
- flute (#1), oboe (#8), clarinet (#9), and bassoon (#12) for the woodwinds ensemble

5.2.3 Procedure

Full ethical consent was sought and approved by Durham University before testing. It is to be noted that this experiment was carried out during COVID-19; thus, specific safety measures were taken into consideration, detailed in section 5.7.2 of the Supplementary Material. The first part of the experiment required participants to answer some demographic questions such as age, gender, and musical expertise. This was administered online via a short survey on Qualtrics. Instructions and a video demonstration for the second part of the experiment (the musical task using *EmoteControl* V2.0) were also presented to the participants online.

The musical task was done in person. Participants were presented with the *EmoteControl* V2.0 interface and instructed to use the seven cues available to change the music given to convey specific emotions designated by the researcher. Overall, all seven musical pieces were altered to convey the seven designated emotions. This yields 49 different musical piece and emotion combinations. As fatigue might set in if the same individual carries out 49 combinations, participants were split into three groups and given a subset of the total combinations. Each group carried out 14 unique piece and emotion combinations consisting of all seven musical pieces to convey two different emotions (14 combinations). In addition, all groups carried out seven more combinations where participants had to portray the emotion already attributed to the different pieces (e.g., the piece composed and validated as conveying anger was altered by participants to express anger) to provide a common frame of reference, totalling 21 combinations. Cue level value alterations made by participants for all seven cues were recorded for each trial. Prior to the musical task, participants were subjected to a practice trial where they changed the cue levels of a musical piece that was not utilised during the actual experiment to get accustomed to the interface and the musical task at hand. The experiment took approximately 30 minutes to complete.

5.2.4 Results

First, we examined the consistency of participant cue usage by calculating the inter-rater agreement within each block of 21 stimuli and emotion combinations across each cue and participant, using Cronbach's alpha (intraclass correlation coefficient). Overall, high consistency in the use of cues was observed, especially in Tempo (α = 0.950-0.957, calculated for the three subsets of the full design), Pitch (α = 0.928-0.936), and Mode (α = 0.894-0.940). The other parameters also had high consistency, Articulation (α = 0.880-0.899), Brightness (α = 0.817-0.900), Dynamics (α = 0.799-0.849), with Instrumentation (α = 0.784-0.841) having the lowest consistency.

An initial exploration of the relationship between the cues and the factors Emotion, Piece, and the factors' interaction (Emotion x Piece) was carried out. First, a linear mixed model (LMM) was applied for each cue individually (using the *lmer* function from the *lme4* package in R), with Participant as the random factor. The factors (Emotion, Piece, and Emotion x Piece) were individually added to the linear mixed model. Generalised linear mixed models (GLMMs) with a binomial distribution were used instead of LMMs for the mode and articulation cues due to their binary nature. Likelihood ratio tests were then run to evaluate whether any of the factors added a statistically significant contribution to the initial model. Table 5.1 presents the results from the likelihood ratio tests between the initial models for each individual cue and with the added factors. The model which included Emotion as a factor differed significantly from the initial model for each of the separate cues. The addition of the Piece factor was of statistical significance for all cue models except the ones pertaining to the brightness and pitch cues. The variance in statistical significance of Piece in relation to the different cues might suggest that the cues were used differently across musical pieces. This is presumably due to the variance in the musical structure of the pieces, which consequentially might affect how the cues are used to portray the same emotion across different pieces. The Piece and Emotion interaction did not add a significant contribution to the models.

Table 5.1. The Chi-squared statistics (χ^2) produced from likelihood ratio tests for separate G/LMM models of tempo, articulation, mode, pitch, dynamics, brightness, and instrumentation cues with and without the factors emotion, piece, and emotion x piece interaction.

	Emotion	**Piece**	**Piece × Emotion**
Tempo	677.04***	31.30***	48.45
Articulation	399.73***	23.70***	36.78
Pitch	223.68***	1.92	47.75
Dynamics	580.69***	27.81***	28.27
Brightness	341.14***	6.65	27.99
Mode	426.42***	17.19**	41.21
Instrumentation	216.77***	23.57***	44.81

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

Since the main aim of this chapter is to better understand the relative contribution of the cues to each of the seven emotions, the rest of the analysis will focus on how the cues were used together to communicate the different emotions. To investigate this, separate LMMs were then calculated for each intended emotion (the independent variable) with respect to the different cues (these being the dependent variables). Piece and Participant were inputted in the models as random factors. As these scores will be compared to Experiment 2 results in a later section, standardised beta scores (Z-scores) were utilised in the calculations rather than the raw scores for easier comparison of the results. Results of all LMM computations are shown in Table 5.2.

The first seven columns in Table 5.2 represent the seven emotions investigated in this experiment. The cues' LMM estimates for all emotions are shown in the rows in Table 5.2. The first four rows of Table 5.2 display the LMM estimates for the continuous cues, tempo, pitch, dynamics, and brightness. The sign (+ or -) of these four cues indicate whether the cue values (via beta coefficients) were positive (+) or negative (-). For example, a positive value for tempo suggests a fast tempo, and a negative value for pitch suggests a low pitch level. Rows 5 – 9 in Table 5.2 represent the estimates for the discrete cues, articulation, mode, and instrumentation. Due to the categorical nature of the instrumentation cue, each instrument option (brass, strings, and woodwinds) was regarded separately for the analysis and thus split into three different rows in Table 5.2. As the LMM estimates of the categorical cues do not represent absolute values, the sign (+ or -) for each of these cues has different meanings. A negative value for the articulation indicates a smooth playing method (*legato*), whilst a positive value indicates a detached playing method (*staccato*). Minor mode is represented by a positive value, whilst a negative value represents major mode. A significant negative value for an instrument indicates that the instrument was specifically not chosen for the intended emotion. A significant positive value indicates that the instrument was explicitly chosen for the particular emotion. A non-significant value for any of the instruments suggests that the particular instrument did not play a role in the communication of the particular emotion.

With regards to cue combinations for the individual emotions, the LMM estimates indicate that a slow tempo, low pitch level, soft dynamics, low brightness level (i.e., a dull sound), *legato* articulation and minor mode were specifically used to communicate *sadness* (Bresin & Friberg, 2000; Hevner, 1936; Thompson & Robitaille, 1992). Furthermore, the woodwinds ensemble was explicitly used to convey sadness, whilst the strings ensemble was specifically not used (Akkermans et al., 2019; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013). The brass ensemble was not a contributing factor when conveying sadness. All cues except for brass instrumentation had significant roles in the communication of joy through the musical pieces. A fast tempo, high pitch level, loud dynamics, rich bright sound, *staccato* articulation, major mode and brass instrumentation were specifically used to communicate *joy* (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Hailstone et al., 2009; Saarikallio, Vuoskoski, & Luck, 2014), whilst strings instrumentation was explicitly not used to convey joy. The woodwinds instrumentation did not have a significant effect on the portrayal of joy.

Table 5.2. Linear Mixed Model (LMM) results for seven emotions across all cues in Experiment 1. The numbers are standardised betas (z-scores) with their 2.5% and 97.5% confidence intervals shown in brackets.

	Sadness	**Joy**	**Calmness**	**Anger**	**Fear**	**Power**	**Surprise**	**Mean Pseudo**
Tempo	-1.44*** (-1.60, -1.27)	0.58*** (0.39, 0.76)	-1.25*** (-1.42, -1.08)	0.89*** (0.71, 1.07)	0.36*** (0.17, 0.54)	0.38*** (0.19, 0.56)	0.47*** (0.28, 0.65)	0.0305
Pitch	-0.75*** (-0.93, -0.57)	0.83*** (0.65, 1.01)	-0.28** (-0.46, -0.09)	-0.19* (-0.38, 0.00)	-0.41*** (-0.59, -0.22)	0.05 (-0.14, 0.23)	0.75*** (0.57, 0.93)	0.0062
Dynamics	-1.03*** (-1.19, -0.87)	0.25** (0.07, 0.42)	-1.22*** (-1.38, -1.07)	0.64*** (0.47, 0.81)	0.21* (0.04, 0.39)	0.80*** (0.63, 0.97)	0.36*** (0.19, 0.53)	0.0077
Brightness	-1.00*** (-1.17, -0.83)	0.72*** (0.54, 0.90)	-0.72*** (-0.90, -0.55)	0.17 (-0.02, 0.35)	-0.34*** (-0.52, -0.16)	0.56*** (0.38, 0.74)	0.62*** (0.44, 0.80)	0.0057
Articulation	-3.58*** (-4.54, -2.82)	0.46* (0.06, 0.87)	-2.56*** (-3.18, -2.01)	1.00*** (0.58, 1.45)	0.51* (0.11, 0.93)	0.97*** (0.55, 1.41)	2.04*** (1.52, 2.63)	0.0103
Mode	1.58*** (1.13, 2.07)	-3.10*** (-3.89, -2.43)	-1.63*** (-2.12, -1.18)	2.32*** (1.78, 2.94)	1.91*** (1.42, 2.45)	0.28 (-0.11, 0.68)	-1.71*** (-2.20, -1.24)	0.0397
Brass Instrumentation	0.03 (-0.43, 0.47)	0.70*** (0.28, 1.11)	-2.16*** (-3.22, -1.34)	0.08 (-0.38, 0.52)	-0.57* (-1.12, -0.07)	0.70** (0.28, 1.12)	0.08 (-0.38, 0.52)	0.0089
Strings Instrumentation	-0.89*** (-1.31, -0.48)	-0.50* (-0.91, -0.11)	-1.37*** (-1.84, -0.94)	1.26*** (0.83, 1.71)	0.97*** (0.56, 1.40)	0.58** (0.18, 0.99)	-0.07 (-0.46, 0.32)	0.0036
Winds Instrumentation	0.94*** (0.54, 1.34)	-0.07 (-0.52, 0.35)	2.44*** (1.99, 2.91)	-2.87*** (-4.07, -1.97)	-0.79** (-1.33, -0.30)	-2.65*** (-3.84, -	0.02 (-0.41, 0.44)	0.0076
Pseudo R^2 (marginal)	0.33	0.19	0.36	0.20	0.12	0.11	0.15	

A slow tempo, low pitch level, soft dynamics, low brightness level (dark timbre), *legato* articulation, major mode, and woodwinds instrumentation significantly portrayed *calmness* (Balkwill & Thompson, 1999; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; K. Watson, 1942). Brass and strings instruments were specifically not chosen to convey calmness. All cues except for brightness and brass instrumentation had a significant impact in communicating anger. A fast tempo, low pitch level, loud dynamics, *staccato* articulation, minor mode, and strings instrumentation were specifically used to express *anger* in the musical pieces (Balkwill & Thompson, 1999; Hailstone et al., 2009; Juslin, 1997b; Scherer & Oshinsky, 1977). The woodwinds ensemble was explicitly not chosen when portraying anger.

Fear was communicated very similarly to anger; however, all cues significantly contributed to the intended emotion. A fast tempo, low pitch level, slightly loud dynamics, dark timbre, *staccato* articulation, minor mode, and strings instrumentation were used to convey *fear* (Behrens & Green, 1993; Hailstone et al., 2009; Juslin, 1997b; Scherer & Oshinsky, 1977), whilst brass and woodwinds instrumentations were specifically not chosen when portraying fear. Tempo, dynamics, brightness, articulation, and instrumentation contributed to conveying power, whilst pitch and mode did not significantly affect this particular emotion. A fast tempo, loud dynamics level, bright sound, *staccato* articulation, and brass and strings instrumentation were particularly chosen to help express *power* in the musical pieces (Scherer & Oshinsky, 1977; Wedin, 1972). The woodwinds ensemble was specifically not used to convey power. Lastly, the combination of fast tempo, high pitch level, loud dynamics, bright sound, *staccato* articulation, and major mode were contributing factors to the communication of *surprise*

(Scherer & Oshinsky, 1977). The instrumentation cue as a whole (i.e., all of brass, strings, and woodwinds ensembles) did not play a role in conveying surprise.

To further explore how effective the cue-emotion models (i.e., the cue combinations used per emotion) were to convey the intended emotions, the Pseudo R^2 value was calculated for each emotion, using the *MuMIn* package in R (Johnson, 2014; Nakagawa, Johnson, & Schielzeth, 2017; Nakagawa & Schielzeth, 2013). The Pseudo R^2 values present the proportion of variability of the emotion explained by the independent variables in the model (i.e., the cue combinations for each emotion), and these can be seen in the last row of Table 5.2.

The marginal Pseudo R^2 values in Table 5.2 suggest that overall, the cue values and combination models used to portray the intended emotions were highly significant. Calmness and sadness cue-emotion models had the highest Pseudo R^2 values of 0.36 and 0.33, respectively. Power (0.11) and fear (0.12) had the lowest Pseudo R^2 values, which suggests that the cue combinations used for these two emotions had the least stable emotion-cue models, compared to the others.

Finally, to investigate the impact of the individual cues on the portrayal of the different emotions, each cue's mean Pseudo R^2 value across all emotions was computed. These values are displayed in the last column of Table 5.2. It is to be noted that the three instrument ensembles (brass, strings, and woodwinds) together make up the instrumentation cue. Thus, to get an accurate mean Pseudo R^2 value, the individual mean Pseudo R^2 values are summed together. Mode (0.0397) and tempo (0.0305) provide the greatest contributions to communicate specific emotions, which is in line with previous findings (Eerola, Friberg, & Bresin, 2013; Juslin, 1997b; Micallef Grimaud & Eerola, 2022). The summation of the instrumentation cue (0.0201) is the third greatest contributor to emotional expression, followed by articulation. Brightness (0.0057) and pitch (0.0062) seem to have the least effect on how emotions are shaped in music.

In summary, the seven cues in question and their combinations mostly had a role in portraying all emotions targeted. Calmness and sadness had the highest overall prediction rates, whereas power and fear had the lowest associations between the cues and emotion via the models. As individual cues, mode and tempo had the most influence on how the different emotions were portrayed, while brightness and pitch were, overall, the least contributing factors in shaping and communicating specific emotions through music.

5.3 Experiment 2: Systematic Manipulation Approach

This experiment aimed to investigate the effect of the same seven cues utilised in Experiment 1 (tempo, articulation, pitch, dynamics, brightness, mode, and instrumentation) on the emotion expressed in music, using a traditional approach. Experiment 2 used a factorial design to systematically manipulate and render variations of the musical stimuli with different cue levels and combinations. The different cue levels were determined based on results from a previous study (Micallef Grimaud & Eerola, 2022) which utilised the same musical stimuli and identified the optimal levels of tempo, pitch, dynamics, brightness, articulation, and mode for each musical stimulus to convey the seven different emotions (sadness, joy, calmness, anger, fear, power, and surprise). These optimal cue levels were utilised as the point of reference and mid-levels for the cue levels in this experiment. The rationale behind the other cue levels will be detailed in the Cue Details section.

5.3.1 Method

A fractional factorial listening experiment was designed where the same seven musical cues studied in Experiment 1 were manipulated on two to three levels across seven different musical pieces.

5.3.1.1 Stimulus Manipulation

Seven cues with 2-3 levels creates a large design matrix in terms of all possible cue and level combinations ($3 \times 2 \times 2 \times 3 \times 3 \times 3 \times 3 = 972$). Due to the size of the design, it would not be feasible for participants to respond to all possible trials. Thus, a fractional factorial design was implemented to systematically reduce the number of trials in the design by omitting trivial effects, whilst ensuring that the main effects are kept (Landsheer & Van Den Wittenboer, 2002). The *optFederov* function from the *AlgDesign* package in R (Wheeler, 2004) was used to calculate the number of trials needed to provide an optimal geometrical design where every main effect and first-order interactions would be balanced (e.g., design symmetry close to 1), without participants having to carry out all 972 trials. Using 36 trials out of the 972 for each musical piece produces an optimal design with a geometrical symmetry of 0.956. This optimal design was employed for each of the seven musical pieces, which totals to 252 different combinations of cue levels across the seven musical pieces (36 trials x 7 musical pieces = 252 combinations).

5.3.1.2 Cue Details

The optimal levels for the seven different cues were calculated for the seven intended emotional expressions, with the optimal cue settings being relative to the distinct emotional expressions.

Tempo. Three levels were computed for the tempo cue. Previous systematic cue manipulation studies (Eerola, Friberg, & Bresin, 2013; Peretz, Gagnon, & Bouchard, 1998; Vieillard et al., 2008) either calculated different levels of notes per second (NPS) across musical pieces or used quantiles to determine different levels which cover a substantial range. In this experiment, the 0.35 and 0.65 quantiles were calculated with respect to the optimal cue level used as the mid-point (0.50 quantile).

Brightness. Like the tempo cue, three brightness levels were computed, utilising the quantiles of 0.35, 0.50, 0.65, relative to the optimal cue levels identified in a previous study (Micallef Grimaud & Eerola, 2022).

Dynamics. Three levels of dynamics were computed for this study. Previous studies have utilised a dynamic range of 20dB (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013), representing the normal range of an acoustic instrument (Eerola, Friberg, & Bresin, 2013; Fletcher & Rossing, 1998). Step-sizes were made in 5dB from -10dB to +10dB, and the dynamics controlled the sample synthesizer rather than the volume output. The stimuli to be utilised in this current study have specific dynamic values attributed to them, which already have been validated by two sets of participants (Micallef Grimaud & Eerola, 2022). Therefore, the pre-established dynamics values of the musical pieces were utilised as the calibrated mid-point and then two levels at a ±5dB difference were calculated.

Pitch. Three levels of pitch were calculated. Previous studies explored different pitch levels an octave (12 semitones) apart (Bresin & Friberg, 2011; Quinto, Thompson, & Taylor, 2014), as well as seven semitones apart (Eerola, Friberg, & Bresin, 2013). In this experiment, pitch levels at ±7 semitones from the point of reference were investigated. The register ranges of the instruments utilised in this experiment and the pitch range of the musical pieces to be used were taken into consideration when deciding on the different pitch levels, ensuring that all pitch levels chosen were in range. The only exception to this methodology was the musical piece composed and validated as conveying anger (Track A1 in Chapter 4), since it originally had a low pitch range. As it would have been impossible for the instruments to play an octave lower than the original pitch, this was used as the lower pitch level, while the higher two levels were computed as +7 semitones and +14 semitones, respectively.

Mode. Two categorical levels of mode were chosen: major and harmonic minor. This cue was controlled using the transposer plug-in in Logic Pro X, the software that was used to render the stimuli for this experiment, that changes the third and sixth scale degree to minor (m3/m6) or major (M3/M6) for minor and major mode, respectively.

Articulation. Two levels of articulation were investigated: *legato* and *staccato*. This cue was controlled through the virtual instrument plug-in (Vienna Symphonic Library) used to export the musical stimuli.

Instrumentation. The instrument groups used were the ones available in the *EmoteControl* V2.0 interface and used in Experiment 1, due to their expressive ranges and varied timbral sound, and ensuring consistency across the two experiments. The brass ensemble consisted of a piccolo trumpet, Vienna horn, euphonium, and trombone. The strings ensemble consisted of a violin, viola, cello, and double bass. The woodwinds ensemble was made up of a flute, clarinet, oboe, and bassoon.

5.3.1.3 Creation of Stimuli

The seven musical pieces utilised in Experiment 1 served as the point of departure for the stimuli in this experiment. Thirty-six iterations for each original musical piece with different cue levels were exported[16]. The Vienna Symphonic Library (VSL) was utilised as the virtual instrument in Logic Pro X to export all tracks.

5.3.1.4 Procedure

The experiment was administered online using the survey tool Qualtrics. The 36 variations for each of the seven musical pieces were put in separate blocks in Qualtrics. To minimise fatigue, each participant was randomly presented two out of the seven blocks, which meant that each participant listened to 72 trials (all variations of two different pieces) out of the 252 total combinations. For each piece, participants rated how much of each of the seven emotions joy, sadness, calmness, power, anger, fear, and surprise they thought the music was conveying. Ratings were done on seven separate five-point Likert scales. A rating of 1 (none at all) indicated that the music did not convey any emotion. A rating of 5 (a lot) indicated that the music strongly conveyed the emotion. Participants carried out a practice trial to familiarise themselves with the music listening task and rating scales. The study took approximately 50 minutes to 1 hour to complete. Participation in this experiment was voluntary. Participants could opt-in a prize draw for two £10 Amazon vouchers at the end of the experiment.

5.3.1.5 Participants

---

[16] The stimuli can be found in the folder Experiment 2: Evaluation Approach, Stimuli following the link: https://osf.io/atxhk/?view_only=6d0ef66b819e476dbfcc40b54ae6c986

Participants were recruited via social media and university communications. 162 participants (51 males, 110 females, one individual preferred not to say) between 18 and 66 years (*M* = 34.22, *SD* = 13.05) 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. Seventy-four of the participants were musicians, and 88 were non-musicians.

5.3.2 Results

The consistency of the participants' ratings was calculated by examining the inter-rater agreement (using Cronbach's alpha) within each emotion scale across each participant and musical piece. High consistency was observed for all rating scales, especially in the sadness rating scale (α = 0.953), joy rating scale (α = 0.944), and calmness rating scale (α = 0.937). The other rating scales also had high consistency (fear α = 0.909, surprise α = 0.906, power α = 0.864), with the anger rating scale having the lowest consistency score α = 0.842.

Similar to Experiment 1, a linear mixed model (LMM) was applied for each emotion with respect to the different cues to identify whether the cues and their combinations had a significant role in conveying the different emotions. Standardised scores (Z-scores) were utilised in the calculations, and the LMM estimates are shown in Table 5.3.

The first four rows in Table 5.3 present the LMM estimates for the continuous variables, tempo, pitch, dynamics, and brightness. Similar to Table 5.2 in Experiment 1, a positive sign (+) represents a relative high cue value, whilst a negative sign (-) denotes a relative low cue value. Rows 5 - 9 in Table 5.3 represent the categorical cues, articulation, mode, and the three types of instrumentation (brass, strings, and woodwinds). The sign (+/-) values for the categorical cues retain the same meaning as in Table 5.2. As these categories represent arbitrary binary values, the sign (+ or -) for each of these cues has different meanings. A negative articulation value indicates *legato*, whilst a positive value suggests *staccato*. Minor mode is represented by a positive value, whilst a negative value represents major mode. A significant negative value for an instrument suggests that the instrument was specifically not chosen for the particular emotion, whilst a significant positive value indicates that the instrument was explicitly chosen. A non-significant value for any of the instruments suggests that the particular instrument did not play a role in the communication of the particular emotion.

Table 5.3. Linear Mixed Model (LMM) results for seven emotions across all cues in the Systematic Manipulation (Fractional Factorial Design) Experiment. The numbers are standardised betas, and their 2.5% and 97.5% confidence intervals are denoted in brackets.

	Sadness	**Joy**	**Calmness**	**Anger**	**Fear**	**Power**	**Surprise**	**Mean Pseudo R^2**
Tempo	-0.21*** (-0.23, -0.18)	0.16*** (0.13, 0.18)	-0.22*** (-0.25, -0.20)	0.03** (0.01, 0.05)	0.05*** (0.03, 0.08)	0.11*** (0.08, 0.13)	0.08*** (0.06, 0.11)	0.0136
Pitch	-0.14*** (-0.16, -0.11)	0.17*** (0.14, 0.19)	0.04*** (0.02, 0.06)	-0.10*** (-0.12, -0.08)	-0.06*** (-0.08, -0.03)	-0.09*** (-0.12, -0.07)	0.07*** (0.05, 0.10)	0.0074
Dynamics	-0.03** (-0.05, 0.01)	-0.02 (-0.04, 0.01)	-0.02 (-0.04, 0.00)	0.00 (-0.02, 0.02)	0.02 (-0.01, 0.04)	0.02 (-0.01, 0.04)	-0.01 (-0.03, 0.02)	0.0002
Brightness	-0.02* (-0.05, 0.00)	0.05*** (0.03, 0.07)	-0.05*** (-0.07, -0.02)	0.01 (-0.01, 0.03)	-0.02 (-0.04, 0.01)	-0.02 (-0.04, 0.01)	0.04*** (0.02, 0.07)	0.0007
Articulation	-0.44*** (-0.46, -0.41)	0.14*** (0.11, 0.16)	-0.27*** (-0.29, -0.24)	0.06*** (0.04, 0.08)	0.01 (-0.01, 0.04)	0.10*** (0.08, 0.13)	0.30*** (0.28, 0.33)	0.0373
Mode	0.27*** (0.24, 0.29)	-0.50*** (-0.52, -0.47)	-0.24*** (-0.26, -0.22)	0.20*** (0.18, 0.22)	0.42*** (0.39, 0.44)	0.04*** (0.02, 0.07)	-0.01 (-0.03, 0.01)	0.0561
Brass Instrumentation	0.01 (-0.05, 0.06)	0.06 (0.00, 0.12)	0.08** (0.03, 0.13)	-0.04 (-0.08, 0.01)	-0.12*** (-0.18, -0.07)	-0.12*** (-0.17, -0.06)	-0.09*** (-0.14, -0.04)	0.0026
Strings Instrumentation	0.10*** (0.04, 0.15)	-0.12*** (-0.18, -0.06)	-0.03 (-0.08, 0.02)	0.11*** (0.06, 0.15)	0.21*** (0.15, 0.26)	0.33*** (0.28, 0.39)	-0.07* (-0.12, -0.01)	0.0038
Woodwinds Instrumentation	-0.10*** (-0.16, -0.05)	0.06* (0.00, 0.12)	-0.05* (-0.10, 0.00)	-0.07** (-0.11, -0.03)	-0.08** (-0.14, -0.03)	-0.22*** (-0.27, -0.17)	0.16*** (0.11, 0.21)	0.0040
Pseudo R^2 (marginal)	0.21	0.20	0.15	0.06	0.12	0.04	0.08	

All cues except for the brass instrumentation had a significant effect in conveying sadness. A slow tempo, low pitch level, soft dynamics, dull sound (low brightness level), *legato* articulation, minor mode, and strings instrumentation were specifically used to convey *sadness* (consistent with Behrens & Green, 1993; Gagnon & Peretz, 2003; Hevner, 1936, 1937; Juslin, 1997b; Krumhansl, 1997; Lindström, 2006; Saarikallio, Vuoskoski, & Luck, 2014; Scherer & Oshinsky, 1977). The woodwinds ensemble was explicitly not used to convey sadness, whilst the brass ensemble was not a contributing factor to the communication of sadness. A fast tempo, high pitch level, bright sound, *staccato* articulation, major mode and woodwinds instrument were used to portray *joy* (Akkermans et al., 2019; Bresin & Friberg, 2011; Gabrielsson & Juslin, 1996; Saarikallio, Vuoskoski, & Luck, 2014; Scherer & Oshinsky, 1977). The strings instrumentation was specifically not chosen to portray joy through music, while dynamics and brass instrumentation did not contribute to the intended emotion. A slow tempo, low pitch level, low brightness level, *legato* articulation, major mode, and brass instrumentation were specifically chosen to communicate *calmness* (Balkwill & Thompson, 1999; Bresin & Friberg, 2011; Kragness & Trainor, 2019). The woodwinds ensemble was explicitly not chosen, while dynamics and strings instrumentation were not contributing factors towards calmness. *Anger* was portrayed by a fast tempo, low pitch level, s*taccato* articulation, minor mode, and strings instrumentation (Hailstone et al., 2009; Kragness & Trainor, 2019; Saarikallio et al., 2019; Scherer & Oshinsky, 1977). The woodwinds ensemble was specifically not chosen to convey anger, whilst dynamics, brightness, and brass instrumentation did not help to express anger through the musical pieces.

A fast tempo, low pitch level, minor mode, and strings instrumentation were strategically utilised to convey the emotion *fear* (Behrens & Green, 1993; Bresin &

Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Hailstone et al., 2009; Scherer & Oshinsky, 1977), whilst brass and woodwinds instrumentations were specifically not chosen. The dynamics, brightness, and articulation cues did not have a significant role in conveying fear through the music. Tempo, pitch, articulation, mode, and instrumentation contributed to conveying *power*, whilst dynamics and brightness did not have a significant effect on communicating power. A fast tempo, low pitch level, *staccato* articulation, major mode, and strings instrumentation were specifically chosen when power was being conveyed (Scherer & Oshinsky, 1977). The brass and woodwinds ensembles were explicitly not used, whilst dynamics and brightness did not significantly affect the portrayal of the power. All cues except for dynamics and mode had a significant effect on the communication of the surprise emotion. A fast tempo, high pitch level, bright sound, *staccato* articulation (Scherer & Oshinsky, 1977), and woodwinds instrumentation contributed to the conveying of *surprise* in the musical pieces, whilst brass and strings instrumentations were explicitly not used.

The Pseudo R^2 marginal values were also computed to investigate how well the cue combinations worked in communicating the intended emotions. Sadness (0.21) and joy (0.20) had the highest Pseudo R^2 values, whilst power (0.04) and anger (0.06) had the lowest scores.

Lastly, to determine the effect size of each cue on the shaping of the emotions, the mean Pseudo R^2 value for each cue across all emotions was computed. These values are displayed in the last column of Table 5.3. Mode (0.0561) and articulation (0.0373) had the biggest effect on the shaping of the different emotions, followed by tempo (0.0136) and the instrumentation cue as a whole (0.0104). Dynamics (0.0002) and brightness (0.0007) were the least effective cues on the emotion expressed in the music. This can also be seen from the LMM estimates of dynamics and brightness in Table 5.3, where dynamics had a significant effect only for sadness, and brightness significantly affected only four (sadness, joy, calmness, and surprise) of the seven expressed emotions.

In summary, most of the seven cues investigated in this systematic manipulation experiment had a significant role in portraying sadness, joy, calmness, anger, fear, power, and surprise. Dynamics had no effect on the intended emotion, except for sadness. As individual cues, mode and articulation had the biggest effect on modelling the desired emotion, while dynamics and brightness were the least contributing factors in shaping specific emotional expressions through music.

5.4 Comparison of Experiments 1 and 2

In this chapter, two different methodologies were used to investigate the role of seven musical cues in shaping seven different emotional expressions in musical pieces. In Experiment 1, a production approach was used, where participants changed musical pieces specifically created for research, to communicate the different emotions via the seven cues using a computer interface. In Experiment 2, a systematic manipulation approach was used to investigate the same cues and emotions used in Experiment 1.

Table 5.4 presents a high-level overview of cue levels used to portray the different emotions in Experiment 1 and 2. Cue levels in a bold, purple font indicate results which adhere to findings from previous literature. Overall, the cues operate similarly in the majority (32/49) of cue-emotion combinations across the two experiments. Only five cue-emotion discrepancies are related differences where both cues are statistically significant and different (as in tempo and anger, +/~).

Additionally, none of the continuous cues demonstrates opposite cue values as all discrepancies are a matter of nuance (as in tempo and anger, where high tempo levels were associated with anger in Experiment 1 but medium levels in Experiment 2). This is not entirely surprising as the two experiments are based on the same underlying music and the manipulation of the same cues, although the actual cue levels are not directly comparable. This internal consistency is nevertheless reassuring and may be interpreted as an internal validation of the methodologies used.

When we compare the cue levels for the different emotions explored in this study to findings of past studies, this mostly supports the previous literature (for an overview see: Cespedes-Guevara & Eerola, 2018; Eerola & Vuoskoski, 2013; Gabrielsson & Lindström, 2001; Juslin & Laukka, 2004); A slow *tempo* is associated with sadness and calmness, a moderate or fast tempo is linked to anger and fear, and a fast tempo is associated with joy, power, and surprise (Balkwill & Thompson, 1999; Bresin & Friberg, 2011; Gagnon & Peretz, 2003; Hevner, 1937; Juslin, 1997b; Kragness & Trainor, 2019; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Thompson & Robitaille, 1992; Wedin, 1972). Furthermore, tempo had a significant effect in shaping all the different emotions in music.

Table 5.4. A summary of cue levels used to shape the different emotions in Experiment 1 and 2.

	Emotion	Sadness	Joy	Calmness	Anger	Fear	Power	Surprise
	Source	E1/E2	E1/E2	E1/E2	E1/E2	E1/E2	E1/E2	E1/E2
Cues	Tempo	**-** / **-**	**+** / **+**	**-** / **-**	**+** / ~	**+** / ~	**+** / **+**	**+** / **+**
	Dynamics	**-** / ~	**+** / [~]	**-** / [~]	**+** / [~]	**+** / [~]	**+** / [~]	**+** / [~]
	Pitch	**-** / **-**	**+** / **+**	**-** / ~	- / -	**-** / **-**	[~] / -	**+** / **+**
	Brightness	**-** / **-**	**+** / **+**	**-** / **-**	[+] / [~]	**-** / [-]	**+** / [-]	**+** / **+**
	Articulation	**L** / **L**	**S** / **S**	**L** / **L**	**S** / **S**	**S** / [S]	**S** / **S**	**S** / **S**
	Mode	**-** / **-**	**+** / **+**	**+** / **+**	**-** / **-**	**-** / **-**	[+] / -	**+** / [+]
	Instrumentation	**W** / **S**	**B** / W	**W** / B	**S** / **S**	**S** / **S**	B, S / S	[ ] / W

Notes. The cue levels of Experiment 1 are denoted first in each cell, followed by Experiment 2 values. Source of data is denoted by the source row, where E1 refers to Experiment 1 and E2 refers to Experiment 2. Results from the two experiments are separated by /, and symbols denote cue levels (+ high level, ~ mid-level, - low level). Cue levels all had a significant effect, except for ones in square brackets [ ]. Articulation cue: L = *legato*, S = *staccato*. Mode cue: - = minor, + = major. Instrumentation cue: B = brass, S = strings, W = woodwinds, [ ] = none of the instruments were significant. Levels in a purple, bold font indicate results that adhere to findings from previous literature.

The *dynamics* cue presented rather diverging results between the two experiments in this study. In Experiment 1, the dynamics cue was a significant contributing factor to all emotions, varying in importance. A low dynamics level was used for sadness and calmness, and a loud dynamics level was used for joy, anger, fear, power, and surprise, which overall complements the existing literature (Bresin & Friberg, 2011; Juslin, 1997b; Kragness & Trainor, 2019; Saarikallio, Vuoskoski, & Luck, 2014; K. Watson, 1942). Dynamics had the least impact, albeit significant, in the communication of fear. Interestingly, in previous studies, it has been suggested that it is possible to convey fear with both a low dynamics level (Akkermans et al., 2019; Gabrielsson & Juslin, 1996) and a high dynamics level (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013).

On the other hand, in Experiment 2, the dynamics cue had the least effect on creating different emotion profiles in music. A low dynamics level was purposely used to portray sadness in Experiment 2, which complements previous research (Eerola, Friberg, & Bresin, 2013; Kragness & Trainor, 2019; Quinto, Thompson, & Taylor, 2014; Wedin, 1972). However, dynamics did not have a significant effect on portraying any of the other emotions. The discrepancy between the non-significant effect of the dynamics cue in Experiment 2 and Experiment 1 findings and previous research is rather notable. Existing literature provides evidence that different dynamics levels have an impact on the emotion being expressed by the music (Bresin & Friberg, 2011; Gabrielsson & Lindström, 2001, 2010; Juslin, 1997b; Kragness & Trainor, 2019), where high activity emotions are usually associated with high dynamics levels, and low activity emotions with low dynamics levels (Gabrielsson & Juslin, 2002). The fact that the three different levels of the dynamics cue used in Experiment 2 were based on quantiles varying in increments of 0.15 means that differences between the levels would be rather subtle, as Experiment

2 focussed on fine-tuning cues, rather than using drastically different upper and lower limits. It is possible that participants did not distinguish between the minor changes between the dynamics cue levels (Cousineau, Demany, & Pressnitzer, 2009), especially since Experiment 2 was an online listening experiment, which might explain why overall, the cue did not register as having a significant impact on the emotional expression. Furthermore, since Experiment 2 was based online, the experiment environment could not be controlled by the researcher, unlike the lab environment used for Experiment 1. Although the instructions in the online experiment informed participants to use headphones, set their volume to a comfortable level and test their sound before the experiment, it is entirely possible that participants used less than adequate headphones, no headphones at all, or changed the volume of their device while carrying out the study. Any of these factors might have had an influence on the dynamics.

For the most part, the *pitch* cue was consistently used across both experiments and mostly had a significant effect on the conveyed emotion. A low pitch level was used in both experiments for sadness, anger, and fear (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Hevner, 1937; Quinto, Thompson, & Taylor, 2014; Saarikallio, Vuoskoski, & Luck, 2014; K. Watson, 1942; Wedin, 1972). A low pitch level was also purposely used for calmness in Experiment 1, while a moderate pitch level was preferred for calmness in Experiment 2. A high pitch level was utilised to convey joy and surprise in both experiments (W. G. Collier & Hubbard, 2001; Eerola, Friberg, & Bresin, 2013; Hevner, 1937; Rigg, 1940b; Scherer & Oshinsky, 1977). A low pitch level was explicitly used for power in Experiment 2, while the pitch level did not significantly affect power in Experiment 1. As with the dynamics cue, conflicting data on which pitch level conveys different emotions exists in current literature. For example, both high and low pitch levels have been used to communicate power, fear, calmness, and anger (Akkermans et al., 2019; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Micallef Grimaud & Eerola, 2022; Saarikallio et al., 2019; Scherer & Oshinsky, 1977).

In both Experiment 1 and 2, a low *brightness* level, i.e., one with few upper harmonics, which creates a dull sound, was used for sadness and calmness (Akkermans et al., 2019; Gabrielsson & Juslin, 1996; Schutz et al., 2008). In Experiment 1, a low brightness level was also used to convey fear; however, in Experiment 2, brightness had a non-significant effect on the portrayal of fear. A high brightness level, i.e., one with multiple harmonics and a bright sound, was specifically used for joy and surprise in both experiments (Akkermans et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Scherer & Oshinsky, 1977). A high brightness level was also used in the portrayal of power in Experiment 1 (Scherer & Oshinsky, 1977), while Experiment 2 produced a non-significant brightness value. In both experiments, the brightness cue did not play a significant role in shaping anger. It has been noted that high brightness sounds are more likely associated with high activity emotions, while low activity emotions are more likely represented by dull and dark sounds (Bresin & Friberg, 2011; Gabrielsson & Lindström, 1995).

The *articulation* cue produced consistent results for all emotions across both experiments. A *legato* articulation was specifically chosen to portray sadness and calmness, while a *staccato* articulation was used to convey joy, anger, fear, power, and surprise (Bresin & Friberg, 2000, 2011; Gabrielsson & Juslin, 1996; Juslin, 1997b; Kragness & Trainor, 2019), consistent with past findings. Furthermore, the articulation cue had a significant effect in conveying all emotions, except for fear in Experiment 2. *Mode* was also used similarly across both experiments and the current literature. A major mode was utilised to convey the positive emotions joy, calmness, and surprise, while a minor mode was chosen for sadness, anger, fear, and power emotions (Costa, Fine, & Ricci Bitti, 2004; Gagnon & Peretz, 2003; Hevner, 1936; Micallef Grimaud & Eerola, 2022; Scherer & Oshinsky, 1977). However, the mode did not significantly affect power in Experiment 1 and surprise in Experiment 2.

Finally, the *instrumentation* cue had quite contrasting results between the two experiments. The instrumentation cue was used similarly only for anger and fear, where a strings instrumentation was specifically chosen to portray the aforementioned emotions (Balkwill & Thompson, 1999; Behrens & Green, 1993; Hailstone et al., 2009) in both experiments. Additionally, the woodwinds instrumentation was specifically not chosen for anger and fear, and brass was also specifically not used to portray fear. In Experiment 1, a woodwinds ensemble was specifically used to convey sadness (Akkermans et al., 2019; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013), while a strings ensemble was distinctively not chosen. The opposite findings can be seen in Experiment 2, where a strings instrumentation was used for sadness (Behrens & Green, 1993; Huron, Anderson, & Shanahan, 2014), and woodwinds instrumentation was specifically not used. Brass instrumentation was chosen for joy in Experiment 1 (Bresin & Friberg, 2011; Hailstone et al., 2009), while a woodwinds ensemble was preferred in Experiment 2. Strings instrumentation was specifically not used to convey joy in both experiments. A woodwinds ensemble was utilised to convey calmness (Balkwill & Thompson, 1999; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013) in Experiment 1, with brass and strings instrumentations specifically not chosen to portray the aforementioned emotion.

On the other hand, a brass ensemble was utilised in Experiment 2, and a woodwinds ensemble was specifically not chosen to portray calmness. A strings instrumentation was preferred for power in both experiments. Additionally, a brass ensemble was also chosen by participants to represent power in Experiment 1, while brass was specifically not chosen in Experiment 2. A woodwinds instrumentation was specifically not used to portray power in both experiments. A woodwinds instrumentation was specifically utilised for surprise in Experiment 2, while brass and strings ensembles were purposely not chosen to portray surprise.

The new data about instrument and emotion association in this chapter partially adheres to a handful of studies that looked at individual instruments rather than instrument ensembles (Akkermans et al., 2019; Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Huron, Anderson, & Shanahan, 2014). However, there are also conflicting results both between the findings of the two current experiments and other previous findings. For example, it has also been reported that sadness is best represented with a trumpet (Hailstone et al., 2009) or French horn (Bresin & Friberg, 2011), and that voice (not investigated here) and strings instruments, such as violin and cello, might have a bigger sadness capacity than other instruments (Behrens & Green, 1993; Huron, Anderson, & Shanahan, 2014). Balkwill and Thompson (1999) reported that the instrument timbre did not have a significant effect on the portrayal of sadness and joy in music. Other studies have reported that fear is best conveyed with a brass instrument, such as the French horn (Bresin & Friberg, 2011), or potentially a trumpet (Eerola, Friberg, & Bresin, 2013). The fact that we used instrument ensembles rather than individual instruments should also be taken into consideration. Although the instrumentation results have been compared to previous findings involving individual instruments from the same instrument families (e.g., a trumpet compared to the brass ensemble), our instrumentation cue consists of instrument combinations rather than individual instruments, which might also affect the perceived emotion. Additionally, most previous studies have investigated the effect of instrument timbre and other cues on monophonic melodies, while this current study explores the effect of instrument ensembles and the other cues on the overall structure of polyphonic music. It has been suggested that music with multiple musical parts likely provides more information than monophonic melodies (Hailstone et al., 2009; Heaton, Hermelin, & Pring, 1999).

It is also worth noting that brightness is one of the major perceptual dimensions of timbre (McAdams, 2019; Saitis & Siedenburg, 2020). Saitis and Siedenburg (2020) found that brightness perception dissimilarity is not distinguished by source-cause categories, i.e., different instrument families. Thus, altering the brightness component together with the instrumentation cue in this work might have affected how participants used these two cues to portray the different emotions. Due to the intrinsic relationship between brightness and timbre, future studies should expand research on how the potential connection between instrument combinations and brightness impacts the perceived emotional expression in music.

The impact of the individual cues across emotions was ranked similarly in Experiments 1 and 2. Pseudo R^2 values show that overall, mode had the highest impact on the different emotion profiles in both experiments (Exp.1 Pseudo R^2 = 0.0397, Exp.2 Pseudo R^2 = 0.0561). In Experiment 1, mode was followed by tempo as the second most influential cue (Pseudo R^2 = 0.0305), which is consistent with the results from two previous studies (Eerola, Friberg, & Bresin, 2013; Micallef

Grimaud & Eerola, 2022). In the existing literature, it has also been suggested that tempo has the greatest impact on emotion shaping (Juslin, 1997b; Scherer & Oshinsky, 1977). The other cues in Experiment 1 were ranked as follows: instrumentation (Pseudo R^2 = 0.0201), articulation (Pseudo R^2 = 0.0103), dynamics (Pseudo R^2 = 0.0077), pitch (Pseudo R^2 = 0.0062), and brightness (Pseudo R^2 = 0.0057). In Experiment 2, articulation (Pseudo R^2 = 0.0373) scored as the second most impactful cue on the expressed emotion, followed by tempo (Pseudo R^2 = 0.0136), instrumentation (Pseudo R^2 = 0.0104), pitch (Pseudo R^2 = 0.0074), brightness (Pseudo R^2 = 0.0007), and dynamics (Pseudo R^2 = 0.0002). Although the cue impact ranking varies slightly between the two experiments, it is overall quite similar. In both experiments, mode had the most effect on shaping different emotions, followed by tempo, instrumentation, and articulation. In both experiments, pitch, dynamics, and brightness had the lowest scores, with a distinct difference between them and the first four cues. It is also worth mentioning that the Pseudo R^2 values for the dynamics (Pseudo R^2 = 0.0002) and brightness (Pseudo R^2 = 0.0007) cues in Experiment 2 are notably smaller than their counterparts in Experiment 1 (dynamics Pseudo R^2 = 0.0077, brightness Pseudo R^2 = 0.0057). The stark difference between the dynamics and brightness cues across experiments may be due to these particular musical cues being less perceptive in an online music listening environment (Experiment 2), whereas participants in Experiment 1 were able to change these cues themselves. The difference in environments might explain the low effect scores of dynamics and brightness in Experiment 2. In addition, it is interesting to note that mode, tempo, and articulation, i.e., the cues flagged as the ones mostly contributing to the communication of the different emotions (Eerola, Friberg, & Bresin, 2013), were the cues used most consistently in both experiments.

When we compare how well the cue combinations represent the intended emotions, we find that overall, the cue-emotion profiles in Experiment 1 scored remarkably higher than the cue combinations used in Experiment 2. This indicates that the cue combinations chosen by participants in the analysis-by-synthesis approach worked better to convey the intended emotions than the fixed cue combinations used in the systematic manipulation design. The fact that in Experiment 1, participants could explore a wider range of the continuous cues (tempo, pitch, dynamics, and brightness) than in Experiment 2 suggest that the methodological limitations inherent in the systematic manipulation of cues approach led to situations where the optimal cue-emotion patterns were probably not always within the pre-defined cue values used in Experiment 2. In Experiment 1, Pseudo R^2 values suggest that the cue patterns used for calmness (Pseudo $R^2 = 0.36$) and sadness (Pseudo $R^2 = 0.33$) were the ones most reliable in conveying the intended emotion, compared to the other investigated emotions; anger (Pseudo $R^2 = 0.20$), joy (Pseudo $R^2 = 0.19$), surprise (Pseudo $R^2 = 0.15$), fear (Pseudo $R^2 = 0.12$), and power (Pseudo $R^2 = 0.11$). These findings support a recent analysis-by-synthesis study that identified calmness and sadness as the two emotions best predicted by a combination of tempo, pitch, dynamics, brightness, articulation, and mode (Micallef Grimaud & Eerola, 2022), and fear and surprise having the least reliable cue-emotion patterns. In Experiment 2, sadness (Pseudo $R^2 = 0.21$) and joy (Pseudo $R^2 = 0.20$) had the highest scores for cue-emotion model reliability, followed by: calmness (Pseudo $R^2 = 0.15$), fear (Pseudo $R^2 = 0.12$), surprise (Pseudo $R^2 = 0.08$), anger (Pseudo $R^2 = 0.06$), and power (Pseudo $R^2 = 0.04$). Across the two experiments, sadness, calmness, and joy emotions seem to have the most reliable cue combinations in shaping the intended emotion. Previous research has proposed that basic emotions, i.e., sadness, happiness, anger, fear, surprise, and disgust (Ekman, 1992), are easier to communicate in music than other emotions (Bigand et al., 2005; Gabrielsson & Juslin, 1996; Laukka et al., 2013; Peretz, Gagnon, & Bouchard, 1998), potentially due to similarities in how basic emotions are expressed in vocal expression and music performance (Juslin, 2013b). Although sadness and joy (basic emotions) were two of the best represented and accurately recognised emotions (Kragness & Trainor, 2019) in this study, this theory does not explain the high ranking of calmness, which is not considered a basic emotion, and how it surpassed other basic emotions, such as anger and fear, which had a high identification rate in previous studies (Balkwill, Thompson, & Matsunaga, 2004; Balkwill & Thompson, 1999; Fritz et al., 2009). Interestingly, sadness, joy, and calmness have been reported as being three of the emotions most often attributed to music (Costa, Ricci Bitti, & Bonfiglioli, 2000; Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Vieillard et al., 2008). An alternative theory to the current findings is that music better expresses emotions, or rather, affective states, that can be explained without having a particular intent, unlike other emotions such as disgust, that are experienced in a specific, intentional situation context (Cespedes-Guevara & Eerola, 2018).

The comparison between Experiments 1 and 2 and the existing literature shows how overall, similar cue patterns were used to express specific emotions in music. The cues registered as the ones contributing most to the different emotion profiles (mode, tempo, and articulation) are used consistently to alter the emotional expression in music, while others sometimes varied between studies, such as the dynamics cue having a significant effect on all emotions in Experiment 1 but having only a significant effect on the portrayal of sadness in Experiment 2. In addition, the fact that different cue values have been reported as significantly affecting the emotional expression suggests that the cue values used are relative to the other cues being used to convey said emotion. Thus, this study provides new evidence that supports the notion that cues work together to communicate the different emotions in their relative context (Argstatter, 2016; Eerola, Friberg, & Bresin, 2013; Hevner, 1936; Lindström et al., 2003). For example, overall, mode had the highest impact on the emotion profiles; however, it did not significantly affect the portrayal of power in Experiment 1. Instead, the power emotion profile was holistically built using the combination of cues.

Most importantly, this work brings a novel contribution to the field, as, unlike previous studies exploring a single cue (Costa & Nese, 2020; Dalla Bella et al., 2001; Hailstone et al., 2009; Hevner, 1935; Lindström, 2006) or a restricted amount of cues and levels, this study extensively explores the effect of a combination of seven musical cues and their multiple cue levels on the emotion expressed in the music, providing insights into the merits of these two methodologies.

5.5 Discussion

This chapter has provided two sets of data investigating the same seven musical cues in relation to seven emotional expressions, using two different approaches: production and evaluation. The findings support the notion that musical cues and their additivity communicate distinct emotions in music (Juslin, 1997a, 2000, 2013b). Overall, five out of the seven cues (tempo, pitch, brightness, articulation, and mode) were utilised in the same manner across the two experiments. The dynamics cue varied between the two experiments, as it significantly contributed to all emotions in Experiment 1 but was not significant for all but one emotion (sadness) in Experiment 2. The instrumentation cue produced the most variance between the two experiments, where the instrument of choice was similar for three (anger, fear, and power) out of the seven emotions.

Overall, the findings from the two experiments are similar, but the differences between the results obtained with the two methodologies raise the question of which one could be deemed more useful. For example, the dynamics cue was the weakest contributing factor (Pseudo R^2 = 0.0002) in the evaluation experiment (Experiment 2). This might be because the differences between the dynamics levels were subtle and perhaps could not be perceived by the participants (Cousineau, Demany, & Pressnitzer, 2009), especially in an online study, where the experiment environment cannot be fully controlled (Dandurand, Shultz, & Onishi, 2008). Therefore, perhaps a different methodology that allows more control over the dynamics cue and research environment might be better suited when investigating dynamics.

Both methodologies have advantages, as well as limitations. The production approach allows for participant engagement and direct user experience, where participants have the opportunity to *show* us how they would change the music to express the intended emotions. Furthermore, the production approach allows for a substantially large parameter space to be explored in a relatively short time, which would not be possible with a systematic manipulation design, where all cue combinations would have to be pre-defined, rendered, and listened to by participants (Micallef Grimaud & Eerola, 2021). One downside to giving the participants free rein of the cue parameter space is the possibility that certain cue levels and combinations might not be explored. Another limitation of the production approach is that it was designed as a lab experiment, where the researcher could have full control of the research environment. This is a limitation that might not necessarily be that restrictive usually. However, due to the COVID-19 pandemic, face-to-face lab experiments have not been possible or limited, and thus, online methodologies that do not require physical contact would be ideal. A solution for this would be adapting the computer interface to an online setting.

One of the biggest advantages of a systematic manipulation approach is that the researcher has total experimental control on the cue combinations explored. This allows studying small differences between cue levels, which might not be explored by participants using a production task. Experiments with systematic manipulations can easily be administered online, making them readily accessible and available to a larger population. Additionally, using online crowdsourcing platforms might make recruiting more diverse samples easier (Aljanaki, Wiering, & Veltkamp, 2016). The downside of systematic manipulation is that it is unfeasible to investigate many cue combinations and levels simultaneously, as it easily leads to participant fatigue and lack of engagement (Lee & Müllensiefen, 2020). There is a possibility of optimising the comparison through fractional factorial designs and dividing subsets of the stimuli across the participants in an optimal fashion, but even these techniques will not remove the combinatorial problems inherent in this approach.

In conclusion, this chapter presented new insights on how combinations of seven musical cues shaped seven different emotion profiles in music, across two different methodologies used in cue-emotion research. Furthermore, utilising both production and systematic manipulation approaches allowed for a critical evaluation of methodologies used for this purpose and confirmed that similar cue-emotion patterns were discovered overall. However, the production approach created cue combinations that were better representatives of the different emotional expressions than the ones produced with the systematic manipulation approach. The production task was also quicker to administer.

Most importantly, this work explored cue-emotion mapping utilising an ambitious number of cues and exceeded past research on the cue-emotion space by using continuous cues with wide ranges and categorical cues with multiple levels. This gives us a glimpse of how real-time interactive approaches (Bresin & Friberg, 2011; Kragness & Trainor, 2019; Micallef Grimaud & Eerola, 2022; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Sievers et al., 2013) may be used to explore numerous, complex cue-emotion mappings that exist in real music. Although the findings from the two experiments in this chapter have shown us that overall, similar results in cue-emotion research are achieved with the two different methodologies, only a sliver of the cue-emotion space can be realistically explored simultaneously using a traditional, systematic approach, like the one used in Experiment 2. Furthermore, real music consists of a considerably larger number of cues and their combinations than explored here and is more complex than the reduced music samples used in systematic manipulation designs (Juslin & Lindström, 2010). Thus, the findings of this chapter suggest that production studies (analysis-by-synthesis) offer a promising way forward in uncovering how the cue-emotion space operates in real music.

In light of this, future studies investigating musical cues and emotion should move away from rigid, systematic manipulation approaches which restrict them to a finite number of cue combinations. Instead, they should focus on alternative ways that tackle the large parameter space more efficiently and further expand the cue-emotion space investigated to include other features that contribute to the emotional expression in music, such as harmony, sound space, articulation, other musical structural organisation principles than mode (other modes, tunings, etc.), panning of sound, and other aspects of timbre, such as spectral shape. It would be worthwhile to explore the variation in the cues across musical genres and traditions. Apart from analysis-by-synthesis methodologies that use computer interfaces or slider apparatus (Bresin & Friberg, 2011; Micallef Grimaud & Eerola, 2021; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014; Sievers et al., 2013), other techniques include the self-pacing method, which allows participants to express emotions in music using a number of expressive cues (Kragness & Trainor, 2016, 2019), the Markov Chain Monte Carlo with People (MCMCP) randomised algorithm which is used to understand participants' representations of perceptual objects and predict their behaviour (Sanborn, Griffiths, & Shiffrin, 2010) or Gibbs Sampling with People which uses a continuous-sampling paradigm of MCMCP and investigates how participants optimise cues for different emotional expressions (Harrison et al., 2020).

5.6 References

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., Busch, V., Lothwesen, K., Elvers, P., Fischinger, T., Schlemmer, K., & 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

Aljanaki, A., Wiering, F., & Veltkamp, R. C. (2016). Studying emotion induced by music through a crowdsourcing game. *Information Processing and Management*, *52*(1), 115–128. https://doi.org/10.1016/j.ipm.2015.03.004

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.

Battcock, A., & Schutz, M. (2019). Acoustically expressing affect. *Music Perception*, *37*(1), 66–91. https://doi.org/10.1525/MP.2019.37.1.66

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., Vieillard, 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. (2000). Emotional Coloring of Computer-Controlled Music Performances. *Computer Music Journal*, *24*(4), 44–63.

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

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., 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

Collier, W. G., & Hubbard, T. L. (2001). Musical scales and evaluations of happiness and awkwardness: Effects of pitch, direction, and scale mode. *American Journal of Psychology*, *114*(3), 355–375. https://doi.org/10.2307/1423686

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., & Nese, M. (2020). Perceived Tension, Movement, and Pleasantness in Harmonic Musical Intervals and Noises. *Music Perception*, *37*(4), 298–322.

Costa, M., Ricci Bitti, E., & Bonfiglioli, L. (2000). Psychological Connotations of Harmonic Musical Intervals. *Psychology of Music*, *28*(1), 4–22. https://doi.org/https://doi.org/10.1177/0305735600281002

Cousineau, M., Demany, L., & Pressnitzer, D. (2009). What makes a melody: The perceptual singularity of pitch sequences. *The Journal of the Acoustical Society of America*, *126*(6), 3179–3187. https://doi.org/10.1121/1.3257206

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.

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, *40*(2), 428–434. https://doi.org/10.3758/BRM.40.2.428

Eerola, T., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Perception*, *30*(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.

Elliott, T. M., Hamilton, L. S., & Theunissen, F. E. (2013). Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. *The Journal of the Acoustical Society of America*, *133*(1), 389–404. https://doi.org/10.1121/1.4770244

Fletcher, N. H., & Rossing, T. D. (1998). *The Physics of Musical Instruments* (2nd Editio). Springer.

Friberg, A., Bresin, R., & Sundberg, J. (2014). Analysis by Synthesis. In W. F. Thompson (Ed.), *Music in the Social and Behavioral Sciences: An Encyclopedia* (p. Vol.1, 44-45). SAGE Publications, Inc. https://doi.org/http://dx.doi.org/10.4135/9781452283012

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & 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., & 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., & Juslin, P. N. (2002). Emotional expression in music. In R. Davidson, K. R. Scherer, & H. Hill Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503–534). Oxford University Press.

Gabrielsson, A., & Lindström, E. (1995). Emotional Expression in Synthesizer and Sentograph Performance. *Psychomusicology*, *14*(1), 94–116.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In *Music and emotion: Theory and research.* (pp. 223–248). Oxford University Press.

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

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

Harrison, P. M. C., Adolfi, F., Marjieh, R., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., Larrouy-Maestri, P., & Jacoby, N. (2020). Gibbs Sampling with People. *ArXiv*.

Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine*, *29*(6), 1405–1410. https://doi.org/10.1017/S0033291799001221

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., Anderson, N., & Shanahan, D. (2014). "You Can't Play a Sad Song on the Banjo:" Acoustic Factors in the Judgment of Instrument Capacity to Convey Sadness. *Empirical Musicology Review*, *9*(1), 29. https://doi.org/10.18061/emr.v9i1.4085

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946. https://doi.org/10.1111/2041-210X.12225

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*, *I*(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., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae*, *5*(Special Issue 2001-2002), 63–122. https://doi.org/10.1177/10298649020050s104

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

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., & Trainor, L. J. (2016). Listeners lengthen phrase boundaries in self-paced music. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1676–1686. https://doi.org/10.1037/xhp0000245

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*. Department of Pscyhology, Keele (CD-rom).

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. In *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* (Vol. 51, Issue 4, pp. 336–353). Canadian Psychological Association. https://doi.org/10.1037/1196-1961.51.4.336

Lahdelma, I., & Eerola, T. (2016). Mild dissonance preferred over consonance in single chord perception. *I-Perception*, *7*(3). https://doi.org/10.1177/2041669516655812

Landsheer, J. A., & Van Den Wittenboer, G. (2002). Fractional Design Wizard: A computer program for cost-effective experimental research design. *Behavior Research Methods, Instruments, and Computers*, *34*(1), 117–127. https://doi.org/10.3758/BF03195430

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*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Lee, H., & Müllensiefen, D. (2020). The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability. *Attention, Perception, and Psychophysics*, *82*(7), 3658–3675. https://doi.org/10.3758/s13414-020-02058-3

Lepa, S., Herzog, M., Steffens, J., Schoenrock, A., & Egermann, H. (2020). A computational model for predicting perceived musical expression in branding scenarios. *Journal of New Music Research*, *49*(4), 387–402. https://doi.org/10.1080/09298215.2020.1778041

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.

McAdams, S. (2019). The perceptual representation of timbre. In K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, & R. R. Fay (Eds.), *Timbre: Acoustics, Perception, and Cognition.* (pp. 23–57). Springer.

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

Micallef Grimaud, A., & Eerola, T. (2022). An Interactive Approach to Emotional Expression Through Musical Cues. *Music and Science*, *5*, 1–23. https://doi.org/10.1177/20592043211061745

Morreale, F., Masu, R., & Angeli, A. De. (2013). Robin: An Algorithmic Composer For Interactive Scenarios. *Sound and Music Computing Conference 2013, SMC 2013*, 207–212.

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134). https://doi.org/10.1098/rsif.2017.0213

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

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., & 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

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. K., & 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

Saitis, C., & Siedenburg, K. (2020). Brightness perception for musical instrument sounds: Relation to timbre dissimilarity and source-cause categories. *The Journal of the Acoustical Society of America*, *148*(4), 2256–2266. https://doi.org/10.1121/10.0002275

Sanborn, A. N., Griffiths, T. L., & Shiffrin, R. M. (2010). Uncovering mental representations with Markov chain Monte Carlo. *Cognitive Psychology*, *60*(2), 63–106. https://doi.org/10.1016/j.cogpsych.2009.07.001

Scherer, K. R., & Oshinsky, J. S. (1977). Cue Utilization in Emotion Attribution from Auditory Stimuli. *Motivation and Emotion*, *1*(4), 331–346.

Schutz, M., Huron, D., Keeton, K., & Loewer, G. (2008). The Happy Xylophone: Acoustics Affordances Restrict An Emotional Palate. *Empirical Musicology Review*, *3*(3), 126–135. https://doi.org/10.18061/1811/34103

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

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, *10*(1), 79–89.

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.

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

Watson, K. (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.

Wheeler, R. E. (2004). *optMonteCarlo. Algdesign. The R project for statistical computing*.

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

5.7 Supplementary Material

5.7.1 Pilot Experiment: Emotional Expressivity of Different Instruments

To determine which instruments had a substantial emotional range and were capable of expressing different emotions, we carried out a pilot study which tested the ability of 27 instruments detailed in Table 5.5, to express the following emotions: sadness, joy, calmness, anger, fear, power, and surprise. The following sequence C3-E3-G3-E3-C3 was exported in Logic Pro X using instruments available in the Vienna Symphonic Library (VSL) sound library. The sequence was exported both in *legato* and *staccato* for instruments that supported the aforementioned articulation methods, and one time for instruments such as celesta, marimba, and xylophone which have one articulation setting. 22 participants in an online study rated how much of the aforementioned seven emotions the different instruments were capable of expressing in the musical sequence. Based on the ratings, the instruments were then ranked with respect to their emotional expressivity range within the articulation. This was attained by calculating the maximum emotional expressivity rating difference between the two articulation types for each instrument (instruments that have only one articulation setting were thus omitted from the ranking). Table 5.5 shows the ranking of the instruments in terms of the emotional expressivity range across the articulation, and the mean rating for each instrument across all seven emotions. Four instruments were selected for each instrument family used in the interface (brass, woodwinds, and strings). We tried to choose instruments that ranked high in the emotional expressivity range. The Vienna horn (ranked #3) and euphonium (ranked #6) were selected as part of the brass ensemble. The cello (#2), violin (#4), and viola (#13) were put forward as part of the strings ensemble. The flute (#1), French oboe (#8), clarinet (#9), and bassoon (#14) made up the woodwinds ensemble.

Table 5.5. Ratings of emotional expressivity ability of instruments.

Ranking	Instrument	Instrument Family	Expressivity Range	Mean
1	**Flute***	Woodwinds	2.15	2.05
2	**Cello***	Strings	1.95	2.14
3	**Vienna Horn***	Brass	1.80	1.93
3	Heckelphone		1.80	2.01
4	Basset		1.75	1.99
4	**Violin***	Strings	1.75	2.11
5	English Horn		1.65	1.90
6	Alto Sax		1.60	2.02
6	**Euphonium***	Brass	1.60	2.02
7	Bass		1.58	2.02
8	**French Oboe***	Woodwinds	1.50	1.92
9	**Clarinet***	Woodwinds	1.45	1.98
10	Soprano Sax		1.40	2.01
11	Trumpet		1.35	2.06
12	**Bassoon***	Woodwinds	1.05	2.02
13	**Viola***	Strings	1.00	2.03
14	Tenor Sax		0.95	2.03
14	Flugelhorn		0.95	2.04
15	Cornet		0.85	1.96
NA	Celesta		NA	2.05
NA	Harp		NA	1.99
NA	Harpsichord		NA	2.02
NA	Marimba		NA	1.89
NA	Organ		NA	2.24
NA	Piano		NA	2.07
NA	Vibraphone		NA	2.02
NA	Xylophone		NA	1.82

Note. * Instruments marked in bold and with an asterisk were taken forward to the study.

5.7.2 Experiment 1 COVID-19 Safety Measures

Experiment 1 was structured in two parts. The first part of the study required participants to answer a number of demographic questions such as age, gender, and musical expertise. This was administered online via a short survey on Qualtrics to minimise participant and researcher interaction due to COVID-19. Instructions and a video demonstration for the second part of the study (the musical task using *EmoteControl*) were also presented to the participants online. The musical task was done in person. Participants could choose where to carry out the experiment from a number of locations, depending on their personal preference and convenience, with each location consisting of a quiet room with a desk on which to put the apparatus. Speakers were utilised rather than headphones, for hygienic and safety purposes due to COVID-19. Speakers were set to the same volume prior to the experiment. A wireless keyboard and mouse were used rather than the laptop's own keyboard and mousepad, which were easier to wipe down. All apparatus was sanitised between each participant. Both researcher and participant wore face masks and sat at a distance from each other.

Chapter 6. General Discussion

Abstract

This final chapter provides a summary of all the work carried out in this thesis and how it addressed previous limitations. The unique contributions produced within this work are highlighted. The implications and uses of the material stemming from this work are discussed within the music and emotion realm and other domains. An overview of limitations of this thesis is given and potential avenues for future work are outlined. In conclusion, the work in this thesis presented an interactive production approach to efficiently explore the cue-emotion space and provided new insights on how different emotional expressions are shaped by cues in music.

6.1 Introduction

The current project aimed to provide a better understanding of how musical cues are used to communicate emotions in music, whilst also addressing multiple limitations present in the current literature. Firstly, the thesis aimed to tackle constraints pertaining to methodological approaches, as typical experiment designs are severely restricted with regards to the possible number of cues and combinations simultaneously investigated, as well as the amount of experimental control over them. This restriction was overcome by creating a novel interactive interface called *EmoteControl* which allows for real-time manipulation of musical cues. Secondly, the thesis addressed limitations pertaining to the musical works used as stimuli in music emotion research. The choice of musical stimuli used in existing literature was predominantly commercial music, which gave rise to familiarity bias, experimental control, and ecological validity concerns. To address these limitations, this thesis presented a new set of specifically composed tonal musical pieces aiming to represent ‘real music’ to be used as stimuli. Thirdly, the project aimed to expand current knowledge on emotions perceived in music by investigating a varied set of emotions outside of any one particular emotion framework. This was also taken into consideration during the creation of the new musical stimuli, where pieces were composed to portray nine different emotional expressions, rather than the usual three (sadness, happiness, and anger) which made up more than half of existing stimuli used in music emotion studies (Warrenburg, 2020a). Lastly, the thesis evaluated the production approach implemented in the current project against a traditional systematic manipulation approach, to assess the approaches’ efficiency and compatibility of results, with the aim of providing insight on the methodological approaches’ suitability for a better exploration of the cue-emotion space in music.

This final chapter recapitulates all the contributions and findings presented in the thesis, contextualising them with regards to existing literature and the music emotion field. The implications of the project are outlined, along with possible uses and future research ideas stemming from the work in this thesis.

6.2 Summary of Findings and Implications

This section will summarise the findings presented in the current project and will discuss consider the successful identification (or lack of) of intended emotion across experiments, the relative contributory weight of the individual cues under investigation, and their role in the cue combinations used to convey the target emotional expressions. For convenience, the three experiments in Chapter 4 (i.e., the 28 musical pieces’ validation experiment, the first *EmoteControl* experiment, and the second validation experiment) will be denoted as Studies 1, 2, and 3, respectively. The two experiments detailed in Chapter 5 (i.e., the second

EmoteControl experiment and the systematic manipulation experiment) will be referred to as Studies 4 and 5.

6.2.1 Emotion Communication in Music

This project set to determine how nine different emotional expressions - sadness, joy, anger, fear, love, longing, power, surprise, and calmness - are communicated in Western tonal musical piece, following the assumption that music may carry emotional meaning (Dowling & Harwood, 1986), and the fact that the emotional expressions under investigation have been previously reported as being expressed by music (Juslin & Laukka, 2004; Kreutz, 2000; Lindström et al., 2003; Zentner, Grandjean, & Scherer, 2008). The different levels of communication accuracy achieved for the different emotions in question will be detailed, followed by potential reasons why not all emotions were accurately recognised in the musical pieces.

Across all empirical work in this thesis, calmness, sadness, and joy were the three most accurately communicated emotions through the music. When looking at recognition accuracy rates from the listeners' perspective (i.e., the decoding of emotional content in the music), findings showed that all the pieces I composed with the intention of conveying the aforementioned emotions, detailed in Chapter 3, were rated highest for their target emotion in Study 1. The original pieces along with the variations created by participants in Study 2 were also rated highest for their target emotion in Study 3. It is also interesting to note how calmness, sadness, and joy emotions were identified as the most straightforward to encode in music from my perspective as the composer, in Chapter 3. Furthermore, the cue-emotion models used by participants in the production studies (Studies 2 and 4) to portray calmness, sadness, and joy in the music had the highest accuracy prediction rates. This means that the target emotions may be predicted the majority of cases with the cue combinations used. For example, the particular calmness cue-emotion model used by participants in Study 2 may accurately predict calmness in music 91% of the time. The sadness cue-emotion model may predict sadness 75.3% of the time, and joy may be predicted 67.4% of the time with the cue-emotion model used by participants. A similar pattern of effectiveness in the cue-emotion models used by participants in Study 4 was observed. The cue-emotion models used for sadness, joy, and calmness in Study 5 also had the highest emotion recognition accuracy ratings amongst all emotion models.

Sadness, joy, and calmness have regularly been reported as being emotions easily recognised in music. Sadness and joy are classified as being basic emotions (Ekman, 1992; Juslin, 2013b; Plutchik, 1994), where the basic emotion theory suggests that this type of emotions may be more easily recognised due to them being core emotions that have an evolutionary origin, and thus are regarded by listeners as common emotions, due to their exposure to them through other communication channels (Juslin, 2019b) and throughout their development (Dolgin & Adelson, 1990; Stachó et al., 2013). However, basic emotion theory does not explain why calmness is consistently found as being one of the most accurately recognised emotions in music, together with sadness and joy (Cespedes-Guevara & Eerola, 2018; Gabrielsson & Juslin, 2003; Thompson & Robitaille, 1992; Vieillard et al., 2008). In fact, in an extensive review on emotions perceived in music carried out by Gabrielsson and Juslin (2003), it was reported that a high level of listener agreement was found for sad/melancholic, happy/triumphant, gentle/relaxing, and angry/violent emotion categories. Other studies have stipulated that the high recognition accuracy ratings of calmness, sadness and joy may be due to these emotions being common in music and frequently perceived during everyday music listening experiences (Juslin & Laukka, 2004; Kallinen & Ravaja, 2006; Lindström et al., 2003).

For the most part, anger and fear emotions were successfully communicated across the different studies. However, the two emotions were sometimes confused with each other. All three musical pieces I composed with the intention of conveying fear were rated as such by participants in Study 1. However, only one of the pieces aiming to convey anger was successfully perceived as such, whilst the other two pieces were interpreted as communicating fear, rather than anger. Anger and fear were sometimes also mixed up in Study 3 where participants rated the original pieces against the variations created by participants with *EmoteControl*. Both pieces aiming to convey fear were correctly identified whilst only the original version of the piece portraying anger was correctly rated highest for the target emotion. On the other hand, the second version of the piece attempting to convey anger was rated highest for fear. The fact that pieces aiming to convey anger and fear were sometimes mixed up concurs with previous studies which have also shown that these two emotions are often confused (Cunningham & Sterling, 1988; Kragness et al., 2021), albeit being basic emotions. The confusion between anger and fear in music may be due to them existing on similar points on the valence-arousal planes (Russell, 1980), which may lead to similar musical features and subtle variations being used to portray these two emotions, which may not be distinctly differentiated by listeners. In fact, distinguishing between anger and fear was also found to be challenging during the compositional process.

Although power may be considered as both a negative- and positive- valenced emotion and thus, interpreted in multiple ways (Kawakami et al., 2013; Zentner, Grandjean, & Scherer, 2008), I chose to interpret power as a negative-valenced emotion, opting to use a minor mode, dissonance, and a relatively low pitch. Although this decision led to distinguishing between power, anger, and fear being challenging during the compositional process, participants in Study 1 successfully decoded power in all pieces when it was the target emotion. Interestingly, participants in Study 2 used a major mode and utilised a relatively higher pitch level than the original piece, which suggests that their interpretation of power was different to mine. However, the cue-emotion model produced by participants only predicts power in the music 49.4% of the time, which suggests that it may not be a reliable way of communicating power through music. This is confirmed in Study 3, where the original piece was rated highest for power, whilst the variation of the piece was misinterpreted as joy. The way power was communicated in Study 4 and perceived in Study 5 also varied, in terms of mode (major mode was used in both, but it did not have a significant effect on shaping power in Study 4), and other cues used. However, the cue-emotion models used to shape power across the experiments all had low emotion accuracy prediction rates. The variations in how power was attempted to be communicated and the inaccuracies in recognition rates show that power was challenging to try and portray and interpret in music, from both the encoders and decoders' perspectives. This may have been due to the possibility of perceiving power as both a positive- and negative-valenced emotion. However, my attempt at portraying power as negative-valenced was successfully recognised – suggesting that this type of power is more easily encoded and decoded in music.

Surprise was also one of the emotions I found most challenging to compose for, due to its potential multiple aspects. The possible different representations of surprise in music are a good example that showcases the multifaceted properties of emotions. Thus, I attempted to express surprise in different ways in the music, also varying in modality. Interestingly, the piece accurately recognised as conveying surprise by participants in Study 1 was the one composed in harmonic minor mode which seemed to express a more negative degree of surprise. Similarly, the participants in Study 3 accurately rated the original piece expressing surprise as its target emotion, whilst the variation produced in Study 2, which was in major mode, was rated highest for joy. Although the cue-emotion models used to communicate surprise in emotions had low accuracy predictive rates, the cues were used quite similarly (except for the change in mode between the original piece composed to portray surprise and its variation) across the different experiments. In fact, surprise was mostly coded in minor mode, with a fast tempo, *staccato* articulation, a high pitch level, and a high brightness level in the majority of experiments in this thesis. This suggests that although surprise may be viewed in different ways (e.g., a happy surprise or a bad surprise), the participants' perception of surprise, both in the production and listening studies, leaned more towards the negative-valenced type of surprise. This might be due to moments of surprise always starting with an initial negative affect as a violation of expectancy occurs (Huron, 2006).

It is interesting to note when power and surprise were attempted to be communicated with a major mode, both emotions were misinterpreted as joy. When positively valenced, power and surprise have similar levels of valence and arousal as joy and are represented with similar musical features, which may lead to listeners being unable to distinguish between the three. Furthermore, listeners may be more likely to interpret the emotional content of the music as joy, since it is usually one of the most commonly present emotions in music (Juslin & Laukka, 2004).

Love and longing emotions were attempted to be conveyed in the original pieces composed in Chapter 3, however, none of the pieces successfully conveyed their intended emotion when rated by participants in Study 1. Interestingly, love and longing pieces were mostly perceived as conveying calmness. It has been reported that recognition accuracy is higher for broad emotion categories or distinct points in the quadrants of the affective circumplex space, and it decreases with the increase of specificity of emotions (Gabrielsson & Juslin, 2003; Juslin, 1997b; Juslin & Timmers, 2010). For example, when participants were asked to determine whether a musical piece was portraying sad-serious or happy, the majority successfully recognised the intended emotional expression of the music. However, accuracy ratings of emotion recognition decreased when participants were asked to identify whether the music was portraying death, sorrow, or religion within the broad category of sad-serious (Rigg, 1942). The confusion of emotional expressions which have similar valence and arousal properties or may exist within the same broad emotion category may explain why anger and fear emotions are sometimes mixed up, as well as other emotions such as calmness, love, longing, and sadness. Timmers and Ashley (2004, 2007) have in fact reported that low-activity emotions love and sadness were not successfully communicated by performers and distinguished by listeners, whilst others have reported similar findings about anger and fear emotions (Cunningham & Sterling, 1988; Kragness et al., 2021).

Additionally, it is very likely that other features in the musical context would have helped communicate particular emotions better in the music. Love and longing have been reported as being frequently portrayed in lyrics (Juslin, 2019b; Kreutz, 2000), whilst single chords in a minor triad root position and major 7^{th} chords in a third inversion position were rated as expressing nostalgia/longing and rated particularly higher as conveying nostalgia/longing when played with a strings timbre as against a piano timbre (Lahdelma & Eerola, 2014, 2015). Timmers and Ashley (2007) investigated how different ornamentations were used by performers to express weak and strong intensity levels of happiness, sadness, anger, and love. They reported that the performers were overall successful in communicating the intended emotions using ornamentations, except for happiness.

Apart from the use of cues and their combinations, the accuracy of emotion recognition in music also depends on the listener and the context of the music (Davies, 2001; Hevner, 1937; Juslin, 2019b). It may be possible that love and longing may have not been recognised due to the listeners having different experiences or interpretations of how they should be expressed in music, as against my own views (or rather, the other way round). Barrett (2004, 2017) presented the concept of emotion granularity which postulates that individuals may identify different nuances of a particular emotion depending on their experience with specific degrees of said emotion. Similarly, it may be that emotion terms hold different meanings for individuals, or certain emotion terms are preferred over others. For example, 'tenderness' and 'love' are two terms that tend to be used interchangeably or denoted together as love/tenderness in studies (Juslin & Laukka 2003). Furthermore, certain emotional expressions might be made up of multiple emotional states, thus making it difficult to reach listener agreement (Eerola & Peltola, 2016; Quinto, Thompson, & Taylor, 2014; Taruffi & Koelsch, 2014; Timmers & Ashley, 2007; Warrenburg, 2020b). Perhaps providing multiple terms that may describe the intended emotional expression would have increased the emotion recognition accuracy in the musical pieces. For example, Hevner (1936, 1937) utilised an adjective circle where emotion terms similar in meaning and valence-arousal properties were grouped together. This adjective circle was later updated by Schubert, which interestingly put terms such as 'calm' and 'tender' in the same word cluster. Similarly, emotion terms denoting high arousal and negative valence ('angry', 'tense', 'restless', and 'agitated') were also grouped together, which may explain why anger and fear tend to also be confused (Schubert, 2003, p. 1121).

6.2.2 Cue utilisation

The use of tempo, mode, articulation, pitch, dynamics, and later, instrumentation, to shape different emotional expressions in music was explored in multiple experiments throughout the thesis. In this section, the effect of the individual cues on shaping different emotions in the music will first be discussed, followed by a summary of how the cues were used together to portray the different emotions, highlighting similarities and differences across experiments.

6.2.2.1 The effect of individual cues on emotion communication in music

The importance of the cues in shaping the different emotions was calculated in Studies 2, 4, and 5. Across the three studies, mode was consistently the cue with the biggest contributory effect to shaping the different emotional expressions in the music.

Tempo, articulation, and instrumentation followed mode in rank, varying in order of importance across the two experiments in Chapter 5. Pitch, brightness, and dynamics were the least effective cues in both production and systematic manipulation approaches. However, it was interesting to note that the dynamics cue in the systematic manipulation experiment only had a significant effect when portraying sadness. Brightness was the second cue with the least significant effect in the latter experiment.

Mode was followed by tempo and articulation, varying in rank across the different studies. Tempo was reported as having the second largest contributory effect in the two *EmoteControl* studies (Studies 2 and 4), whilst articulation ranked second and tempo came in third in the manipulation study (Study 5). Mode and tempo have been consistently reported as being salient in shaping emotions in music, albeit sometimes alternating in rank (tempo being first, followed by mode) (Battcock & Schutz, 2019, 2021; Eerola, Friberg, & Bresin, 2013; Hevner, 1935, 1936, 1937; Juslin, 1997a; Peretz, Gagnon, & Bouchard, 1998; Schutz, 2017). The significant effect of mode on shaping emotions in music may also explain why the variations created to portray power and surprise by participants in the first *EmoteControl study* were perceived differently than the originals – where the biggest change in musical features was in fact the mode, going from minor to major. Consequentially, both surprise and power variations were perceived as portraying joy. Other studies which looked at cues that did not include mode also denoted tempo as the most powerful cue in shaping different emotions within their selection of cues being investigated (Juslin, 1997b; Scherer & Oshinsky, 1977).

It is also interesting to note that mode and tempo are said to carry emotional meaning due to different underlying mechanisms. Mode is a music-specific cue which may not exist in other channels of communication, such as speech or movement. Moreover, it is a cue that exists within a Western music framework, making it culture-specific (Balkwill & Thompson, 1999). In light of this, it is theorised that emotional meaning may be perceived in particular modes, depending on their regular association with specific events, such as music played at a funeral typically being in minor mode. It is important to note that since mode is a culture-specific cue that individuals use in a Western culture to discriminate between emotions (Dalla Bella et al., 2001), and it may be perceived differently in music from other cultures (Lahdelma, Athanasopoulos, & Eerola, 2021). However, in this particular context, with the music being Western tonal music and listeners being mostly from a Western culture and thus, familiar with the style of music, the mode was the most significant factor conveying emotional meaning in the music. On the other hand, tempo, is said to be a universal cue, where it is likely to be employed similarly around the world, due to its evolutionary origin (Balkwill & Thompson, 1999; Fritz et al., 2009; Juslin, 1997a, 2019b). Tempo has been reported as being used similarly across other communication channels such as speech (Juslin & Laukka, 2003) and movement (Davies, 2001), and it is thus considered to be a salient cue in communicating different emotional expressions in music.

Unlike mode and tempo, articulation is regarded as an expressive cue (Gabrielsson, 2002; Livingstone & Brown, 2005), and it is regarded as one of the primary expressive cues in defining the emotional expression in music (Juslin & Timmers, 2010). Similar to emotion communication in tempo, articulation is perceived as able to communicate emotional meaning in music due to its physiological origin in vocal communication, and thus mimicking acoustic cue-emotion patterns (Juslin, 1995; Juslin & Laukka, 2003).

Brightness, dynamics, and pitch were found to have the least overall contributing effect to the expressed emotion in the musical excerpts across the three studies, with brightness being at the bottom for both *EmoteControl* studies, and dynamics being last for the systematic manipulation study. Contrasting findings on the contributory effect of the dynamics cue on the emotion conveyed have been reported in previous literature. On one hand, the dynamics cue having a small effect on the emotion expressed complements previous studies which also found that the dynamics cue ranked low with regards to its effect on the emotion conveyed, within the cue combinations being investigated (Juslin, 1997b; Scherer & Oshinsky, 1977). However, other studies have reported the dynamics cue as being a prominent cue in shaping different emotion profiles (Juslin & Laukka, 2003; Juslin & Timmers, 2010), particularly since it is also regarded as a universal cue, due to its similar uses in emotion communication in speech. Similarly, pitch has been said to be salient in emotion communication through music, and it has been closely linked to speech, where the same patterns have been used to communicate different emotions in both speech and music (Juslin & Laukka, 2003).

This discrepancy in the contributory effect of the cues on the emotional expression might be explained by the fact that the functionality of a cue is relative to the other cues under investigation and also depends on the particular musical pieces being used (Gabrielsson & Juslin, 2003; Hevner, 1937; Juslin, 1997b; Juslin & Lindström, 2010; Laukka et al., 2013). In fact, the findings of this project also revealed that the relative contributory weight of the cues varied across emotional expressions communicated.

For example, although mode had the overall strongest effect on shaping the different emotions in music, it did not offer a significant contributory effect on the portrayal of power in Study 4 and in shaping surprise in Study 5. Similarly, tempo did not significantly contribute to the shaping of power in Study 2, and articulation was not significant in the portrayal of fear in Study 5. On the other hand, dynamics and brightness were overall, the least contributing cues to emotional expression across the studies in question. However, they had a significant role in conveying power when mode and tempo did not.

6.2.2.2 Cue combinations used to portray different emotions in music

The findings from the empirical experiments in this thesis identified how the cues were used in combination to portray the different emotions in the music. The production studies (i.e., the two *EmoteControl* experiments) presented cue combinations used by participants to portray their interpretation of sadness, calmness, joy, anger, fear, power, and surprise in the musical pieces. The systematic manipulation study provided another set of cue combinations which were perceived as conveying the same target emotions. Overall, the empirical studies showed that the cues had a significant main effect on the different emotions trying to be conveyed. This means that for the most part, the cue combinations used for the particular emotional expressions were not by chance, but rather, were used in a specific way to communicate the target emotions in the music. The two validation studies (Studies 1 and 3) that assessed the original pieces and the variations of the pieces (only the second validation study assess the variations) also helped determine how the cue combinations used in the original musical pieces and their variations fared in communicating the intended emotions. In total, four different sets of cue combinations were created for each of the emotional expressions under investigation, which will be discussed holistically in this sub-section.

The findings revealed how certain emotional expressions were consistently shaped with the same cue levels and combinations throughout the thesis, whilst others varied across studies. A similar cue combination was observed for the portrayal of sadness in music. Across all studies and variations of the pieces, sadness was communicated with a slow tempo, minor mode, *legato* articulation, a low pitch level, soft dynamics (Gagnon & Peretz, 2003; Hevner, 1936; Lindström, 2006; Scherer & Oshinsky, 1977; Thompson & Robitaille, 1992), varying slightly in brightness levels across studies, and in the instrumentation used (when given the choice). Similarly, calmness was mostly portrayed the same way across the different studies. A major mode, a slow tempo, *legato* articulation, and a low brightness level were used to express calmness in all studies, whilst the pitch varied from a low to a moderate level. Subtle differences were also present in the dynamics cue, varying between a soft to moderately soft level across studies. Furthermore, the use of dynamics was cue did not have a significant effect on the portrayal of calmness in Study 2. The instrumentation cue brought the most variation, as a woodwinds ensemble and a brass ensemble were chosen in different studies. Joy was communicated by a fast tempo, relatively high pitch level, a high brightness level, *staccato* articulation, and a major mode (Akkermans et al., 2019; Bresin & Friberg, 2000, 2011; Gabrielsson & Juslin, 1996; Gabrielsson & Lindström, 1995; Hevner, 1937; Peretz, Gagnon, & Bouchard, 1998; Quinto, Thompson, & Taylor, 2014) across all studies. The use of the dynamics cue varied in between studies whilst it did not offer a significant contribution to the portrayal of joy in Study 5. The instrumentation cue was also used differently, with a brass ensemble being used in Study 4 (Bresin & Friberg, 2011) as against a woodwinds ensemble in Study 5.

The consistent cue patterns and the fact that all pieces aiming to convey calmness, sadness, or joy were correctly identified, implies that the individuals who took part in these studies have a rather distinct idea of how calmness, sadness, and joy should sound like in music. This was also confirmed with the high accuracy prediction rates calculated for the cue-emotion models in question. The newly composed musical excerpts and the outcomes from the rating studies suggest that calmness, sadness, and joy may have more distinctly identifiable properties in music while others are less well defined.

Surprise was also consistently communicated across studies with a fast tempo, a high pitch level, loud dynamics, *staccato* articulation, and a high brightness level. A minor mode was mostly used, apart from Study 2 where a major mode was utilised in an attempt to convey surprise. However, as discussed in the previous section, the piece in major mode was not successfully interpreted by participants in Study 3. Furthermore, the instrumentation cue had the most variance in the last two studies, where it did not offer a significant contribution to conveying surprise in Study 4, whilst a woodwinds ensemble was used in Study 5.

Anger and fear emotions tended to also be portrayed with roughly similar cue combinations across studies. Similar musical features and subtle variations were used to portray these two emotions, and I also found it challenging to distinguish between anger and fear in the music during the compositional process. A minor mode, fast tempo, and *staccato* articulation were used across all studies to portray anger. A low pitch level was mostly used, except in Study 2 where a high pitch level was chosen. When given the option (Studies 4 and 5), a strings ensemble was chosen to portray anger. The brightness cue did not offer a significant contribution in any of the studies, whilst the dynamics cue had the most variation across studies; with soft dynamics used in Study 2, loud dynamics used in Study 4, and having a non-significant effect on anger in Study 5. Mode and tempo were the two cues most consistent in the portrayal of fear, with a minor mode and fast tempo being used across studies. A strings ensemble was consistently chosen when given the option, and a low brightness level was used across studies, although the brightness cue was non-significant in the portrayal of fear in Study 5. Articulation, pitch, and dynamics varied in levels and significance across studies. The dynamics cue varied from moderately soft (Study 2) to moderately loud (Study 4) and was non-significant in Study 5, whilst the pitch cue varied from low (Studies 4 and 5) to high (Study 2). *Detaché* was chosen in Study 2 to portray fear, whilst *staccato* articulation was used in the rest of the studies. It is to be noted that the *detaché* articulation option was only available in Study 2 (when a strings virtual instrument was the default instrument and only option). Since it is an articulation technique exclusive to string instruments, it was not made available in Studies 4 and 5, where participants could select between a strings, brass, or a woodwinds ensemble.

The cue combinations attempting to portray power in music featured the most variation across studies. The piece portraying power was originally composed in minor mode; however, a major mode was used in Studies 2 and 5. Furthermore, mode was not significant in the shaping of power in Study 4. *Staccato* articulation was used in Studies 4 and 5, whilst *detaché* was used in Study 2. The tempo, pitch, dynamics, and brightness cues varied from low to high levels and also varied in significance (i.e., sometimes being non-significant) across the studies. Lastly, power was attempted to be represented with a brass and/or strings ensembles in Studies 4 and 5. Although the cue combinations aiming to convey power and surprise differed across studies, most of the cue levels used had a significant effect on portraying the target emotion, which implies that the intended emotion was specifically shaped in different ways, supporting the notion that the same emotional information may be coded differently in music, potentially due to the cues being partially redundant and overlapping in some information (Juslin, 2000).

This thesis set out to investigate how seven musical cues would be used to convey different emotional expression in music, under the assumptions that musical cues may hold emotional meaning (Dowling & Harwood, 1986; Juslin & Sloboda, 2012), that the probabilistic and partially redundant cues would be able to communicate particular emotional expressions within their combinations in the music (Juslin, 1997a), which may be successfully decoded by listeners, subject to them being familiar with the style of music and capable of understanding it (Davies, 2001; Hevner, 1937; Juslin, 2019b).

The findings of this thesis confirmed that calmness, sadness, joy, anger, fear, surprise, and power emotions were able to be encoded in the Western tonal musical pieces and decoded by participants, varying in accuracy rates. The findings support the notion that cues are probabilistic and partly redundant as individual cues, which cannot pinpoint to a specific emotional expression on their own. However, the layering of the cues helped convey particular emotional expressions (Eerola, Friberg, & Bresin, 2013; Juslin, 2000; Juslin & Lindström, 2010). The three cues that had the most significant contributory effect on shaping the different emotions in the music (mode, tempo, and articulation) were the ones used most consistently across studies (e.g., a slow tempo, minor mode, and *legato* articulation was consistently used to portray sadness across studies). This suggests that mode, tempo, and articulation were the most salient cues in shaping the different emotional expressions in the music and thus, providing the core emotional coding in the music, whilst the remaining cues, such as pitch and instrumentation added extra information which may help in conveying the target emotion. Moreover, dynamics and brightness were found to have the least importance in shaping the different emotions in music.

The fact that certain cues such as dynamics and pitch, which have been reported in previous literature as being salient cues in emotion communication in music, had a small contributory weight in emotion communication in this work, supports the theory that the effect of the individual cues are relative to the other cues in the cue contribution under investigation (Hevner, 1936, 1937). This may also explain why the brightness cue was non-significant or offered a small significant contributory weight on the emotion communicated, compared to the other central cues. Furthermore, it is also worth noting that brightness is a component of timbre (McAdams, 2019; Saitis & Siedenburg, 2020) which may also have dependencies on pitch and dynamics. Although brightness may be closely related to dynamics and pitch, the cues sometimes varied in significance across emotional expressions. For example, when portraying anger in Studies 2 and 4, the brightness cue levels were non-significant in their contribution, however, the dynamics levels were. Thus, it would have been impossible to determine the cues' independent contribution to the emotional expressions if potentially related cues were coupled up. In fact, the inability to decouple expressive cues in performance studies in order to determine the effect of the cues on the listeners' emotion judgments has been reported as a limitation of performance studies (Juslin & Timmers, 2010).

These findings support the notion that the context within which the musical experience exists, together with the previous knowledge of the listener also affect the accuracy of successfully emotion recognition in music (Hevner, 1937; Juslin, 2019b). If this was not the case, mode would not have had the most significant impact on shaping the different emotional expressions in the musical pieces, since it is a culture-specific (and music-specific) cue. Furthermore, the findings parallel previous literature which suggests that emotional expression may be conveyed in music through the layering of different types of codes (e.g., iconic, associative) (Juslin, 2013b, 2019e).

The studies also showed that some cues, albeit having an overall large effect on the expressed emotion (e.g., mode and tempo) did not contribute to the target perceived emotion in certain instances, providing evidence that the emotional expressions are shaped by the summation of the cues (Argstatter, 2016; Eerola, Friberg, & Bresin, 2013; Gabrielsson, 2008; Quinto, Thompson, & Taylor, 2014; Ramos, Bueno, & Bigand, 2011), thus supporting the notion of Brunswik's lens model which states that emotional expressions are encoded in music using a combination of cues, which in turn, are also used by listeners to accurately decode the intended emotion (Brunswik, 1956; Juslin, 1997b, 2000). These findings thus encourage future studies to shift the focus from investigating the effect of individual or a small number of cues on the perceived emotion, to looking at how a larger number of cues are used in combination to shape the different emotional expressions in music.

6.3 Potential Uses of the Research

6.3.1 The Implications and Uses of *EmoteControl* for Music Emotion Research and Beyond

The current project saw the creation of *EmoteControl*, its formal evaluation, and implementation as a tool allowing a larger area of the cue-emotion space to be explored simultaneously in a shorter amount of time than in a traditional listening experiment with systematic manipulation of the cues. These findings imply that interactive, user-centred production approaches provide an efficient way of probing the large cue-emotion space. Furthermore, allowing participants to take control of the cue-emotion space themselves might produce more informative results than when employing a forced-choice paradigm. More specifically, apart from confirming whether the cues were utilised in a similar manner to previous literature, production studies allow us to also identify distinct mean cue combinations utilised across participants to convey particular emotional expressions (Bresin & Friberg, 2011). Furthermore, the data gathered using this approach presents the opportunity to look at potential differences in cue usage across individuals (e.g., looking at musicians vs non-musicians, participants with different age levels, and individuals from different cultures).

Aside from having methodological implications for an efficient exploration of the cue-emotion space, the usability and learnability properties of *EmoteControl* and similar tools create opportunities for researchers to tap into the perception of individuals from various population samples which might not have been probed as much due to potential accessibility limitations of certain methodological approaches. For example, the portable version of *EmoteControl* could be used to carry out cross-cultural studies in areas where people might not have technological devices at hand and thus, not be able to participate in online experiments. Another potential use of *EmoteControl* would be to investigate children’s perception of emotional expressions in music. Existing literature on children’s emotional development shows that children can distinguish different emotions in music at an early age, such as the basic emotions sadness and happiness (Dolgin & Adelson, 1990; Kastner & Crowder, 1990), as well as anger and fear, albeit the last two tending to be mixed up (Boone & Cunningham, 2001).

Children's emotion perception in music has been investigated by administering listening experiments consisting of systematically manipulated musical pieces, varying in structural cues (Dalla Bella et al., 2001), performance cues (Stachó et al., 2013), or both (Dolgin & Adelson, 1990). Previous studies suggest that children's emotion decoding accuracy and which cues within the music they use to decode the intended emotion may vary depending on their age (Dalla Bella et al., 2001; Oura & Nakanishi, 2000; Stachó et al., 2013). However, a limitation identified in studies was not controlling for the cues in the music used (Saarikallio et al., 2019; Stachó et al., 2013), making it unclear which musical cues had an effect on the children's perception of emotions.

In light of this, Saarikallio et al. (2019) used a production approach to investigate how 3- and 5- year old children used three musical cues (tempo, pitch, and loudness) in real-time to express happiness, sadness, and anger in monophonic melodies from well-known children's songs. The study confirmed that both groups of children used the three cues to differentiate between the emotional expressions, whereas previous studies were unclear on whether children could distinctly express anger in music (Boone & Cunningham, 2001; Thompson, Schellenberg, & Husain, 2004). Similarly, Kragness et al. (2021) asked children to express joy, sadness, anger, and peace in chord sequences taken from Bach chorales by pressing one key on a MIDI keyboard which controlled tempo, loudness, and articulation. The use of *EmoteControl* to investigate how children might use both structural and expressive cues simultaneously to communicate a selection of emotions in music would be a natural continuation to these mentioned previous studies using interactive paradigms. Following on Saarikallio et al.'s study (2019), providing a visual representation of the emotions under investigation and asking the children to draw depictions of the target emotions would help verify that the children understand which emotions are to be communicated through the music. Since musical terms currently present in the interface may be too advanced and thus not be helpful for young children, the cue labels could potentially be changed to icons or pictures which serve as a visual aid to describe the cues. For example, the minimum value of the tempo cue (i.e., at the bottom of the tempo slider) could be denoted with a person walking and the maximum value of the tempo cue (i.e., at the top of the tempo slider) could be depicted with a person running. A picture of a snail and a cheetah running could be another example of visual representations of a slow and fast tempo, respectively. Carrying out an experiment to assess the viability of potential pictures/icons would be useful to determine which visual representations would be most appropriate for the cues in question.

Apart from being used as an investigative tool, *EmoteControl* may also be employed as a medium for musical engagement and emotional communication in a game-based learning educational context (Anastasiadis, Lampropoulos, & Siakas, 2018; Nouwen et al., 2016). The portable version of the interface has already been successfully used in an emotion recognition activity with young teenagers between the ages of 14 and 16, as described in **Chapter 2**. The activity was presented as a game, where the teenagers attempted to convey a specific emotion through the music to their team members. For each correct guess, the team would be awarded a point. The team with the most points won. The teenagers were mostly successful in conveying the intended emotional expression to their team members. Thus, this game-based musical activity could be employed in an educational environment as an interactive way to learn about how emotions are communicated through music and as a means to learn music by creative learning through digital music games (Nouwen et al., 2016).

EmoteControl could also be implemented in a therapeutic setting. For example, the interface may be used as a tool to aid mood regulation, where individuals manipulate music in *EmoteControl* to their liking, allowing for personalisation of musical stimuli. This would be a beneficial way of utilising the tool since music which may be soothing to one individual may be aggravating to another (Guzzetta, 1991; Saarikallio, 2011; Saarikallio & Erkkilä, 2007). Another potential use of *EmoteControl* may be as a mode of communication for non-verbal patients (Silverman, 2008). For example, the interface could be fed information on a possible range of cue combinations which have been reported in the literature as likely used when conveying a particular emotion. When individuals modulate the musical piece via the cues in a way that matches one of the cue combinations, the interface would alert the user that the music is likely conveying the emotion in question. Thus, this may be used to assess how non-verbal patients are feeling and assist interactions between patients and music therapists (Krøier, Stige, & Ridder, 2021; Ridder et al., 2013). Similarly, *EmoteControl* may assist people with limited movement control by helping them create and express music, which may be beneficial (Magee & Burland, 2008; Partesotti, Peñalba, & Manzolli, 2018).

In particular, *EmoteControl* could serve as a musical activity for children with Autism Spectrum Disorder (ASD) to engage with music and create their own musical variations in social play, as per the PRESS-Play framework (Lense & Camarata, 2020) and the Sounds of Intent framework (Lisboa et al., 2021; Voyajolu & Ockelford, 2016; Welch et al., 2009), which posit that musical engagement activities may provide support for social development in children with learning difficulties. Along the lines of these musical activities, another potential implementation of *EmoteControl* would be as a tool for dynamic music in a game-based environment. In an interactive environment such as that of a video game, the narrative of the game changes depending on decisions taken by the player. Therefore, the music accompanying the visuals needs to allow for dynamic adaptations, in order to match the changes in the game's narrative. To do so, dynamic music systems are being implemented in video games and have been reported to enhance the player's immersive experience in the game (Gasselseder, 2014; Livingstone & Brown, 2005; Scirea et al., 2014). Moreover, *EmoteControl* could also be implemented as an audio-based game (Pires et al., 2013) where players predominantly use their auditory sense to tackle interactive challenges and progress in the game.

EmoteControl could also be applied as an automated dynamic music system, both in video games, and also other sectors which require continuous and changing music, such as the film industry, where music is regularly used to express emotions and emphasise specific elements in the film (Cohen, 2001; Parke, Chew, & Kyriakakis, 2007; Reymore, 2018), and influence how the audience perceives a movie scene (Boltz, 2001) and particular characters (Chełkowska-Zacharewicz & Paliga, 2020; Hoeckner et al., 2011). Music is also used in films as a means of certain events or characters' actions (Tan, Spackman, & Bezdek, 2007). Digital systems such as One Control by Infinity Audio[17] and OSC/PILOT[18] are compatible with music production software and help composers and performers modify their music in real-time. Similarly, the interface may be used as a composer's aid when creating music for films, by using the recording function in the Logic Pro X platform which would be already running whilst utilising *EmoteControl* and changing the desired cues to slowly make dynamic changes.

---

[17] https://www.infinity.audio/one-control
[18] https://oscpilot.com

6.3.2 A New Set of Rated Musical Excerpts

Aside from the use of these specifically composed musical excerpts as unfamiliar, new music which emulates a real musical context in this thesis, the collection of excerpts has other applications in various research domains. Firstly, the musical excerpts expand the current repository of music material specifically created to investigate emotion perception in music (Dolgin & Adelson, 1990; Gosselin et al., 2005; Hailstone et al., 2009; Vieillard et al., 2008) and represent a larger number of emotional expressions which may be communicated through music (Juslin, 2013b; Juslin & Laukka, 2004; Lindström et al., 2003; Zentner, Grandjean, & Scherer, 2008). Secondly, the excerpts may be used independently from *EmoteControl* as stimuli in future studies, irrespective of the chosen methodology. For example, the excerpts may be used as material for other interactive paradigms which allow for real-time cue manipulations such as *Music Box* (Bresin & Friberg, 2011; Saarikallio et al., 2019; Saarikallio, Vuoskoski, & Luck, 2014), other rule-based systems that change structural cues such as Robin (Morreale, Masu, & Angeli, 2013; Morreale & De Angeli, 2016) and EMS (Wallis et al., 2011), expressive cues (Friberg, 2006; Juslin, Friberg, & Bresin, 2002), or systems that investigate emotional expression in music and movement (Sievers et al., 2013), to name a few. The excerpts would be ideal candidates as stimuli for systematic manipulation experiments investigating multiple cues in detail (Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010), since they were composed to allow for rigorous manipulations of cues. In this work the musical excerpts were used in both production and systematic manipulation studies, in which they were manipulated multiple times with regards to their tempo, mode, pitch, articulation, dynamics, brightness, and instrumentation properties. Furthermore, since most manipulations are done within the musical excerpts themselves by changing the

MIDI information, rather than superimposing effects on an audio file, the excerpts retain a real music feel as they morph from their original version to other defined variations. Another potential avenue for the new set of musical pieces would be to be used as stimuli in studies looking at how different population samples perceive emotions in music, such as children (Kragness et al., 2021; Saarikallio et al., 2019; Stachó et al., 2013), musical experts as against non-musicians (Battcock & Schutz, 2022b; Kragness & Trainor, 2019; Morreale et al., 2013), non-Western populations (Athanasopoulos et al., 2021; Lahdelma, Athanasopoulos, & Eerola, 2021; Laukka et al., 2013), and clinical patients (Gosselin et al., 2005; Järvinen et al., 2010, 2012; Peretz, Gagnon, & Bouchard, 1998). Outside the context of music emotion research, the pieces which have been validated as strong candidates of their intended emotion may also be used as non-verbal representatives of said target emotion in a music therapy environment, potentially as a communication and engagement tool for non-verbal patients (Silverman, 2008) or children with learning difficulties (Ockelford, 2018).

6.3.3 Application of Empirical Findings

The empirical findings of this thesis have applications for industries that also utilise music as a communicative medium of emotions and strategically change the music through the cues to alter the perceived emotion. A sector that would find the findings of this thesis beneficial is the marketing industry, where music has been used as a means of communicating a brand's identity in advertisement (Lepa et al., 2020; Oakes, 2007) or as a tool for cross-modal influence on the perceived identity of a non-auditory object, such as the taste of wine (North, 2012). For example, music has been paired with flavours, referred to as "sonic seasoning", as a way of amplifying the multisensory experience of a particular product, such as beer or coffee (Holt-Hansen, 1968, 1976; Spence, 2021; Spence et al., 2021). On a similar note, the cue levels and combinations identified in the mapping of emotional expressions may be useful for the music information retrieval (MIR) field which uses computational models to analyse common music features used across songs to determine their classification with respect to genre (Azevedo & Bressan, 2018), for example. In particular, the emotion recognition (MER) subfield of MIR investigates which features of a music song are relevant to convey a particular emotion, with the aim of classifying the music songs by the expressed emotion (Chowdhury et al., 2019; Goméz-Cañón et al., 2021; Panda, Malheiro, & Paiva, 2018; Yang, 2021). These methods and information are used by music streaming services such as Spotify and Pandora to curate personalised playlists for their users (Panda et al., 2021).

6.4 Limitations of the Research

Certain limitations pertaining to specific parts of the project have already been discussed within the relevant chapters. In this section, an overview of limitations of this thesis will be addressed.

Although *EmoteControl* is a good first effort at providing users with an interface that lets them change tonal music in real-time, there are several improvements which may be taken into consideration for future research. Increasing the range of the pitch cue to more than the present ±2 semitones would allow researchers to investigate more perceptible register changes and better determine the cue's effect on the emotion conveyed. Another potential limitation is that not all cues in *EmoteControl* were calibrated. Tempo, mode, articulation, brightness, and instrumentation cues all overrode the incoming MIDI data from the inputted musical piece and changed the relative settings of the in the same way, irrelevant of the piece. That is, if the original musical piece had a tempo of 120bpm, was in major mode, and played with a woodwinds instrument, these would be overridden to the default starting points of the cues, which was always the same (e.g., the starting point for tempo was always 80bpm). However, this was not the case for pitch and dynamics. The pitch cue worked by transposing the pitch of the inputted piece up or down two semitones. Although the starting point was always the mid-point of its respective range, the sounding pitches depended on the pitch of the notes in the musical piece. That is, the original pitches of the piece would be transposed up or down 2 semitones, and thus, the actual pitch transpositions were relative to the inputted musical piece. The dynamics also made relative changes depending on the inputted piece, since the dynamics varied the volume level within the instrument, but this may vary across pieces depending on the velocity information of the MIDI notes in the piece, since the dynamics cue would not override the velocities. Hence, calibrating these two cues would ensure that finite cue ranges are being used across all possible inputted pieces.

It should be noted that in this current project *EmoteControl* was consistently used with music material specifically composed for the interface and although *EmoteControl* allows all MIDI files to be inputted in the interface, certain musical pieces might not be fully compatible. In particular, music material which comprises pitches outside of a particular key signature might affect the accuracy of the mode cue when the music is changed from major to minor and vice-versa. Therefore, further investigation is needed to determine how other music material outside of the excerpts introduced in this thesis, such as excerpts from popular music (Hung et al., 2021) are used with *EmoteControl.*

The fact that the musical pieces presented in this thesis had to adhere to certain requirements in order to be fully compatible with *EmoteControl* may also be regarded as a limitation. It is possible that if there were no restrictions on how the musical excerpts were composed, the pieces' abilities to communicate their intended emotion might have increased. For example, pieces conveying anger and fear tend to be composed with heavy dissonance and *staccato* articulation. However, I could not modulate out of a particular key signature due to the mode parameter. Therefore, diatonic dissonance was utilised in order to keep the notes within the specific key. Without the mode restriction, I might have been able to create a higher degree of dissonance by using note clusters outside of the key signature. I also had to keep in mind that the interface allows for articulation changes. Therefore, the note durations used could not be too short to begin with, in order to allow for the articulation changes to be effective. The use of one timbre for all parts of the musical pieces might also be considered a limitation. Perhaps the use of multiple timbres (e.g., a violin timbre for the melody and a piano timbre for the accompaniment) would have allowed for more emotional colouring than using one timbre. Therefore, the musical pieces might have been better representatives of their intended emotion if certain restrictions did not need to be implemented. On a related note, the pieces were originally composed and recorded as performances on the piano, with the aim of portraying ecologically valid music. Since the pieces were not composed in a deadpan condition, the participants using *EmoteControl* would not have been able to control any residual expressive cues which were encoded in the music during the performances (e.g., variation of MIDI velocities). However, creating the pieces in a deadpan condition, which has been considered as being rather difficult to do (Quinto, Thompson, & Taylor, 2014; Ramos & Mello, 2021; Shoda & Adachi, 2012; Timmers & Ashley, 2007) would have forfeited the naturalistic quality of the music. Therefore, manipulating the performed pieces albeit potentially losing some control over the music seemed like a logical compromise to find a balance between experimental control and ecologically valid music.

Although some of the cues available in *EmoteControl*, such as tempo and dynamics have been reported as being common across cultures due to being core psychophysical elements, making them universal (Athanasopoulos et al., 2021; Balkwill, Thompson, & Matsunaga, 2004; Balkwill & Thompson, 1999; Fritz et al., 2009), other cues may be culture-specific. For example, mode is associated with a traditional Western musical context, therefore, making it unclear how non-Western music would be modulated in the interface. Consequentially, the emotional expressions investigated with the current cues available in *EmoteControl* were mapped in a Western context. In addition, the experiments carried out in this thesis mostly dealt with participants that fit the WEIRD (Western, Educated, Industrialised, Rich, and Democratic) criteria, which represents only 12% of the world's population (Henrich, Heine, & Norenzayan, 2010) and thus, findings may not be generalised across populations. Therefore, it would be beneficial to alter and extend the selection of cues in *EmoteControl* to include ones which are not culture-specific, and carry out similar investigations in a cross-cultural context, in order to explore how cues are utilised to communicate emotional expressions in music across cultures.

The fact that the production study in **Chapter 5** was carried out in a lab setting as against the systematic approach study being run online may have also affected the results and might have explained some discrepancies between studies, such as dynamics not contributing to the emotion conveyed in the online experiment. Reproducing the two studies both in a lab setting or in an online setting would serve as a validation of results produced in **Chapter 5**. Perhaps placing attention checks in the online experiment may have helped have a more controlled online environment. Finally, participants that carried out the *EmoteControl* experiments (Experiment 2 in **Chapter 4** and Experiment 1 in **Chapter 5**) may have used the cues in a particular way depending on their perception of the cues (i.e., what they think the different cues do) rather than their perception of the emotional expressions. Future studies should gather post-task feedback which may help determine the participants' motivations during the cue manipulation task.

6.5 Recommendations for Future Directions

Based on the work detailed in this project and the limitations presented above, a number of recommendations are proposed for future investigations on music emotion research. Firstly, future research should steer towards using interactive production approaches in studies focussing on musical cues and emotional expression. The work in this thesis showed that using an interactive paradigm such as *EmoteControl* allows for a substantially large area of the cue-emotion space to be efficiently explored in a relatively short time, considering the number of possible cue combinations the participants could navigate through in real-time. It would not be possible to investigate the same number of cue combinations simultaneously utilising a traditional systematic manipulation approach since the experiment design would be too large and thus, unviable (Juslin, 2000). Therefore, this serves as motivation for future studies to focus on using interactive methodologies (Bresin & Friberg, 2011; Friberg, 2006; Kragness et al., 2021; Kragness & Trainor, 2016, 2019; Morreale, Masu, & Angeli, 2013; Saarikallio et al., 2019; Sievers et al., 2013) which allow for a direct user interactive approach to self-report studies.

Furthermore, carrying out a direct user approach with similar interfaces lets the participants, irrespective of their musical background (if they have one or not), have a first-hand experience in creating different emotional expressions in musical pieces, and consequentially revealing how they make use of the available cues to convey the intended emotions. Moreover, these interactive paradigms would allow to investigate cue usage in relation to emotional expressions across different population samples, if the paradigm is easy to use and does not require any particular expertise. For example, Saarikallio et al. (2019) and Kragness et al. (2021) have already successfully utilised interactive paradigms to investigate how children use three musical cues to communicate a small selection of emotions in music.

Another potential avenue would be to explore whether cue usage in portraying different emotions and accuracy of emotion communication vary depending on musical expertise. A review on emotional communication through music performance studies showed that professional musicians were mostly highly successful in communicating a selection of basic emotions (happiness, anger, sadness, fear, and tenderness) through music to the listeners (Juslin & Laukka, 2003). Furthermore, the same authors found that the emotion recognition accuracy may be affected by the performers' musical expertise, where amateur musicians may be less successful in communicating emotions due to inconsistent use of musical features (Juslin, 2019a; Juslin & Laukka, 2000). The findings of Study 3 in this thesis showed that musical variations attempting to convey calmness, fear, joy, and sadness created by participants (with differing levels of musical skill) in Study 2 (the first *EmoteControl* study) were successfully recognised by listeners. However, the pieces aiming to convey anger, power, and surprise were perceived as other emotions. On the other hand, the target emotion in all original pieces (composed by myself) was correctly identified in Study 3. These findings may suggest that overall, most of the basic emotions were successfully communicated through the music by both the composer and the participants differing in musical training. On the other hand, the emotions noted as being more ‘complex’ were only successfully communicated by the composer in that particular study. This may suggest that my musical training may have helped to encode power and surprise in the original pieces during the compositional process, however, this is unclear. Furthermore, it is good to note that the musical variations in Study 2 were created using mean cue levels averaged across participants, and thus, the resulting cue combinations for power and surprise may not have been successful in conveying the intended emotion due to there being lower agreement on how complex emotions may be communicated in music (Gabrielsson & Juslin, 1996), especially due to power and surprise being considered as both negative and positive valenced.

Some literature has reported that musical expertise may affect a listener’s decoding accuracy of perceived emotions in music (Akkermans et al., 2019; Castro & Lima, 2014). Other researchers have suggested that it is unclear whether formal training improves the decoding accuracy of emotions (Bigand et al., 2005; Juslin, 1997a). When musicians were asked to assess the emotional content of performances, results showed that tempo was overall, one of the more important cues in decoding the different emotions in the music, whilst the importance of other cues varied depending on the emotion being judged (Juslin, 1997b; Juslin & Lindström, 2010). Eerola, Friberg, and Bresin (2013) found that mode and tempo were the two cues most prominently contributing to emotion judgements made by participants described as mostly having ‘an extensive musical background’ (Eerola, Friberg, & Bresin, 2013, p. 5). However, in a study by Scherer and Oshinsky

(1977) where untrained participants rated the emotional content of synthesised sequences, tempo was also reported as being the most powerful predictor in emotion judgement. Battcock and Schutz (2019) showed that timing and mode accounted for most of the variance in valence, intensity, and arousal judgements made by non-musicians, whilst pitch height was the least contributing cue to participants' judgements. A following study by the authors reported that musically-trained listeners relied more on mode than non-musicians when making valence judgements (Battcock & Schutz, 2022a). Nevertheless, mode was perceived as a strong predictor of valence, intensity, arousal, irrespective of musical training.

In this thesis, mode and tempo were consistently found as being the two most important cues in shaping the different emotional expressions in music, both in the production studies (Studies 2 and 4) and the systematic manipulation study (Study 5), suggesting that these cues have importance during both the encoding and decoding processes of emotion. Kragness and Trainor (2019) investigated whether the level of formal musical training had a significant effect on how non-expert performers used tempo, articulation, and dynamics cues in an interactive paradigm to 'perform' different emotions in chords from Bach chorales. Their findings showed that musical expertise had no significant effect on the cue usage. In fact, results showed that the cue patterns used were nearly identical across participants (Kragness & Trainor, 2019, p. 8). A study by Saarikallio et al. (2014) assessed how adolescents made use of tempo, loudness, pitch, articulation, and timbre to express happiness, sadness, and anger in instrumental music. Although participants were not assessed on their musical expertise, the overall cue combinations used mostly complimented previous studies with adult participants (Bresin & Friberg, 2011; Juslin & Laukka, 2003, 2004; Juslin & Timmers, 2010).

Overall, the production studies (Studies 2 and 4) in this thesis also produced results similar to ones employing musical experts as their participants, bar for some anomalies such as anger being conveyed with a soft dynamics level in Study 2 (Bresin & Friberg, 2011; Eerola, Friberg, & Bresin, 2013; Juslin & Lindström, 2010). Further probing this line of enquiry from a production paradigm's perspective would help provide more data on whether musical expertise plays a role in the emotion communication process in music.

The use of interactive paradigms has also been employed in other areas, such as ethnomusicology. For example, Arom and colleagues (1997) employed an interactive experimental procedure to investigate the musical scales and pitches used in Central Africa and Java. They created a device linked to a digital synthesiser that could simulate different traditional instruments. The researchers then asked native musicians, instrument makers, and tuners to retune the synthesised simulations of their traditional instruments by altering pitches on the device, to determine the scales and intervals used in traditional music of Central Africa and Java. On a related note, with the creation of a new interactive paradigm, its usability, learnability, and suitability for its intended research purpose come into question. An evaluation study of the system would help assess the system's functionality as a tool for research and its suitability for its purpose (Seiça et al., 2020; Wanderley & Orio, 2002), whilst validating the results produced with the system.

Secondly, future studies should strive towards ecologically sound musical excerpts as stimuli, to try emulating a real musical context. However, studies should consider moving away from utilising audio files of pre-existing commercial music, since this limits the amount of experimental control one has on the music to not compromise its ecological validity. Work stemming from the music information retrieval (MIR) field has been attempting to deal with this limitation by curating music databases, such as EMOPIA (Hung et al., 2021) and VGMIDI (Ferreira & Whitehead, 2019) which include representations of musical excerpts taken from video game soundtracks or piano covers of popular songs in MIDI format, which would allow less restrictions on experimental control than audio files, since the cue manipulations are implemented within the properties of the music. Furthermore, these musical excerpts are accompanied by emotion annotations of the excerpts on valence and arousal dimensions, thus also providing information on their perceived emotional content. Another optimal approach would be to utilise polyphonic musical excerpts specifically composed for music and emotion perception research. Furthermore, future research may make use of existing stimuli sets in this domain which have also been evaluated on their emotional content, such as the new set of pieces presented in this thesis.

Alternatively, future work could expand the current stimuli available by creating new musical pieces for music emotion research to represent other emotional states related to music, such as tenderness and humour. Moreover, any newly composed excerpts aiming to convey specific emotions should undergo a validation study, to assess whether the intended emotion is successfully being conveyed through the music and recognised by the listeners. My own experience of composing the new musical excerpts for this current project is a prime example of why an evaluation of newly composed stimuli is beneficial. Although I have experience conveying emotions in music as a composer, it does not necessarily mean that I will always succeed in conveying the intended emotional content to the listener. The validation experiment described in **Chapter 4** is proof of this, since 12 of the composed pieces were not strong representatives of their target emotion. Therefore, if these pieces had not been pre-assessed and instead used as stimuli under the pretence that they were representatives of a distinct emotion, my results might have been skewed. Warrenburg (2020a) noted that less than 1% of the stimuli used in studies spanning from 2010 to 2018 reviewed in the PUMS database (Warrenburg, 2021) carried out pilot testing on the stimuli created prior to their use. Hence, pre-validating the excerpts would confirm whether the pieces are good representatives of their target emotion and may be used as such.

Thirdly, the work carried out in this thesis could be taken outside of a Western musical context to determine whether participants from different cultures would utilise the cues available in *EmoteControl* in a similar way to portray the investigated emotions in excerpts of their own music as well as music outside of their culture. Existing literature has proposed that on some level, emotions may be communicated in music across cultures (Balkwill & Thompson, 1999; Fritz et al., 2009). Fritz et al. (2009) carried out a study where Western participants attempted to identify the emotion expressed through unfamiliar Mafa music, while participants from the Mafa ethnic group carried out the same procedure with unfamiliar Western music. Findings reported that both groups of participants successfully identified happiness, sadness, and fear in the heard stimuli. Balkwill and Thompson (1999) also reported that Western participants were able to differentiate between ragas expressing joy, sadness, and anger, even though the participants were unfamiliar with the Hindustani tonal system.

Aside from extending the production studies detailed in this thesis by carrying out the studies with participants and musical pieces from differing cultures, the studies may also be expanded by adding more musical cues available for manipulation. Although certain cues, such as tempo, loudness, and acoustic roughness have been noted as behaving similarly across music from different cultures (Athanasopoulos et al., 2021; Balkwill & Thompson, 1999; Laukka et al., 2013), there still exists a myriad of cues which have not been explored in a cross-cultural context or need further investigation to determine whether they are culture-specific or common across musical cultures. For example, Fritz et al. (2009) reported that both Western and Mafa participants made use of tempo and mode when making their emotion judgments. More recently, Lahdelma et al. (2021) discovered that mode had a contrasting effect between Western and Northwest Pakistani participants. Major mode was associated with a more positive valence than a minor mode by Western participants, whilst the opposite was true for the non-Western participants. Athanasopoulos et al. (2021) also noted that mode held more importance in communicating emotions in Western music, whilst Khow and Kalash (Northwest Pakistani tribes) music tended to rely more on other psychophysical cues (e.g., tempo, loudness) for emotion communication in music. These findings thus give motivation for future studies to further investigate how different musical cultures make use of the various cues and determine whether the usage of distinct cues is cultural-specific or universal.

Fourthly, future research should aim to increase the number of cues investigated simultaneously, since previous evidence and the work in this thesis have shown that the contributory weight of the different cues varies depending on the rest of the cues within the combination, as well as the emotion being expressed (Hevner, 1937; Juslin, 1997b; Juslin & Lindström, 2010). Two potential cues to investigate would be harmony and rhythm. It has been reported that a regular rhythm may be representative of joy, peacefulness, majesty and seriousness (Lindström, 2006; Watson, 1942), whilst an irregular or complex rhythm may be indicative of uneasiness and anger (Gundlach, 1935; Lindström, 2006; Thompson & Robitaille, 1992). Similarly, simple harmonic structures have been attributed to happiness, serenity, and also solemnity and majesty (Hevner, 1936; Rigg, 1939; Watson, 1942), whilst complex harmony has been linked with tension and other high arousal emotions such as anger and fear (Krumhansl, 1996, 1997; Lindström, 2006). It would be interesting to explore how the addition of rhythm and the ability to change the harmonic content of the pieces (from simple and consonant to complex and dissonant) would affect how the different emotions are shaped in the music and whether there would be a significant change in the way the other cues are used.

A possible expansion to the current work would be to concurrently investigate a balanced number of structural and expressive cues as both types of cues are deemed responsible for emotional communication in music, albeit to different degrees (Gabrielsson, 2003; Juslin & Laukka, 2004; Juslin & Timmers, 2010). For example, Quinto, Thompson, and Taylor (2014) discovered that fear was best communicated through the investigated structural cues, whilst anger was best communicated with the use of expressive cues. Joy and sadness were best expressed using a combination of both expressive and structural cues. However, across all emotional expressions investigated, the researchers determined that overall, the use of both structural and expressive cues together resulted in a slightly better accuracy of emotion recognition than with the use of one type of cues.

Another possible expansion to the *EmoteControl* interface would be to create presets containing context-specific cues depending on the genre of the musical piece inputted in the interface. For example, the major/minor mode function could be omitted for non-tonal music, whilst a specific setting to manipulate the rhythm of a drum machine could be created for hip-hop music. The required preset would then be selected prior to administering any cue manipulations to the music.

Finally, future work should consider moving away from the use of one specific emotion framework since it restricts the number of emotional expressions investigated. Having a range of emotion as target expressions was useful in this work, as the findings showed that the success of emotion communication varied across emotions, which was not always in favour of the basic emotions. Solely looking at one emotion framework also prevents us from gaining information on other less explored emotional expressions and nuances of emotion states within them which may also be communicated through music. Therefore, future studies should focus on investigating how a broad range of emotional expressions are expressed through music, irrespective of which emotion framework they belong to.

6.6 Conclusions

The work in this thesis set off to explore the role of distinct musical cues in shaping different emotional expressions in music whilst tackling numerous limitations identified in the existing literature. A new interactive interface called *EmoteControl* was specifically created which allowed users to personally change musical excerpts and their emotional content via a large selection of cues. Furthermore, this interactive production approach eliminated experimental design constraints present in previous traditional approaches which limited the number of possible cue combinations explored together. This thesis also contributed specifically composed novel musical excerpts to extend the current, limited repository of stimuli used in music emotion research. These new polyphonic excerpts provide the research field with unfamiliar, ecologically valid music which can withstand cue manipulations, giving the researcher ample experimental control. The work in this thesis also saw the first utilisations of the *EmoteControl* interface and a subset of the musical excerpts in multiple empirical studies. The use of an interactive approach allowed the simultaneous investigation of six and later, seven cues in relation to seven emotional expressions, and thus, the exploration of a substantially large cue-emotion space in real-time, compared to previous research. The findings of the studies presented particular cue combinations used by participants when portraying different emotional expressions through the same musical pieces. These results also produced new information on the contributory effect of the different cues within their combination, with mode and tempo being overall the most important across emotions, whilst dynamics and brightness having small effects on the conveyed emotional expression. However, the weighting of the cues also changed depending on the emotion being conveyed. Most importantly, these studies allowed us to tap into the participants' perception of different emotional expressions in music by letting them show us how they would use the musical cues to shape the intended emotions in the music. The findings of this thesis showed that calmness, sadness, and joy emotions were the three most consistently represented and accurately recognised in music, love and longing were not successfully communicated, and emotions with similar valence-arousal properties tended to be mixed up. Finally, the work in this thesis has provided an evaluation of two methodological approaches (systematic manipulation and production approaches) to musical cues and perceived emotion research. The evaluation revealed that overall, similar results may be achieved with both approaches. However, the production approach created cue combination models which were better representatives of their intended emotion than the ones produced with a systematic manipulation approach. Moreover, the production approach allowed the participants to efficiently navigate through a huge number of possible cue combinations in real-time, which was not feasible with the traditional systematic manipulation approach. In conclusion, this thesis has presented a new interactive production approach which allowed participants to directly interact with a large cue-emotion space. This provided new knowledge on how participants thought emotions should sound like in music, allowing for a better understanding of how emotional expressions are communicated in music.

6.7 References

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., Busch, V., Lothwesen, K., Elvers, P., Fischinger, T., Schlemmer, K., & 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

Anastasiadis, T., Lampropoulos, G., & Siakas, K. (2018). Digital Game-based Learning and Serious Games in Education. *International Journal of Advances in Scientific Research and Engineering*, *4*(12), 139–144. https://doi.org/10.31695/ijasre.2018.33016

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

Arom, S., Léothaud, G., & Voisin, F. (1997). Experimental Ethnomusicology: An Interactive Approach to the Study of Musical Scales. In I. Deliège & J. A. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 3–30). Hove (UK).

Athanasopoulos, G., Eerola, T., Lahdelma, I., & Kaliakatsos-Papakostas, M. (2021). Harmonic organisation conveys both universal and culture-specific cues for emotional expression in music. *PLoS ONE*, *16*(1), 1–17. https://doi.org/10.1371/journal.pone.0244964

Azevedo, B. C. ., & Bressan, G. M. (2018). A comparison of classifiers for musical genres classification and music emotion recognition. *Advances in Mathematical Sciences and Applications*, *27*(2), 241–262.

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.

Barrett, L. F. (2004). Ratings of Experienced Emotion. *Journal of Personality*, *87*(2), 266–281.

Barrett, L. F. (2017). How emotions are made: The secret life of the brain. In *How emotions are made: The secret life of the brain.* Houghton Mifflin Harcourt.

Battcock, A., & Schutz, M. (2019). Acoustically expressing affect. *Music Perception*, *37*(1), 66–91. https://doi.org/10.1525/MP.2019.37.1.66

Battcock, A., & Schutz, M. (2021). Individualized interpretation: Exploring structural and interpretive effects on evaluations of emotional content in Bach's Well Tempered Clavier. *Journal of New Music Research*, *50*(5), 447–468. https://doi.org/10.1080/09298215.2021.1979050

Battcock, A., & Schutz, M. (2022a). Emotion and expertise: how listeners with formal music training use cues to perceive emotion. *Psychological Research*, *86*(6), 66–86. https://doi.org/10.1007/s00426-020-01467-1

Battcock, A., & Schutz, M. (2022b). Individualized interpretation: Exploring structural and interpretive effects on evaluations of emotional content in Bach's Well Tempered Clavier. *Journal of New Music Research*. https://doi.org/10.1080/09298215.2021.1979050

Bigand, E., Vieillard, 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

Boltz, M. G. (2001). Musical Soundtracks as a Schematic influence on the Cognitive processing of Filmed Events. *Music Perception*, *18*(4), 427–454.

Boone, R. T., & Cunningham, J. G. (2001). Children's expression of emotional meaning in music through expressive body movement. *Journal of Nonverbal Behavior*, *25*(1), 21–41. https://doi.org/10.1023/A:1006733123708

Bresin, R., & Friberg, A. (2000). Emotional Coloring of Computer-Controlled Music Performances. *Computer Music Journal*, *24*(4), 44–63.

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

Brunswik, E. (1956). Perception and the representative design of psychological experiments. In *Perception and the representative design of psychological experiments, 2nd ed.* University of California Press.

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception*, *32*(2), 125–142. https://doi.org/10.1525/MP.2014.32.2.125

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

Chełkowska-Zacharewicz, M., & Paliga, M. (2020). Music emotions and associations in film music listening: The example of leitmotifs from The Lord of the Rings movies. *Roczniki Psychologiczne*, *22*(2), 151–175. https://doi.org/10.18290/rpsych.2019.22.2-4

Chowdhury, S., Vall, A., Haunschmid, V., & Widmer, G. (2019). Towards explainable music emotion recognition: The route via Mid-level features. *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019*, 237–243.

Cohen, A. J. (2001). Music as a source of emotion in film. In *Music and emotion: Theory and research.* (pp. 249–272). Oxford University Press.

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.

Davies, S. (2001). Philosophical perspectives on music's expressiveness. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 23–44). Oxford University Press.

Dolgin, K. G., & Adelson, E. H. (1990). Age Changes in the Ability to Interpret Affect in Sung and Instrumentally-Presented Melodies. *Psychology of Music*, *18*(1), 87–98. https://doi.org/10.1177/0305735690181007

Dowling, W. J., & Harwood, D. L. (1986). *Music Cognition*. NY: Academic Press.

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., & Peltola, H. R. (2016). Memorable experiences with sad music-reasons, reactions and mechanisms of three types of experiences. *PLoS ONE*, *11*(6). https://doi.org/10.1371/journal.pone.0157444

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3/4), 169–200.

Ferreira, L. N., & Whitehead, J. (2019). Learning to generate music with sentiment. *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019*, 384–390.

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

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A.

D., & 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/https://doi.org/10.1177/10298649020050S105

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, *31*(3), 221–272. https://doi.org/10.1177/03057356030313002

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/oxfordhb/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., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503–534). Oxford University Press.

Gabrielsson, A., & Lindström, E. (1995). Emotional Expression in Synthesizer and Sentograph Performance. *Psychomusicology*, *14*(1), 94–116.

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

Gasselseder, H.-P. (2014). Dynamic music and immersion in the action-adventure an empirical investigation. *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound (AM'14)*, *28*, 1–8.

Goméz-Cañón, J. S., Cano, E., Eerola, T., Herrera, P., Hu, X., Yang, Y.-H., & Goméz,

E. (2021). Music Emotion Recognition. *Intelligent Signal Processing for Affective Computing*, *November*, 106–114. https://doi.org/10.1201/b10731

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., & Samson, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*(3), 628–640. https://doi.org/10.1093/brain/awh420

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

Guzzetta, C. (1991). Music therapy: nursing the music of the soul. In D. Campbell (Ed.), *Music Physician for Times to Come* (pp. 146–166). Quest Books.

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

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

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.

Hoeckner, B., Wyatt, E. W., Decety, J., & Nusbaum, H. (2011). Film Music Influences How Viewers Relate to Movie Characters. *Psychology of Aesthetics, Creativity,* *and the Arts*, *5*(2), 146–153. https://doi.org/10.1037/a0021544

Holt-Hansen, K. (1968). Taste and Pitch. *Perceptual and Motor Skills*, *27*, 59–68. https://doi.org/10.1111/j.1834-7819.1968.tb02238.x

Holt-Hansen, K. (1976). Extraordinary experiences during cross modal perception. *Perceptual and Motor Skills*, *43*, 1023–1027. https://doi.org/10.2466/pms.1976.43.3f.1023

Hung, H.-T., Ching, J., Doh, S., Kim, N., Nam, J., & Yang, Y.-H. (2021). EMOPIA: A Multi-Modal Pop Piano Dataset For Emotion Recognition and Emotion-based Music Generation. *Proceedings of the 22nd International Society for Music Information Retrieval Conference*.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001

Järvinen, A., Dering, B., Neumann, D., Ng, R., Crivelli, D., Grichanik, M., Korenberg, J. R., & Bellugi, U. (2012). Sensitivity of the autonomic nervous system to visual and auditory affect across social and non-social domains in Williams syndrome. *Frontiers in Psychology*, *3*(September). https://doi.org/10.3389/fpsyg.2012.00343

Järvinen, A., Vines, B. W., Hill, K. J., Yam, A., Grichanik, M., Mills, D., Reiss, A. L., Korenberg, J. R., & Bellugi, U. (2010). Cross-modal influences of affect across social and non-social domains in individuals with Williams syndrome. *Neuropsychologia*, *48*(2), 456–466. https://doi.org/10.1016/j.neuropsychologia.2009.10.003

Juslin, P. N. (1995). Emotional communication in music viewed through a Brunswikian lens. In G. Kleinen (Ed.), *Music and expression: Proceedings of the conference of DGM and ESCOM* (pp. 21–25). University of Bremen.

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*, *l*(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. (2019a). Music, Experience, and Affect. In *Musical Emotions Explained*. Oxford University Press. https://doi.org/10.1093/oso/9780198753421.003.0004

Juslin, P. N. (2019b). The Voice of Angels? Iconic Expression. In P. N. Juslin (Ed.), *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (pp. 156–C11.N5). Oxford University Press. https://doi.org/10.1093/oso/9780198753421.003.0011

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae*, *5*(Special Issue 2001-2002), 63–122. https://doi.org/10.1177/10298649020050s104

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae*, *4*, 151–183.

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 P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 452–489). Oxford University Press.

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191–213. https://doi.org/10.1177/102986490601000203

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.

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, *4*(June), 1–15. https://doi.org/10.3389/fpsyg.2013.00311

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. (2016). Listeners lengthen phrase boundaries in self-paced music. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1676–1686. https://doi.org/10.1037/xhp0000245

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*. Department of Pscyhology, Keele (CD-rom).

Krøier, J. K., Stige, B., & Ridder, H. M. (2021). Non-Verbal Interactions between Music Therapists and Persons with Dementia. A Qualitative Phenomenological and Arts-Based Inquiry. *Music Therapy Perspectives*, *39*(2), 162–171. https://doi.org/10.1093/mtp/miab008

Krumhansl, C. L. (1996). A Perceptual Analysis of Mozart's Piano Sonata K. 282: Segmentation, Tension, and Musical Ideas. *Music Perception*, *13*(3), 401–432. https://doi.org/10.2307/40286177

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. In *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* (Vol. 51, Issue 4, pp. 336–353). Canadian Psychological Association. https://doi.org/10.1037/1196-1961.51.4.336

Lahdelma, I., Athanasopoulos, G., & Eerola, T. (2021). Sweetness is in the ear of the beholder: chord preference across United Kingdom and Pakistani listeners. *Annals of the New York Academy of Sciences*, *1502*(1), 72–84. https://doi.org/10.1111/nyas.14655

Lahdelma, I., & Eerola, T. (2014). 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

Lahdelma, I., & Eerola, T. (2015). Theoretical Proposals on How Vertical Harmony May Convey Nostalgia and Longing in Music. *Empirical Musicology Review*, *10*(3), 245–263. https://doi.org/10.18061/emr.v10i3.4534

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*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Lense, M. D., & Camarata, S. (2020). PRESS-Play: Musical Engagement as a Motivating Platform for Social Interaction and Social Play in Young Children with ASD. *Music and Science*, *3*, 1–13. https://doi.org/10.1177/2059204320933080

Lepa, S., Herzog, M., Steffens, J., Schoenrock, A., & Egermann, H. (2020). A computational model for predicting perceived musical expression in branding scenarios. *Journal of New Music Research*, *49*(4), 387–402. https://doi.org/10.1080/09298215.2020.1778041

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.

Lisboa, T., Shaughnessy, C., Voyajolu, A., & Ockelford, A. (2021). Promoting the Musical Engagement of Autistic Children in the Early Years Through a Program of Parental Support: An Ecological Research Study. *Music and Science*, *4*, 1–24. https://doi.org/10.1177/20592043211017362

Livingstone, S. R., & Brown, A. R. (2005). Dynamic Response: Real-Time Adaptation for Music Emotion. *2nd Australasian Conference on Interactive Entertainment*, 105–111.

Magee, W. L., & Burland, K. (2008). Using electronic music technologies in clinical practice: opportunities, limitations and clinical indicators. *British Journal of*

Music Therapy, *22*(1), 3–15.

McAdams, S. (2019). The perceptual representation of timbre. In K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, & R. R. Fay (Eds.), *Timbre: Acoustics, Perception, and Cognition.* (pp. 23–57). Springer.

Morreale, F., & De Angeli, A. (2016). Collaborating with an Autonomous Agent to Generate Affective Music. *Comput. Entertain.*, *14*(3). https://doi.org/10.1145/2967508

Morreale, F., Masu, R., & Angeli, A. De. (2013). Robin: An Algorithmic Composer For Interactive Scenarios. *Sound and Music Computing Conference 2013, SMC 2013*, 207–212.

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). The Effect of Expertise in Evaluating Emotions in Music. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion (ICME3)*. University of Jyväskylä, Department of Music.

North, A. C. (2012). The effect of background music on the taste of wine. *British Journal of Psychology*, *103*(3), 293–301. https://doi.org/10.1111/j.2044-8295.2011.02072.x

Nouwen, M., Schepers, S., Mouws, K., Slegers, K., Kosten, N., & Duysburgh, P. (2016). Designing an educational music game: What if children were calling the tune? *International Journal of Child-Computer Interaction*, *9–10*(March 2015), 20–32. https://doi.org/10.1016/j.ijcci.2016.10.001

Oakes, S. (2007). Evaluating Empirical Research into Music in Advertising: A Congruity Perspective. *Journal of Advertising Research*, *47*(1), 38–50. https://doi.org/10.2501/S0021849907070055

Ockelford, A. (2018). *Tuning In Music Book: Sixty-Four Songs for Children with Complex Needs and Visual Impairment to Promote Language, Social Interaction and Wider Development*. Jessica Kingsley Publishers.

Oura, Y., & Nakanishi, R. (2000). Enso no jodohyogen no kaidoku-gijutsu no hattatsu. [How do children and college students recognize emotions of piano performances?]. *Journal of Music Perception and Cognition*, *6*(1), 13–29.

Panda, R., Malheiro, R. M., & Paiva, R. P. (2018). Novel audio features for music emotion recognition. *IEEE Transactions on Affective Computing*, *11*(4), 614–626. https://doi.org/10.1109/TAFFC.2018.2820691

Panda, R., Redinho, H., Gonçalves, C., Malheiro, R., & Paiva, R. P. (2021). How does the Spotify API compare to the music emotion recognition state-of-the-art? *Proceedings of the 18th Sound and Music Computing Conference*, 238–245.

Parke, R., Chew, E., & Kyriakakis, C. (2007). Quantitative and visual analysis of the impact of music on perceived emotion of film. *Computers in Entertainment*, *5*(3). https://doi.org/10.1145/1316511.1316516

Partesotti, E., Peñalba, A., & Manzolli, J. (2018). Digital instruments and their uses in music therapy. *Nordic Journal of Music Therapy*, *27*(5), 399–418. https://doi.org/10.1080/08098131.2018.1490919

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(3), 111–141.

Pires, D., Furtado, B., Carregã, T., Reis, L., Pereira, L. L., Craveirinha, R., & Roque, L. (2013). The blindfold soundscape game: A case for participation-centered gameplay experience design and evaluation. *Proceedings of the 8th Audio Mostly Conference (AM'13)*, 1–7. https://doi.org/10.1145/2544114.2544122

Plutchik, R. (1994). *The Psychology and Biology of Emotion*. Harper Collins College Publishers.

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

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

Ramos, D., & Mello, E. K. S. (2021). Communication of emotions in music between Brazilian composer, performers, and listeners. *Psychomusicology: Music, Mind, and Brain*, *31*(1), 1–17. https://doi.org/10.1037/pmu0000269

Reymore, L. (2018). Musical affect and embodiment: fear , threat , and danger in the music of The Lord of the Rings. In R. Parncutt & S. Sattmann (Eds.), *Proceedings of ICMPC15/ESCOM10* (Issue December, pp. 378–383). Centre for Systematic Musicology, University of Graz.

Ridder, H. M. O., Stige, B., Qvale, L. G., & Gold, C. (2013). Individual music therapy for agitation in dementia: An exploratory randomized controlled trial. *Aging and Mental Health*, *17*(6), 667–678. https://doi.org/10.1080/13607863.2013.790926

Rigg, M. G. (1939). What features of a musical phrase have emotional suggestiveness? *Bulletin. Oklahoma A & M College*, *36*, *13*, 38.

Rigg, M. G. (1942). The expression of meanings and emotions in music. In F. P. Clarke (Ed.), *Philosophical essays in Honor of Edgar Arthur Singer, Jr.* (pp. 279–294). University of Pennsylvania Press.

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. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, *39*(3), 307–327. https://doi.org/10.1177/0305735610374894

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109. https://doi.org/10.1177/0305735607068889

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. K., & 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

Saitis, C., & Siedenburg, K. (2020). Brightness perception for musical instrument sounds: Relation to timbre dissimilarity and source-cause categories. *The Journal of the Acoustical Society of America*, *148*(4), 2256–2266. https://doi.org/10.1121/10.0002275

Scherer, K. R., & Oshinsky, J. S. (1977). Cue Utilization in Emotion Attribution from Auditory Stimuli. *Motivation and Emotion*, *1*(4), 331–346.

Schubert, E. (2003). Update of the Hevner Adjective Checklist. *Perceptual and Motor Skills*, *96*(3), 1117. https://doi.org/10.2466/PMS.96.3.1117-1122

Schutz, M. (2017). Acoustic constraints and musical consequences: Exploring composers’ use of cues for musical emotion. *Frontiers in Psychology*, *8*(NOV), 1–10. https://doi.org/10.3389/fpsyg.2017.01402

Scirea, M., Nelson, M. J., Cheong, Y. G., & Bae, B. C. (2014). Evaluating musical foreshadowing of videogame narrative experiences. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2636879.2636889

Seiça, M., Roque, L., Martins, P., & Cardoso, F. A. (2020). Contrasts and similarities between two audio research communities in evaluating auditory artefacts. *ACM International Conference Proceeding Series*, 183–190. https://doi.org/10.1145/3411109.3411146

Shoda, H., & Adachi, M. (2012). The Role of a Pianist's Affective and Structural Interpretations in his Expressive Body Movement: A Single Case Study. *Music Perception: An Interdisciplinary Journal*, *29*(3), 237–254. https://doi.org/10.1525/mp.2012.29.3.237

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

Spence, C. (2021). Sonic Seasoning and Other Multisensory Influences on the Coffee Drinking Experience. *Frontiers in Computer Science*, *3*(April), 1–19. https://doi.org/10.3389/fcomp.2021.644054

Spence, C., Wang, Q. J., Reinoso-Carvalho, F., & Keller, S. (2021). Commercializing Sonic Seasoning in Multisensory Offline Experiential Events and Online Tasting Experiences. *Frontiers in Psychology*, *12*(September), 1–21. https://doi.org/10.3389/fpsyg.2021.740354

Stachó, L., Saarikallio, S., Van Zijl, A. G. W., Huotilainen, M., & Toiviainen, P. (2013). Perception of emotional content in musical performances by 3-7-year-old children. *Musicae Scientiae*, *17*(4), 495–512. https://doi.org/10.1177/1029864913497617

Tan, S. L., Spackman, M. P., & Bezdek, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception*, *25*(2), 135–152. https://doi.org/10.1525/mp.2007.25.2.135

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS ONE*, *9*(10), 1–17. https://doi.org/10.1371/journal.pone.0110490

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, *10*(1), 79–89.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Timmers, R., & Ashley, R. (2004). *Communicating emotions through ornamentation*. *January*.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, *25*(2), 117–134. https://doi.org/10.1525/mp.2007.25.2.117

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

Voyajolu, A., & Ockelford, A. (2016). Sounds of Intent in the Early Years: A proposed framework of young children's musical development. *Research Studies in Music Education*, *38*(1), 93–113. https://doi.org/10.1177/1321103X16642632

Wallis, I., Ingalls, T., Campana, E., & Goodman, J. (2011). A Rule-Based Generative Music System Controlled by Desired Valence and Arousal. *Proceedings of the 8th Sound and Music Computing Conference, SMC 2011*.

Wanderley, M., & Orio, N. (2002). Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*, *26*(3), 62–76. https://doi.org/10.1162/014892602320582981

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). People Experience Different Emotions from Melancholic and Grieving Music. *Music and Science*, *3*, 1–23. https://doi.org/10.1177/2059204320977384

Warrenburg, L. A. (2021). The PUMS Database: A Corpus of Previously-Used Musical Stimuli in 306 Studies of Music and Emotion. *Empirical Musicology Review*, *16*(1), 145–150. https://doi.org/10.18061/emr.v16i1.7208

Watson, K. (1942). The nature and measurement of musical meanings. *Psychological Monographs*, *54*(2), i–43. https://doi.org/10.1037/h0093496

Welch, G., Ockelford, A., Carter, F. C., Zimmermann, S. A., & Himonides, E. (2009). "Sounds of Intent": mapping musical behaviour and development in children and young people with complex needs. *Psychology of Music*, *37*(3), 348–370. https://doi.org/10.1177/0305735608099688

Yang, J. (2021). A Novel Music Emotion Recognition Model Using Neural Network Technology. *Frontiers in Psychology*, *12*(September), 1–9. https://doi.org/10.3389/fpsyg.2021.760060

Zentner, M. R., Grandjean, D., & Scherer, K. R. (2008). Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494