A comparative investigation of emotional expression in Baroque and modern music: perspectives from cognitive psychology

Author: Joshua Benjamin Frank

Thesis presented in fulfilment of the requirements for the degree of Master of Music in the Faculty of Arts and Social Sciences at Stellenbosch University.

Supervisor: Dr Carina Venter

Co-supervisor: Dr Barry Ross

March 2020

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Acknowledgements

I would like to thank my supervisors, Dr Carina Venter and Dr Barry Ross, for all the effort they have put into this thesis. Without their guidance and insight, this would have been an impossible task. My appreciation also goes out to my friends, who supported me through the trying times with uplifting words and fresh coffee, and my family, for the constant background motivation. Special thanks are due to Caryn Pretorius, for her unshakeable enthusiasm.

Abstract

An experiment was conducted to test listeners' emotional responses to Baroque and modern musical excerpts written to portray the same affective content. A dimensional (valence and arousal) model was used to capture and assess listener responses. The Baroque excerpts were drawn from Johann David Heinichen's treatise *Der General-Bass in der Composition.* Modern excerpts were selected from a set designed by Vieillard *et al.* (2008). It was found that responses differed between music from the two time periods. Modern music was better decoded than Baroque music. Patterns of acoustic-structural cue utilisation corresponding to listener ratings were examined, and found to agree with the existing literature. The cues discussed are tempo, mode, pitch height, interval size, and texture. Tempo was seen to have a positive effect on both valence and arousal. The major mode was associated with positive valence, the minor mode with negative valence. Interval size was positively correlated with valence. No other effects of cue levels on valence and arousal were clear. Baroque excerpts were found to have higher values for tempo and pitch height than modern excerpts for all emotion portrayals. A comprehensive literature review detailing significant research and theoretical frameworks in the fields of emotion research in general and music and emotion in particular was conducted prior to experimentation.

Contents

Abstract i
Contents ii
List of figures iv
List of tables v
CHAPTER 1. INTRODUCTION 1
1.1 Research aims 2
1.2 Research questions 3
1.3 Overview of thesis 3
CHAPTER 2. RESEARCHING EMOTION: THEORY AND METHODOLOGY 5
2.1 Discrete emotion theory 7
2.2 Dimensional theory 12
2.3 Component Process theory 18
2.4 Prototype theory 21
CHAPTER 3. MUSIC AND EMOTION 23
3.1 The cognitivist position 23
3.2 The emotivist position 28
CHAPTER 4. ACOUSTIC CUES AND EMOTIONAL RESPONSES TO MUSIC 33
4.1 Discrete emotion research 36
4.2 Dimensional research 42
CHAPTER 5. BAROQUE CONSIDERATIONS 48
5.1 Descartes and the passions 48
5.2 The Doctrine of the Affections 53
5.3 Selection of Baroque excerpts – Heinichen's *Der General-Bass in der Composition* 56
5.4 Selection of modern excerpts – Vieillard *et al.* (2008) 59
CHAPTER 6. EXPERIMENTAL DESIGN 62
6.1 Research questions 62
6.2 Selection and implementation of the dimensional theory 62
6.3 Positions of discrete emotion labels on the VA space 63
6.4 Selection and analysis of experimental stimuli 67
CHAPTER 7. EXPERIMENTAL PROCESS, STATISTICAL PROCEDURE, AND RESULTS 71
7.1 Experimental Process 71
7.2 Results 75
CHAPTER 8. COMPOSITIONAL CUES AND DISCUSSION 85

8.1 Cue levels 85
8.1.1 Valence 87
8.1.2 Arousal 88
8.1.3 Quadrants 88
8.1.4 Individual cues 91
8.2 Comparison with Vieillard *et al.* (2008) 94
8.3 Comparison with Descartes' predictions 95
8.4 Possible explanations for results 95
CHAPTER 9. CONCLUSION 99
9.1 Assessment of principal research questions 99
9.2 Reflection and suggestions for further research 99
9.3 Concluding remarks 102
BIBLIOGRAPHY 103
Appendix A – Musical excerpts used in the present experiment, cue levels, and Heinichen's descriptions 113
Appendix B – Full results of statistical computations 119
Appendix C – Data gathering materials 129
Questionnaire 129
Rating sheet 131

List of figures

Figure 4.1.1. Hypothesised cross-modal encoding patterns for acoustic cues..... 34

Figure 6.3.1. Quadrant division of the VA space..... 64

Figure 7.2.1. Scatterplot of VA ratings given for Baroque and modern excerpts..... 76

Figure 7.2.2. Scatterplot of VA ratings for Happy, Peaceful, and Angry emotion portrayals 77

Figure 7.2.3. Scatterplot of VA ratings for Musicians and Non-musicians 78

Figure 8.1.1. Scatterplot of VA ratings per excerpt 85

List of tables

Table 4.2.1. Summary of findings of discrete emotion studies 45

Table 4.2.2. Summary of findings of dimensional studies 46

Table 6.3.1. Positions of discrete emotion labels per quadrant of the VA space 65

Table 6.3.2. Groupings of related emotion words 66

Table 6.3.3. Quadrant positions for discrete emotion labels 67

Table 7.1.1. Results of MANOVA with valence and arousal as the dependent variables 75

Table 7.2.1. Results of factorial ANOVA with valence as the dependent variable 79

Table 7.2.2. Results of factorial ANOVA with arousal as the dependent variable 81

Table 8.1.1. Levels of compositional cues for each excerpt 86

Table 8.1.2. Summary of ratings by quadrant for the excerpts 88

Table 8.1.3. Summary of findings of dimensional studies (reprint) 91

Table 8.2.1. Comparison of VA ratings from the present experiment with Vieillard *et al.* 93

CHAPTER 1. INTRODUCTION

Emotion is a central facet of musical experience. Philosophical speculation on the emotional content of music can be traced back to the earliest of times, and continues unabated to this day. Scientific research in this area has been extensive, and has a long and storied history. The present thesis aims to expand this research in the historically comparative domain, by comparing the emotional portrayals of Baroque and modern musical excerpts, each written in accordance with the predominant theories of musical emotion of their respective times.

Before the beginning of the Baroque period around 1600, musical composition was primarily concerned with questions of internal structural relations. Vocal music was the style of the day. Thick polyphonic textures, layering voices against each other with overlapping words, frequently distorted the text to the point of incomprehensibility (Rochberg, 1997:171-172). The pure aesthetic beauty of the counterpoint was prized above all else (Palisca, 2001). Emotional expression was far from the forefront of the minds of most composers in the 16th century. This was not to last.

The radical break in musical style around 1600 is often referred to as the *seconda pratica*. The term was coined at the time of this new development, representing a conscious break with inherited tradition – it may however do more work in modern parlance than was originally intended (Carter & Chew, 2001). The new style, championed most famously by Claudio Monteverdi, held as a key ideal the clear and intelligible setting of text to music (Rochberg, 1997:171). This consideration gave composers licence to break with the old rules of counterpoint to more closely adhere to the expressive intention of the text, laying the foundations for the rapid rise of expression as the primary purpose of music.

A number of tangible musical changes occurred in rapid succession at the beginning of the 17th century. Symmetry was eroded, dissonance saw freer use, and harmony began to develop in the direction of modern tonality. Melody took pride of place in musical composition, written always in accordance with the expressive intentions of its text (Rochberg, 1997:170-171). For the present purpose, these developments are most interesting insofar as they led to the development of Baroque theories of emotional expression.

A group of late-16th century music theorists known as the Florentine Camerata was responsible for an early and influential theory in this vein. They were of the opinion that music may achieve emotional expression through imitation of the human voice. This was codified in terms of a number of technical correspondences to the natural movement of the human voice when roused

to a state of high emotion (Kivy, 1980:18-20). This theory was immediately applicable to composers of the time looking to imbue their music with specific expressive properties.

The theory of musical emotional expression that served as a foundation for Baroque composers was put forward after the time of the Camerata. This was René Descartes' theory of the passions, most notably expounded in his 1649 treatise *The Passions of the Soul*. His highly technical theory was adopted and adapted by many music theorists, most famously Johann Mattheson. Mattheson placed great emphasis on the centrality of emotion in music, seeing it in essence as music's *raison d'être* (Buelow, 2001a). These attempts by Baroque writers to lay out technical bases for emotional expression came around the turn of the 20th century to be referred to as the "Doctrine of the Affections" (Buelow, 2001b). This diverse yet united body of literature forms the historical point of comparison for the present research, as it represents one of very few (at least partly) formalised theories of musical expression present in the history of Western Art Music.

Modern developments in the theory of musical emotional expression have been swift and wide-reaching since the early 20th century. Core debates still rage strong: can music truly cause listeners to feel emotion? By what mechanism do composers imbue their works with emotional content, and how do listeners recognise this content? Are the techniques of musical emotional expression invariant and universal phenomena, or are they specific to time and place? Although many such issues remain unresolved, remarkable progress has been made, particularly in the investigation of acoustic-structural cues and their relation to emotion portrayal in music.

The present research is situated at the intersection of musicology and cognitive psychology. It utilises an empirical methodology, for two principal reasons. First, there is strong precedent in the literature for the use of such a framework, and the format derived from it will allow the present research to enter directly into conversation with this literature. The second justification for the use of an empirical methodology is based on the two primary research questions of this thesis, which are best approached by means of the experimental generation of data.

1.1 Research aims

The present research aims to determine whether contemporary listeners recognise the intended emotional content in Baroque music. This is assessed quantitatively and comparatively, by gathering data on listeners' recognition of emotion in samples of Baroque and modern music. Three emotional portrayals are examined, each embodied in one excerpt from each time period. By comparing listeners' responses to each emotional portrayal across the two time periods, the

similarity of their recognition accuracy can be determined. The patterns of acoustic-structural cues present in the present musical excerpts are examined. There is a large body of literature pertaining to the effects of these cues on emotional portrayal and recognition. The second aim of this research is to examine whether the results found in the literature are applicable to the musical excerpts used in this experiment. The results generated by the present experiment represent an exploration of the variability of musical emotional expression over time.

The choice of Baroque music as a historical point of comparison is motivated by the present author's interest in Baroque music performance. This research aims to explore certain questions relating to emotional expression in Baroque music, in order to create scope for further research. This line of research is hoped to yield information useful to performers of Baroque music, which can be applied at the practical level.

1.2 Research questions

1. Are there differences in contemporary listeners' perceptions of emotion in Baroque and modern (academically conceived, 21st-century Western tonal) music, where that music was designed to portray the same emotion?
2. How do compositional-structural parameters contribute to the portrayal of emotion in this music?

These questions are particularly amenable to empirical investigation. In their generalised form, they are here taken to amount to two empirical questions: (1) holding the intended emotion constant, do judgements of emotion by listeners (both musicians and non-musicians) differ for Baroque, as opposed to modern, musical excerpts? (2) Can patterns of acoustic-structural cues in the excerpts can be used to explain similarities and differences in emotion recognition by listeners?

The first research question will be investigated by an inferential-statistical assessment of data generated by an experiment. The second will be explored by means of appeal to compositional cues present in the musical materials used in that experiment, and will be discussed with reference to a statistical description of compositional cues.

1.3 Overview of thesis

This thesis is broadly divided into two sections. The first section (chapters 2 to 5), which comprises a comprehensive literature review, is structured as follows. First (chapter 1), a large body of literature pertaining to the general empirical study of emotions will be reviewed. The

focus of this chapter is on exploring the two predominant general theories of emotion: discrete emotion theory and dimensional theory. Second (chapter 2), theories of musical emotion will be explored and assessed on their relative levels of empirical support. The question of whether music can induce emotion or only be recognised as portraying emotion is discussed. Third (chapter 3), prior research on acoustic-structural cues related to the portrayal of emotion in music will be examined. Finally (chapter 4), the literature review will be taken full circle, and Descartes' theory of the passions will be discussed, and traced through some of its musical developments. This chapter also contextualises the selection of excerpts representative of the Baroque period. Throughout the literature review, the foundations of an experimental model will be laid.

The second section of the thesis (chapters 6 to 8) describes and reports on experimentation. First (chapter 6), the design of the experiment will be explained in detail. Second (chapter 7), the experimental procedure and results will be presented. Finally (chapter 8), these results will be discussed in terms of the two principal research questions stated above, and related to the existing literature. A brief concluding section (chapter 9) will summarise the results, reflect on the experiment, and outline important considerations for future research.

CHAPTER 2. RESEARCHING EMOTION: THEORY AND METHODOLOGY

Emotion is central to human life, and has been a subject of intellectual consideration since the earliest of times. Despite the ubiquity of emotional experience across cultures and ages, there is a serious and pervasive problem of definition associated with formal and informal investigations in this area (Kleinginna & Kleinginna, 1981; Scherer, 2000; Juslin & Sloboda, 2001, Fehr & Russell, 1984). For the purposes of this chapter, the early theoretical formulations of emotion will be disregarded (although some sources from the 16th to 18th centuries will be examined in chapter 5). The theories under review here will pick up from the time of Darwin, and continue through to contemporary developments in psychology and neurology. These are the theories underpinning, in one way or another, the modern field of emotion research.

A seminal attempt to define emotion was made in 1981 by Paul and Anne Kleinginna (Sloboda & Juslin, 2001:75). They based their definition on a review of 92 prior definitions from various sources, which they compared to find points of agreement (Kleinginna & Kleinginna, 1981:345). The definition they reached reads as follows:

> Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labelling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal-directed, and adaptive. (Kleinginna & Kleinginna, 1981:355)

The drive towards a consensual definition of emotion gained impetus over the following decades. Scherer (2000:138) notes that along with the wide adoption of a generally "multicomponential" view of emotion, there is also broad agreement as to the nature of at least three of these components, often referred to as the "reaction triad" of emotion. These are: physiological arousal, motor expression, and subjective feeling. Most participant-based empirical studies on emotion, as well as those specifically investigating music and emotion, aim to examine one or more of these components.

Alongside the "reaction triad", a number of other emotion components and criteria have been proposed, with varying degrees of support. A few of the most influential include: cognitive appraisals (Sloboda & Juslin, 2001:75, Scherer, 2000:138-139), action tendencies (Scherer,

2000:138-139; Frijda *et al.,* 1989:213), and at least some degree of dynamic synchronicity between component processes (Scherer, 2000:138).

Having laid out some necessary criteria of a definition of emotions, it is possible to differentiate emotions from other affective phenomena. This will avoid confusion among the different types of affective phenomena that have relevance to music, an issue that has been noted as pervasive in the music and emotion literature (Scherer & Zentner, 2001:362; Konečni, 2008:115-116). Scherer and Zentner (2001:363) identify six types of affective phenomena. These are: preferences, emotions, mood, interpersonal stances, attitudes, and personality traits. All of these phenomena may be incorporated into studies of music and emotion. Of these, mood is the most susceptible to confusion with emotion.

In contrast to emotions as defined above, moods may be defined as: "[a] diffuse affective state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause (*cheerful, gloomy, irritable, listless, depressed, buoyant*)" (Scherer & Zentner, 2001:363). The difference is clear – whereas emotion represents changes in a number of components, mood is predominantly centred on a single component, namely subjective feeling. Furthermore, the intensity and duration of moods contrast with those of emotions. Scherer and Zentner (2001:363) define these characteristics for emotions as being of high intensity and relatively brief. They also suggest that moods have a low level of component synchronisation as compared to emotions (Scherer & Zentner 2001:363). It is important to keep in mind the distinction between mood and emotion when surveying the literature, and when considering the conceptual object of certain models (e.g. the dimensional model).

The modern approach to the study of emotions has yielded two primary schools of thought, or general theories of emotion. These are *discrete emotion theory* and *dimensional theory* (Sloboda & Juslin, 2001:76-79; Vieillard *et al*, 2008:722-723; Ekman, 2016). Arguable for inclusion as a theory in its own right is Scherer's component process theory (e.g. Scherer, 2000). However, given that, as Scherer states, the majority of modern theorists define emotions in terms of multiple components (Scherer, 2000:138), it seems more appropriate to subsume this approach into the general framework of the other theories. Despite this qualification, Scherer's definition will be given separate treatment below, along with the two general theories. The section will conclude with a discussion of a final useful framework, which is compatible with both discrete emotion theory and dimensional theory, namely prototype theory (Rosch, 1978; Fehr & Russell, 1984).

2.1 Discrete emotion theory

It will be useful first to clear up some issues of definition. Discrete emotion theory works on the assumption of a set of "basic" emotions, which serve vital functions in everyday life. The literature takes up any of three main senses of the word "basic" as applied to emotions: conceptual, biological, and psychological (Scarantino & Griffiths, 2011). The terms discrete, basic, and primary will be used interchangeably here. One of these senses of basicness, the biological sense, is by far the most common and will form the bulk of this section. In order to sidestep potential confusion, as well as delineate some useful frameworks through which to view the idea of discrete emotions, all three senses of the term will be discussed here.

Conceptual basicness is defined by the position of a concept in the basic level of a conceptual taxonomy. The basic level is defined as containing categories which "are the most abstract categories of which a representative image can be formed, the categories named more quickly by adults, the categories used most often by parents to teach a language to their children, the categories designated by shorter names, and the categories learned first by children" (Scarantino & Griffiths, 2011:446). Investigation into the conceptual basicness of emotion categories has been carried out by Fehr and Russel (1984) who identified *anger, fear, love, happiness, sadness,* and a number of other "middle-level" categories as conceptually basic. The middle level is defined as falling below the superordinate category "emotion", and above the subordinate types of each emotion, such as wrath, annoyance, rage, and fury or anger (Fehr & Russell, 1984:467). Shaver et al (1987) conducted similar research, cautiously limiting Fehr and Russell's list to just five categories, *anger, sadness, love, fear,* and *joy* (Shaver *et al.*, 1987:1067-1068).

The idea of conceptual basicness may not have application in the affective sciences (at least, the classical tradition of such), due to the difference between *folk* and *scientific* definitions and categorisations of emotional phenomena. Folk and scientific definitions draw their lines in terms of different criteria – where the folk definition is fairly loose and based on "family resemblances" (attributed to Wittgenstein), the scientific definition has historically been tied to the notion of natural kinds, and draws firm lines in terms of the potential of these kinds for inductive and explanatory generalisations. For example, the folk definition of an emotion such as "anger" may include states that the scientific definition would exclude (Scarantino & Griffiths, 2011:447). Folk definitions are at stake in the discussion of conceptual basicness, and the concept has no direct relation to the scientific definitions of emotions as phenomena.

However, the question of conceptual basicness may be useful in examining lay conceptions of affect, a question which will become relevant in the following section.

The most commonly used idea of basicness in discrete emotion theory is that of biological basicness, and this is the sense in which discrete emotion theory is invoked in the rest of this thesis, unless specifically stated otherwise. Emotions are considered to be biologically basic in as far as they have developed due to their adaptive value for the survival of the organism, and fulfil a number of specific related criteria. The idea of biological basicness of emotions incorporates cultural influences as well as biological ones; as much as emotions are seen as relying on innate (biological) affect programmes, cultural factors are seen as influencing the expression and realisation of these on a per-organism basis (Scarantino & Griffiths, 2011:447-448).

Biologically basic emotions are assumed to have adaptive value in a Darwinian-evolutionary sense (Scherer, 2004:246). In this way, discrete emotion theory is intimately tied to the psychological paradigm of functionalism (not to be confused with functionalism in the philosophy of mind; e.g. Turing, 1950). Functionalism in psychology may be defined as the use of a framework of adaptive value, based on evolutionary principles, to explain human functioning (Juslin, 1997a:385). Interpreted through the functionalist paradigm, the phenomenon of emotion has adaptive significance. Theorists working in this view proposed that the full range of complex emotional states may be derived from a much smaller set of primary emotions, which have adaptive value as time-efficient ways of dealing with life's exigencies (Juslin, 1997a:385-386). Juslin sets out seven main criteria by which a primary emotion may be distinguished from a complex emotional state:

> Basic emotions (a) have distinct functions that contribute to individual survival, (b) are found in all cultures, (c) are experienced as unique feeling states, (d) appear early in the course of human development [ontogeny], (e) are associated with distinct autonomic patterns of physiological changes, (f) can be inferred in other primates, and (g) have distinct emotional expressions… (Juslin, 1997a:386)

The adaptive value of emotions, according to a functionalist perspective, can be divided into a number of interrelated effects, which ought to co-occur in unique patterns for each discrete emotion. These tend to have a high degree of concord with the reaction triad of emotions. In a 2011 meta-study of emotion elicitation research utilising the discrete emotions model, Lench

et al. (2011:834-835) identified five levels of effects, namely cognition, judgement, experience, behaviour, and physiology. From this set, experience, behaviour, and physiology correspond to the reaction triad (subjective feeling, motor expression, and physiological arousal). A secondary finding of the same study was that these three effects had a far higher level of covariance than the cognition and judgement effects, lending support to the idea of the reaction triad as comprising the "components essential to emotion" (Lench *et al.*, 2011:849-850).

While there are disagreements between theorists of discrete emotion as to which emotions qualify as biologically primary, with anywhere from seven to fourteen emotions being suggested to fall into this category (Scherer, 2004:246), there is also some broad level of agreement. Within the functionalist approach to emotion, five primary emotions appear to be held in general consensus. These are: happiness, sadness, anger, fear, and disgust (Sloboda & Juslin, 2001:77). Lench *et al.* indicate conditions which may give rise to the first four of these, stating that: "*happiness is generally considered to arise from a success…sadness from failure or loss with no hope of reinstating the goal…anger from failure with the possibility of goal reinstatement (i.e., the goal is still possible with additional effort…), and anxiety [the term they use for what is sometimes called fear] from the anticipation of threats to important goals…"* (Lench *et al.*, 2011:835). It is worth noting that these four correspond to four of the five emotions identified by Shaver *et al* (1987: 1067-1068) as occupying conceptually basic status. The emotion of *love* from Shaver *et al.*'s list is replaced here by *disgust*.

Scarantino and Griffiths give a definition of psychological basicness: "an emotion is psychologically basic if it does not have other emotions as constituents and it is not a species of another emotion" (Scarantino & Griffiths, 2011:451). The idea of "fundamentality", or not being constituted of parts which are themselves emotions, is a core component in the formulation of the discrete emotions model. While discrete emotion theory focuses on a small set of primary emotions, non-primary or complex emotions are also accounted for. A common paradigm used to explain the existence of emotions outside of the small primary set is "emotion blending", analogous to the manner in which the blending of primary colours forms the other colours on the spectrum. Another idea is cognitive elaboration, whereby basic emotions are altered through combination with other mental states (not emotions; Scarantino & Griffiths, 2011:451). This fits with the idea of emotional species: the subordinate species of a basic emotion may be determined by the context of the emotion's elicitation (as assessed using cognitive processes), as well as the level of intensity of the emotion. For example, *excitement* could be considered to be a species of the basic emotion *joy* (Shaver *et al*, 1987:1067-1069).

Support for the theory of basic emotions has come from a number of fields, with cross-cultural research yielding by far the strongest evidence. The early stages of this research focused on basic emotions as unitary constructs – however, since the advent of component process theory in the 1980s, basic emotions have been examined in terms of each of their components (Mesquita, Frijda, & Scherer, 1997:260). Note that these components are not themselves emotions, meaning that basic emotions viewed in these terms still satisfy the criterion for psychological basicness. The component-process approach has gained traction in cross-cultural studies, as it does not assume that there exist a set number of predefined basic emotions *per se* (Mesquita, Frijda, & Scherer, 1997:261), leaving room for some different emotions to manifest as basic in different cultures. It focuses on the presumed universality of emotion *components*, which tend to co-occur in similar patterns to form what are defined as the basic emotions. This allowance for some level of difference in basic emotion categories has proven profitable in cross-cultural research.

Components that have been examined cross-culturally include: emotion antecedents (the preceding events linked to an episode of emotion), appraisal, physiological reactions, action tendencies, emotional expression, and subjective experience. Emotional expression and recognition by facial expression form the greatest and most convincing portion of the data in this area (Mesquita, Frijda, & Scherer, 1997; Matsumoto *et al*, 2002). Across all components, similar patterns have been found which correspond to the generally agreed-upon basic emotions, to differing degrees of certainty per component. Mesquita, Frijda, and Scherer (1997) make two important observations to situate this evidence. Firstly, the existing research has focused almost exclusively on emotional *potential*, rather than *practice*. While the data reliably indicate that the potential for emotional reactions is similar across cultures, the ecology of emotional life in different cultures has not been explored thoroughly, and is likely to show a higher level of differentiation. This level of differentiation seems to decrease as the level of description becomes more abstract, and increase as the level of description becomes more concrete. For instance, all cultures may share a reaction pattern to a type of eliciting condition; for example, danger eliciting fear. Per-culture differences exist to a great degree in the specific nature of the danger perceived (e.g. a sandstorm vs a tidal wave) and the kinds of events appraised as dangerous (e.g. heavy rain may be appraised as dangerous or life-saving depending on the climate and soil conditions of a region). This level of cultural specificity can quite happily coexist with the idea of universality in emotional *potential*. Both areas are valid

and valuable fields for further research (Mesquita, Frijda, & Scherer, 1997; Matsumoto *et al*, 2002).

When considering emotional responses to music in terms of the discrete emotions model, a serious problem arises. The functionalist conception of emotion contains a requisite condition for defining a state as an emotion – that state must have adaptive value for the individual (Juslin, 1997a:386). This means that, in order to be classed as a discrete emotion, a contending state must contribute to the survival of the individual, in relation to existential situations in everyday life. This condition goes hand-in-hand with the physiological and behavioural manifestations of emotions. Music, as a candidate stimulus for inducing emotions, falls short here – there are few if any music-listening or -performing situations which have immediate existential consequences. A number of approaches have been taken in addressing this problem.

One potential solution to the problem of adaptive value in emotional responses to music may be to delineate *aesthetic emotions* as distinct from *basic emotions*. Taken to the extreme, this idea can lead to the total reclassification of affective phenomena in response to aesthetic stimuli; for instance, Konečni (2008:117) argues that musical "emotions" should not be called *emotions* at all, with that term being reserved for the basic emotions, due to their "biological significance". The fact remains, though, that phenomena defined as "emotional" are reported in response to musical stimuli. This has led Konečni to develop his Aesthetic Trinity Theory (ATT), which puts forward the idea of *aesthetic awe, being moved,* and *thrills/chills* as the components of aesthetic "emotional" experience, explaining their labelling in terms of common emotional categories as either linguistic/cultural bias (against reporting a null result in an experimental setting), the confusion of induction and representation of emotion, or a product of mediating factors (e.g. memories and other mental imagery; Konečni, 2008:122-123).

In terms of the present research, discrete emotion theory is not an ideal framework. As mentioned above, it utilises a set of primary/basic emotions, with other emotions being explained primarily through the blending of these. There are two potential problems with applying a discrete methodology to musical emotions. First, such a model leads to difficulty in identifying emotions when the stimulus is not *strongly* or *unambiguously* indicative of that emotion, as music is likely to be in many cases (e.g. Eerola & Vuoskoski, 2011:40). Second, the model is far stronger when used with broad, basic emotion labels (e.g. happy, sad, angry) than more specific labels (which may fall under one of the broader categories; e.g. Juslin, 1997c). The emotional labels attributed freely to music are varied and complex (Juslin,

1997c:88-94). One possibility for bypassing this second problem would be to limit the experimental targets and response categories to a set of basic emotions. This would lead, however, to a highly restricted set of target emotions. Of the generally accepted basic emotion categories mentioned above, happy, sad, angry, fearful, and tender/loving music have shown to be more or less discriminable using discrete methodologies (Quinto *et al*, 2014:513; Juslin & Laukka, 2003). Disgust has seldom been examined in music (cf Juslin & Laukka 2003:778-785 – compare Tables 2 and 3, lists of emotions studied in vocal expression and musical expression studies respectively).

Discrete emotion theory is amenable to a data collection procedure utilising rating scales (e.g. Mehu & Scherer, 2015:805; Mikels *et al*, 2005), which is the preferred tool for the present research. This would be easy to implement; most discrete studies utilise no more than seven scales (Juslin, 1997c:79). Participants should be able to understand these scales without difficulty - discrete emotion labels are usually not explained to participants, likely because participants are assumed to be intimately familiar with them. These scales seem to be interpreted similarly by participants, judging from the overall consistency in the data. A small set of discrete emotion labels is, however, very limiting, as the total number of emotions that may be represented in the data is limited to the number of labels provided.

The association of a set of acoustic cues with a discrete emotion term is easy to interpret. The set of cues is taken to constitute a greater or lesser part of the acoustic profile of the emotion in question. In discrete methodologies, a valid form of conclusion is that an ambiguous set of experimental results indicates a hybrid/blended emotion consisting of two or more basic emotions (e.g. Mikels *et al*, 2005). This approach is not appropriate for music, as it is desirable to capture finer-grain differentiation in emotion judgements than levels of association with discrete labels can provide.

2.2 Dimensional theory

The most significant early formulation of the dimensional model was provided by Wilhelm Wundt, shortly after its original proposal by Herbert Spencer (Sloboda & Juslin, 2001:77; see e.g. Spencer, 1895). Wundt suggested that three bipolar dimensions underlie the subjective experience of emotion; he labelled these pleasantness-unpleasantness, rest-activation, and tension-relaxation (Scherer, 2004:247). Scherer explains that these dimensions were posited by Wundt specifically as underlying the *qualia* (singular *quale*; the subjective experience) of emotion, as opposed to the emotion as a whole, of which any quale can only be a component

(Scherer, 2004:247-248). Three-dimensional models more or less similar to Wundt's remain popular (e.g. Scherer, 1972). Somewhat more commonly used, however, are two-dimensional models (e.g. Vieillard *et al*, 2008). These are usually based on the work of James Russell, who demonstrated the utility of his bipolar, 2-dimensional model in a landmark study in 1980. This model comprises the dimensions of *valence* (positive-negative) and *arousal* (high-low activation; Russell, 1980).

Russell (1980:1162) draws evidence for his model from studies dating as far back as 1952, chosen because they dealt specifically with people's internal conceptions of affective states. The studies asked participants to make affect judgements based on either observed behaviour in others, or verbal affect terms. Both of these methodologies allow for testing of the participants' cognitive framework for conceptualising emotion – they must rely on this knowledge rather than on introspection, unlike if they were asked to self-report on their own affective states (Russell, 1980:1162). Across the studies cited, Russell finds evidence, particularly in the data generated by multidimensional scaling, of the suitability of a two-dimensional model for representing lay people's cognitive framework of emotion, and in particular a bipolar two-dimensional model, with affect terms forming a circular pattern around the centre, in the two-dimensional space generated by this model (Russell, 1980:1163).

While Russell's work was influential in the widespread adoption of the two-dimensional model, specifically with his conceptualisation of the two component dimensions (discussed below), he was by no means the first to suggest such a model, or even to provide it with empirical backing. The earliest study cited by Russell, conducted by Harold Schlosberg (1952), may have been the first research in which a 2-dimensional model was demonstrated experimentally. Numerous early findings (e.g. Schlosberg, 1952; Abelson & Sermat, 1962; Schaefer & Plutchik; 1966), taken alongside the results of Russell's own experiments (utilising labelled and unlabelled category sorting tasks, self-report, and principal-components analysis), led Russell to conclude that the two dimensions of pleasure-displeasure and degree-of-arousal are sufficient to map both the domains of self-reported emotion and external emotion judgements (Russell, 1980:1176-1177; see throughout for Russell's empirical justification). These two dimensions will from here on be referred to as Valence (V) and Arousal (A) respectively, with their combination being labelled VA.

Recent and influential support for the VA model is provided by Lench *et al.* in their 2011 meta-study of emotion elicitation experiments. They used their collected data on the efficacy of

emotion elicitations in terms of five main component effects, in combination with the theoretical predictions of different general emotion theories, to compare the levels of evidence provided for these theories. While their primary goal was to examine evidence of the discrete emotions model, they discuss the VA model at some length. Two predictions associated with the VA model were examined: the prediction that anger and anxiety, both being of negative valence and high arousal, would exhibit relatively similar component effects; and the prediction that some amount of "difference" in discrete emotion category labels may be explained by people's beliefs and/or expectations, rather than neuro-biological factors (Lench *et al.*, 2011:850).

It was found that the effect size of the comparison of anger and anxiety was smaller than for other comparisons (Lench *et al.*, 2011:850). This result presents evidence in relation to an influential critique of the VA model, namely that it places emotions that are qualitatively very different close to one another in the affective space. A finding of high differentiation among the other non-experiential components of the emotions in question would have supported this critique; the findings of Lench *et al.*, however, weaken this critique (as do their findings relating to the second prediction of the model).

For the prediction that some difference in emotion labels may be explained by beliefs or expectations, it was found that only the component of self-report showed effects across all emotion comparisons. Comparisons between emotions with similar positions on the VA space (anxiety and anger, sadness and anger, sadness and anxiety) showed only small effects (Lench *et al.*, 2011:850). The finding that self-report (experience) was the only measure in which emotions falling relatively close on the VA space (more specifically, the valence axis) reliably differed, while other (physiological, behavioural) measures showed no significant difference, lends credibility to the idea that subjective beliefs about emotions played a part in their categorical differentiation in self-reported experience (Lench *et al.*, 2011:850). Thus, the VA model is supported as a valid representation of emotional states across component processes, adding further to its explanatory value (at first cautiously limited by Russell to self-reported feeling and categorical judgements) (Russell, 1980).

Up to this point, Russell's VA model has been identified with the idea of underlying dimensions of emotion, as it is by far the most commonly used formulation of the dimensional approach in the emotion literature. There are a number of alternative dimensional approaches, however, which have merit and experimental validity. One such example, which has been used

in a number of studies, is the Positive and Negative Affect Schedule (PANAS) of Watson, Clark, and Tellegen (1988). This approach was developed on the back of Russell's VA model, and has been shown to be valid in experimental settings.

The dimensional model is very commonly used in research on music and emotion (Vempala & Russo, 2018:2). The VA dimensions as proposed by Russell, however, are never presumed to represent the entirety of musical emotional experience. An example of a formulation using VA dimensionality as a base is Eerola's integrated framework (Reybrouck & Eerola, 2017:8). This model takes VA as the basis of an emotional experience, or *core affect*. This core affect is subjected to conscious interpretation, presumably based on a number of innate, cultural, and environmental factors, and recognised in terms of the framework of basic emotion labels. Further conceptual work involving higher-level functions may take place, leading to recognition in terms of aesthetic emotion labels. The results of these last two processes may then influence the lower-level processes, creating a feedback system within the model (Reybrouck & Eerola, 2017:8).

As indicated above, dimensional models do not always take *emotions*, in the sense outlined at the beginning of this chapter, as their conceptual object. The framework is versatile – depending on the research question, a dimension approach may be applied to many different emotion-like phenomena. For instance, the conceptual object of PANAS is *mood* (Watson, Clark, & Tellegen, 1988:1063), and Eerola's integrated framework limits the utility of a dimensional approach to the level of *core affect* (Reybrouck & Eerola, 2017:8). As one moves to more highly-differentiated types of emotion-like states, the predictions of the dimensional model require more qualification.

Musical emotions, and "aesthetic emotions" in general, are complex, subtle, and varied emotional phenomena. The dimensional model is therefore unable to account for every minute detail of such an emotional response, and does not endeavour to do so. It is, however, a useful and valid indication of the basic stage of even the most complex emotional phenomena. It has predictive power, and its predictions have been tested to an extent, with confirmatory results (e.g. Lench *et al.,* 2011). On top of its validity, the dimensional model, especially a two-dimensional version such as Russell's, provides a clear and simple representation of emotional states in terms of easily understood and experimentally implementable dimensions. The present research, aiming as it does to identify general principles rather than provide rich descriptions,

will make use of the dimensional model as a core conceptual framework for experimental design.

Dimensional theory is better suited to the requirements of the present research. Specifically in its iteration as a 2-dimensional (VA) model, it has proven to be highly effective in enabling ratings of perceived musical emotions (e.g. Eerola & Vuoskoski, 2011; Schubert, 2004). It is able to account for less extreme stimuli more comfortably than discrete emotion methodologies (e.g. Eerola & Vuoskoski, 2011:40). The inclusion of complex target emotions is theoretically unproblematic in a dimensional approach, as it makes no assumptions of the primacy of any particular emotional states. A number of discrete emotion labels have been plotted onto VA (or similar) dimensions with more or less consistency (e.g. Mehu & Scherer, 2015; Smith & Ellsworth, 1985; Russell, 1980), forming a body of literature which allows for fairly secure interpretation of dimensional results.

The present research uses stimuli written to portray three discrete emotions. Discrete emotion theory would be sufficient to determine whether or not these emotions are portrayed successfully. However, it is likely that there will be at least some small discrepancy between the stimuli in terms of their emotion portrayals; stimuli that are theorised to portray the same emotion are unlikely to produce the exact same ratings, be they discrete or dimensional (e.g. Vieillard *et al.*, 2008:744-752; see VA ratings and best label percentages in Appendix 2). These discrepancies will be more effectively captured by a dimensional framework, as this allows for quantitative differentiation between rating clusters (for each excerpt) in terms of the two dimensions (e.g. Vieillard *et al.*, 2008:728). A discrete model would facilitate differentiation by the extremity of ratings for each emotion label individually, but it would be more difficult to quantify the *direction* of deviations from typical ratings, as each emotion is theoretically distinct in a discrete approach (Ekman, 1992:170). Studies of musical emotion utilising a discrete framework give their results in the form of accuracy scores for the categorisation of stimuli (e.g. Balkwill & Thompson, 1999:53). While this allows for comparison in terms of which emotion labels caused confusion in an ambiguously rated stimulus (e.g. Banse & Scherer, 1996:622), an underlying mechanism for this confusion in terms of patterns of cue utilisation is difficult to find. Through the use of emotion dimensions, the present research allows for comparison of cue utilisations in terms of specific cues' effects on each individual axis, as has been done in much prior research (e.g. Ilie & Thompson, 2006:324; Leman *et al.*, 2005:54; Goudbeek & Scherer, 2010:1333).

A dimensional approach is well-suited to the use of rating scales corresponding to the dimensions under investigation (e.g. Scherer & Oshinsky, 1977). Feasibility is not an issue; a VA-dimensional approach would lead to only two rating scales being required. Even using only two rating scales, a dimensional approach is far less limiting than a discrete one, as a far greater (theoretically infinite) number of emotional states may be represented in the dimensional space. The VA model allows for greater differentiation between rating clusters, using a minimal number of rating scales. The scales themselves are easy to explain to participants; no studies reviewed over the course of this research project reported difficulty in explaining at least the dimensions of valence and arousal. The overall consistency in the data generated using a VA-dimensional approach indicates that participants interpreted these scales consistently.

In a dimensional approach, acoustic cues can be examined in two complementary ways. First, a set of cues can be associated with a given area of VA space, creating cue profiles for broader or narrow regions. Second, individual cues can be assessed in terms of their impact on one or both dimensions. A dimensional approach also goes some way towards minimising the impact of emotion words themselves on listeners' judgements of musical emotions. Emotions can be reported and analysed with more nuance than is allowed by forcing discrete labels onto emotional portrayals. In the course of a human life, one is likely to experience and perceive a great many emotions for which one does not have a ready label; the number of possible emotional states may be near-infinite (see Banse & Scherer, 1996:616 – footnote 3). The accumulation of dimensional data can help us to learn about emotions which occur quite normally, but which are not easily able to be put into words.

The dimensional approach is more suited to the requirements of the present research. It is likely to yield greater accuracy of ratings that do not fit neatly into primary labels, while posing no problem for those that do. It will be at least as simple to implement methodologically as the discrete approach, and is highly likely to be equally easy to explain effectively to participants. In terms of interpreting the experimental results, the dimensional approach offers an advantage in terms of the deliberate "fuzziness" of the labels applied to the stimuli. The stimulus labels will be plotted as fuzzy areas around a central point (as identified in previous research or imputed based on the available evidence), and thereafter be identified with this area, rather than with the label itself. The plotting of an area (as opposed to a point) makes some degree of allowance for the almost-certain historical and translational drift in the meaning of the emotion label – as long as the participants' ratings are clustered, and that cluster falls within the

generous area assigned to the target emotion, the ratings will be interpreted as representative of that emotion.

2.3 Component Process theory

Component process theory is discussed here on the basis of its inclusion in most current research on emotion (Scherer, 2000:138). It must once again be emphasised that this theory is not a *general theory* of emotions in its own right; rather, it aims to explicate the complexity and interrelatedness of the subcomponents of what we understand to be an emotional episode.

At its most basic, component process theory makes itself known in the concept of the reaction triad of emotion, as discussed above. The reaction triad is constituted of the three most widely accepted component processes of emotion: physiological arousal, motor expression, and subjective feeling (Scherer, 2000:138). A number of researchers assume two additional processes: action tendencies (approach-avoidance), and cognitive reactions to stimuli (Scherer, 2000:138-139). According to component process theory, an emotion may be defined in terms of these processes, and more importantly, in terms of their *coordination*. It is generally held that it is not enough to observe a change in one component process and conclude on this basis that an emotion is present. Emotions are, to some extent, holistic and dynamic. Therefore, "*interdependent and synchronized changes* in component processes are required as a necessary condition for the definition of emotion". (Scherer, 2000:138, emphasis original). The five major component processes will now be outlined in brief.

Physiological arousal is a widely used indicative measure of emotional phenomena. Particular types of physiological arousal which may be tested for include changes in: internal body temperature, respiration rate, heart rate, and skin conductivity (Lench *et al*, 2011:840). There are fairly well-established tools for these measurements. Intuitively, it seems sensible and highly desirable to find unconscious physiological correlates of emotions; however, the findings in this area to date are by no means conclusive.

Motor expression (behaviour) subsumes a number of forms – gestures, posture, facial expression, and vocalisations may all occur as part of emotional episodes (Scherer, 2004:240). As well as simply being involuntary reactions to stimuli and their experiential correlates, motor expressions are assumed to play a role in the *communication* of emotional states. As they are able to be perceived by other individuals, it has been hypothesised that these expressions play a role in the phenomenon of emotional contagion (Neumann & Strack, 2000:211).

Action tendencies are considered by some theorists to be a component process of emotion, lending the phenomenon an additional function of motivation for the organism. These tendencies are considered to be fairly broad and vague, readying the organism for immediate response and priming various behaviours that are likely to be necessary, to be chosen from once more information has been made available or processed (Scherer, 2004:241). Examples of such tendencies include "moving towards" (approach), "moving away" (avoidance), and "moving against" (attack; Frijda *et al*, 1989:213).

Cognitive reactions to emotion-inducing stimuli are considered to consist primarily of the appraisal of these stimuli in terms of their likely impact on the survival and goal-achievement of the organism. As emotions have been shown to modulate cognition in several areas (e.g. attention, memory, problem-solving), the cognitive reaction process is considered to include feedback systems – appraisal plays a role in giving rise to an emotional response, which affects subsequent appraisal, and so forth (Scherer, 2004:241).

The final component process to be discussed here is subjective feeling. This is the qualitative aspect of emotional phenomena, and is usually accessed through self-report. A definition of this component is provided by Scherer: "the feeling component of emotion can be conceptualized as a reflection of *all* changes in components during an emotion episode" (Scherer, 2004:241). This suggestion involves the combination in conscious experience of the experiential correlates of cognitive processes and action tendencies, as well as of motor expression and physiological arousal in the form of proprioception (Scherer, 2004:241).

Various self-report methodologies may be used to access the component of subjective feeling when testing for emotion. Theorists in the discrete emotions tradition tend to use checklists or rating scales consisting of basic emotion categories; dimensional theorists use rating scales on a number of dimensional axes (usually two, VA; Scherer, 2004:240). A less structured approach is sometimes used, in which terms for inclusion on checklists or rating scales are chosen by researchers on a case-by-case basis (Scherer, 2004:248).

As mentioned above, component process *coordination/synchronisation* has been held up as the gold standard for identification of an emotional episode, especially when working in a discrete emotion framework. However, two critiques of this assumption (taking discrete emotion theory in terms of a theory of biological basicness of emotions) are relevant; one of the relation between self-reported emotions and other components, and one of the level of coordination between component processes generally. These are termed the No One-to-one Correspondence

(NOC) thesis and the Low Coordination (LC) thesis, respectively (Scarantino & Griffiths, 2011:448).

Scarantino and Griffiths draw these critiques from Ortony and Turner (1990). This earlier article gives convincing evidence of the lack of consistency in formulations of basic emotions models and the inconclusive nature of data examining their claims. Throughout their critique, the authors outline a model in which component processes may be seen as basic (primarily in the biological sense, but also the psychological one), even referencing Scherer, whose work on component process theory was relatively new at this time (Ortony & Turner, 1990:327). They dismiss the idea of *emotions* in the fully-fledged sense as being candidates for "basic" status in any but the conceptual sense. In terms of component process theory, their critiques of existing basic emotions models are outlined and responded to below under the aforementioned headings of NOC and LC (although these labels are not used in the original article), drawing on Scarantino and Griffiths (2011; Ortony & Turner, 1990)

NOC states that proposed biologically basic emotions lack a one-to-one correspondence with any component responses (physiological, neurobiological, expressive, behavioural, or phenomenological). LC states that these component responses (as well as a cognitive response component) do not show significant coordination in instances of proposed basic emotions (Scarantino & Griffiths, 2011:448). Scarantino and Griffiths discuss these critiques at length, acknowledging them as valid and accurate reflections of the empirical data. Rather than abandoning the notion of biological basicness, they suggest two revisions to the theory. The first suggestion is the development of specific scientific emotion terms, in contrast to the folk terminology currently employed. Because the folk definitions of emotions follow different rules to the scientific ones, instances that fall into the folk category may be cited as examples of the NOC or LC thesis – conflating two distinct ideas (Scarantino & Griffiths, 2011:449-450).

The authors' second suggestion is that affective science take up an antiessentialist approach to emotion classification. They suggest Boyd's theory of Homeostatic Property Clusters as a viable candidate: in this theory, members of a kind are determined by a cluster of properties which regularly co-occur, and by the mechanisms underlying that co-occurrence (Scarantino & Griffiths, 2011:445). This approach allows for the variability in correspondence and coordination seen in the data, thus giving the NOC and LC theses pause for thought. While the authors take the view that the notion of basic emotions is a valid one, they acknowledge that

the data are inconclusive, and the utility of the concept for scientific generalisation remains an open empirical question (Scarantino & Grifiths, 2011:447-450).

2.4 Prototype theory

The theory of emotion prototypes is strongly antiessentialist. It argues that "emotion" is not a classical concept, in the sense of definition by "individually necessary and jointly sufficient attributes" (Fehr & Russell, 1984:464). Rather than a binary attribute, a proposed emotion's occupation of the "emotion" domain is a matter of degree, and the boundaries of the domain are fuzzy (Fehr & Russell, 1984:464).

Prototype theory is not limited to the domain of emotion. Eleanor Rosch (1978:259) defines prototypes as "the clearest cases of category membership defined operationally by people's judgements of goodness of membership in the category". For instance, to ascertain the identity of prototypical members of the category "buildings", one might ask a number of people to write down the types of buildings they feel to be the clearest category members, or to select, rate, or order members from a list (see e.g. Shaver *et al*, 1987:1064-1065). It is clear that one can expand this definition to any number of categories. Prototypicality thus defined has been shown to have a useful structural feature: degree of prototypicality of a category members correlates positively with the number of attributes shared with other category members, and negatively with the number of attributes shared with members of contrasting categories (Rosch, 1978:260). In terms of emotion research, an approach to the classification of emotions into related groups can utilise this property. Ratings of prototypicality, alongside correlations in structural features, can be used to find relations between both emotion concepts at the *folk* level of definition, and emotional phenomena according to a *scientific* description.

A concept closely allied to prototype theory, which has relevance to the present research, is cue validity. This can essentially be defined as a function of the association of a given item's constituent cues with the predictor cues of its category and non-association with predictor cues of other categories (Rosch, 1978:254). For example, the cue of *having wheels* increases the probability of membership of the item *hatchback* in the category *cars*; and the cue of *having wings* decreases the probability of membership of the item *aeroplane* in the same category.

Prototype theory also entails the idea of fuzzy hierarchies (Rosch, 1978:255). Russell and Feldman Barrett (1999) suggest a way in which these fuzzy hierarchies may be reconciled with both the discrete and dimensional theories of emotion. They suggest that emotion labels occupying similar areas of a dimensional space may be traced upwards to a shared

superordinate category, and downwards to finer-grained subordinate categories (Russell & Feldman-Barrett, 1999:808). This brings discrete emotions and dimensional emotion spaces together systematically, maintaining useful elements of both.

The field of emotion research is broad, complex, and full of unanswered questions and conflicting theories. A broad survey of these has been provided. The next chapter addresses questions of the relation between music and emotions.

CHAPTER 3. MUSIC AND EMOTION

It is taken as true *a priori* by lay people and musicians alike that music has emotional content. Upon closer examination, this claim becomes somewhat unclear—what do we mean by "emotional content"? This definitional gap is filled by two predominant schools of thought: emotivism and cognitivism. Emotivism holds that music listening *induces* emotions, while cognitivism holds rather that listeners *recognise* emotions in the music. Each of these paradigms allows for the investigation of certain questions regarding music and emotion, and requires a specific methodological approach to match. In this chapter, cognitivism and emotivism will be defined separately, and examples of their application and methodologies will be given.

3.1 The cognitivist position

Cognitivism states that musical expression may take place without emotion actually being felt by the listener. This occurs through a mechanism of *perception* of expressive features in the music, which listeners recognise and interpret, but which do not necessarily elicit emotion (Kim, 2013:163). The systematic perception of emotional expression in music listening has been widely confirmed (e.g. Hevner, 1935b, 1936, 1937; Juslin & Laukka, 2003; Quinto *et al.*, 2014; Schubert, 2004). There are three core points to discuss regarding cognitivism: first, why it may be asserted that emotions are perceived in music; second, how they get there; and third, how cognitivism may account for listeners' self-reports of emotion induction.

Studies within a cognitivist framework always seek to uncover listeners' *judgements of emotional portrayal/quality* in the music used as stimuli. Self-report is the most commonly used measure in this research, and can take a number of forms. Listeners may be asked to give individual descriptions of their judgements of the music's emotional content (e.g. Juslin, 1997c:89), to select appropriate adjectives from a list (e.g. Hevner, 1935b:108; Juslin, 1997c:81), to make ratings on scales of emotion terms (e.g. Eerola *et al.*, 2009:622) or dimensions (e.g. Vieillard *et al.*, 2008:726; Gagnon & Peretz, 2003:30), or to group stimuli according to judgements of similarity (e.g. Vieillard *et al.*, 2012:429). These methodologies are also frequently employed for the self-report components of research in the emotivist paradigm; the difference then lies in the specific wording of the instructions given to participants. Cognitivist research will always emphasise to participants that they should report the emotions they *perceive* in the music (or similar instructions, e.g. Vieillard *et al.*, 2008:725),

while emotivist research will instruct participants to report the emotions they *experience* in response to the music (e.g. Bigand *et al.*, 2005:1119).

Quinto *et al.* (2014) conducted a study in which listeners attempted to decode emotional portrayals in three musical conditions: composition (musicians wrote melodies to portray specific emotions, which were played to listeners in MIDI format), performance (musicians performed melodies composed to be emotionally neutral, with the intention of portraying specific emotions in their performance – these performances were recorded and played back to listeners), and combined (musicians performed their own melodies, aiming to enhance the composed emotion portrayal in their performance – these performances were recorded and played back to listeners) (Quinto *et al.*, 2014:508-510). Listeners rated the emotional expression they perceived in the music using an adjective checklist (forced-choice) (Quinto *et al.*, 2014:503). The lowest level of decoding accuracy they found was 24.11% (for the emotion of anger, in the composition condition). This seems low, but given that chance would yield only 16.6% accuracy under the conditions of their experiment, it is in fact in line with the idea that musical emotional portrayals may be decoded systematically. They found higher accuracy for all other emotions and conditions, with the highest being 71.43% for happiness in the combined condition (Quinto *et al*, 2014:511-512). The fact that all emotions in all conditions were decoded with significantly above-chance accuracy is strong support for the systematic perception of emotion in music.

Gagnon and Peretz (2003) tested listeners' judgements of emotional expressions along the dimension of happiness-sadness using a rating scale paradigm. They found that listeners judged this dimension reliably in accordance with the intended emotional portrayals of their excerpts (Gagnon & Peretz, 2003:31-32). The accuracy of listeners' ratings despite the simplicity of the stimuli used in this experiment, indicates just how powerfully the ability to identify emotional portrayals in music manifests in humans.

As part of the experiment reported in Vieillard *et al.* (2008), listeners rated the emotional expression they perceived in a set of musical excerpts. A rating scale paradigm was used, with 20 listeners making ratings for perception of four emotion labels, and 20 making ratings for valence and arousal. Another group rated their *experience* of emotions according to the four emotion labels – their ratings were similar enough to those made for emotion *perception* that the authors felt justified in collapsing results across this divide. Their findings in both the emotion labels and VA conditions were systematic. In the emotion labels condition, all excerpts

received higher ratings for their intended emotion portrayal than for the alternatives. In the VA condition, excerpts formed distinct clusters by emotion portrayal, with clusters forming around areas of the VA space corresponding to the appropriate emotion label (Vieillard *et al.*, 2008:725-729). This further supports the systematic and reliable perception of emotional expression in music.

Peter Kivy offers a speculative explanation of how music, specifically melody, may be expressive of emotion. The phrase "be expressive of" is important – it is meant to contrast with the similar term "express". For some object to *express* an emotion, a necessary condition is taken to be that the object *feel* that emotion. For example, if I feel angry, and bare my teeth because of this, I may be said to be *expressing* my anger. For an object to *be expressive of* an emotion, no such condition is required. For example, the face of a St Bernard *is expressive of* sadness, regardless of the dog's own emotional state. It *is expressive of* sadness in virtue of containing certain features which would be typical of a human *expressing* sadness – hence its recognisably sad quality. Music, being non-sentient, can never be said to *express* an emotion, but only to *be expressive of* one (Kivy, 1980:12-17).

The explanation Kivy offers is that melody may be expressive of emotions in virtue of its resemblance to the normal (non-musical) behavioural expressions of those emotions in humans. These expressions may range from vocal, through motor, to postural, and indeed any archetypal behavioural expression within a given culture. Kivy traces this explanation, at least in simple form, back to the theoretical and practical musical innovations of the Florentine Camerata around the turn of the 17th century. It was held by these theorists that music may be expressive of emotions insofar as it is designed to mimic the natural vocal inflections accompanying impassioned human speech. Their theory did not extend to other forms of expressive behaviour. Composers who followed this school of thought designed their melodies to follow the rough contours of the human voice when roused to expression of felt emotional states, and believed this to impart in their music a recognisable expression of the chosen emotion. For example, sadness may be imbued into the melody through a general downward contour, and a preponderance of "sigh motifs" (semitones, usually descending) (Kivy, 1980:18-20). Kivy's ideas have been empirically tested, at least in terms of the correspondence between musical and vocal emotional expression, with positive results.

Juslin and Laukka (2003) studied the similarities between vocal and musical expression and perception of emotions. Their findings indicated a large degree of overlap between the two

channels, in terms of both encoding and decoding of emotional content (Juslin & Laukka, 2003:805). They were able to draw up a fairly extensive list of characteristic patterns of cue utilisation common to both channels for five emotional portrayals (Juslin & Laukka, 2003:802 – see figure 4.1.1 in the next chapter of this literature review). Their results support the idea of music's expressiveness being at least in part a function of its resemblance to vocal emotional expressions.

The theory that emotional expression in music is based on emotional expression in the vocal domain leads to the consideration of the cultural specificity of musical emotional expression. There have been findings of vocal emotional expressions being cross-culturally decodable (e.g. Laukka *et al.*, 2013). In the musical domain, some evidence of this phenomenon has also been found. Balkwill and Thompson investigated Western listeners' ability to decode emotional portrayals in North Indian music, with mixed results. They found that listeners were able to accurately decode the emotions of joy and sadness, but their results for anger and peace did not show reliable decoding (Balkwill & Thompson, 1999:52-53).

Balkwill and Thompson (1999:43) proposed that listeners utilise two theoretically distinct sources of information in forming emotion judgements of music. On the one hand, they draw on cultural conventions of emotional expression in individual tonal systems. On the other, they draw on *universal* psychoacoustic parameters of the sound. They found some degree of support for the role of psychoacoustic universals. The potentially universal psychoacoustic cues they examined were tempo, rhythmic complexity, melodic complexity, pitch range, and timbre. Across the four emotions studied, listeners' judgements of emotional portrayal were significantly associated with several of these cues (Balkwill & Thompson, 1999:54-56). Their finding of some accurate decoding of music across cultural boundaries hints at the existence of universal aspects of musical emotion portrayals.

Cognitivism could stand perfectly comfortably without ever mentioning the induction of emotion by music. However, it is not afforded this luxury. It is an indisputable fact that people at least *report* felt emotional states in response to music, both in experimental contexts (e.g. Evans & Schubert, 2008; Kallinen & Ravaja, 2006) and in more ecologically valid settings (e.g. Gabrielsson & Lindström Wik, 2003). Cognitivism may offer a number of explanations for this, covering a wide range of theoretical ground. Two major examples of these explanations will now be discussed.

The first of these explanations is well-known, under the name of the "our song" phenomenon. It holds that music may induce genuine emotional states in listeners through the past association of a piece of music with an emotionally salient event. Upon hearing the music once more, the listener is reminded of the event it is associated with, and the recollection of this event causes an emotional response (Kivy, 2002:111-112). The appeal of this explanation for cognitivists is that the emotional object is not the music, but the memories associated with it. It accounts for many real-world cases, and will likely be known to most people from personal experience. However, it cannot be the only explanation for music's ability to induce emotions; if it were, listeners' self-reported emotional responses to music would be functionally random. In order to explain the fact that many emotional responses to the same music are described in intersubjectively similar, if not identical, emotional terms, another explanation will be needed.

The second explanation is that listeners self-reporting a specific felt emotional state (*specific* in terms of labels used to denote everyday emotions) are simply wrong. That is, the label they choose is a misattribution to the emotional phenomenon they are in fact undergoing. Kivy believes that music is not able, through structural properties alone, to induce such emotions in listeners. He holds rather that music *moves* listeners, in virtue of its aesthetic beauty, to a state of awe or rapture. This is certainly an "emotional" state, but it is not, by our modern, functionalist definition, an *emotion*. The aesthetic beauty of a work of art may be constituted of any of a number of factors, structural or otherwise. One of these factors, according to Kivy, is the emotional expression embodied in music. He says that, in cases where such expression is the object of aesthetic awe, listeners may label their emotional state according to the perceived nature of the *object*, rather than the nature of their subjective feeling (Kivy, 2002:130-134).

Kivy himself adduces no evidence as to the existence of this effect. However, some preliminary support for this hypothesis may be found in studies investigating musically induced emotions using self-reported subjective feeling alongside other emotion components. Music has been found to effect physiological changes in listeners (e.g. Rickard, 2004; see also Scherer & Zentner, 2001). Physiological response is part of the reaction triad of emotion, constituting a core component process of emotional states. The reality of self-reported induced emotional states in response to music has been assessed by examining physiological responses along with self-report. Krumhansl (1997:338-339) cites two studies in which musical manipulations were associated with a change in self-reported subjective feeling state, but were not associated with any changes in physiological measures. She also mentions one in which there were changes in

both subjective feeling and physiology, but in opposite directions. This could potentially indicate that participants were misattributing the emotions they *perceived* in the music as being *felt*, rather than introspecting successfully.

Konečni (2008:122) also mentions this effect as a possible explanation for the self-report of felt emotional states in response to music. Another idea he puts forward is that the belief in a link between music and emotion is propagated in common language use, and that this informs participants' responses in experimental settings. Given the semi-axiomatic status of this link, participants may "feel foolish circling 'zero'" when asked by an experimenter to provide a rating of their emotional response to music (Konečni, 2008:122).

Cognitivism provides a good account of how people may *recognise* emotional portrayals in music. In terms of accounting for the phenomenon of self-reporting of emotional states in response to music, it answers some cases, but it cannot be claimed with any certainty that it accounts for all. Assuming, as for the moment there is no compelling reason not to, that self-reported felt emotion in response to music is not entirely fictitious, and cannot be fully displaced onto mediating effects, the competing paradigm of emotivism will now be discussed to shed light on this phenomenon from another angle.

3.2 The emotivist position

Emotivism stands in contrast to cognitivism in claiming that music can indeed induce genuine emotional states in listeners. Various explanatory mechanisms are offered to ground this claim. These may be split into two kinds: first, those that are based on a *direct* induction of emotion by music, and second, those that make allowance for cognitive processing as a component of the explanatory framework, or rely on other mechanisms. Before this, however, research which claims to support the emotivist position will be discussed, in order to assess the level of empirical evidence for its claims.

In a questionnaire given to 141 participants (roughly equally split between musicians and non-musicians), Juslin and Laukka (2004) gathered strong evidence of the induction of emotion in everyday music listening. Participants were asked a number of questions regarding musical emotion induction. In response to the question of how often during music listening they felt strong emotions, no participants indicated that they had never experienced this phenomenon. The mean response was 55% of total listening time (Juslin & Laukka, 2004:230). This suggests that emotional experience in response to music does indeed occur, and is by no means a rare phenomenon. Another question asked participants to rate the relative importance of three

contributing factors to musical emotion induction: the listener's psychological state, the music, and the listening situation. All three were rated as important, but music received the highest mean rating (Juslin & Laukka, 2004:231). Although this rating was not an extreme stand-out in comparison with the other two, it supports the idea of the emotional object of music listening being the music rather than other factors. The role of lyrics in emotional induction was not investigated in this paper, and it cannot be ruled out that this was a contributing factor to the questionnaire respondents' emotional experiences.

Bigand *et al.* (2005:1119) conducted an experiment in which two groups of participants (musicians and non-musicians) listened to a set of musical excerpts and grouped them in terms of the similarity of the emotional experience each excerpt induced. This experiment was later repeated with the same participants and stimuli to facilitate assessment of the within-subjects consistency in grouping over time. Their results were encouraging: both musicians and non-musicians produced similar numbers of groups, and there was a high degree of consistency between experimental sessions (Bigand *et al.*, 2005:1121). This indicates that there was a real and systematic mechanism underlying the participants' ratings of felt emotion.

The studies discussed above provide some useful data in support of the emotivist position. However, as they rely exclusively on self-report, their findings may fall victim to the problem of the definition of emotions, as discussed in chapter 2 of this literature review. Specifically, participants may have been making reports based on an internal understanding of emotion in line with *folk definitions* (see Scarantino & Griffiths, 2011). In order to ground self-report in terms of a scientific definition of emotion, it is desirable to include other measures alongside self-report.

Rickard (2004) studied the effects of emotionally powerful music on a number of physiological processes in listeners. Rather than asking participants to self-report their states using emotion terms, an indirect measure was used, with participants rating the stimuli (four types, of which emotionally powerful music was one) on the following scales: *moving, involving, interesting, absorbing, concentration*. High ratings on these scales were taken to mean that the stimulus had had an emotional impact on the participant (Rickard, 2004:378-379). Having used these measures to establish the validity of the emotionally powerful music in inducing emotional experience, the physiological results were examined. A significant increase in skin conductance and frequency of chills was observed when participants were listening to the emotionally powerful music. None of the other physiological measures (heart rate, skin temperature,

electromyography, salivary cortisol) differed in this treatment condition (Rickard, 2004:381-382). It is important to note that the excerpts were familiar to participants (2004:378), and as such it cannot be ruled out that the emotional response was a product of mediating factors (e.g. memories). The association of subjective emotional experience in response to music (as measured indirectly by the scales mentioned above) with physiological changes provides support for the emotivist position using a scientific (as opposed to folk) definition of emotion.

A similar experiment was conducted by Grewe *et al.* (2007a). They tested responses across self-reported subjective feeling, physiological arousal, and motor reactions (the reaction triad of emotion). Participants (professional musicians, amateur musicians, and non-musicians) listened to seven pieces of music. There were low, but frequently non-zero ratings of familiarity with several of these pieces across different participants (Grewe *et al.*, 2007a:780). Subjective feeling was reported using a computerised version of the VA space (the EMuJoy software). Physiological effects were assessed with a measure of skin conductance, and motor reactions were assessed using facial electromyography (Grewe *et al.*, 2007a:777). In stark contrast to Rickard's (2004) study, no systematic relationship between the three components was found (Grewe *et al.*, 2007a:784). The authors acknowledge that they were using a highly conservative definition of emotion to inform their criteria, which to their knowledge had never been met in an experimental setting (Grewe *et al.*, 2007a:787). Recall the Low Coordination thesis discussed in chapter 2 of this literature review – the synchronous occurrence of component processes has not been shown conclusively in any line of experimental emotion research (Ortony and Turner, 1990; Scarantino and Griffiths, 2011).

The data are inconclusive as to whether or not music can induce genuine emotional states. A significant portion of the blame for this ambiguity lies with the lack of consensus on a scientific definition of emotion, and its systematic implementation in experimental research. This complexity disallows a clear judgement on the validity of the emotivist position. That being the case, it will at the least be informative to examine some theories of how emotional induction may take place through music, if indeed it can.

The strongest form of the emotivist position claims that music induces emotions directly through its interaction, as physical stimulus, with our human sensory and wider auditory mechanisms (Kivy, 1990:31-32). This position does not seem too drastic of a claim, until one considers that it reduces the entirety of music's emotive power to a purely physical system at the biological, unconscious level, excluding any possible role of the cognitive faculties (thus

this effect could theoretically be generalised to non-musical auditory stimuli). This is a bold claim (at least, so it seems today), and seldom appears in any meaningful way in the modern literature. Its most notable historical proponent was René Descartes, whose views will be described in brief here. Following this, some later forms of the emotivist position will be discussed.

Descartes' emotivism was based on a speculated interaction between the musical stimulus and internal forces within the body. Music, for Descartes, is capable of arousing different emotional states in virtue of its resemblance to the normal motion of these forces for each emotion (Kivy, 1990:32-36). Descartes' view, though now rejected alongside his theory of the passions in general (and much of his whole psychology), held sway in various forms in much of the music-philosophical thought of his own century, and the next. This will be discussed further in chapter 5 of this literature review.

An interesting theory of musical emotion induction is put forward by Scherer (2004). Aesthetic norms and standards, as socially/culturally determined, may be seen as a reference point against which a piece of music may be compared. This comparison constitutes a cognitive appraisal process (one of the component processes of emotion). If a piece of music violates these norms, it may induce negative emotions in listeners, such as anger or disgust. Scherer (2004:245) adduces the scandalous premiere of Stravinsky's *Le Sacre du Printemps* as an example of this process. Such a process could likely account for positive emotions as well – a piece that conforms to social/cultural norms may induce happiness through this same appraisal process. This line of reasoning is very promising, and certainly deserves further research.

Another candidate theory rests on the assumption that a change in one component process of emotion can lead to changes in others. There is some degree of empirical evidence for this claim (Scherer, 2004:245). In this theory, music would cause a change to a specific component process (motor expression and/or physiological processes), which could then lead to the matching of this change by other component processes. Music could affect the motor expression component, for instance, by inducing foot tapping or head movement. A potential physiological effect would be a change in respiration as a function of rhythmic elements in a piece of music. It is theoretically possible that such individual changes could, through this mechanism, lead to the induction of an emotional state (Scherer, 2004:245).

A theory relying on a process similar to that described in the previous paragraph claims that music may facilitate in a listener the expression of already existing, but suppressed, emotions.

A person may suppress an emotional reaction due to any number of social/cultural norms, or for strategic reasons. It is possible that music can work towards bypassing these regulatory efforts. This could occur through music causing a change in one or more component processes of emotion, as described above. If these changes reinforce the pre-existing but suppressed tendencies in component processes corresponding to the emotion being kept in check by the listener, it may become more difficult to continue suppressing the emotional reaction as a whole. There is anecdotal evidence of this phenomenon in the context of film music – the addition of music to an emotionally powerful scene is often reported to lead to crying in cases where tears were being held back, representing a weakening of emotional control (Scherer & Zentner, 2001:372). Film music occurs in a highly designed context, and is often written in such a way as to correspond precisely to the emotional tone of the scene in question. However, it is not unreasonable to think that this phenomenon could occur in less structured contexts. The essential elements are a subject attempting to control an emotional reaction, and a musical stimulus that matches this emotion in terms of any of the effects on component processes it induces in the subject. Such a situation is likely very common. Two conceptual objections may be raised to this theory in as far as it claims to support emotivism. First, if the emotional reaction in the subject prior to the musical stimulus may already be called an emotion, albeit one that is subjected to regulatory efforts, then this is not truly a case of emotional *induction* by music, as the music is in no way *creating* an emotion. Second, if the emotional reaction prior to the musical stimulus *cannot* yet properly be called an emotion, then the question remains as to whether the music may in any way be considered to be the *object* of the emotion, once the proto-emotion's realisation as an emotion proper has been facilitated. Notwithstanding these objections, this theory certainly has explanatory power in terms of real-world phenomena.

The emotivist theory is by no means conclusively proven. Its proposed underlying mechanisms provide fertile ground for further empirical research. Due to the prevalence of lay ideas about music's ability to produce emotional reactions, it will be of great value to have more information in this area. The present research, however, is situated within a cognitivist framework, focusing on the *recognition* of emotional portrayals in music. The next chapter discusses empirical research on cue utilisation in such portrayals.

CHAPTER 4. ACOUSTIC CUES AND EMOTIONAL RESPONSES TO MUSIC

Taking as a starting point Peter Kivy's (1980:18-20) assertion that at least one way in which music's sonic structure may be expressive of emotions is through similarity to the human voice under the influence of the corresponding emotion, it is possible to flesh out this theory beyond the contour-similarity that Kivy proposes. This requires an operationalisation of vocal expression during emotional episodes, in terms of various components of the sonic structure. Just such a project was undertaken in 1986 by Klaus Scherer.

Scherer's paper deals only with the vocal affect cues that may be brought about by *physiological* changes due to emotional state (Scherer, 1986:146). It is important to note that these represent one of two types of factor which may influence vocal expression, namely "internal push" factors. The other type, not studied by Scherer in his paper, are "external pull" factors, representing social and cultural norms and judgements which inform the emotional display of the individual (Scherer, 1986:145-146).

The vocal affect cues Scherer studied were: perturbation, mean, range, variability, contour, and shift regularity of fundamental frequency, mean of first and second harmonics, bandwidth of first harmonic, formant precision, mean, range, and variability of intensity, frequency range, high-frequency energy, spectral noise, speech rate, and transition time. Predictions relating to these cues were given under conditions of the following emotional states: enjoyment/happiness, elation/joy, displeasure/disgust, contempt/scorn, sadness/dejection, grief/desperation, anxiety/worry, fear/terror, irritation/cold anger, rage/hot anger, boredom/indifference, and shame/guilt (Scherer, 1986:158).

Scherer's model was utilised by Juslin and Laukka (2003) as a comparison point for their findings from a metastudy of research dealing with the expressive cues utilised in both vocal and musical expression. Their results, generated from an expansive review of studies (104 for vocal expression, 41 for music), matched very well with Scherer's predictions (though they did not make comparisons with the full set of predictions) (Juslin & Laukka, 2003:801). Juslin and Laukka aimed to study the relationship between vocal expression of emotion and expression of emotion in musical *performance* (as distinct from compositional elements; this division is discussed below). They found a very high degree of agreement between vocal and musical/performance expression, both in terms of the acoustic cues used in forming the

communicative code, and in terms of patterns of decoding accuracy (Juslin & Laukka, 2003:797).

Based on their findings across the vocal and musical expression modalities, Juslin and Laukka developed a cross-modal (vocal and musical) acoustic cue profile for encoding of each of the five emotions they studied (anger, fear, happiness, sadness, and tenderness – see figure 4.1.1). They also listed the prevalence of cue utilisation in encoding of musical expression alone. Table 4.2.1 (discussed in detail later) contains this information in summary, as well as summary versions of the findings of other studies (which were not included in Juslin & Laukka's metastudy) in terms of patterns of cue utilisation for discrete emotions in music, for ease of comparison.

Emotion	**Acoustic cues (vocal expression/music performance)**
Anger	Fast speech rate/tempo, high voice intensity/sound level, much voice intensity/sound level variability, much high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and microstructural irregularity
Fear	Fast speech rate/tempo, low voice intensity/sound level (except in panic fear), much voice intensity/sound level variability, little high-frequency energy, high F0/pitch level, little F0/pitch variability, rising F0/pitch contour, and a lot of microstructural irregularity
Happiness	Fast speech rate/tempo, medium-high voice intensity/sound level, medium high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and very little microstructural regularity
Sadness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contour, slow voice onsets/tone attacks, and microstructural irregularity
Tenderness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contours, slow voice onsets/tone attacks, and microstructural regularity

Figure 4.1.1. Hypothesised cross-modal encoding patterns for acoustic cues, reprinted from Juslin & Laukka (2003:802). Note: F0 = fundamental frequency.

While Juslin and Laukka (2003) were aiming to study the relation of elements of music *performance* to elements of vocal expression, some of the cues they identified can play a role in musical *composition* as well. For instance, tempo, sound level, pitch level, and articulation as performance features overlap with compositional elements that may be indicated in a musical score. The separation of acoustic cues into compositional and performance categories is an important step for the present research, and is no trivial endeavour.

The boundary between composition and performance is, in most musical genres, a fairly soft one. Even limiting the scope of inquiry to Western Art Music, one finds the history of this relationship to be complex and subject to change. The general trend over the past 400 years of this music has been towards skewing the balance of influence over the sonic product in favour of composition. Composers of the Baroque era generally left much to the discretion of the performer – the oft-cited examples are ornamentation, with performers expected to provide their own embellishments to the written notes, and basso continuo notation, which left the majority of sounded notes to the discretion of the keyboard player.

Throughout the 18th and 19th centuries, the liberties afforded to performers by composers were pared down. Dynamics, articulation, and tempo indications began to be written into scores, accompanied by a greatly diminished emphasis on ornamentation. Where cadenzas were once improvised by performers, it became standard for these to be written out by composers, to be reproduced note-for-note in performance (Badura-Skoda *et al.*, 2001). In the 20th century, this prescriptivist trend reached its logical extreme; the aesthetic ideology of total serialism did its utmost to turn the performer into an obedient vessel for the reproduction of scores, with composers attempting to indicate every facet of the sonic product in the written score. When this proved difficult to implement, composers of this school renounced their reliance on performers entirely, and wrote music to be played back electronically.

The history of this Western Art Music illuminates a broad spectrum of possible relations between composition and performance. This spectrum is by no means universally exhaustive – the practices of structured and unstructured improvisation in jazz, as well as modes of musical creation in non-Western genres, offer novel insights into the roles of composers and performers, and opportunities to critique the global application of this dichotomy. The present research aims to illustrate the role of *compositional* cues in emotion judgements, and will, as far as possible, draw a meaningful distinction between composition and performance. The approach to this is explored below.

OVERVIEW OF ACOUSTIC CUES IN THE LITERATURE

Selection of the acoustic cues to place under consideration in chapter 8 of this paper must be done in accordance with the demands of all the excerpts used in the experiment. That is to say, cues must be selected so as to have meaningful analytic application to all the musical excerpts, Baroque and modern. In order to make a selection, it is necessary to have an idea of the cues analysed in the existing literature on music and emotion. The following section will survey the

work in this area, with the aim of compiling a comprehensive list of acoustic cues relevant to the compositional encoding of emotion. The relation between these cues and emotion judgements will be examined, to provide a framework for the prediction of experimental emotion judgements based solely on analysis of the music.

Two distinct approaches have historically prevailed in the study of the relationship between acoustic cues and emotion, in line with the two most prevalent theories of emotion discussed in chapter 2 of this review. The research in this area can thus be divided fairly neatly into the camps of discrete emotions and emotional dimensions, though there are some instances of combined methodologies. As far as is possible, these two bodies of literature will now be examined separately. Where necessary, the approaches to the delineation of cues into compositional and performance categories used by different researchers will be noted.

4.1 Discrete emotion research

Exploration of the relationship between specific structural elements of music (both as score and as sound) and emotion have been undertaken since the early part of the 20th century. One of the early pioneers, whose work remains foundational in the study of music and emotion today, was Kate Hevner. In the 1930s, she conducted a series of experiments to investigate the roles played by various musical cues in the communication of emotion (Hevner, 1935b; Hevner, 1936; Hevner, 1937), which represents the first detailed and systematic exploration in this field (Schubert, 2004:562). Her research will serve to begin an examination of the work that has been done in this area, in terms of both findings and methodology.

Hevner's research covered six acoustic parameters over a series of three experiments. These were: mode, melodic direction, rhythmic profile, harmonic consonance, pitch height, and tempo. Her methodology was broadly consistent across experiments. Pieces of Western Art Music were selected for use as stimuli. These were then recomposed, with only the individual parameter under investigation being altered, while all other factors were (as far as possible) held constant. All the original pieces and altered versions were learned by a pianist, and selections were performed live for participants. Participants made selections from adjective lists (from the second experiment onward, these were organised into eight groups of related words) to report their emotional characterisations of the music (Hevner, 1935b; Hevner, 1936; Hevner, 1937). The statistical figure used to represent the results per adjective group was a ratio of the difference between proportions of votes for that group under the two experimental

conditions (original and altered excerpts) and the probable error of this difference (Hevner, 1937:623; Edgerton & Paterson, 1926).

The methodology employed by Hevner has some flaws, and she was the first to point them out, albeit arguably without sufficient consideration of their seriousness. Due to the limitations on sound recording and production technology at the time, the presentation of stimuli had to be done through live performance, which leaves room for slight variation between performances. Hevner was aware of this, and made some attempt to monitor and minimise these differences (Hevner, 1935b:108; Hevner, 1936:256; Hevner, 1937:623), but there is still the very real possibility of small discrepancies between performances influencing the results. Moreover, while some of the acoustic parameters were easy to manipulate without changing any other elements of the music (e.g. pitch height), some parameters could not be manipulated in isolation without either making the music "unmusical" (Hevner, 1936:258), or effecting accidental and serious changes to other parameters. When these difficulties arose, Hevner always opted to make calculated and minor changes to other parameters along with the parameter under investigation, so as to preserve the "musicality" of the music and avoid more serious discrepancies (Hevner, 1935b:105-108; Hevner, 1936:258-259). Despite all efforts to minimise these changes, it remains possible that they affected the experimental results. In addition, in the study of tempo, the contrasting tempi did not accord very closely (or consistently) to the intended rule of one tempo being roughly twice as fast as the other (Hevner, 1937:622). Despite the potential sources of error in Hevner's experimental design, her results are worth noting, and will be discussed below for each acoustic parameter studied in her research.

For the most part, Hevner's results line up with traditional "common sense". In her first experiment, the minor mode was found to be associated mostly with adjectives of negative and ominous connotation, such as melancholy, gloomy, depressing, plaintive, mysterious, restless, and dark. The major mode was associated with more positive adjectives, such as happy, sprightly, bright, playful, graceful, and quaint (Hevner, 1935b:111-112). Hevner's findings on mode predated her organisation of the adjectives she used into eight categories (Hevner, 1936:249). These categories were used in the rest of her experiments.

Hevner's second experiment investigated rhythm, melodic direction, and harmonic consonance/simplicity. The two rhythmic categories were *firm* (full chord on every beat) and *flowing* (chordal notes spread out in an arpeggio-like fashion across the duration of the beat) (Hevner, 1936:256-257). Firm rhythms were associated with adjectives such as vigorous and

dignified, flowing rhythms with happy, graceful, dreamy, and tender. Complex/dissonant harmonies were associated with excitement, agitation, vigour, and to a lesser extent sadness, simple/consonant harmonies with happiness, grace, serenity, and lyricism. Melodic direction did not have any clear significance in listeners' judgements of emotion, but descending melodies tended towards both exhilaration and serenity, while ascending melodies tended towards dignity and solemnity (Hevner, 1936:268).

Hevner's final experiment, dealing with pitch height and tempo, found strong effects of both. High pitch was associated most strongly with humour and playfulness, low pitch with sadness, dignity, and vigour (Hevner, 1937:625-626). Fast tempo had its influence most strongly on happiness and excitement, and also affected grace and vigour. Slow tempo was strongly associated with dignity, calmness, tenderness, and sadness (Hevner, 1937:624).

Based on her results, Hevner listed the six parameters she had studied in order of the strength of their contributions to the perceived musical emotion. The list is ordered as such: tempo, mode, pitch height, harmony, rhythm, and melodic direction (Hevner, 1937:625-626). In the broadest terms, Hevner's findings on the affective characters of these parameters, as well as their relative strength, have been largely echoed in later research. These six parameters are fairly unambiguously *compositional* cues, at least as they were manipulated in Hevner's experiments (tempo especially may be said to overlap conceptually with the *performance* category, as performers rarely adhere strictly or consistently to the composer's intended tempo, where such is indicated). While imperfect, Hevner's minimisation of the interference from performance cues (by pursuing "identical" performances by the experimental pianist) shows a strong commitment to the investigation of *compositional* cues as defined within the boundaries of the musical score.

Direct and detailed comparisons using Hevner's research are difficult, as most later research has used different adjective categories to those she designed. A number of the adjectives she used define concepts that are not generally considered either discrete emotions or emotion dimensions (e.g. lyrical, dark, heavy) or that could be considered in terms of a discrete emotions paradigm, but only at a higher level, rather than as the basic-level emotions usually studied (e.g. frustrated, triumphant, mournful, as opposed to angry, happy, sad, etc.) (Hevner, 1936:249). This difficulty is by no means a cautionary tale about a problem now-fixed. As will be seen throughout the rest of this section, even within overarching theories of emotion, different

schools of thought may be found, resulting in experimental data which are often only partially comparable across even a modest number of studies.

Notably, while Hevner was interested in participants' selection of specific adjectives/adjective groups, the set of eight groups she developed has dimensional characteristics. Groups lying at opposite positions were designed, in essence, to form poles of an axis (Hevner, 1936:250). A quick assessment of the grouping (Hevner, 1936:249, see Fig. 1) reveals what may be interpreted as valence and arousal axes lying at the vertical and horizontal extremes of the arrangement, respectively. This will be discussed again at the end of this section, using Schubert's (1999) work as support for the dimensional re-interpretation of Hevner's results.

Since the discrete emotions paradigm began to be formalised around the 1960s, steps have been made towards studying a unified set of emotions, though discrepancies do remain. The uniqueness of Hevner's emotion categories can be attributed to the time of her research – taking place before a formal theory of discrete emotions was in common use. More recent research has focused on "basic" emotions. For instance, Juslin and Laukka (2003) examined five basic emotions: happiness, sadness, anger, fear, and tenderness.

These same five emotions (along with a neutral category) were used in a study by Quinto *et al.* (2014). This study explored the relation of acoustic parameters to emotion in terms of a fairly conservative definition of compositional and performance cues. The cues they deemed compositional were the mode, mean F0 (pitch height), average interval size, and range of a melody line (Quinto *et al*., 2014:514). They defined articulation, intensity (sound level), intensity variability, and high-frequency energy (a contributing factor to perceived brightness) as performance cues (2014:515). Tempo was defined as both a compositional and a performance cue (Quinto *et al*., 2014:505). Their experimental design involved eight musicians using three channels to convey emotion – composition, performance, and a combination of both. For the composition experiment, the musicians composed short melodies in accordance with a prescribed emotion. These melodies were played back digitally to listeners, to avoid the variability of performance cues in this channel. The performance experiment utilised melodies pre-composed by the researchers to be as emotionally ambiguous as possible. The musicians were then instructed to perform the melodies to convey a specific emotion (thus keeping the compositional cues constant). In the combined experiment, musicians performed the melodies they had written themselves, with the intention of conveying the intended emotion of the composition, maximising the cues available to them. The results of the three listening

experiments, as well as the compositions and performances used as stimuli, were subjected to principal components analysis in order to determine which cues were associated with the encoding and decoding of specific emotions (Quinto *et al*., 2014:503).

The results from Quinto *et al.*'s principal components analysis are included in table 4.2.1 in the next section (note that their neutral category has been omitted). Before giving these results, they reported the associations of the levels of individual cues with different emotions (also indicated in part in table 4.2.1). The major mode was associated with happiness and tenderness, the minor mode with anger, fear, and sadness. Low pitch height was associated with anger, while the other emotions showed association with higher pitch. Interval size and melodic range did not show significant differences per emotion. Legato articulation was associated with sadness and tenderness, while staccato articulation was associated with happiness, anger, and fear. High intensity was associated with anger and happiness, low intensity with sadness, fear, and tenderness. High intensity variability was associated with fear, sadness, tenderness, and happiness, while low intensity variability was associated with anger. High-frequency energy did not differ significantly per emotion. Fast tempo was associated with happiness, intermediate tempo with anger, and slow tempo with fear, sadness, and tenderness (Quinto et al, 2014: 514-516).

Quinto *et al.* report the efficacy of communication of different emotions across the three channels they investigated (composition, performance, combined). They found the following: anger was communicated more clearly using performance cues (as they were defined in this paper) than compositional cues, fear was communicated more effectively through compositional cues, happiness could be communicated through both (and so was most accurately decoded in the combined condition), sadness and tenderness had a heavier leaning on the performance cues, but also likely utilised compositional cues to some extent (Quinto et al, 2014:519-520). These findings have direct relevance to the present research, as certain emotions may be more communicable through certain cues than others, which must be taken into account in designing and interpreting the results of the experiment. This will be returned to in chapter 7, when discussing the cues used in musical analysis of the excerpts.

Scherer and Oshinksy (1977) report an experiment using synthesised stimuli with systematic variation in seven 2-level factors: amplitude variation, pitch level, pitch contour, pitch variation, tempo, envelope (attack/decay ratio: low or equal), and filtration cut-off level (number of harmonics). They also tested two other stimulus types, but only in an exploratory fashion, with

too small a number of stimuli to make confident generalisations of results. These latter types explored the effects of tonality and rhythm in combination with the previously mentioned factors (Scherer & Oshinsky, 1977:333). The cues chosen for study in these experiments represent both compositional and performance parameters. Pitch level, pitch contour, pitch variation, and tonality may be considered purely compositional cues. Envelope and number of harmonics are purely performance cues (unless one considers the overtone profile of different instruments, in which case instrumentation may be seen as a compositional level of the "number of harmonics" cue). Tempo and rhythm both overlap with composition and performance, as performers deviate from the score in terms of these parameters as a matter of course (depending on the level of detail at which analysis is made, one could consider rhythm to be a purely compositional cue in terms of broad rhythmic profiles). The authors did not make the distinction between compositional and performance cues in their paper, treating all the cues in the same manner.

This study is interesting for its combined testing over emotion dimensions and discrete emotion categories. Their results in terms of dimensions will be examined in the section on dimensional research below. The discrete emotions tested were anger, boredom, disgust, fear, happiness, sadness, and surprise, using a simple yes/no selection of appropriateness to the stimulus (Scherer & Oshinsky, 1977:334). Their findings are summarised in table 4.2.1 in the next section in terms of the emotion portrayals relevant to the present experiment, excluding envelope and filtration cut-off level, which cannot be applied to the present experimental context. The following results were reported (with associations listed in decreasing order of strength per cue). Small amplitude variation was associated with happiness, large amplitude variation with fear. Small pitch variation was associated with disgust, anger, fear, and boredom, large pitch variation with happiness and surprise. A downward pitch contour was associated with boredom, pleasantness, and sadness, an upward contour with fear, surprise, and anger. Low pitch level was associated with boredom and sadness, high pitch level with surprise, anger, and fear. Slow tempo was associated with sadness, boredom, and disgust, fast tempo with surprise, happiness, fear, and anger. A round envelope was associated with disgust, sadness, fear, and boredom, a sharp envelope with happiness and surprise. A low number of harmonics was associated with boredom, happiness, and sadness, a high number with anger, disgust, fear, and surprise (Scherer & Oshinsky, 1977:339).

Balkwill and Thompson (1999; as briefly discussed in the previous section) investigated Western listeners' use of cues to decode emotion in North Indian music. They investigated four

emotion portrayals: joy, sadness, anger, and peace, utilising five cues: tempo, rhythmic complexity, melodic complexity, pitch range, and timbre. Unusually for research in this paradigm, their quantification of all of the cues except timbre was based on listeners' reported perception of these cues, given in conjunction with their emotion ratings. Emotion ratings were given both by selecting from a list of the four emotion portrayals the emotion most appropriate to the piece, and by rating on a nine-point scale the degree to which each emotion was present (Balkwill & Thompson, 1999:51-52). Joy was found to be associated with fast tempo, low rhythmic complexity, and low melodic complexity; sadness with low tempo, high rhythmic complexity, high melodic complexity, and small pitch range; anger with the timbre of string instruments; and peace with slow tempo, low rhythmic complexity, low melodic complexity, and the timbre of the bansuri flute (Balkwill & Thompson, 1999:54-56). These findings are complicated by the perception measure used to quantify cue levels, and by the unequal distribution of timbres within emotion portrayals (Balkwill & Thompson, 1999:51). However, the results of this study are in line with the other data discussed above, at least for tempo (the other cues were not examined in the studies already mentioned).

4.2 Dimensional research

Schubert (2004) investigated the role of acoustic cues in emotion judgements through a dimensional framework. Using a 2-dimensional emotion space like that of Russell (1980), the contributions of acoustic cues to the dimensions of arousal and valence were recorded. Schubert utilised a continuous sampling methodology, using a computer program which represented the emotion space visually, with participants moving a cursor to indicate their emotion judgement in real time (Schubert, 2004:564-565). The granularity of sampling was one sample per second (Schubert, 2004:561). This continuous response methodology was utilised to facilitate testing of the lag between presentation of an acoustic cue and the manifestation of its effect on emotion judgement (Schubert, 2004:561). The present research does not investigate the time-scale of emotional responses to acoustic cues, and so this aspect of Schubert's research will not be discussed here. The cues Schubert investigated were melodic contour, tempo, loudness, texture, and sharpness of timbre (measured as spectral centroid) (Schubert, 2004:566, 571).

A modified linear regression model was used to associate changes in cue with changes in reported emotional judgement (Schubert, 2004:572). Over the four pieces used, it was found that loudness and tempo correlated positively to arousal ratings, with loudness having the

strongest effect. A positive correlation was found between arousal and melodic contour, as well as an ambiguous correlation with texture and timbral brightness (one positive, one negative correlation for each); however, the author states that these effects are likely due to musical context rather than the cue itself, specifically a conflation of the extrapolated effects with concurrent effects of loudness. Loudness, tempo, and melodic contour were found to correlate positively with valence, with loudness again being a stronger contributor than tempo. A further positive correlation between valence and texture was found, though the author cautions that it is likely that harmony (which was not examined in this experiment) played some role in this correlation. Schubert also notes that there may be an interaction effect whereby the articulation, mode, and tempo of a piece affects whether changes in tempo, texture, and loudness manifest effects in the arousal or the valence dimension. These interactions are only suggested as a potential explanation for observed effects – they were not measured statistically (Schubert, 2004:572-579).

The dimensional results from the paper by Scherer and Oshisnky (1977, discussed above) were presented in terms of three dimensions: pleasantness, activity, and potency. These were rated on 10-point bipolar scales (Scherer and Oshinsky, 1977:334). Their results, again in decreasing order of strength per cue, were as follows. Small amplitude variation was associated with pleasantness and activity. Likewise, large pitch variation was associated with pleasantness and activity. A downward pitch contour was associated with pleasantness, an upward contour with potency. Low pitch level was associated with pleasantness, high pitch level with potency and activity. Fast tempo was associated with activity, pleasantness, and potency. A round envelope was associated with potency, a sharp envelope with pleasantness and activity. A low number of harmonics was associated with pleasantness, a high number with potency and activity (Scherer & Oshinsky, 1977:339). In their second experiment, they found preliminary evidence of an association between the major mode and pleasantness, and of rhythmic tone sequences with activity (Scherer & Oshinsky, 1977:341). Their operationalisation of rhythm is unclear, and their results for this cue will not be included in Tables 1 and 2.

The relationship of tempo and mode to the valence dimension of emotion was further examined by Gagnon and Peretz (2003). Using simple melodic stimuli over a range of three tempi and three modalities (major, minor, whole-tone/neutral), they performed two experiments to test the influences of these factors individually and in combination. They found major mode and fast tempo to be indicative of positive valence. Tempo was found to be more powerfully predictive of emotion attribution than mode when the two cues were manipulated in

combination (Gagnon & Peretz, 2003:32). This finding was replicated using less extreme tempo values – reducing the range of tempo values did reduce the predictive power of tempo, but it remained higher than that of mode (Gagnon & Peretz, 2003:36).

Ilie and Thompson (2006) studied the influence of intensity, rate of presentation (tempo), and pitch height on emotional expression in music and speech. Their combined findings will be used below in comparing the patterns of cue use in emotional expression across both music and other modalities. For now, their findings for music will be examined. Their study was based on a 3-dimensional model of emotion consisting of valence, energy arousal, and tension arousal (Ilie & Thompson, 2006:320). Data were collected from 27 listeners using a 5-point bipolar rating scale for each dimension (Ilie & Thompson, 2006:322). Because the effects they found were always in the same direction for the two arousal dimensions, or featured an insignificant effect on one of the two dimensions (Ilie & Thompson, 2006:324), these two dimensions will be treated as a single arousal dimension in the present analysis. Intensity was found to be negatively correlated with valence, and positively correlated with arousal. Tempo had no significant effect on valence, and correlated positively with arousal. Pitch height was negatively correlated with valence, and positively correlated with arousal (Ilie & Thompson, 2006:324). Additionally, Ilie and Thompson found interaction effects among the cues they studied. Intensity had a greater effect on valence under the condition of high pitch than low pitch. Fast tempo led to higher ratings of valence for low than for high pitch; no significant effect of pitch height was found under conditions of slow tempo (Ilie & Thompson, 2006:323).

Tables 4.2.1 and 4.2.2 present a summary of the cues studied in the research discussed above. Only cues relevant to the present research are included; i.e. compositional cues, excluding timbre. This provides a useful point of reference in selecting cues for study in the present research. In order to tie in to the existing literature, only cues that have been studied previously will be taken as eligible for use here. Two further considerations will be used in making the final selection. First, the cues to be studied must have shown systematic relation in the literature to all three of the emotion portrayals embodied in the excerpts chosen for the experiment (happiness/joy, anger/fear, and peacefulness/tenderness). Second, the excerpts must be clearly differentiable in terms of the chosen cues. This aspect of the selection process is discussed in chapter 7.

	Juslin & Laukka 2003	Quinto *et al.* 2014	Scherer & Oshinsky 1977	Balkwill & Thompson 1999
Tempo	X (A+, H+, P-)	X (A=, H+, P-)	X (A+, H+, /)	X (/, H+, P-)
Intensity	X (A+, H=, P-)	X (A+, H+, P-)	/	/
Intensity variability	X (A+, H-, P-)	X (A-, H+/=, P+)	X (/, H=, /)	/
Pitch height	/	X (A-, H+, P+)	X (A+, /,/)	/
Pitch contour	X (A+, H+, P-)	/	X (A+, /, /)	/
Mode	/	X (A-, H+, P+)	X* (A-, H+, /)	/
Interval size	/	X (/,/,/)	X* (A-, H+, /)	/
Range	/	X (/,/,/)	/	X (/, /, /)
Rhythmic complexity	/	/	N/A	X (/, H-, P-)
Melodic complexity	/	/	/	X (/, H-, P-)

Table 4.2.1. Summary of findings of discrete emotion studies for the three musical emotion portrayals relevant to the present research's experimental component. An "X" represents the inclusion of cue in a study. The brackets next to the "X" indicate for which of the three emotion portrayals being considered in this research significant associations were found. These portrayals are happiness/joy (H), anger/fear (A), and peacefulness/tenderness (P). The "+" or "-" signs indicate the direction of the effect (e.g. fast tempo (+) is an indication of anger, but slow tempo (-) is an indication of peacefulness). An "=" sign indicates an effect of the cue at a moderate level, where such was included in the study. Note: for pitch contour, "+" represents up, "-" represents down. For mode, "+" represents major, "-" represents minor. "X" in the mode and interval size rows for the findings of Scherer & Oshinsky (1977:341) is due to the low number of stimuli used to demonstrate this effect – the authors remark that further replication studies are needed to establish the effect they found.*

The findings reported in table 4.2.1 generate patterns of cue use per emotion label. Where an equal number of effects are found in opposite directions, no prediction is made.

Happiness/joy: fast tempo, high/moderate intensity, moderate intensity variability, high pitch, ascending pitch contour, major mode, large interval size, low rhythmic complexity, low melodic complexity.

Fear/anger: fast/moderate tempo, high intensity, ascending pitch contour, minor mode, small interval size.

Peacefulness/tenderness: slow tempo, low intensity, high pitch, descending pitch contour, major mode, low rhythmic complexity, low melodic complexity.

The results of dimensional studies are displayed in table 4.2.2. A "+" represents a positive correlation between the cue and the dimension (e.g. increasing tempo increases arousal), and a "-" represents a negative correlation (e.g. decreasing pitch height increases valence).

	Scherer & Oshinsky 1977		Schubert 2004		Gagnon & Peretz 2003		Ilie & Thompson 2006		Hevner 1935b, 1936, 1937	
	V	A	V	A	V	A	V	A	V	A
Tempo	+	+	+	+	+			+	+	+
Intensity			+	+			-	+		
Intensity variability	-	-								
Pitch variation	+	+								
Pitch height	-	+					-		+	-
Pitch contour	-		+	+						-
Mode					+				+	
Texture			+	+/-						
Rhythm	N/A	N/A							-	+
Harmonic complexity									-	+

Table 4.2.2. Summary of findings of dimensional studies for the three musical emotion portrayals relevant to the experimental component of the present research. Note: for pitch contour and mode, indications are as in table 4.2.1. For texture, "+" represents a thick texture, "-" a thin texture. For rhythm, "+" represents a firm rhythm, "-" a flowing rhythm.

Hevner's experimental results were parsed into a dimensional model by Schubert (1999:55). Two pairs (4) of Hevner's eight adjective groups are used to ground poles of the VA axes, with the other groups falling in the spaces between. The valence axis is anchored by the happy/bright and sad/heavy groups, the arousal axis by the majestic/vigorous and serene/lyrical groups. As this arrangement does not represent a perfect mapping (Schubert, 1999:57), these results are included in table 4.2.2 purely for comparative purposes. Hevner's findings are only listed where differential effects were found for the two groups grounding opposite poles of a dimension.

The VA space may be split into four quadrants, each bordered by the two axes and representative of a unique combination of dichotomous levels valence and arousal.

Happiness/joy is defined as occupying quadrant 1 (high valence, high arousal), anger/fear as occupying quadrant 2 (low valence, high arousal), and peacefulness/tenderness as occupying quadrant 4 (high valence, low arousal). From this, table 4.2.2 can be used to further infer cue profiles for each label. Hevner's results were not included in forming the cue profiles. Cues in brackets represent findings that suggest an effect on only one dimension, but in the appropriate direction for the label.

Happiness/joy: fast tempo, large interval size, ascending pitch contour, thick texture, high intensity, low intensity variability, (major mode), (low pitch height), (descending pitch contour), (thin texture).
Fear/anger: high pitch, fast tempo, thin texture, (high intensity), (ascending pitch contour), (minor mode), (high/low intensity variability), (large/small interval size).
Peacefulness/tenderness: low pitch, slow tempo, thick texture, (low intensity), (descending pitch contour), (major mode), (high/low intensity variability), (large/small interval size).

The emotion labels of anger/fear and peacefulness/tenderness are defined by contrasting levels over the two axes (low valence and high arousal, and high valence and low arousal, respectively). Where results from the studies in table 4.2.2 show an equal balance of effects in similar directions over the axes (e.g. a positive correlation with both valence and arousal), these are included in the profiles for these two emotions. This is indicated by giving both dichotomous options in the profile (e.g. large/small interval size). The results from the present experiment can be used in comparison with these mixed predictions to provide support for the effect direction on one axis or the other.

CHAPTER 5. BAROQUE CONSIDERATIONS

This chapter serves to contextualise the Baroque excerpts chosen for the experimental component of the present research, as well as to provide some background to 17th- and 18th-century conceptions of music and affect. First, Descartes' theory of the passions will be covered in more detail, and its correspondence with certain elements of modern theories of emotion noted. Second, the "Doctrine of the Affections" will be examined. Third, Johann David Heinichen's treatise *Der General-Bas in der Composition* will be discussed, as it is the source from which Baroque excerpts have been drawn for the present experiment. This will be followed immediately by a brief description of the study by Vieillard *et al.* (2008) from which the modern excerpts were drawn.

5.1 Descartes and the passions

René Descartes' (1596-1650) influence on 17th-century thought was profound, and echoes of his ideas have resonated through the ages, in some cases even remaining with us today. His psychology, which includes a well-developed theory of the passions, was widely accepted by his contemporaries and immediate successors. The foundation of his psychological theorising is the notion of *dualism*: the irreducibly distinct nature of the body and the mind/soul, and their union in the human being (Duncan, 2000:488). His theory of the passions, laid out in *The Passions of the Soul* (1649), takes as its basis this dualist framework.

On reading Descartes' treatise, a great many similarities with the modern conception of emotions, especially the functionalist conception, become apparent. Throughout the following discussion, these similarities will be pointed out and discussed in brief. They include: the notion of basic emotions, the utility of emotions in promoting the survival and wellbeing of the individual, the concept of action tendencies, cognitive appraisal, and component processes in general, and the hierarchical division of emotions into species.

Some foundational and terminological points must be established before laying out precisely what Descartes thought the passions were. His definition rests on an understanding of the following: dualism, the mechanics of bodily motion and physiology, the functional nature of the soul, the nature of perceptions, and the manner of mediation between body and soul. These points will be outlined with as much brevity as possible, leading up to a definition of the passions. After this, the nature of specific passions will be discussed. All information is drawn from the primary source (*The Passions of the Soul*), as translated by Robert Stoothoff (1985).

Dualism, to Descartes, was a means of reconciling the lawful nature of physical reality with the dogma of human freedom. To this end, he ascribes as the sole domain of the human *body* any human phenomena that may come about in inanimate objects (Decartes, 1985:329-§3). This neatly circumscribes the entirety of the known laws of physics and physiology at the time, parsing them into one element of the dualist model. The *soul*, conversely, is to be regarded as the source of all other phenomena, which cannot be attributed to inanimate objects (1985:329). The only phenomenon which Descartes believes is necessarily attributed to the soul is thought, of which he delineates two kinds: actions (volitions), and passions (perceptions; 1985:335-§17).

Volitions and perceptions may both arise in two forms, having either the body or the soul as their object. The role of the object in *volitions* is being a point of termination: a volition may be restricted in its impact entirely to the soul (Descartes gives the example of the will to love God), or it may transfer its impact to the body, as in any number of purposeful bodily movements (Descartes, 1985:335-§18). As regards *perception*, the object is a causal one. A perception may be caused entirely by the soul, as in perceiving a volition or a willed thought (1985:335-§19). Importantly, this idea is expanded upon to include all perceptions for which a proximate cause is not known; Descartes includes among such perceptions "the feelings of joy, anger and the like" (1985:337-§25). Perceptions causally attributable to the body are furthermore divisible into two kinds. One kind refers to the perception of internal bodily sensations, such as hunger and pain (1985:337-§24). The other is also a function of the body; however, its cause may be further referred outward to an external object. This object causes a reaction in at least one sense organ of the human body, and this reaction travels through the nerves to cause changes in the state of the brain, which translate into a sensory perception by the soul of the external object (1985:337-§23). The manner in which a physical phenomenon may make an impression upon the soul (and vice versa) will be discussed below; for now it will suffice to say that this relationship is immediate and correlative.

Descartes' conception of bodily motion and sensation relies on a medium which he calls the animal spirits (*esprits animaux*). These are "a certain very fine air or wind" (1985:330-§7), made up of a large number of miniscule individual bodies. The animal spirits constitute the motive element of physical movement. They travel to all parts of the body – they are transported by the blood to the brain, and from there by the nerves into the muscles and organs (1985:331-332-§10). By flowing into a muscle and out of its opposed muscle, the animal spirits may cause it to contract (1985:332-§11). A combination of such individual contractions can constitute any range of bodily movements. The animal spirits also facilitate sensory perception,

by suspending nerve fibres in such a state that any motion on the nerve ending will reflect upon the nerve's origination point in the brain (1985: 332-333-§12), a prerequisite for the formation in the soul of a perception of sensation.

The soul and the body, for Descartes, constitute two parts of a totally unified whole. The mechanism of this unification is proposed to be the pineal gland, or as Descartes puts it, "a little gland in the brain where the soul exercises its functions more particularly than in the other parts of the body" (1985:340-§31). He justifies this proposition by stating that all human sense organs come in pairs, as do all parts of the brain except the pineal gland. As the soul receives a *single* impression from any pair of sense organs, there must exist a place where the duplicate impressions generated by these organs are reduced (1985:340-§32). As the pineal gland is the only eligible bodily element Descartes believes to be singular, it is the obvious (and only viable) candidate for this role. The pineal gland is proposed to be suspended within brain cavities filled with the animal spirits. The spirits can exert their influence on the gland in any way, which will reflect as a perception in the soul. The soul itself can likewise exert an influence directly on the pineal gland, in line with any volition. This movement of the gland caused by the soul will excite the animal spirits to motion in line with the nature of the volition, moving them through the brain's pores, and from there via the nerves to the muscles, to induce any manner of motion desired (1985:341-§34).

Enough groundwork has now been laid to discuss the passion themselves. Descartes provides the following definition of the phenomenon which is the subject of his treatise: "the passions of the soul… [are] those perceptions, sensations or emotions of the soul which we refer particularly to it [the soul], and which are caused, maintained and strengthened by some movement of the spirits" (1985:338-339-§27). This broad definition essentially captures all phenomena of the soul which are not volitions, or sensations with a physical object; in effect, all passive phenomena of the soul which have the soul as their object. They arise through the body, and are reflected on the soul via the pineal gland. The animal spirits constitute the physical medium for their existence at the embodied level. The passions are "[referred] particularly to [the soul]" to differentiate them from other mere sensations. The sensations which give rise to the passions are distinct from the passions themselves – these former are referred to the body. The passions are *caused* by the body, but are only meaningfully called passions at the experiential level – the level which is referred to the soul (1985:339-§28-29).

In the second part of his treatise, Descartes lists and describes a number of passions. As an example, he describes love and hatred as follows: "… when our esteem or contempt is directed upon some other object that we regard as a free cause capable of doing good and evil, esteem becomes veneration and simple contempt becomes scorn" (1985:350-§55). The modern notion of cognitive appraisal is paralleled in this description – there is a clear evaluative element to these emotions thus described. Specifically, the dimension of agency/responsibility, which is a widely used appraisal criterion in modern research, is evident here (van Reekum & Scherer, 1997:259-260). Good and evil play a crucial role in Descartes' theory of the passions, and are intimately linked to the concept of agency. It is consistent with his overall ethics that the nature of a passion should change depending on whether or not its object is considered to be a free agent; he usually applies this criterion of agency to the self. He states to this effect that as long as a person never fails to pursue virtue, the passions which may impinge upon their daily life for external reasons (outside of their control as a free agent) will pale in comparison to the surety of purpose related by the soul to itself (Descartes, 1985:381-382-§147-148).

While many passions are listed in Descartes' treatise, he makes a distinction between these and the "simple and primitive" passions, of which he states there are precisely six. These are wonder, love, hatred, desire, joy, and sadness (1985:353-§69). As discussed earlier in this literature review, the idea of basic/primary emotions is crucial to the theory of discrete emotions. Some degree of overlap can be seen between Descartes' primary emotions and those commonly assumed in modern functionalist theories. Recall, for instance, the five emotions listed by Sloboda and Juslin (2001:77) as being held in general consensus: happiness (joy), sadness, anger (hatred), fear, and disgust. Of these, the first three correspond to three of Descartes' proposed primary emotions. His list bears even more similarity to Shaver *et al.*'s (1987:1067-1068) list of conceptually basic emotions: anger (hatred), sadness, love, fear, and joy. Of these, only fear does not appear in Descartes' list.

A further similarity between Descartes' emotional typology and that of modern discrete emotion theory is found in the idea of emotion blending. This is a common explanation given by discrete emotion theorists for the existence of complex emotional states (Scarantino & Griffiths, 2011:451). Descartes gives two explanations for the multiplicity of emotional states evident in human life, the first of which is that all non-primitive emotions are "composed from some of these six" (Descartes, 1985:353-§69).

The second explanation Descartes offers for the non-primitive emotional states is that they may be species of the primary emotions (1985:353-§69). This is stated explicitly: "…the six primitive passions […] are, as it were, the genera of which all the others are species" (1985:383-§149). He makes use of this idea when discussing, for example, regret and cheerfulness: "… a past good [meaning a good thing which has departed] gives rise to regret, which is a kind of sadness; and a past evil [an evil which has departed] gives rise to cheerfulness, which is a kind of joy" (1985:352-§67). This definition works in a very similar fashion to the concept of emotion species used by Shaver *et al.* (1987:1067-1069), in which the species of a basic emotion may be determined by a cognitive appraisal of the context of its elicitation.

Descartes provides a detailed discussion of his six proposed basic emotions. One of the points that he discusses is the set of bodily effects that constitute part of the emotion. The parallel here with component process theory is striking. He describes physiological effects, such as pulse rate, strength, and regularity, feelings of heat in the chest and body, and changes in digestive functioning (1985:363-§97-101). Facial expressions are linked to emotions, with particular focus on the region around and including the eyes: "[t]here is no passion which some particular expression of the eyes does not reveal" (1985:367-368-§113). Finally, Descartes speaks of the association of certain passions with corresponding action tendencies. For instance, in discussing repulsion, he states that this passion is "ordained by nature to represent to the soul a sudden and unexpected death" (Descartes, 1985:359-§89). He argues that in repulsion, an unexpected sensation "produces a sudden agitation which leads the soul to do its utmost to avoid so manifest an evil [the perceived threatening sensation]". He states that this desire to avoid certain death may also be called "avoidance" (1985:360-§89). Taken in sum, Descartes has now discussed the passions in terms of action tendencies, motor expression, physiological arousal, cognitive appraisal, and subjective feeling (this last is inherent is his definition of a passion as being a phenomenon of the soul, as the soul is the seat of experience). These are the five component processes most commonly studied in modern emotion research (Scherer, 2004:240).

Another point addressed by Descartes is the manner in which the passions may be useful, as well as harmful, for the person in whom they manifest. For Descartes, the passions are useful in terms of aiding the soul in preserving "good" thoughts. They are harmful insofar as they perform this function with too great an intensity or for too long a period, or when the thoughts they facilitate are not good ones (Descartes, 1985:354-§74). He uses as an example the passion of wonder. The utility of wonder, he states, is removing ignorance by facilitating learning and

retention in memory of new things, which would not be guaranteed without the help of this passion (1985:354-355-§75). While there are some differences, this approach of assessing the utility of emotions for the individual foreshadows in spirit the framework of adaptive value employed in the functionalist conception of emotions (e.g. Juslin, 1997a).

The precise manoeuvring of the animal spirits in the body in different passions is detailed at some length. Descartes theorises patterns of movement for the spirits, which always conform to a similar procedure. First, the spirits are excited to directed motion by the object of the passion. They move through the body to specific muscles and organs, and perform certain functions at these locations. The manner of interaction of the spirits with these parts of the body has an effect on the quality of the blood that flows to the heart. This quality has an effect on the "heat" of the heart, which is considered essential to life (Descartes, 1985:366-§107). This effect of the blood on the "heat" is mirrored on the blood being pumped through the heart's chambers. The quality of this blood dictates the quality of the animal spirits as they flow up from it to the brain, and thereby affect the ideas present in the brain (1985:364-365-§102-106).

Having outlined in broad terms the principles of the passions according to Descartes, later developments which were indebted to his work can now be discussed. Descartes himself died in 1650, shortly after the publication of *The Passions of the Soul*. His work was foundational in musical thinking during the Baroque era. *The Passions of the Soul* was of great relevance to musicians interesting in portraying affect through music, as it provided a psychological groundwork upon which to base their theories and experimentations (e.g. Mattheson, 1981). These theories, under the broad title of the Doctrine of the Affections, will now be discussed.

5.2 The Doctrine of the Affections

The Doctrine of the Affections is a nebulous body of ideas, whose specifics are difficult to pin down. In general terms, it was an approach to music that promoted the use of rhetorical techniques to move the affections of listeners. The concept of the affections was based on Cartesian philosophy, and its scope was determined by the mechanics of this philosophy (Whitfield, 2010:16-17). It is often asserted that Descartes believed that only a single passion could hold sway in a person at one time, with rapid change also being unlikely (LeCoat, 1972:220), and that for this reason a single musical movement was crafted to embody only a single affection (Whitfield, 2010:17). It is unclear why this is asserted; Descartes explicitly states that several passions can coexist simultaneously. In a concluding section for a discussion of the primitive origins of the passions *in utero*, he writes: "For each passion, however, I have

noted solely what can be observed when it is the only one present […] I have yet to deal with the many external signs which usually accompany the passions – signs which are much better observed when several [passions] are mingled together, as they normally are, than when they are separated" (Descartes, 1985:367-§112). Regardless, it is certain that Descartes' influence was strong in the minds of musicians and theorists of the Baroque (see e.g. Mattheson, 1981:104).

The Cartesian model was not the only one informing the Doctrine, or at least the approach to affective music-making of Baroque musicians. This area of musicianship has a long history, with roots at least as far back as Classical Greece (Dissmore, 2017:2). Some of these influences will be mentioned below, alongside those of Descartes.

In addition to the overarching imperative of affective content, the Doctrine of the Affections encompasses technical devices by which composers could imbue their music with the necessary affections. These were grounded in Descartes' work on music, specifically his *Compendium Musicae* (Compendium of Music) of 1618 (Vlock, 1998:521). His theoretical writings do not enter into great depth on technical procedures, but some are given, specifically for the parameters of pitch and time (Dissmore, 2017:1). He states, for example, that tempo is proportional to the "speed" of affect – slower tempo giving rise to "quieter" affects (e.g. sadness), and fast tempo to "faster" affects (e.g. joy) (Vlock, 1998:524). Of intervals, Descartes wrote that wider intervals should be put to use for happy emotions, and narrow intervals for sadder ones (Vlock, 1998:521). It is intriguing to notice the similarities between these simple and broad statements of Descartes and the findings of modern empirical studies (e.g. Juslin & Laukka, 2003:802; Scherer & Oshinksy, 1977:334).

The key of a piece of music was also seen to contribute to its affective potential. Jean-Philippe Rameau (1683-1764) contributed notably in this area. He offered recommendations of keys suited to expressing among others the affects of joy, fury, tenderness, and melancholy (Dissmore, 2017:7). There was no unanimously agreed-upon set of affective qualities attributed to keys; various opinions abounded among theorists and composers (Hall, 2017:57). It is worth noting that in the Baroque era, different keys contained intervals of different qualities, as a function of the system of non-equal temperament used at the time. In modern, equal-tempered music, all keys contain intervals of precisely the same quality.

It is not entirely clear to what extent the "affections" or affects which formed the object of the Doctrine of the Affections resemble our modern conception of emotion, in either the folk or

the scientific sense of the term. Whitfield (2010:17) cites Lorenzo Giocomini's definition: "[an affection is] a spiritual movement of operation of the mind in which it is attracted to or repelled by an object it has come to know". The attraction/repulsion dichotomy brings to mind the paradigmatic cases of approach and avoidance in the conception of action tendencies (e.g. Frijda *et al.*, 1989:214), a commonly studied component process of emotion (Scherer, 2004:240). The affects may also be viewed as "rationalized states of passion". This rationalisation refers to the systematic study to which the affects as phenomena were subjected (Whitfield, 2010:17).

The nature of the affects is intimately tied to the concept of musical expression. Vincenzo Galilei encourages musicians to inform their expression by observing human behaviour in specific contexts. For instance, he mentions the differences in expression embodied in the address of a prince to his subjects, a man reacting in anger, the act of lament, the mannerisms of timidity, and exultant joy (LeCoat, 1972:221). Some of these contexts bear direct relation to emotions, in the modern conception – anger, lament (sadness), timidity (fear), and joy. The example of the prince can easily be read as pride.

Galilei further recommends specific cues to observe in the contexts he mentions. Ascending or descending tone of voice, loudness of speech, and speech rate are all mentioned (LeCoat, 1972:221). The contexts noted by Galilei (which can easily be read as emotional) together with the cues he singles out as noteworthy (which are among the acoustic cues studied in modern empirical research on vocal and musical emotional expression e.g. Juslin & Laukka, 2003:802 – reprinted in figure 4.1.1, chapter 4), can lead to a reading of his conception of affect as being at least broadly similar to modern definitions of emotion.

The influence of the Doctrine of the Affects was evident to a great extent in vocal music. During the Baroque era, music was largely considered to be subservient to written text, aiming to express the affective content of the text and thereby further enhance it (LeCoat, 1972:221). An instrumental work or movement, lacking text from which to draw affective content, took on the task of meditating on a single affection (Wilson, 1989:101). The entirety of affective expression in instrumental music had to be accomplished through musical-expressive devices. Composers were able to draw on general guides to such devices, such as the work of Johann Mattheson (which was based on Cartesian ideas). Composers exercised their own discretion as well as following the prescriptions of theorists, leaving space for individual styles and methods (Dissmore, 2017:5-6).

The principal purpose of a composition informed by the Doctrine of the Affections was to move its listeners to states of passion. This represents a shift in the orientation of the music; whereas in the preceding centuries it had taken its text as object, the listener now became the object (Hall, 2017:55). In line with Descartes' conception of the passions as arising through the motions of the animal spirits within the body, the approach taken to moving the listener was entirely mechanistic. The listener was to react directly and physically to the sonic stimulus, rather than coming to an emotional state through the use of their conscious faculties (Hall, 2017:54-55). For Descartes, a passion could easily be aroused in such a fashion by an external object's action upon the sensory organs (Descartes, 1985:332-333-§12), and music took on the role of this object. In order to move the listener to an affection, the appropriate affection had also to be *expressed* in the music (Lenneberg, 1958:47).

The Doctrine of the Affections may be said to have been the dominant musical ideology of the Baroque, but the specific praxis inspired by it was by no means universal. Various theorists and composers wrote about their particular methods, notably Johann Mattheson (1681-1764) and Johann David Heinichen (1683-1729) (Whitfield, 2010:17). The latter will be discussed in some depth here, as it is from his treatise, *Der General-Bass in der Composition* (The Thorough-Bass in Composition, published 1728), that the musical excerpts used to represent the category of Baroque in the present experiment are drawn.

5.3 Selection of Baroque excerpts – Heinichen's *Der General-Bass in der Composition*

Finding suitable Baroque excerpts proved to be a greater challenge than expected. The name most commonly associated with Baroque theories of affect (or the floating signifier, "the Doctrine of the Affections") is Johann Mattheson, who formed the first point of inquiry. Unfortunately, Mattheson does not in fact give concrete musical examples of his theories. He does, however, make mention of such examples, from Heinichen's *Der General-Bass in der Composition*, specifically in the *Einleitung* (Introduction) to this treatise. This set of examples was chosen as the pool from which to draw the second class of experimental stimuli, for two reasons. First, the present author was unable to find any clearer examples in the literature which could be used instead. Second, drawing examples from multiple sources would necessitate a comparison of rating accuracy between theorists and detailed discussion of their minor differences, which is outside the scope of the present research. It is assumed that, despite there being no universal application of the Doctrine of the Affections, at least in terms of practical

compositional rules, any one theorist's ideas will have enough broad overlap with those of other theorists to justify a working generalisation.

Heinichen's name is little-known today outside the context of his theoretical writings. In his own day, he was highly regarded as a composer, receiving lofty praise from such figures as Johann Scheibe, Charles Burney, and Johann Mattheson. He was the Capellmeister at the Dresden court, on top of his compositional and theoretical endeavours (Buelow, 1986a:1). The best-known of his theoretical writings is *Der General-Bas in der Composition*, a technical manual expounding his view on the fundamentals of thorough-bass (basso continuo) realisation. The main body of this treatise, while no doubt a valuable tool for continuo players, is not relevant to the present discussion. In the *Introduction* to this treatise, however, Heinichen delves deeply into the manner of portraying the affects in music.

Heinichen's *Introduction* contains the most extensive collection of practical examples of affective portrayal in Baroque music-theoretical writing (Buelow, 1986b:307). The demonstration of affective compositional techniques therein is grounded in a discussion of the *loci topici*, a fundamental tool of rhetorical invention which Heinichen applies to music in the context of text-setting (Heinichen, 1986:330). He demonstrates in detail the manner in which a composer may draw inspiration for affective composition from a text, be it rich in expression or utterly bland. Most importantly in the present context, he provides concrete musical examples of the affective portrayals he describes, a surprising rarity among Baroque writers on musical affect.

The examples Heinichen provides are related to specific texts, which he gives in conjunction with them. He outlines the application of the *loci topici* to these texts, drawing from them affective ideas suitable for musical portrayal (Heinichen, 1986:332). The examples he gives are short, being only fragments of hypothetical arias based on the corresponding texts. They bear the hallmarks of a theorist well-versed in composing such arias. In total, he adduces fifteen such examples (Heinichen, 1986:333-375). Each example is prefaced by a brief description of the expressive intention of the excerpt. These prefaces formed the basis for the pre-selection of candidate Baroque excerpts for the present experiment. Unfortunately, there is not a clear and singular affective label attached to each piece; rather, Heinichen discusses its expressive intention at length. These prefaces were examined in depth, and only those from which affective words or phrases could be drawn with some confidence had their corresponding excerpt selected as a candidate stimulus. An example of one such description is as follows:

> *This text, beautiful in itself, on first inspiration creates in our imaginations six, eight, or more entirely different inventions, if one will give some consideration to the words and circumstances or, as I would say, the* ***loci topici****. In general, one could seize upon the tenderness of the affection; and with this in mind, a* ***Siciliana*** *(a form of composition willingly expressing languid thoughts) might suggest the following invention among others…* (Heinichen, 1986:356, emphasis as italics in original)

From this description, the relevant affective word is *tenderness*. Heinichen follows this description directly with the musical example, a 10-bar Siciliana written for continuo (unfigured), unspecified melody instrument, and voice. This was one of the excerpts considered for inclusion in the experiment; however, it was excluded on other grounds later in the selection process. Heinichen present texts which were interpreted in terms of the following affective labels: fury, pomposity, love, playfulness, tenderness, and anxiety. In stating these ideas, a vast compound problem presents itself – namely, the problem of meaning.

Heinichen's treatise was originally written in German. All translations necessarily suffer from inaccuracies, usually due to the lack of direct and unambiguous one-to-one correspondences between words across language boundaries. Alongside this problem, the question of historical terminology comes into play – how are we to know that a word, even in the same language, has the same meaning (denotative and/or connotative) now as it did in 1728? These questions cannot be fully addressed in the present research. A pragmatic approach will be applied to circumvent this problem. The discrepancies in terminological boundaries across languages and centuries may be greatly minimised by adopting a *fuzzy* conception of the terminology in question. That is to say, a word such as "love" will not be taken solely to mean the strict English dictionary denotation of love. Instead, it will be redefined to incorporate the semantic space surrounding this denotative core; rather than meaning one highly specific thing, the term will be used as a marker for a range of affective concepts that share elements of the core. Thus, "love" will no longer represent a *point* on a semantic space, but rather an *area* (as discussed in the previous chapter). It is hoped that this conceptualisation will limit the impact of translation across languages and times; the meaning of a word may well be expected to shift under these transformations, but it is likely to remain within a broadly conceived semantic space. This approach is not perfect. In the case that the experimental results show a mismatch in affect ratings for the Baroque excerpts, the issue of trans-linguistic and -historical meaning will

contend as possible explanations. If the experimental results show a good match with the terms, then it is likely that these issues did not manifest significantly in this specific case.

5.4 Selection of modern excerpts – Vieillard *et al.* (2008)

The modern excerpts for use in the present experiment were drawn from Vieillard *et al.*'s (2008) paper, *Happy, sad, scary and peaceful musical excerpts for research on emotions*. They designed a set of 56 musical excerpts, with fourteen for each of their four emotion categories. A detailed list of structural consideration that were used in composing the excerpts is not provided. It is stated in the paper that the happy excerpts were composed with a fast tempo, in a major mode; sad excerpts with a slow tempo in a minor mode, scary excerpts with an intermediate tempo in a minor mode, and peaceful excerpts with a slow tempo in a major mode (Vieillard et al., 2008:723-724). Their expectations for valence and arousal ratings were broadly in line with those provided in table 4.2.2 above. However, they hypothesised that sadness may be judged as pleasant rather than unpleasant, placing it in quadrant 4 rather than 3. While sadness proper belongs in quadrant 3, musical portrayals of sadness may be perceived as melancholic, giving rise to a higher valence rating (Vieillard et al., 2008:723-724).

Three experiments are reported in the paper, of which only the first is relevant here. This was an experiment utilising listener ratings, to establish the clarity of portrayal of the intended emotion, as well as valence and arousal ratings, for each excerpt. The second and third experiments were to determine the length of time needed for confident and accurate recognition of emotion, and to further examine emotional portrayals without verbal labels, respectively (Vieillard et al., 2008:720).

Three groups of listeners performed the rating task, each under different instruction conditions. One group (N=20) rated each of the excerpts on 10-point scales for each of the four emotions (happy, sad, scary, peaceful – listeners were allowed to make ratings on as many scales as they wished). This group was instructed to rate the excerpts based on the emotions they recognised in the music. The second group (N=19) performed the same task, but were instructed to make ratings according to the emotions they experienced while listening to each excerpt. The third group (N=20) rated each excerpt for valence and arousal. The instructions listed for this last group are unclear in terms of whether listeners were expected to make ratings of perceived or felt emotion (Vieillard et al., 2008:725-726).

To assess the clarity of the emotion portrayal for each excerpt, the researchers examined the level of overlap between 90% confidence intervals around the mean rating of each excerpt on

each emotion rating scale. An excerpt was classified as having a clear emotional portrayal if the mean on one emotion rating was higher than all of the others, and there was no overlap between the confidence intervals for that emotion rating and for the others. They found that all the excerpts achieved clarity according to this measure (Vieillard et al., 2008:726). They also assessed the percentage of listeners who gave the highest rating on the appropriate scale for each excerpt, a measure which they termed best label (Vieillard et al., 2008:726-727).

An analysis of variance (ANOVA) was conducted to determine whether or not the instruction conditions (recognition vs experience of emotions) had a significant effect of ratings. The ratings used for this analysis were the mean best label values for each emotion. A small but significant effect was found, with the emotion category appropriate to each excerpt receiving a higher rating in the experience condition. The researchers isolated this further, finding a significant effect for the sad emotion label. The effect of instruction condition was small and general, and it was decided to average the data across the two conditions for further analysis (Vieillard et al., 2008:727).

Of the four emotion portrayals examined, the happy, sad, and scary labels received high ratings for mean best label (i.e. these emotion portrayals were well-discriminated by listeners). The peaceful label received a lower rating, although it was still very clearly portrayed. The researchers state that the peaceful excerpts were more often rated ambiguously (receiving a join-highest rating on two emotion portrayals by a listener). The majority of this ambiguity was due to high ratings of sadness, meaning that the peaceful excerpts portrayed/induced sadness quite frequently (Vieillard et al., 2008:727-728).

In terms of valence and arousal, the excerpts formed distinct clusters by intended emotion portrayal. Happy excerpts fell into quadrant 1 (high valence, high arousal). Sad excerpts fell into quadrant 4 (high valence, low arousal). Scary excerpts were located in quadrant 2 (low valence, high arousal). The peaceful excerpts fell into quadrant 4 (high valence, low arousal; Vieillard et al., 2008:728). These results match well with the data presented in table 6.3.1 (see next chapter; the placement in VA space of these excerpts was typical), with the exception of sadness. This could suggest that the sad excerpts portrayed/induced a more positive species of sadness, such as melancholy, as hypothesised by the researchers (Vieillard et al., 2008:723-724). Sadness is not included in the set of emotion portrayals under investigation in the present research.

Having given a general background to the theory of emotions which held sway in Baroque musical thought, and its manifestation both in general and in a specific music-theoretical instance, it is time now to turn to the experimental aspect of the present research. This literature review will be used as a point of reference throughout, to support theoretical assumptions, to ground necessary choices, and to place the present findings in conversation with the existing literature.

CHAPTER 6. EXPERIMENTAL DESIGN

This chapter serves to explain the various considerations that went into designing the present experiment, and the decisions to which these considerations led. First, the research questions will be listed. Four important points will then be discussed: the selection and implementation of the dimensional theory of emotion as a theoretical framework, the plotting of discrete emotion labels onto the VA space, the selection and analysis of experimental stimuli, and the experimental process proper.

6.1 Research questions

The present experiment addresses two primary research questions. These are:

1. Are there differences in listeners' perceptions of emotion in Baroque and modern music, where that music was designed to portray the same emotion?
2. How do compositional-structural parameters contribute to the portrayal of emotion in this music?

The first question examines whether Baroque and modern excerpts that have been given similar emotion labels by their composers portray similar emotions to contemporary listeners. The null hypothesis for this question is that there is so significant difference between rating clusters for Baroque and modern excerpts with the same theorised emotion portrayal. The second question builds on the first by explaining the similarities and/or differences in VA ratings between Baroque and modern excerpts in terms of their individual patterns of cue utilisation.

6.2 Selection and implementation of the dimensional theory

The dimensional theory of emotion has been selected as the framework for the present research. The use of this theory allows for more fine-grain comparison of listener ratings than would be possible within a framework of discrete emotion theory, rendering it particularly suited to the two research questions this experiment aims to address. There may be small but significant differences between excerpt ratings in a certain direction on the VA space, which can be explained in terms of differences in cue profiles between these excerpts. The correspondence between listener ratings and the intended affective content of the excerpts will be assessed by interpreting the VA space in terms of discrete labels *at the quadrant level*. Each of the three emotional portrayals in the present set of excerpts is assigned to a quadrant, in accordance with prior research, and to facilitate a more lenient interpretation of the ratings.

This allows the accuracy of emotional communication of the excerpts to be assessed, while leaving space for comparison between excerpts falling in the same quadrant.

A rating scale paradigm will be used to gather data from listeners. The two ratings scales will correspond to the dimensions of the VA model, valence and arousal. This will allow listeners to quickly and easily make judgements of the emotional content of the excerpts. No difficulty is expected in explaining the nature of the rating scales to listeners. The valence dimension will be anchored with the words “happy” and “sad” at the positive and negative poles, respectively. The arousal dimension will be anchored at the positive pole with “energetic” and at the negative pole with “sleepy”. This will help to facilitate consistent interpretation of the rating scales between listeners.

6.3 Positions of discrete emotion labels on the VA space

A number of researchers in several fields (see table 6.3.1) have presented studies in which discrete emotion categories were plotted onto the VA space. The findings of these studies are compared in table 6.3.1. This will be used as a basis for plotting the musical stimuli used in the present experiment onto the VA space, which is an important step in experimental design.

Findings are summarised in table 6.3.1 in terms of the location of discrete emotion labels *per quadrant of the VA space*. The four quadrants comprising the space will be labelled as follows: Q1 (positive valence, high arousal), Q2 (negative valence, high arousal), Q3 (negative valence, low arousal), and Q4 (positive valence, low arousal). Figure 6.3.1 illustrates this division of the VA space.

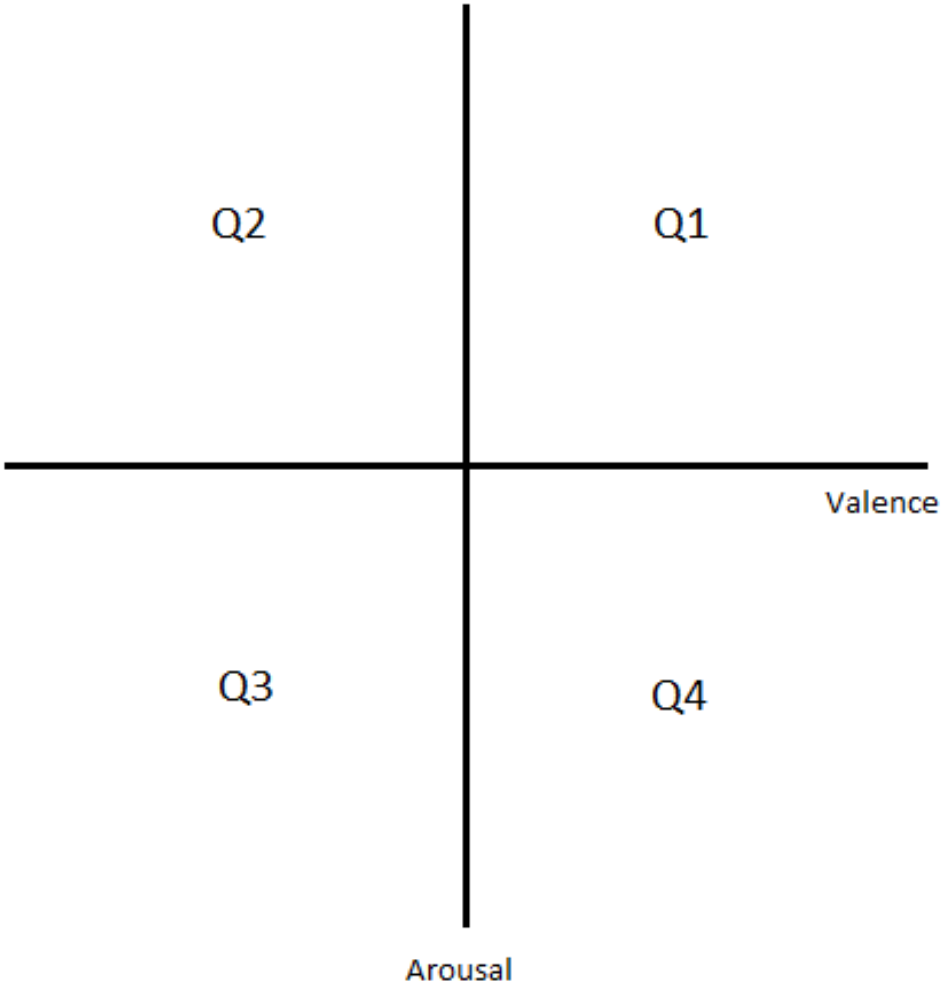


Figure 6.3.1. Quadrant division of the VA space.

	Mehu & Scherer 2015	Eerola & Vuoskoski 2011	Hupont *et al.* 2013	Vieillard *et al.* 2008	Fontaine *et al.* 2007	Richins 1997
Q1	elation, amusement	happiness	surprise, joy, amusement, cheerful	happiness	love, surprise, joy, happiness, pleasure	proud, joyful, pleased, amazed
Q2	panic fear, hot anger, despair, anxiety/worry	anger, fear	anger, disgust, fear, nervous	scary	anxiety, anger, irritation, hate	frustrated, irritated, angry, worried, disgusted, furious
Q3	interest, sadness, cold anger/irritation	sadness	sadness, worried, bored	/	sadness, despair, disgust, contempt, disappointment	sad, threatened, afraid, tender
Q4	pride, relief, sensual pleasure	tenderness	calm, satisfied	sadness, peacefulness	pride, contentment, interest, compassion	loving, calm, peaceful, comforted, relieved
Field	*Facial expressions*	*Music*	*Emotion words*	*Music*	*Emotion words*	*Consumption behaviour*

	Russell 1980	Morgan & Heise 1988
Q1	excited, delighted, glad, happy, pleased	excited, glad, happy, pleased
Q2	alarmed, afraid, angry, annoyed, distressed	angry, annoyed, disgusted, frustrated, furious
Q3	sad, tired, bored, depressed	afraid, depressed, sad, distressed
Q4	satisfied, content, calm, relaxed	calm, contented, melancholy
Field	*Emotion words*	*Emotion words*

Table 6.3.1. Positions of discrete emotion labels per quadrant of the VA space.

Table 6.3.1 can be reworked to show the positions of the emotion labels in terms of their frequency of placement in quadrants of the VA space. In order to simplify this reworking, related labels will be grouped under a single heading, with reference to the model proposed by Shaver *et al.* (1987:1067, 1074-1076). Some words which were not included in their model are given here as separate categories. This is illustrated in table 6.3.2.

Happiness	joy, elation, amusement, cheerful, pleased, sensual pleasure, glad, excited, delighted, pride, contentment, relief, satisfied
Sadness	despair, disappointment, depressed, melancholy, tired
Fear	nervous, anxiety, worried, panic fear, alarmed, distressed, afraid, threatened
Anger	irritation, hate, frustrated, furious, scary, annoyed, disgust, contempt
Bored	/
Peacefulness	calm, relaxed, comforted, love, compassion, tenderness
Interest	/
Surprise	amazed

Table 6.3.2. Groupings of related emotion words.

Table 6.3.3 shows the results of this arrangement per quadrant of the VA space. It is clear that quadrant 1 is dominated by happiness, quadrant 2 by fear and anger, quadrant 3 by sadness, and quadrant 4 by peacefulness (with a fair showing from happiness). Note that for all emotions, there are no occurrences in quadrants not adjacent to the principal quadrant for that emotion (Q1 and Q4 are adjacent).

	Q1	Q2	Q3	Q4
Happiness	X (8)			X (6)
Sadness		X (1)	X (7)	X (2)
Fear		X (6)	X (3)	
Anger		X (8)	X (2)	
Bored			X (2)	
Peacefulness	X (1)		X (1)	X (7)
Interest			X (1)	X (1)
Surprise	X (3)			

Table 6.3.3. Quadrant positions for discrete emotion labels. An "X" marks the inclusion of a group within a quadrant by a study. The number in brackets next to the "X" indicates the number of studies which have placed this group (either the head term or any number of its constituents) in this quadrant.

In terms of their position on the VA space, it is clear from table 6.3.3 that happiness and peacefulness form clusters in quadrant 1 (with overlap into 4) and 4 respectively. Fear and anger cannot be differentiated in VA space at the quadrant level, both falling into quadrant 2. The quadrant into which each emotion label most commonly falls will be taken as the fuzzy boundary of that emotion. The ratings from the experiment will be compared to these predictions at the quadrant level. Three emotion labels are under consideration in this experiment, corresponding to three separate quadrants: Happiness (Q1), Peacefulness (Q4), and Anger (Q2).

6.4 Selection and analysis of experimental stimuli

The musical excerpts to be used in the experiment have been selected in line with the first of the principle research questions of the present paper. This is: what discrepancies, if any, exist between listener ratings of two classes of stimuli. The first class includes modern experimental stimuli composed in accordance with a subset of the general rules of acoustic cues as identified in prior empirical research. The second class includes excerpts written by a Baroque composer and music theorist. These were written to exemplify affective portrayal according to his understanding (and, by inference, broadly conforming to aspects of the shared understanding of European music theorists at the time). This paper's second research question (what is the relationship between musical structural parameters and portrayed affect) will then be used as a framework for comparative analysis of the excerpts. This comparison will help to ground speculation as to the relative historical stability of the use of structural parameters in musical affect portrayal.

Heinichen's treatise was chosen as the pool from which to draw the second class of experimental stimuli, for two reasons. First, the present author was unable to find any clearer examples in the literature which could be used instead. Second, drawing examples from

multiple sources would necessitate a comparison of rating accuracy between theorists and detailed discussion of their minor differences, which is outside the scope of the present research.

The use of a broad, per-quadrant definition of emotion labels in the VA space helps to ameliorate the effects of translation and time on the meaning of Heinichen's descriptions of his excerpts. This represents a broadening of the semantic concept in question, to include related terms that fall close to the original on the VA space. That is to say, a word such as "love" will not be taken solely to mean the strict English dictionary denotation of love. Instead, it will be redefined to incorporate the semantic space surrounding this denotative core; rather than meaning one highly specific thing, the term will be used as a marker for a range of affective concepts that share elements of the core. Thus, "love" will no longer represent a *point* on a semantic space, but rather an *area*. It is hoped that this conceptualisation will limit the impact of translation across languages and times; the meaning of a word may well be expected to shift under these transformations, but it is likely to remain within a broadly conceived semantic space. This approach is not perfect. In the case that the experimental results show a mismatch in affect ratings for the Baroque excerpts, the issue of trans-linguistic and -historical meaning will contend as possible explanations. If the experimental results show a good match with the terms, then it is likely that these issues did not manifest significantly in this specific case.

As already mentioned, the modern excerpts were drawn from a paper by Vieillard *et al.* The experimental data they generated (2008:728) show three clear groupings across the four emotion portrayals ("happiness", "sadness", "peacefulness", "scary") they studied. These correspond to three of the four quadrants of the VA space. As the present research aims to investigate the effectiveness of emotion portrayal in a sample of the modern excerpts designed for use in Vieillard *et al*'s paper and a sample of Baroque excerpts, the decision was made to select one excerpt of each type which has a similar emotion label. It can be expected, if the Baroque excerpts are rated in accordance with their labels by contemporary listeners, that they will occupy similar places on the VA space to their more recently-composed counterparts. The modern excerpts for Happiness, Peacefulness, and Anger were found by Vieillard *et al.* to fall into quadrants 1, 4, and 2, respectively.

A number of factors were involved in the selection of the excerpts of both types. First, the excerpts from Heinichen's *Einleitung* were assessed in terms of their prefatory material to determine their affective portrayals. In some cases, there was no clear statement of an affect to be depicted by the excerpt (Heinichen also wrote excerpts here to demonstrate compositional

and rhetorical techniques other than affective portrayal). These excerpts were discarded. The remaining nine excerpts could be identified by the following affective descriptions, respectively: 1 – Fury, 2 – Rage, 3 – Amorous, 4 – Love, 5 – Flirtatious, 6 – Tenderness, 7 – Anxious, 8 – Playful, 9 – Love.

At the time of making the categorisation and selection of the stimuli, the present author was not aware that Mattheson (1981:106) provided a list of emotional expressions in terms of which he interprets Heinichen's excerpts. He lists these as follows: rage, quarrelsomeness, majesty, fear, play, strife, unity, happiness, volatility, sorrowfulness, love, fieriness, yearning, sighing, flirtatiousness, and umbrageousness. The direct correspondence between each of these expressions and specific excerpts from Heinichen's *Einleitung* is unfortunately not clearly stated. However, the fact that all nine of the affective categorisations made by this author correspond either directly or very closely (in terms of the groups constructed in table 6.3.2) to those of Mattheson lends credibility to the present method of categorisation.

Vieillard *et al*'s (2008) results were compared to the VA spaces depicted in other research (see table 6.3.1 above). These results matched well for the emotion labels of "happiness", "scary", and "peacefulness". These three emotion labels were used to make selections from Heinichen's excerpts. The words selected as descriptors for Heinichen's excerpts were compared to the extensive taxonomy of Shaver *et al.* (1987:1067, 1074-1076). Where the exact or similar words were included as species of the emotion category corresponding to those of Vieillard *et al.*, these were kept. Grouped by the three categories of Vieillard *et al.*, the candidate Baroque emotion labels now read as follows:

These excerpts were now grouped by their similarity to the emotion terms of Vieillard *et al* (2008). This resulted in the following three groups:
HAPPINESS: 8 – Playful
PEACFULNESS: 3 – Amorous, 4 – Love, 5 – Flirtatious, 6 – Tenderness, 9 – Love
SCARY: 1 – Fury, 2 – Rage, 7 – Anxious

The HAPPINESS category only had one eligible Baroque excerpt, so this was selected to be its representative. For the other categories, the procedure of selection continued with a preliminary musical analysis. In order to keep the counterpart excerpts of a similar duration (no more than two seconds' difference), the durations in seconds of the modern excerpts were calculated using the formula No. of beats / BPM * 60. BPM prescriptions (beats per minute, used to measure tempo) were provided by the authors (Vieillard *et al*, 2008:744-752). In order

to sidestep any potential effects of cadentiality on emotion perception, it was decided to cut each excerpt to end on a pre-dominant function. The modern excerpts were harmonically analysed, starting at the end of each and working in reverse, to find an appropriate ending.

The candidate Baroque excerpts underwent the same procedure for duration calculation, using BPM indications provided by Martin Berger, a local Baroque expert and choral conductor. As the Baroque excerpts were all considerably longer than the modern excerpts, the musical portions of each excerpt corresponding to the timings of (the minimum duration of the corresponding modern excerpt group) -2 to (the maximum duration of the corresponding modern excerpt group) +2 seconds were subjected to harmonic analysis. Possible (pre-dominant function) ending points within this range were noted, along with their time-coded values (calculated with the same formula as previously used for duration). These were then compared with durations of the modern excerpts up to their various potential ending points, to find all possible pairs within emotion classes that differed in duration by no more than two seconds.

This left a number of possible counterpart pairs. The next heuristic applied was a measure of the multivariate centrality of each of the modern excerpts in terms of its VA ratings within its emotion class cluster. This was calculated using a scatterplot and the Mahalanobis function for calculating multivariate outliers in R's StatMatch package. The number of outliers to be indicated was incremented by one with each iteration of the function, and the order of identification of the excerpts was noted. This yielded a ranked list of the excerpts in each emotion class in order of centrality. The counterpart pairs containing the most central of the remaining modern excerpts per emotion class were retained, the others discarded. This left one pair of excerpts per emotion label, which were used as stimuli in the experiment.

CHAPTER 7. EXPERIMENTAL PROCESS, STATISTICAL PROCEDURE, AND RESULTS

Having discussed the selection of experimental stimuli, the experimental procedure and results will now be reported. This chapter addresses the first research question of the present thesis: what differences, if any, can be found between listener ratings of Baroque and modern musical excerpts, where those excerpts were designed to portray the same emotional content?

7.1 Experimental Process

Participants

Participants (N=30, 14 females) were recruited from the undergraduate student population of the University of Stellenbosch. The mean age of participants was 22, SD=7.14, range=39 (18-57). Participation was voluntary. Participants received no payment.

Half of the participants (N=15) were musicians, and the other half non-musicians. Participants were classified as musicians if they had received formal training in music, and currently played an instrument or sang at a serious level. As the musician participants were primarily recruited from within the University of Stellenbosch Music Department, these criteria were easily satisfied.

Non-musicians were defined as those who had received no formal training in music, and did not currently play a musical instrument or sing seriously. This criterion was relaxed somewhat to include participants who had sung in a choir but did not play an instrument at a serious level.

Recruitment

Participants were recruited through two channels. For the musician category, a mass email was sent to students within the University of Stellenbosch Music Department, detailing the broad scope of the study and the experimental procedure (listening experiment), and inviting students to participate. To acquire participants for the non-musician category, the present researcher requested permission from lecturers in other departments of the University of Stellenbosch (Philosophy and Visual Arts) to make a bid to students during a lecture. Students were presented again with the broad scope of the research and the experimental procedure, and invited to participate.

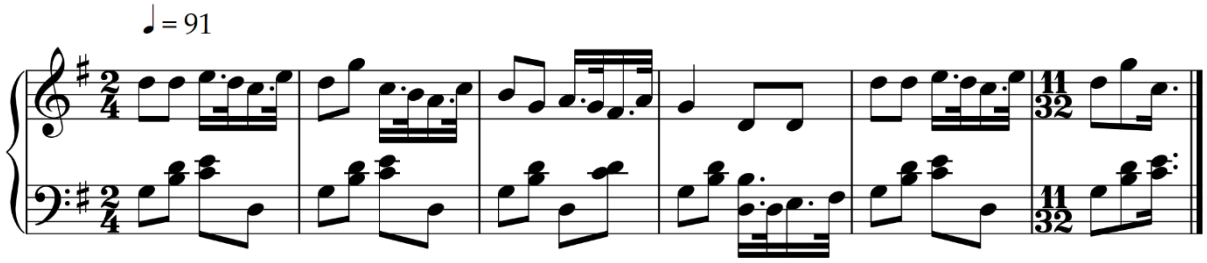
Ethics

Ethical clearance for this study was acquired through the University of Stellenbosch Research Ethics Committee: Humanities. Gatekeeper (institutional) permission for the inclusion of University of Stellenbosch students in the present experiment was acquired through the University of Stellenbosch Division for Information Governance. All clearance was acquired prior to recruitment and experimentation. The relevant documentation is available from the author upon request.

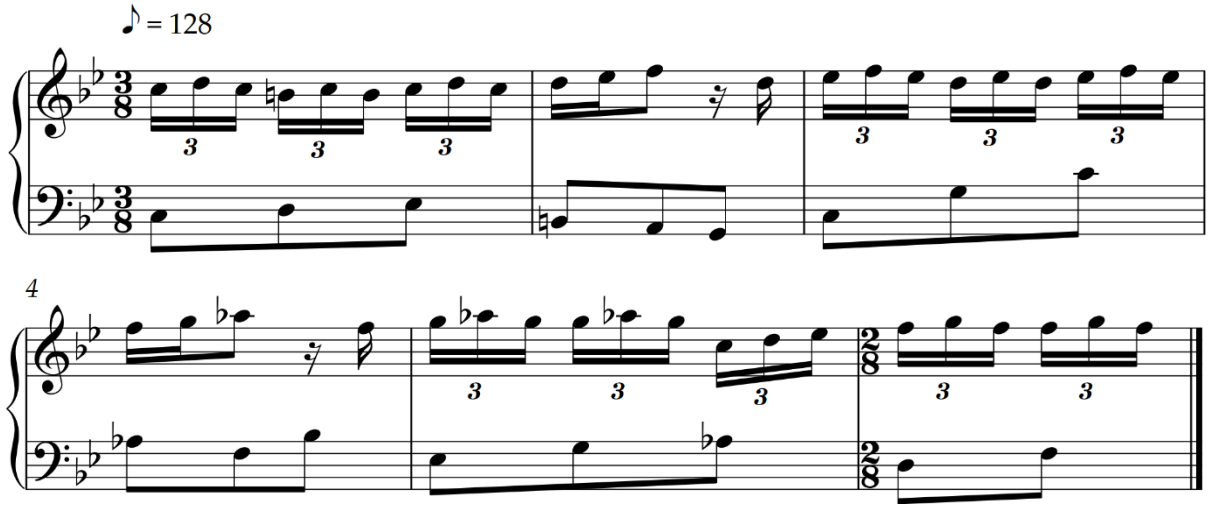
Procedure

Participants were tested individually. They were first asked to fill out a questionnaire detailing their age, sex, musical experience, and so on (reprinted in Appendix C). After completing the questionnaire (with the exception of the last question, which was optional and only to be completed at the end of the experiment), the participant was given a copy of the rating sheet. It was explained to the participant that six short excerpts of music would be heard, with each excerpt being repeated twice. For the first hearing of each excerpt, the participant was instructed to concentrate on the *happy-sad* quality of the music, and to indicate their judgement on the first rating scale. On the second hearing, the subject was to concentrate on the *energetic* quality of the music (high energy-low energy), and indicate judgement on the second scale. On completion of each rating, the participant was instructed to give a thumbs-up sign to the experimenter, at which point the excerpt would be repeated or the next excerpt played, as appropriate. Participants were given the chance to ask questions about the procedure before listening to the excerpts.

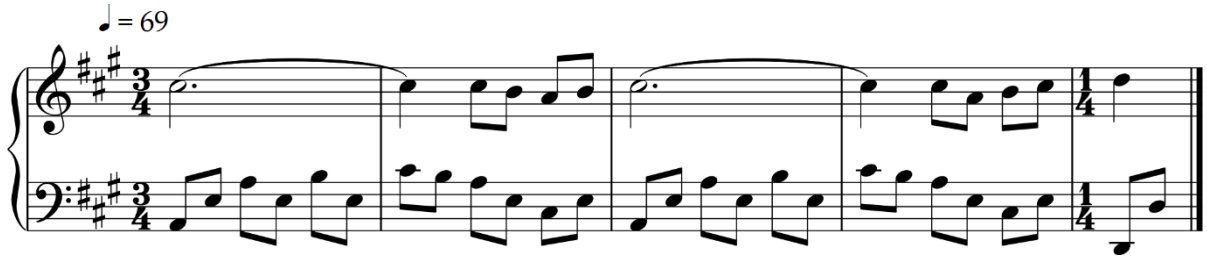
The excerpts were played back as .WAV files, exported from Sibelius in 32-bit quality. They were played off a Dell Latitude 7280 laptop computer, through Sennheiser HD 419 over-ear headphones. The volume was set to a comfortable level, which was kept constant across participants. Once the ratings were completed, participants were debriefed, and were offered the chance to ask questions and fill in the comments section of the questionnaire. The full procedure took roughly 10-15 minutes. The excerpts used are printed below.



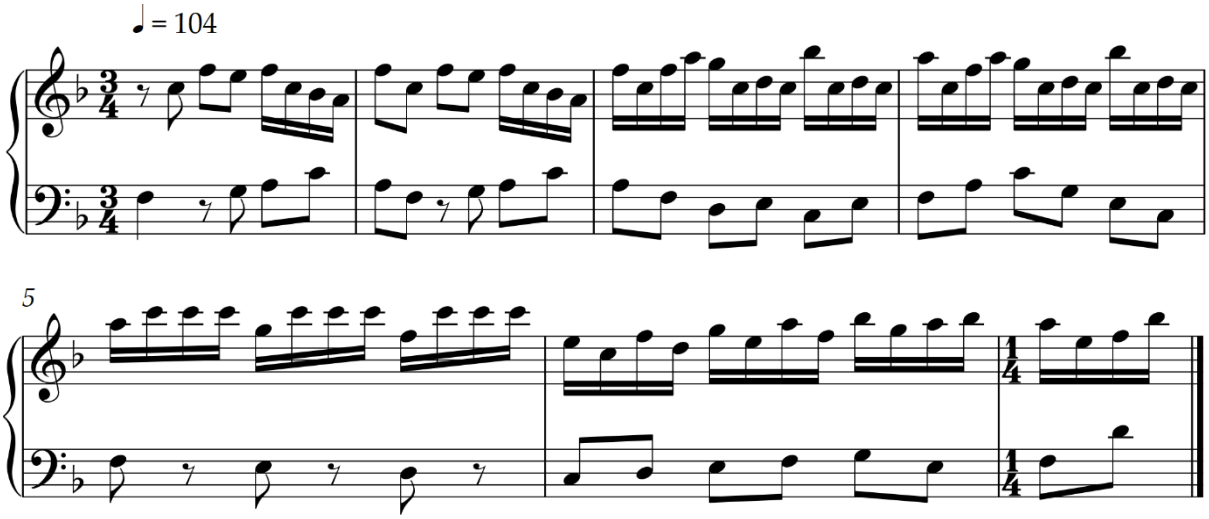
Excerpt 1: Happy, modern



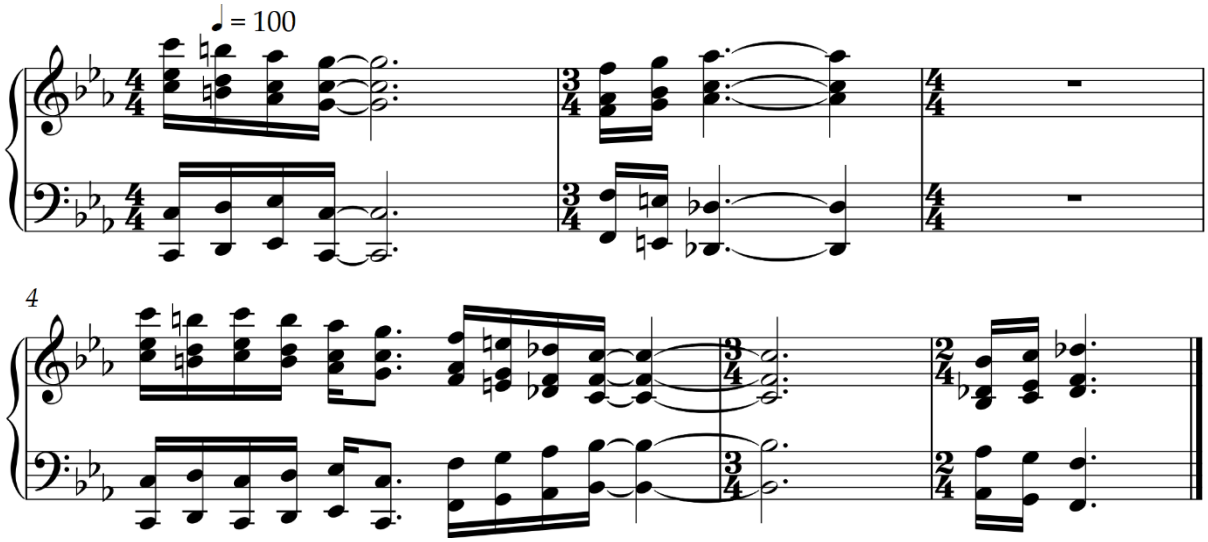
Excerpt 2: Happy, Baroque



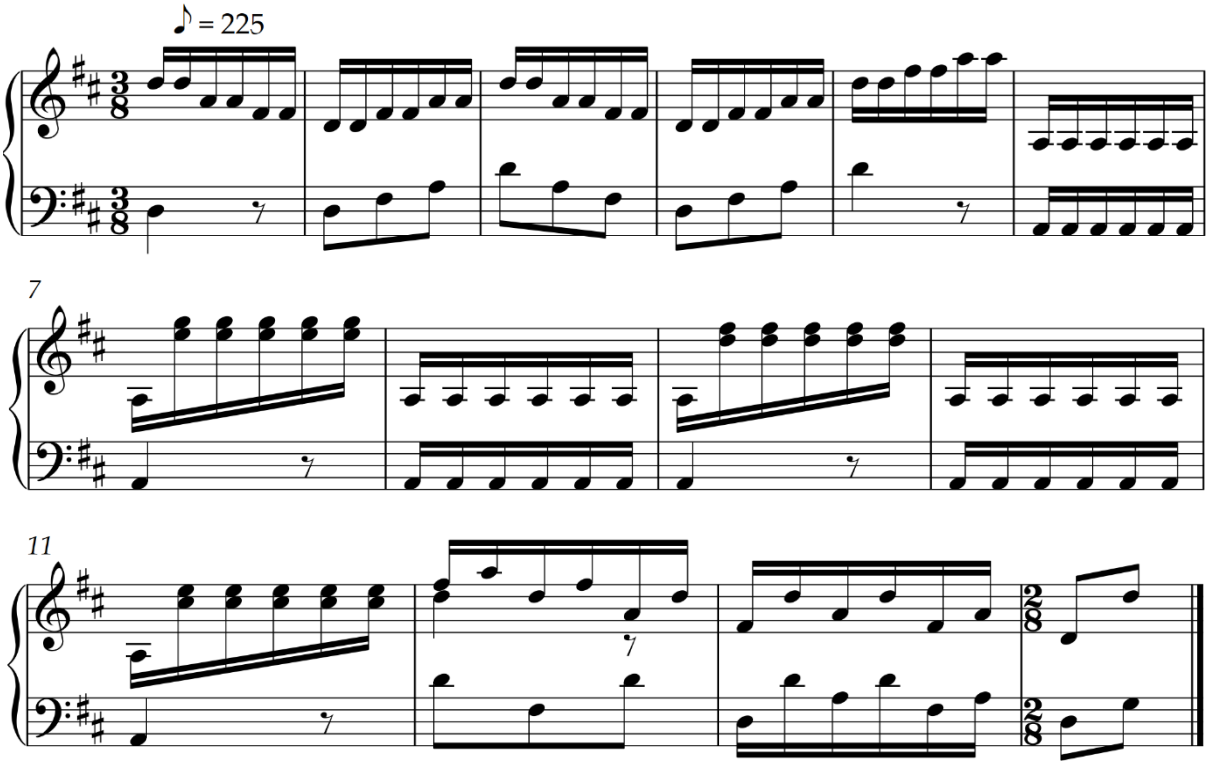
Excerpt 3: Peaceful, modern



Excerpt 4: Peaceful, Baroque



Excerpt 5: Angry, modern



Excerpt 6: Angry, Baroque

7.2 Results

The necessary assumption tests for a multivariate analysis of variance (MANOVA) were computed using R. These all gave appropriate results, with the exception of the assumption of multivariate normality, which was assessed using Mardia's test, and Box's M test for equivalence of covariance matrices (see Appendix B for full results of all statistical computations, including complete results of assumption tests). The ratings were grouped by emotion label for Mardia's test. Normality was reported for the ratings of happy and peaceful excerpts. For angry excerpts, Mardia's test reported a non-normal distribution. However, on visual assessment of the Q-Q plot for this group (included in Appendix B), it was decided to continue with the MANOVA procedure, as the plot did not indicate a particularly high degree of deviation from a normal distribution. Box's M test is highly sensitive to non-normality, and thus a significant result was not surprising given the previous result for angry excerpts. Nonetheless, a judgement to continue with a parametric multivariate test was made.

The use of MANOVA, as opposed to factorial ANOVA on each dimension (valence and arousal) individually, is based on the assumption of a medium degree of correlation between the dimensions (French *et al.*, 2008). If the dimensions were entirely uncorrelated, there would be no advantage to treating them together rather than separately. If they were (almost) perfectly correlated, this would indicate redundancy in the model, as the two dimensions would not embody distinct concepts. The level of correlation between valence and arousal was assessed using Pearson's *r*, yielding $r(178)=.52$, $p<.0001$. This indicates a degree of correlation between the valence and arousal ratings, justifying the use of the MANOVA procedure. Note that in all primary computations, the required significance level was set at .05.

Before the primary MANOVA was run, three individual MANOVAs were computed to assess the impact of other, non-theoretically important factors on the ratings. One test each was run for the factors of Age, Sex, and Presentation Order (see Appendix B). In all cases, results did not achieve significance ($p>.05$). This justified the exclusion of these factors from the main model.

A MANOVA was run with valence and arousal as the dependent variables, and Label, Time Period, and Musicianship as factors (as well as all possible interactions between those factors; the formula in R was VA ~ Label*Time period*Musicianship). Significant main effects and interaction effects were found, summarised in table 7.1.1 below.

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
Time period	1	0.34302	43.596	2	167	5.831e-16 ***
Label	2	0.40206	21.135	4	336	1.467e-15 ***
Musicianship	1	0.04506	3.940	2	167	0.021282 *
Time period:Label	2	0.82101	58.495	4	336	< 2.2e-16 ***
Time period:Musicianship	1	0.00190	0.159	2	167	0.853478
Label:Musicianship	2	0.04737	2.038	4	336	0.088756 .
Time period:Label:Musicianship	2	0.10060	4.449	4	336	0.001622 **
Residuals	168					

Table 7.1.1 Results of MANOVA with valence and arousal as the dependent variables.

A significant main effect was found for each of the three factors: Time period ($F(2, 167)=43.60$, $p<.0001$), Label ($F(4, 336)=21.14$, $p<.0001$), and Musicianship ($F(2, 167)=3.94$, $p<.05$). The two-way interaction between Time period and Label was significant ($F(4, 336)=58.50$, $p<.0001$). The other two-way interactions (Time period and Musicianship, $F(2, 167)=0.16$, $p=.85$; Label and Musicianship, $F(4, 336)=2.04$, $p=.08$) did not reach the critical p-value of .05. The three-way interaction between Time period, Label, and Musicianship was significant, $F(4, 336)=4.45$, $p<.005$.

Figure 7.2.1 illustrates the VA ratings given by participants for the Baroque and modern excerpts. Modern excerpts are seen to tend towards the bottom left of the VA space compared to Baroque excerpts. Based on examination of this figure 7.2.1, it was suspected that modern excerpts would score lower on both valence and arousal than Baroque excerpts.

The VA ratings by participants for the three emotion portrayals (Happy, Peaceful, Angry) are shown in figure 7.2.2. Examination of this figure led to the expectation that Peaceful excerpts would have lower arousal ratings than both Happy and Angry excerpts. No other predictions were made based on figure 7.2.2. No prediction was made for the effect of Musicianship, shown in figure 7.2.3.

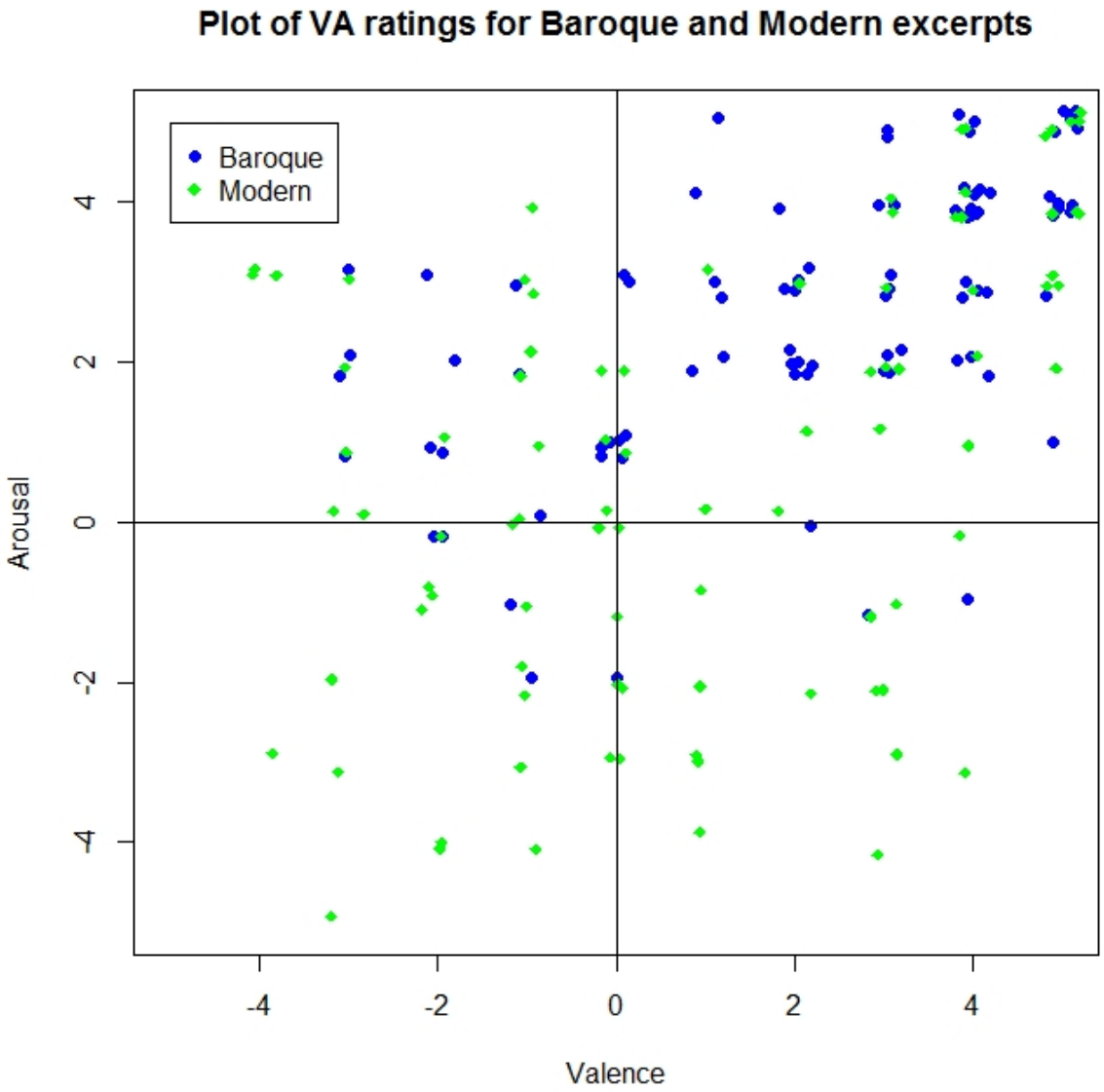


Figure 7.2.1. Scatterplot of VA ratings given for Baroque and modern excerpts.

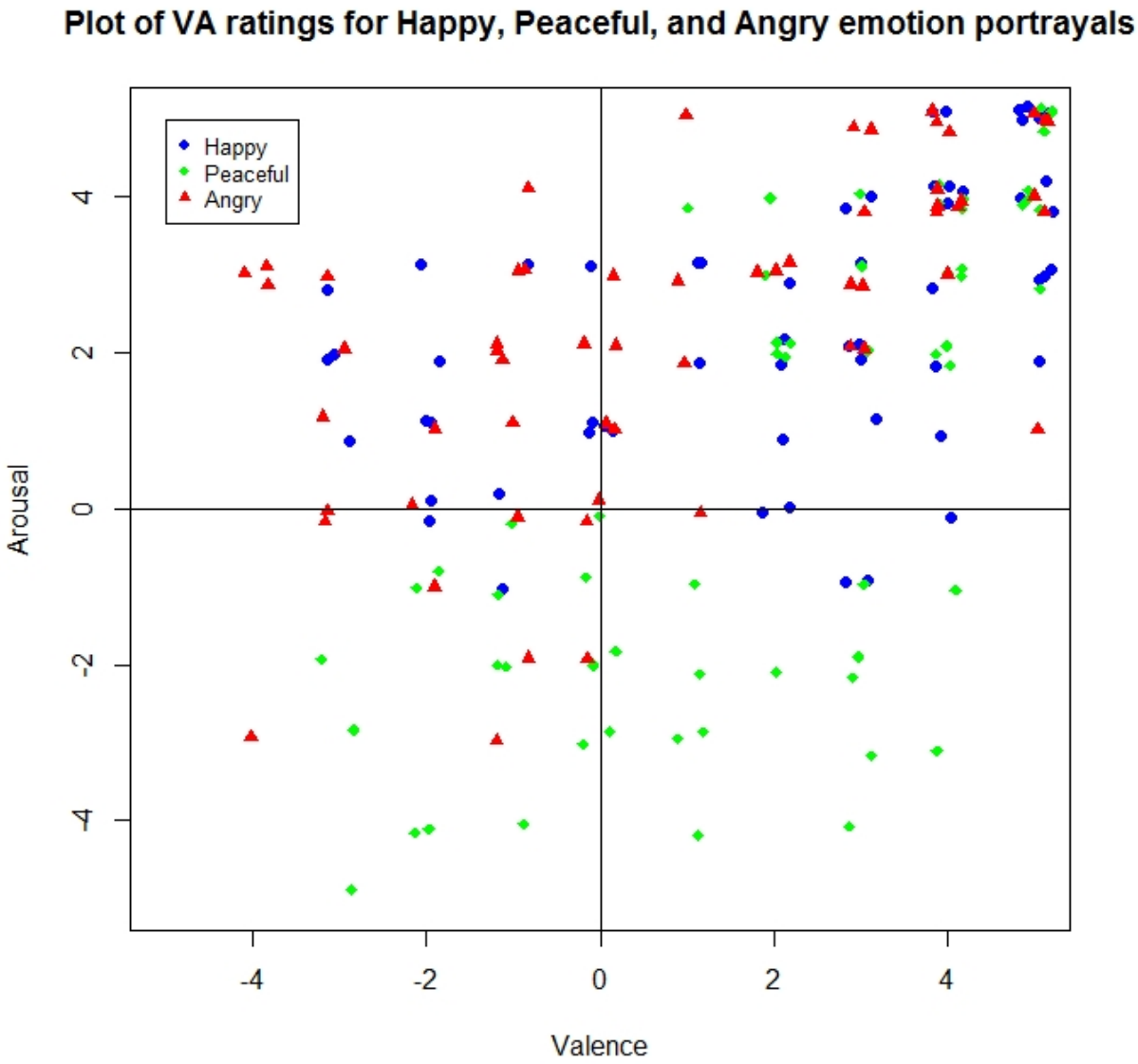


Figure 7.2.2. Scatterplot of VA ratings for Happy, Peaceful, and Angry emotion portrayals.

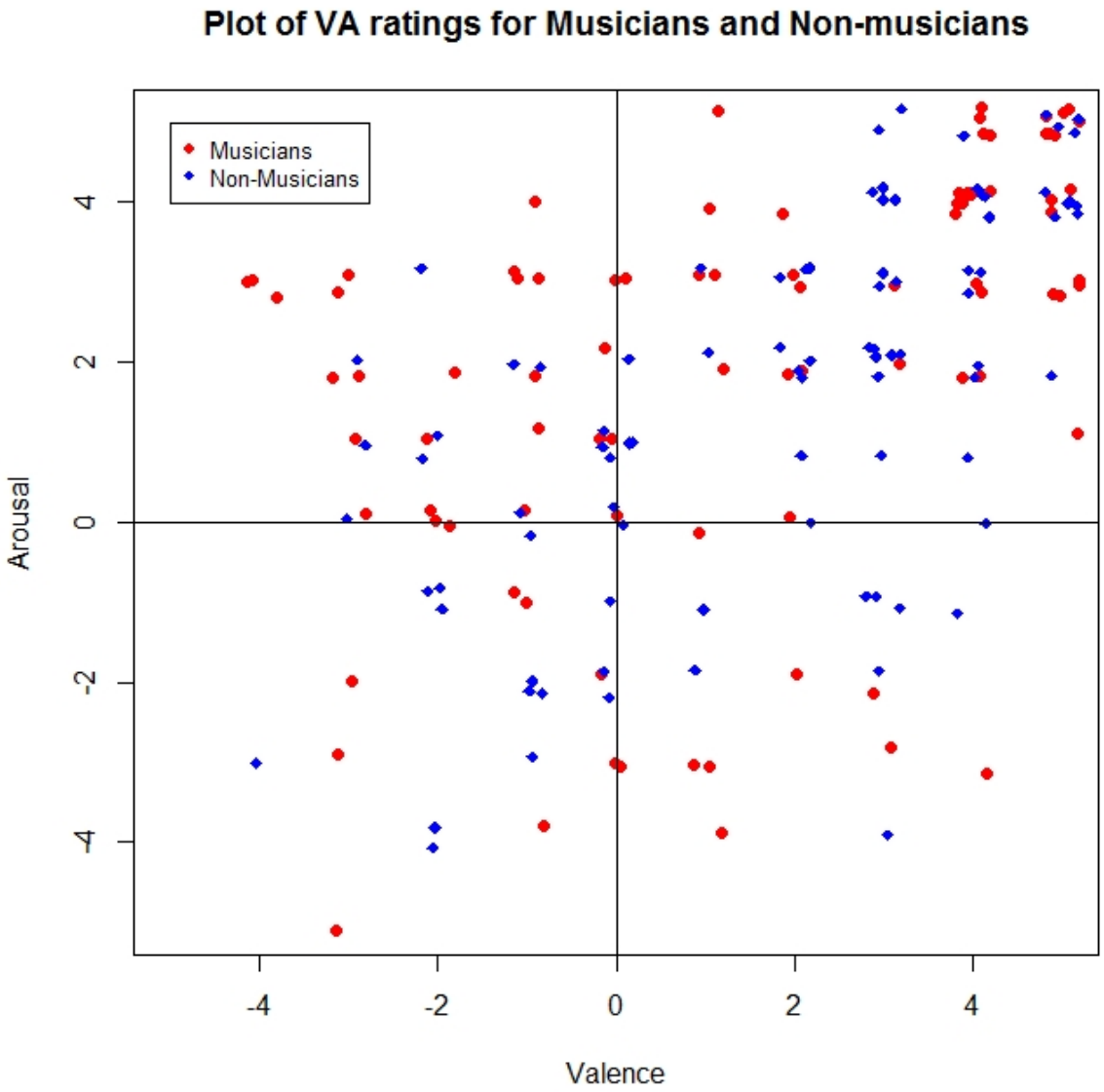


Figure 7.2.3. Scatterplot of VA ratings for Musicians and Non-musicians.

The results of the MANOVA were examined in more detail by conducting two factorial ANOVAs, one for each of the dependent variables (valence and arousal) separately. The factorial ANOVAs were followed up using post hoc Tukey honestly significant difference (HSD) tests. The model used for each ANOVA was similar to that used in the MANOVA above; valence was taken as the sole dependent variable in the first model, and arousal in the second. The results of each ANOVA are presented and explored below.

ANOVA with valence as the dependent variable

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time period	1	64.8	64.8	22.169	5.2e-06 ***
Label	2	40.0	20.0	6.848	0.00138 **
Musicianship	1	4.4	4.4	1.490	0.22391
Time period:Label	2	635.7	317.9	108.740	< 2e-16 ***
Time period:Musicianship	1	0.6	0.6	0.190	0.66342
Label:Musicianship	2	4.2	2.1	0.720	0.48809
Time period:Label:Musicianship	2	23.5	11.7	4.016	0.01978 *
Residuals	168	491.1	2.9		

Table 7.2.1. Results of factorial ANOVA with valence as the dependent variable.

For the ANOVA with valence as the dependent variable, main effects of Time period (F(1, 168)=22.17, p<.0001) and Label (F(2, 168)=6,85, p<.005) showed significance. There was no significant main effect of Musicianship (F(1, 168)=1.49, p=.22). There was a significant two-way interaction between Time period and Label (F(2, 168)=108.74, p<.0001). The other two-way interactions (between Time period and Musicianship (F(1, 168)=0.19, p=.66), and between Label and Musicianship (F(2, 168)=0.72, p=.49)) were not significant. The three-way interaction between Time period, Label, and Musicianship was significant (F(2, 168)=4.02, p<.05).

The significant effect of Time period on valence supports the expectation mentioned above that modern excerpts would score lower on valence than Baroque excerpts. Figure 7.2.1 illustrates this effect. Means and SDs were calculated for valence across the two Time periods. Modern excerpts had a mean valence rating of 0.97 (SD=2.79), while Baroque excerpts had a mean valence rating of 2.17 (SD=2.39). This shows a clear and sizeable difference between valence ratings across the Time periods, with modern excerpts scoring lower than Baroque excerpts. The effect size (Cohen's d) associated with this difference was d=0.46.

The significant effect of Label on valence ratings was not initially suspected from assessment of figure 7.2.2. This effect, as well the interaction effects, was examined in detail with Tukey HSD tests (p-values were adjusted automatically by the function in R). The full results of these tests are included in Appendix B.

The results of the Tukey HSD test for the effect of Label on valence showed that ratings for the Happy and Angry labels differed significantly (p adj<.005, *d*=0.37), as did ratings for the Peaceful and Angry Labels (p adj<.01, *d*=0.37). Peaceful and Happy labels did not differ significantly (p adj=.99). Means and SDs were calculated for each label. Happy excerpts had a mean valence rating of 1.92 (SD=2.64). Peaceful excerpts had a mean valence rating of 1.88 (SD=2.41). Angry excerpts had a mean valence rating of 0.90 (SD=2.83). These results show that Angry excerpts were rated significantly lower on valence than Happy and Peaceful excerpts.

The Tukey HSD test was used to explore the two-way interaction between Time period and Label beyond what had already been established by examining the main effects. Happy excerpts were found to differ significant as a function of Time period (p adj<.0001, *d*=2.53), as did Peaceful (p adj<.0001, *d*=3.11) and Angry excerpts (p adj<.0001, *d*=2.22). Within the Happy label, the mean valence score for the modern excerpt was 3.97 (SD=1.10), and the mean valence score for the Baroque excerpt was -0.13 (SD=2.06). Within the Peaceful label, the mean valence score for the modern excerpt was 0.13 (SD=2.01), and the mean valence score for the Baroque excerpt was 3.63 (SD=1.19). Within the Angry label, the mean valence score for the modern excerpt was -1.20 (SD=1.94), and the mean valence score for the Baroque excerpt was 3.00 (SD=1.84). These results show that the mean valence ratings of modern excerpts were lower than Baroque excerpts for the emotion portrayals of Peacefulness and Anger, but were higher than the Baroque excerpt for Happiness. The main effect of Time period identified above thus did not cause an effect in a consistent direction.

Despite the significant three-way interaction effect found for Time period, Label, and Musicianship, none of the results of the Tukey HSD test in which Time period and Label were constant between the two levels of Musicianship reached significance.

ANOVA with arousal as the dependent variable

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time period	1	184.0	184.02	84.163	< 2e-16 ***
Label	2	147.0	73.49	33.610	5.27e-13 ***
Musicianship	1	8.9	8.89	4.065	0.0454 *
Time period:Label	2	375.2	187.62	85.809	< 2e-16 ***
Time period:Musicianship	1	0.1	0.09	0.041	0.8405
Label:Musicianship	2	10.8	5.42	2.480	0.0868 .
Time period:Label:Musicianship	2	17.9	8.96	4.096	0.0183 *
Residuals	168	367.3	2.19		

Table 7.2.2. Results of factorial ANOVA with arousal as the dependent variable.

For arousal, main effects of Time period ($F(1, 168)=84.16$, $p<.0001$), Label ($F(2, 168)=33.61$, $p<.0001$), and Musicianship ($F(1, 168)=4.07$, $p<.05$) were significant. The two-way interaction between Time period and Label ($F(2, 168)=85.81$, $p<.0001$) was significant. The other two way interactions, between Time period and Musicianship ($F(1, 168)=0.04$, $p=.84$), and Label and Musicianship ($F(2, 168)=2.48$, $p=.09$) proved non-significant. The three-way interaction between Time period, Label, and Musicianship ($F(2, 168)=4.10$, $p<.05$) was statistically significant.

It was suggested, based on figure 7.2.1, that modern excerpts would have a lower mean arousal rating than Baroque excerpts. The main effect of Time period found immediately above supports this assertion. Modern excerpts had a mean arousal rating of 0.67 (SD=2.75), and Baroque excerpts had a mean arousal rating of 2.69 (SD=1.69). This result confirms that modern excerpts received lower arousal ratings than Baroque excerpts. The effect size for this comparison was $d=0.89$.

Tukey HSD tests were used to investigate the remaining effects. For the main effect of Label, Happy excerpts (mean arousal=2.30, SD=1.69) were found to differ from Peaceful excerpts (mean arousal=0.40, SD=3.03, $p<.0001$; $d=0.77$). Angry excerpts (mean arousal=2.33, SD=2.07) also differed from Peaceful excerpts ($p<.0001$; $d=0.74$). Happy and Angry excerpts did not differ significantly ($p=.99$). This supports the assessment of figure 7.2.2 made above; Peaceful excerpts were rated lower for arousal than other emotion portrayals.

The significant effect of musicianship on arousal ratings was unexpected. The mean arousal score for musicians was 1.90 (SD=2.61), and the mean arousal score for non-musicians was 1.46 (SD=2.36). The effect size for this difference was small, $d=0.18$. This shows that musicians reliably produced slightly higher arousal ratings than non-musicians.

In terms of the two-way interaction effect between Time period and Label, all three emotion portrayals differed by Time period. For the Happy label, modern excerpts (mean arousal=3.10, SD=1.71) had a higher mean arousal score than Baroque excerpts (mean=1.50, SD=1.25; p<.0001, *d*=1.07). The same patterns were found between modern (mean arousal=-2.33, SD=1.27) and Baroque (mean arousal=3.13, SD=1.28; p<.0001, *d*=4.28) excerpts for the Peaceful label, and between modern (mean arousal=1.23, SD=1.72) and Baroque (mean arousal=3.43, SD=1.81; p<.0001, *d*=1.25) excerpts for the Angry label. These results show that modern excerpts had a lower arousal rating than Baroque excerpts for all emotion portrayals.

The three-way interaction between Time period, Label, and Musicianship was examined next. Again, despite the significant effect found in the ANOVA, no significant differences were found between levels of the Musicianship factor when Time period and Label were constant.

The first research question of this thesis aimed to investigate whether there is a difference in listeners' ratings of Baroque and modern musical excerpts, where those excerpts were designed to portray the same emotional content. The null hypothesis is that there is no difference in these ratings as a function of Time period. Ratings for all three emotion labels were found to differ between the Baroque and modern time periods. Baroque excerpts had higher mean valence scores than modern excerpts when written to portray Peacefulness and Anger, and lower mean valence scores for Happiness. Baroque excerpts had higher arousal ratings than modern excerpts within all emotion portrayals. Based on these results, the null hypothesis was rejected.

There was a main effect of Label on both valence and arousal ratings. Mean valence scores differed between the Peaceful and Angry labels, and between the Happy and Angry labels. The mean valence scores of Happy and Peaceful excerpts did not differ significantly. Mean arousal scores differed between the Happy and Peaceful labels, and between the Angry and Peaceful labels. The Happy and Angry labels did not differ significantly in their mean arousal ratings. These results are in line with the existing literature on these three emotion portrayals. Happiness is generally positioned in quadrant one of the VA space (high valence, high arousal), Peacefulness in quadrant 4 (high valence, low arousal), and Anger in quadrant 2 (low valence, high arousal; Russell, 1980: 1167; Eerola & Vuoskoski, 2011:34). In accordance with this, Anger should be differentiated from both Happiness and Peacefulness on the valence dimension, and Peacefulness should be differentiated from both Anger and Happiness on the arousal dimension. The present results are in agreement with this theoretical prediction, both

in terms of differentiation and direction of effects; Angry excerpts had a low mean valence score, and Peacefulness had a lower mean arousal score.

The effect of Musicianship on emotion recognition in music has not received much empirical attention. Martynuska & Horabik (2015:238) give a cautious indication that musical training may play a role in emotion recognition. The present results indicate a small but significant effect of Musicianship, with Musicians giving slightly higher arousal ratings than Non-musicians overall. Further studies with larger sample sizes will be needed to confirm this effect.

CHAPTER 8. COMPOSITIONAL CUES AND DISCUSSION

In this chapter, the excerpts used in the experiment are given detailed statistical descriptions in terms of a number of compositional cues, which are then tied to the existing literature on cue utilisation in emotion portrayal. The relevance of the information in this chapter is its accordance or lack of accordance with the prior findings of inferential studies in the literature. The process of statistical description here detailed was not controlled for or operationalised in the present experiment. The cues that are identified were chosen based on their prior study in the literature, and their applicability to the present excerpts, as played back electronically – i.e. performance cues were not included. The purpose of this chapter is to address the second research question of this thesis: how do compositional-structural parameters contribute to the portrayal of emotion in the music used in the present experiment? This will be examined by summarising the effects of cues at three levels. First, patterns of cue utilisation corresponding to levels of valence and arousal will be discussed. Second, the positions of excerpts in quadrants of the VA space will be shown, and explained in terms of cue utilisation. Finally, the specific effects seen will be summarised per cue. Throughout, reference will be made to the literature to situate the present results.

This chapter will conclude with three short discussions. First, the correspondence between the present experimental results for the modern excerpts and the results generated by Vieillard *et al.* (2008) using the same excerpts will be examined. Accordance between the present results and elements of Descartes' predictions relating to musical emotion will then be noted. Finally, possible explanations for the difference in ratings between Baroque and modern excerpts will be discussed.

8.1 Cue levels

The compositional cues chosen for examination in this chapter are tempo, mean pitch height, mode, mean interval size, and texture. Tempo was treated as two separate cues: BPM (beats per minute) as indicated in the score or provided by the Baroque expert, and rate of event presentation (events per second). This latter (REP) was calculated by summing the number of discrete events (pitch onsets) in an excerpt and dividing this by the duration of the excerpt in seconds. Mean pitch height was calculated by summing the height of each pitch (measured in semitones, using C4 as the zero point), and dividing by the number of pitches. Mode was assessed as a binary major/minor variable. Mean interval size was calculated by summing the absolute values of the size (in semitones) of melodic intervals for all voices in the excerpt, and

dividing by the number of intervals. Where a voice split into two (or two voices were collapsed to one), the larger interval size was used. Texture represents the maximum number of voices present in the excerpt. The values of these cues for each excerpt are shown in table 8.1.1 below. The label, time period, and mean ratings of valence and arousal for each excerpt are included, as well as mean values for VA and all cues. Figure 8.1.1 presents a plot of the VA ratings for all excerpts, for ease of comparison.

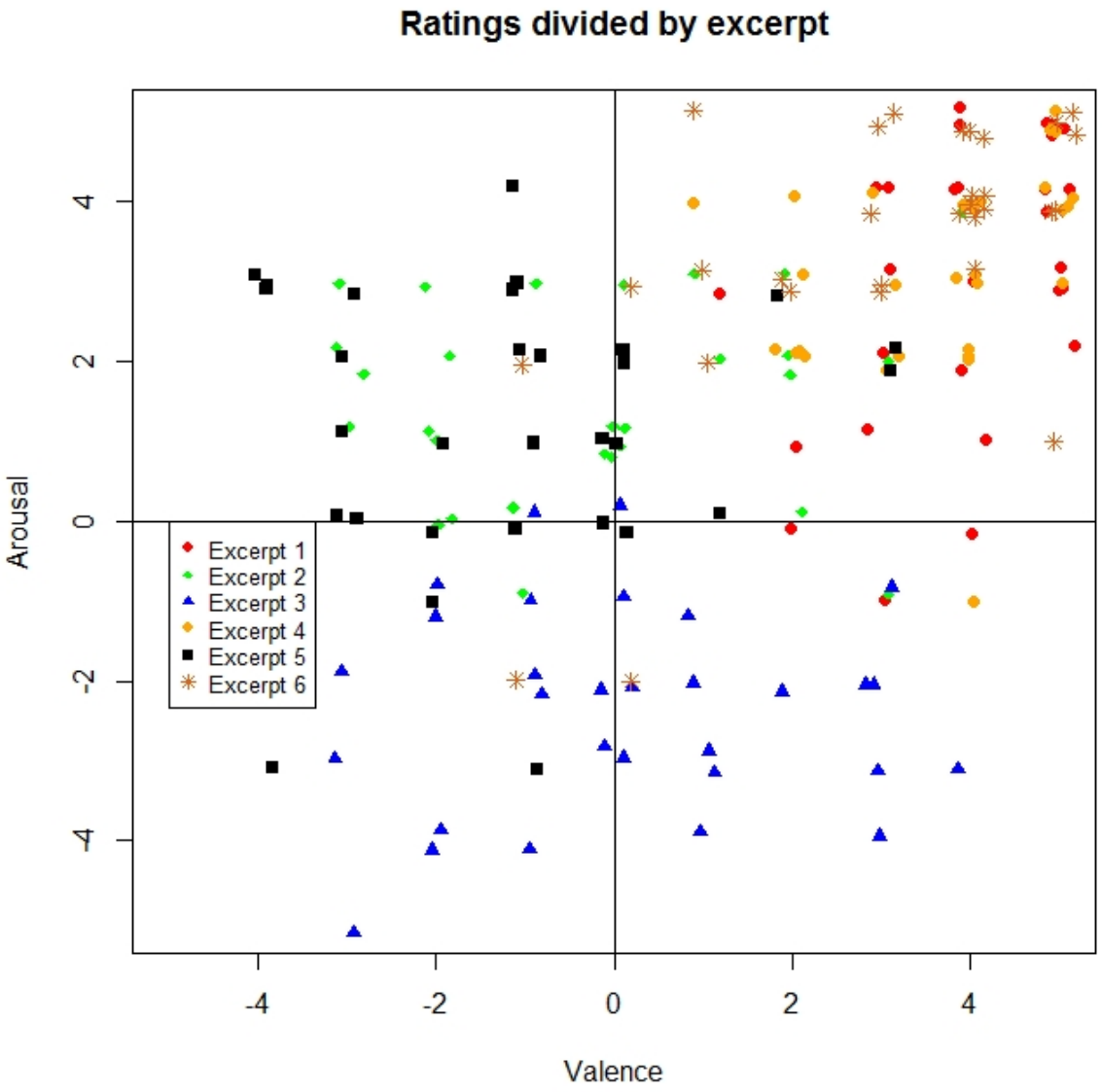


Figure 8.1.1. Scatterplot of VA ratings per excerpt.

	Label	Time period	BPM	Mean pitch height	Mode	Mean interval size	Texture	REP	Mean V	Mean A
Excerpt 1	Happy	Modern	91	4.39	Major	4.57	3	4.40	3.97 (1.10)	3.10 (1.71)
Excerpt 2	Happy	Baroque	128	8.81	minor	2.33	2	5.38	-0.13 (2.06)	1.50 (1.25)
Excerpt 3	Peaceful	Modern	69	-1.16	Major	4.85	2	2.30	0.13 (2.01)	-2.33 (1.27)
Excerpt 4	Peaceful	Baroque	104	9.69	Major	3.88	2	6.18	3.63 (1.19)	3.13 (1.28)
Excerpt 5	Angry	Modern	100	1.27	minor	1.92	5	1.67	-1.20 (1.94)	1.23 (1.72)
Excerpt 6	Angry	Baroque	225	1.98	Major	7.39	3	7.27	3.00 (1.84)	3.43 (1.81)
AVG			119.5 (55.11)	4.16 (4.33)	N/A	4.16 (1.98)	2.83 (1.17)	4.53 (2.20)	1.57 (2.22)	1.68 (2.17)

Table 8.1.1. Levels of compositional cues for each excerpt. Standard deviations for mean valence and arousal per excerpt, as well as for mean values of each cue (AVG row), are shown in brackets.

8.1.1 Valence

Excerpts 1 (Happy, modern), 4 (Peaceful, Baroque), and 6 (Angry, Baroque) were rated higher than the average value for valence. All three of these excerpts were written in the major mode. This suggests that the major mode is a probable predictor for higher valence ratings. On the other hand, Excerpts 2 (Happy, Baroque), 3 (Peaceful, modern), and 5 (Angry, modern) were rated lower than the average value for valence. Two of these excerpts were written in the minor mode – the excerpt in the major mode (excerpt 3, Peaceful, modern) had the highest valence rating of the three, which fell on the positive side of the scale (although only just). Again, this suggests that mode (major vs. minor) is likely to have an impact on valence ratings. Support for this position can be found with Hevner (1935:111-112), and Gagnon and Peretz (2003:33), who found a similar effect in more controlled circumstances.

Excerpts 2 (Happy, Baroque) and 5 (Angry, modern) were rated below the neutral (zero) value for valence. They both had interval sizes smaller than the average (the smallest two values), and were both written in the minor mode. The other four excerpts were rated above neutral for valence. All four were written in the major mode. Three of the four had mean interval sizes above the average, and the other (excerpt 4, Peaceful, Baroque) was not far below the average. The impact of mode on valence has already been discussed. Scherer and Oshinsky (1977:339) found that interval size had a positive impact on valence; none of the other literature surveyed in this research found investigated this effect. The present findings support those of Scherer and Oshinky (1977).

8.1.2 Arousal

Excerpts 1 (Happy, modern), 4 (Peaceful, Baroque), and 6 (Angry, Baroque) had higher than average arousal values, as they did for valence. These excerpts were written in the major mode. Again as for valence, excerpts 2 (Happy, Baroque), 3 (Peaceful, modern), and 5 (Angry, modern) had lower than average arousal scores. Two of these were written in the minor mode. This general pattern suggests that the major mode is associated with higher arousal ratings; however, the major mode excerpt rated below average for arousal had by far the *lowest* arousal rating. The present results are not clear with regards to the possible effect of mode on arousal. No studies surveyed in this research found a clear correlation between mode and arousal.

Only excerpt 3 (Peaceful, modern) was rated below the neutral value for arousal. This excerpt had by far the lowest mean pitch height and BPM. The other five excerpts, rated above the neutral value for arousal, shared no relations besides having a higher mean pitch and BPM than excerpt 3. Excerpt 6 (Angry, Baroque) had the highest arousal rating, as well as the highest REP, BPM, and mean interval size. Taking BPM and REP as measures of tempo, this positive relationship with arousal is completely expected. Tempo was found to contribute positively to arousal by Hevner (1937:626), Scherer and Oshinsky (1977:339), Schubert (2004:573), and Ilie and Thompson (2006:324). The present finding of a positive relationship between mean pitch height and arousal is more equivocal in the literature. Hevner (1937:626) found a negative relationship between pitch height and arousal, while Scherer and Oshinsky (1977:339) found a positive relationship, in line with the present results. Only one study (Scherer & Oshinsky, 1977:339) examined the effects of interval size (pitch variation) on arousal, and found a positive relationship between the two, as hinted at in the present results.

8.1.3 Quadrants

In terms of quadrant placements in the VA space, excerpts 1 (Happy, modern), 4 (Peaceful Baroque), and 6 (Angry, Baroque) were situated (on average) in quadrant 1 (high valence, high arousal). The only element common to these three excerpts was being written in the major mode. Excerpts 2 (Happy, Baroque) and 5 (Angry, modern) were situated in quadrant 2 (low valence, high arousal). These excerpts were both written in the minor mode, and had the lowest mean interval sizes of all excerpts. Excerpt 3 (Peaceful, modern) was situated in quadrant 4 (high valence, low arousal). As stated above, this excerpt had the lowest mean pitch height and BPM. None of the average VA ratings situated any excerpt in quadrant 3.

The "success" of each excerpt at communicating its intended emotional portrayal was assessed. Success was defined by a greater number of ratings falling within the VA quadrant appropriate to the intended emotion portrayal than in any other quadrant. Ratings falling at a zero-point on one dimension (i.e. on the boundary between two quadrants) were counted as 0.5 for both quadrants. A neutral rating on both dimensions was assigned as 0.25 to each quadrant. The results per quadrant per excerpt are summarised in table 8.1.2 below.

	Emotion label	Time period	Expected quadrant	Q1 ratings	Q2 ratings	Q3 ratings	Q4 ratings
Excerpt 1	Happy	Modern	Q1	28	0	0	2
Excerpt 2	Happy	Baroque	Q1	12	12.5	2.5	3
Excerpt 3	Peaceful	Modern	Q4	0.25	0.75	14.25	14.75
Excerpt 4	Peaceful	Baroque	Q4	29	0	0	1
Excerpt 5	Angry	Modern	Q2	6	17.5	5.5	1
Excerpt 6	Angry	Baroque	Q2	26.5	1.5	1.5	0.5

Table 8.1.2. Summary of ratings by quadrant for the excerpts.

Excerpts 1 (Happy, modern) and 5 (angry, modern) clearly portrayed their intended emotions at the quadrant level (happiness and anger, respectively). Excerpts 4 (Peaceful, Baroque) and 6 (Angry, Baroque) had clear emotion portrayals at the quadrant level, but these did not match up to their intended portrayals (both portrayed Happiness instead of their intended emotions). The results for excerpts 2 (Happy, Baroque) and 3 (Peaceful, modern) are more ambiguous. Excerpt 2 portrayed anger slightly more than its intended emotion of happiness. Excerpt 3 technically satisfied the criterion for successful portrayal, but its intended emotion of peacefulness was assigned so marginally higher a rating than sadness that the portrayal cannot be said to be clear. In sum, two of the modern excerpts (portraying happiness and anger) clearly portrayed their intended emotions at the quadrant level. All other excerpts did not succeed in this regard.

For all three emotion portrayals examined, a significant difference was found between the Baroque and Modern excerpts. In terms of accuracy (position in an area of the VA space corresponding to the intended emotion portrayal, assessed in the secondary descriptive analysis of cues), only two excerpts clearly achieved accuracy. These were both Modern excerpts, for the emotion portrayals of Happiness and Anger. The Modern excerpt with the intended emotion

portrayal of Peacefulness was a borderline case. While it technically satisfied the accuracy criterion, it inhabited the area of VA space identified with Sadness to almost the same degree as its intended emotion portrayal. This is in line with the findings of Vieillard *et al.* (2008:727-728) over their Peaceful excerpts. They found that the emotion of Peacefulness was quite often confused with Sadness by their listeners. Quinto *et al.* (2014:519-520) found in their listening experiment that the portrayals of Sadness and Tenderness (included as part of the Peacefulness category in table 6.3.2 above) had a higher reliance on performance cues than compositional cues for their communication. The present experiment only investigated compositional cues; the MIDI rendering and playback of excerpts effectively controlled for the effects of performance cues. If Sadness and Tenderness are reliant on performance cues for their discrimination by listeners, the levels of compositional cues which constituted the present excerpts may be common to both emotion portrayals (at least in terms of their valence and arousal components). This could explain the Modern Peaceful excerpt's high degree of confusion with sadness in the present results.

None of the Baroque excerpts had a clear and accurate emotion portrayal. The Happy Baroque excerpt came close – it inhabited the area associated with its intended emotion portrayal to only a slightly lesser degree than the adjacent area associated with Anger. This may be explained by examining the patterns of cue utilisation of the two Happy excerpts, as well as that of the Angry Modern excerpt, all three of which fell clearly into the area associated with Anger. The arousal rating of the Happy Baroque excerpt was appropriate for its intended emotional portrayal; the ambiguity was located on the valence dimension, where this excerpt was rated slightly negative (its intended emotional portrayal would have had a positive valence rating). The Happy Baroque excerpt shares two cue levels with the Angry Modern excerpt which it does not share with the Happy Modern excerpt. These excerpts both had very small mean interval sizes (2.33 and 1.92, respectively), and were both written in the minor mode. These cue levels were found to be the only two which, together, could differentiate excerpts with a negative value on the valence dimension (see table 8.1.1 above). These findings tie in to the literature: Quinto *et al.* (2014:514) as well as Scherer and Oshinsky (1977:341) found that a minor mode corresponded to portrayals of Anger. Scherer and Oshinsky (1977:339) also found that small interval size was associated with Anger. Hevner (1935:111-112), and Gagnon and Peretz (2003:33) found that the minor mode was negatively associated with valence, and that interval size was positively associated with valence. It is possible that the Happy Baroque excerpt was rated highly in the Angry quadrant as a function of the effect of these cues on its perceived valence.

The higher (near-neutral) valence rating of the Happy Baroque excerpt as compared to the Angry Modern excerpt may be due to its less extreme levels of these two cues (minor mode, small interval size). These two excerpts had by far the smallest mean interval sizes, but this was more pronounced for the Angry Modern excerpt. The same can be said of the mode of the two excerpts – mode was operationalised as a binary variable, but analysis of the score for the two excerpts indicates that the Angry Modern excerpt has a greater percentage of minor harmonies than the Happy Baroque excerpt. As these two cue levels corresponded to negative values for valence, less extreme levels would likely lead to less extremely negative valence ratings, explaining the less-negative valence of the Happy Baroque excerpt in comparison with the Angry Modern excerpt.

The other two Baroque excerpts (for the Peaceful and Angry emotion portrayals) showed clear emotion portrayals which were in the *incorrect* quadrant. Both of these excerpts fell into the area associated with Happiness (quadrant 1). The levels of cue utilisation that these excerpts share with each other and with the Happy Modern excerpt (which clearly and accurately fell into quadrant 1) can be examined to explain this finding. Both the Peaceful and Angry Baroque excerpts have BPM and REP values higher than their Modern counterparts (which both met the criterion for accuracy, albeit only clearly in the latter). They were also written in the major mode. They share this set of levels with the Happy Modern excerpt. The combination of medium or high tempo (as BPM and/or REP) and major mode seems to be a sufficient condition for the ratings of an excerpt falling into quadrant 1 (high valence, high arousal; see table 8.1.1). This echoes the brief description given by Vieillard *et al.* (2008:723-724) of the levels of compositional cues used in composing their excerpts – they composed the Happy excerpts with a fast tempo, in a major mode. Support for the positive association between the portrayal of Happiness and both high tempo and major mode is provided by Quinto *et al.* (2014:514-516) and Scherer and Oshinsky (1977:339-341).

8.1.4 Individual cues

Some correspondence is observed between the present results and the findings reported in table 4.2.2 (reprinted here as table 8.1.3 for the convenience of the reader) in terms of the impact of individual cues on overall valence and arousal ratings. The effects of cues are examined at the broad level of positive versus negative mean ratings on each dimension. Hevner's (1935b, 1936, 1937) results from her series of early experiments in this field, which were parsed into a dimensional model by Schubert (1999:55), will be included in this comparison.

	Scherer & Oshinsky 1977		Schubert 2004		Gagnon & Peretz 2003		Ilie & Thompson 2006		Hevner 1935b, 1936, 1937	
	V	A	V	A	V	A	V	A	V	A
Tempo	+	+	+	+	+			+	+	+
Intensity			+	+			-	+		
Intensity variability	-	-								
Pitch variation	+	+								
Pitch height	-	+					-		+	-
Pitch contour	-		+	+						-
Mode					+				+	
Texture			+	+/-						
Rhythm	N/A	N/A							-	+
Harmonic complexity									-	+

Table 8.1.3. Summary of findings of dimensional studies for the three musical emotion portrayals relevant to the experimental component of the present research. Note: for pitch contour, "+" indicates an ascending contour, "-" a descending contour. For mode, "+" represents the major mode, "-" the minor. For texture, "+" represents a thick texture, "-" a thin texture. For rhythm, "+" represents a firm rhythm, "-" a flowing rhythm.

Tempo

The mean BPM of excerpts with a positive mean valence rating was marginally higher than for those with a negative mean valence ratings (122.25 [SD=70.01] vs 114 [SD=19.80]). The same pattern is observed for REP (5.04 [SD=2.17] vs 3.53 [SD=2.62]). The positive association seen in the present results between tempo (BPM and REP) and valence supports the findings of four out of five studies which examined this effect: Hevner (1937:626), Scherer and Oshinsky (1977:339), Gagnon and Peretz (2003:33), Schubert (2004:573). Ilie and Thompson (2006:324) found no association of tempo with valence in music. Overall, the present result is in line with observations made in the literature.

The only excerpt with a negative mean arousal score had by far the lowest BPM, and had an REP score of 2.30, lower than the average for the other excerpts (4.98, SD=2.13). The findings of the four studies in table 8.1.3 that reported a positive effect of tempo on arousal (Hevner,

1937:626, Scherer and Oshinsky, 1977:339, Schubert, 2004:573; Ilie & Thompson, 2006:324), are supported by the present results.

Interval size

Mean interval size (pitch variation) was only assessed by one study reported in table 8.1.3 (namely, Scherer & Oshinsky, 1977). They reported a positive association between interval size and both valence and arousal. This pattern was seen in the present results for positive versus negative valence (mean interval sizes of 5.17 [SD=1.53] vs 2.13 [SD=0.29], respectively). For arousal, the mean interval sizes associated with positive vs negative arousal (4.02 [SD=2.18] and 4.85, respectively) did not match Scherer and Oshinsky's results.

Pitch height

Mean pitch height was on average lower for excerpts rated on the positive side of the valence dimension than the negative side (3.73 [SD=4.58] vs 5.04 [SD=5.33]), supporting the findings of two of the three studies in table 8.1.3 that assessed this cue for valence. Ilie and Thompson (2006:324) found a negative association between pitch height and valence, as did Scherer and Oshinsky (1977:339). Hevner (1937:626) found the opposite effect. Overall, the present findings are taken to fit in with those of the literature. For arousal, mean pitch height was markedly higher over excerpts with a positive than a negative arousal rating (5.23 [SD=3.86] vs -1.16). This supports Scherer and Oshinsky's (1977:339) results, but not Hevner's (1937:626).

Mode

Mode was assessed by only two studies reported in table 8.1.3, and results can only be compared for the valence dimension. Both studies (Gagnon & Peretz, 2003:33; Hevner, 1937:626) associated the major mode positively with valence, and the minor mode negatively. The present results support this unambiguously: the excerpts with positive mean valence scores were all written in the major mode, and those with negative mean valence were written in the minor mode. No clear association between mode and mean arousal rating can be seen in the present results. Texture did not show any clear patterns of use over the valence or arousal dimensions.

The accuracy of emotion portrayals across the excerpts has been discussed, as have the patterns of cue use that may account for the ratings gathered in the experiment. The next chapter will discuss the results of the experiment in general.

8.2 Comparison with Vieillard *et al.* (2008)

The modern excerpts' valence and arousal ratings in the present experiment can be compared to those generated in Vieillard *et al.*'s paper, where these excerpts were first tested. Excerpts 1, 3, and 5 from the present experiment are labelled in Vieillard *et al.*'s paper as *G03*, *A07*, and *P07*, respectively (see Vieillard *et al*, 2008:744-752 – Appendix 2). They utilised a 10-point rating scale for their experiment, ranging from 0 to 9, whereas the present experiment was carried out with an 11-point scale, from -5 to 5. Vieillard *et al.*'s results were transformed to a rating in the format of the present experiment (11-point, -5 to 5 scale), for ease of comparison. The result is shown in table 8.2.1. Note that Vieillard *et al.* did not provide standard deviations for their mean valence and arousal values per excerpt.

	Valence		Arousal	
	Present Results	Vieillard *et al.*, 2008	Present Results	Vieillard *et al.*, 2008
Excerpt 1 (Happy)	3.97 (SD=1.10)	2.44	3.10 (SD=1.71)	3.56
Excerpt 3 (Peaceful)	0.13 (SD=2.01)	2.67	-2.33 (SD=1.27)	-2.78
Excerpt 5 (Angry)	-1.20 (SD=1.94)	-1.22	1.23 (SD=1.72)	3.22

Table 8.2.1. Comparison of VA ratings from the present experiment with Vieillard et al.*'s (2008) results. All results are expressed in an 11-point (-5 to 5) rating scale format.*

As mentioned above, all three of the Modern excerpts achieved accuracy in the present experiment (at least *technically*), and so the mean valence and arousal ratings match with those of Vieillard *et al.* in terms of sign (indicating which side of the neutral point the mean value fell). Two mean ratings stand out, however. Mean valence for Excerpt 3 and mean arousal for Excerpt 5 differed by near to two points between the present results and those reported in Vieillard *et al.*'s paper. It is not clear why this was the case. Perhaps the sheer number of excerpts listeners were exposed to in Vieillard *et al.*'s experiment (all listeners heard the full set of 56 excerpts – Vieillard *et al.*, 2008:725-726), as well as the character of these excerpts, caused a shift in their baselines for valence and arousal such that it differed from listeners' baselines in the present experiment.

The large difference in arousal ratings for Excerpt 5 is particularly interesting in light of the effect of musicianship on arousal ratings found in the previous chapter. Vieillard *et al.* (2008:724-726) did not control for musical training in the selection of their participants. They did collect data on this, however, reporting that 32 out of their total of 59 participants had no musical training. No further mention is made of which of these participants performed the VA

rating experiment (20 of the 59 participants took part in this experiment). It was found in the previous chapter that musicians made higher ratings of arousal than non-musicians, overall. It can be speculated that this effect may have been particularly pronounced in Excerpt 5, in which case the difference in arousal values for this excerpt between the present experiment and that of Vieillard *et al.* could be due to their possible inclusion of a higher number of musicians than non-musicians in their VA rating task.

8.3 Comparison with Descartes' predictions

The present results can be compared with Descartes' broad predictions of the effects of tempo and interval size on musical affect. He states that slower tempo gives rise to "quieter" affects, which he defined as including languor (interpreted here as close to peacefulness), sadness, and fear. He links a faster tempo with "faster" emotions, of which the example he gives is joy (Vlock, 1998:524). Of intervals, a wider size is associated by Descartes with happy emotions, and a narrow size with sad emotions (Vlock, 1998:521). There is some correspondence between his statements and the present results. The mean BPM of the excerpts falling in quadrant 1 (joy) was higher than that of the others (140 [SD=79.90] vs 99 [SD=29.51]). The same was true for mean REP (5.95 [SD=1.45] vs 3.12 [SD=1.99]). These two measures of tempo therefore accord in the present results to Descartes' predictions. This holds as well if "faster" and "quieter" emotions are interpreted as positive and negative arousal values, respectively – the only excerpt from the present experiment which had a negative arousal value (Excerpt 3 – Peaceful, modern) had the lowest BPM and second-lowest REP. To assess Descartes' predictions for interval size, "happy emotions" are taken to mean emotions portrayals falling on the positive side of the valence dimension, and "sad emotions" to correspond to the negative side of the same dimension. The average mean interval size for excerpts with a positive mean valence score was higher than for those with a negative mean valence score (5.17 [SD=1.53] vs 2.13 [SD=0.29]), supporting Descartes' assertion. It is interesting to note that correspondence with both of Descartes' predictions is clearer in the Baroque than the Modern excerpts.

8.4 Possible explanations for results

As discussed above, none of the Baroque excerpts were rated in the expected area of the VA space (although the Happy excerpt came very close). There are a number of possible reasons why the Baroque excerpts were not rated accurately, two of which stand out. The first is the necessarily imperfect method used to identify the emotional portrayals of these excerpts. Unlike the excerpts presented by Vieillard *et al.*, the Baroque excerpts were not labelled with

a clear and unambiguous term to describe their intended emotional portrayal. In describing his excerpts, Heinichen makes use of various often-conflicting concepts, and the selection process may have drawn out the incorrect one. Selection was made based on a close reading of the descriptive text, *without reference to the written score*, so as to avoid bias from modern musical ideas.

The second contending explanation is that Heinichen was simply writing for different ears. We have no way of looking into the mind of a music listener in the 18th century; we cannot know what emotional connotations this music held for such a listener. It is quite possible that an 18th-century listener would have rated the Baroque excerpts in the expected areas. There are further, historically grounded factors to consider in understanding the musical soundscape of the Baroque. Important structural factors have changed in Western music over the past few hundred years. Standard pitch height was highly variable in the Baroque era: the historical tuning which is most commonly known today is 415Hz (as opposed to our modern 440Hz, at which the MIDI excerpts were rendered). Other tunings both above and below our current standard were frequently employed (Haynes & Cooke, 2001). We do not know for which standard pitch height Heinichen intended these excerpts. Pitch height has been shown to have an effect on musical emotion recognition, *when the musical material is held constant* (e.g. Quinto *et al.*, 2014; Hevner, 1937). This may theoretically be generalised to standard tunings – in the case of a musical excerpt played at different standard tunings, the musical content remains invariant, but the mean pitch height will differ. Although the prior research has shown this effect only for much larger pitch differences (see e.g. Hevner, 1937:623, where differences were a fifth and an octave; a standard tuning of 415Hz would only create a difference of one semitone in comparison to the modern 440Hz standard), this factor may go some way towards explaining the present results.

Another example of a structural parameter that has changed in Western music since the 18th century is temperament. Equal temperament is now the Western standard, and listeners are highly accustomed to this through everyday exposure. In the 18th century, many different tuning systems were in common use in Europe, each with its own set of acoustic parameters totally unique from equal temperament (Lindley, 2001). It is theoretically possible that temperament may modulate emotion judgments. Heinichen does not indicate a preferred temperament for his excerpts. They were rendered in equal temperament for the present experiment. It is *possible*, though no empirical evidence exists to support this claim, that this

difference in temperament may have skewed the ratings by deviating from the sonic baseline Heinichen had in mind when composing the excerpts.

Timbre may also have played a role in the mismatch of Baroque excerpts with their expected areas of VA space. All the excerpts were rendered using Sibelius' piano timbre. Vieillard *et al.* (2008:725) originally tested their excerpts as rendered in a piano timbre. Heinichen indicates instrumentation for his excerpts inconsistently. Timbre has sometimes been found to play a role in musical emotion portrayal (e.g. Balkwill & Thompson, 1999; Juslin & Laukka, 2003). It is possible that, despite not explicitly prescribing instrumentation for all of his excerpts, Heinichen may have had a specific set of timbres in mind when composing them. The use of a timbre not intended for the Baroque excerpts may have impacted on the VA ratings ascribed to them by listeners.

A further consideration, as already mentioned, is the effect of translation and time on the meaning of words. The intended emotional portrayals as extracted from Heinichen's descriptions for the present research may have misrepresented his expressive intentions, due to limitations in translation. The descriptions were being assessed in a different language to that in which they were originally written, and with a time difference of nearly 300 years between the date of publication of *Der General-Bass in der Composition* and the present research. It is difficult to determine the extent of possible errors in translation, and it cannot be known whether they did in fact impact the present results. The possibility, at least, is present.

The present experiment was designed to investigate compositional cues only. Performance cues were controlled for in virtue of the excerpts being rendered in MIDI. This was done in part as a practical measure – Vieillard *et al.* wrote their excerpts to be played back electronically, without any performance cues included, and Heinichen also does not systematically indicate performance cues in his excerpts. The exclusion of performance cues leads to a historical disconnect, and thus a possible violation of ecological validity. That is, there are important expressive cues that have been eliminated from this (and the other quoted) studies, in favour of experimental control. While he does not specifically indicate these in his excerpts, Heinichen may well have assumed that any performer playing his excerpts would have understood the emotional content, and made appropriate use of performance cues to enhance its portrayal. Quinto *et al.* (2014:519-520) assessed the relative reliance on performance and compositional cues of different emotional portrayals. They found that anger had a heavier leaning on performance cues, as did tenderness (but this also made use of compositional cues). Happiness

could be communicated well through both. Of the three Baroque excerpts used in the present experiment, the Happy excerpt had mean VA values closest to its intended quadrant (though it was still inaccurate by 0.13 [SD=2.06] on one dimension). The Peaceful and Angry Baroque excerpts both deviated from their intended quadrant (each on only one dimension) by a far greater margin than this (3.13 [SD=1.28] and 3.00 [SD=1.84], respectively). This lends support to the necessity of performance cues in accurate communication of Heinichen's excerpts – the emotion portrayal with the highest leaning on *compositional* cues was the best-decoded of the three.

The results of the experiment have now been presented and discussed. The final chapter summarises the results in terms of the primary research questions, and reflects on the experiment while offering suggestions for further research.

CHAPTER 9. CONCLUSION

9.1 Assessment of principal research questions

This thesis set out to investigate two principal research questions: first, whether Baroque and modern musical excerpts are rated differently by listeners within emotion categories; and second, how the compositional cues of tempo, pitch height, mode, interval size, and texture relate to these ratings, both in terms of emotion labels and the dimensions of valence and arousal. It was found that the excerpts did indeed differ within emotion categories as a function of time period. This was explained in terms of the method of categorisation of the Baroque excerpts within emotion categories, the likely role of performance cues in the communication of emotion in Baroque music, and the difference in historical context between the time in which these excerpts were written and the present day. Issues of translation and shifting meanings of words over time may also have played a role. In terms of the second question, it was found that the levels of compositional cues within the present excerpts related to emotion categories and VA ratings in ways mostly supported by the existing literature.

Interpreted very generously, the results of the present experiment provide a preliminary indication that emotional expression in music was conceived of differently in the Baroque era, and is therefore variable across time even within the narrow confines of Western music. However, the above-mentioned issues of categorisation, historical context, and translation of words (both between languages and over time) strongly discourage a generalisation of the present results in this context. These interfering factors present sizeable obstacles to comparative research on emotional portrayal in Baroque music. Baroque theorists did not make the task easy, giving lengthy and detailed description of emotional portrayals rather than a simple label. Extracting a single emotion label from such a description is a difficult and highly subjective process. It is worth questioning whether this process is even valid – perhaps the extraction of simple emotion labels from these in-depth and rich descriptions is inappropriate to the way in which theorists at the time thought about musical emotion. It is entirely possible that, if we were able to ask him, Heinichen would not be able to summarise his music-emotional ideas in a word; or if he could, perhaps he would be unwilling, as such a word would not fully capture the nuance of his intended emotional portrayal.

9.2 Reflection and suggestions for further research

The utility of this research in terms of Baroque performance practice lies primarily in the understanding of the difference in listener contexts between the modern concert setting and the

one for which the music was originally composed. It is not enough to acknowledge this difference without attempting to bridge the gap. Historically informed performance practitioners tend to have no difficulty in identifying the intended affective content of a piece of music, but bringing this content across to the listener is a more complex task. By exploring in concrete terms the relationship between Baroque and modern affective portrayals, it is possible to equip performers with a set of tools to use in "updating" the music to have its intended effect on contemporary ears. For instance, if a Baroque work that was intended to sound *angry* contains a set of compositional-structural cues more likely to sound *joyful* to contemporary listeners, performers could, if they wished, compensate for this disconnect through the targeted use of performance cues.

Heinichen's *Einleitung* was chosen as the source of the Baroque excerpts as it was expected to most clearly represent examples of emotional portrayal in Baroque music. This idea was, however, hamstrung by the complexity of Heinichen's descriptions of his excerpts. An alternative approach would perhaps be to find examples in the vocal music of other Baroque composers. Much vocal music of the time was written to express and enhance the affective portrayal of the text (LeCoat, 1972:221). A highly proficient linguist and expert on Baroque rhetoric could potentially extract from the text its primary expressive idea – it would be necessary to locate appropriate texts, in which a single and clear emotional idea is dominant. Taking this interpretation to represent the intended affective portrayal of the music, the present experimental design could be replicated, perhaps yielding results grounded in a less equivocal interpretation of intended affective portrayal.

Another option would be to vastly increase the number of stimuli used. This could be done in accordance with the above-mentioned procedure for the categorisation of affective portrayal, with a slight modification. Instead of necessitating the extraction of a single emotional idea from a given text, a large pool of stimuli could be assembled in which each excerpt was associated with a list of potential emotional portrayals. This would allow for the use of music written to texts that do not have a single, clear emotional idea at their core. By conducting cue analysis on these excerpts, and comparing the results with the large set of affective words identified over the excerpts, correlations could be traced between cue levels and certain words. For instance, if a large interval size were present to a statistically significant degree in excerpts from which the emotion idea of "jubilation" had been extracted (among other words), and this cue level was not significantly associated with any other idea in this manner, it could be inferred that that particular idea was characterised in part by that level of cue utilisation. If this

procedure proves successful, the extracted cue profiles could be compared to those of modern music written to portray similar emotions. This would facilitate a direct comparison of the compositional means employed to portray different emotions across these two time periods, which was not possible in the present research.

The question of the exact nature of "affects" as the object of musical expressive intention in the Baroque is one that requires further investigation. If affects do not fully correspond to any modern definitions of emotions, the scope of the comparison between Baroque and modern music will have to be expanded, or the comparison may have to be abandoned altogether. The present research has assumed that there is enough correspondence between affects and emotions to allow for both to be worked with in the same terms. This assumption was made predominantly based on the high degree of correspondence between modern definitions of emotions and Descartes' definition of the passions, which greatly informed the Doctrine of the Affections, upon which much of Baroque musical expression was built (Whitfield, 2010:16-17). It is hoped that this assumption holds up. If not, alternatives will need to be sought if comparative research on Baroque musical emotion is to continue.

The literature on cue utilisation in musical emotional portrayals is a rich area of knowledge, with a great deal of research having been conducted on both compositional and performance cues. The present research did not aim to assess cue utilisation inferentially, only using the literature as a point of comparison against which to describe the present results. In retrospect, it would have been profitable to control for these cues, and design an inferential model to examine their use. The results of the present experiment are for the most part in line with the literature – however, the literature itself is not without disagreements, and the present results could have contributed further data to this area. Given the small pool of Baroque excerpts from which experimental stimuli were able to be drawn, however, it would have been difficult, if not impossible, to produce such a design without looking elsewhere for further stimuli.

One might argue that the choice of "happy" and "sad" as anchoring words for the valence dimension in the present experimental instrument may be unwise. The majority of the extant literature defines the dimension as a continuum of "pleasant-unpleasant", and the terms "happy" and "sad" have often appeared firmly in quadrants 1 and 3 respectively, rather than falling on the valence axis, though they are often close (e.g. Russell, 1980). It is unclear whether the present terms facilitate a similar mode of judgement in participants in comparison to the standard conventions. This is a question that would need further investigation, possible in the

form of an experiment utilising differently anchored valence dimensions for self-report to the same musical stimuli by different participants. The broad level of correspondence, at least at the quadrant level, between the present results and those of Vieillard *et al.* (2008, who used "pleasant" and "unpleasant" as valence anchors), suggests that there was a high degree of correspondence between the rating scales. However, it should be noted that the mean difference, assessed between the present results and those of Vieillard *et al.*, in the mean ratings given for each excerpt, was higher for valence (1.36) than for arousal (0.97). The present experiment anchored the arousal dimension using the terms "sleepy" and "energetic", while Vieillard *et al.* used "relaxing" and "stimulating". If the dimensions are not equivalent, the generalisability of the present results in terms of the existing literature is limited. Its value would then lie primarily in demonstrating the discrepancy in listener ratings between the Baroque and Modern excerpts in terms of the rating scales used, and the descriptive analysis of cue profiles for levels on these scales.

9.3 Concluding remarks

The present research represents a tentative foray into comparative analysis of emotion portrayal in Baroque and modern music. This area of research has the potential to reveal data which would be of great practical use to performers of Baroque music, as well as of interest to music historians. An empirically-grounded understanding of techniques of emotional portrayal in Baroque music could assist performers in making historically appropriate musical choices. Once such an understanding has been established, further research could be conducted to investigate the ways in which contemporary listeners interpret music written in line with these techniques. There is likely to be both similarity and difference in the way this music is interpreted today. An understanding of this phenomenon would be of use to performers in navigating the complex intersection of historical authenticity and expressive efficacy in the modern musical context.

BIBLIOGRAPHY

Abelson, R. P., and Sermat, V. 1962. Multidimensional Scaling of Facial Expressions. *Journal of Experimental Psychology*, 63(6):546-554.

Badura-Skoda, E., Jones, A. V., and Drabkin, W. Cadenza. In D. Root (ed.) *Grove Music Online*. [Online]. Accessed 10 October 2019. Available: https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000043023.

Balkwill, L., and 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.

Banse, R. and Scherer, K. R. 1996. Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, 70(3):614-636.

Bartel, L. R. 1992. The Development of the Cognitive-Affective Response Test – Music. *Psychomusicology: A Journal of Research in Music Cognition*, 11(1):15-26.

Bigand, E., and Poulin-Charronnat, B. 2006. Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100(1):100-130.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, F., and 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.

Blasius, L. D. 1996. *Schenker's Argument and the Claims of Music Theory*. Cambridge: Cambridge University Press.

Buelow, G. J. 1986a. Biography. In Heinichen, J. D. *Thorough-Bass Accompaniment according to Johann David Heinichen*. G. J. Buelow (ed.). Michigan: UMI Research Press:1-15.

Buelow, G. J. 1986b. Heinichen's *Einleitung* to the *General-Bass* Treatise: A translation. In Heinichen, J. D. *Thorough-Bass Accompaniment according to Johann David Heinichen*. G. J. Buelow (ed.). Michigan: UMI Research Press:307-308.

Buelow, G. J. 2001a. Mattheson, Johann. In D. Root (ed.), *Grove Music Online*. [Online]. Accessed 26 October 2019. Available:

www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000018097.

Buelow, G. J. 2001b. Affects, theory of the. In D. Root (ed.), *Grove Music Online*. [Online]. Accessed 26 October 2019. Available: www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000000253.

Carter, T., and Chew, G. 2001. Monteverdi [Monteverde], Claudio. In D. Root (ed.), *Grove Music Online*. [Online]. Accessed 26 October 2019. Available: www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000044352.

Coorevits, E. and Moelants, D. 2016. Tempo in Baroque Music and Dance. *Music Perception: An Interdisciplinary Journal*, 33(5):523-545.

Descartes, R. 1985. The Passions of the Soul. In J. Cottngham, R. Stoothoff, and D. Murdoch (trans.), *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press:326-404.

Dissmore, J. L. 2017. Baroque Music and the Doctrine of Affections: Putting the Affections into Effect. *The Research and Scholarship Symposium*. [Online]. Accessed 9 October 2019. Available: http://digitalcommons.cedarville.edu/research_scholarship_symposium/2017/podium_presentations/18.

Duncan, G. 2000. Mind-Body Dualism and the Biopsychosocial Model of Pain: What Did Descartes Really Say? *Journal of Medicine and Philosophy*, 25(4):485-513.

Edgerton, H. A., and Paterson, D. G. 1926. Table of standard errors and probable errors of percentages for varying numbers of cases. *Journal of Applied Psychology*, 10(3):378-391.

Eerola, T., Larillot, O., and Toiviainen, P. 2009. Prediction of Multidimensional Emotional Ratings in Music from Audio using Multivariate Regression Models, in *10th International Society for Music Information Retrieval Conference*. Kobe, Japan:621-626.

Eerola, T. and Vuoskoski, J. K. 2011. A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1):18-49.

Eerola, T, and 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.

Ekman, P. 1992. An Argument for Basic Emotions. *Cognition and Emotion*, 6(3/4):169-200.

Ekman, P. 2016. What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*, 11(1):31-34.

Evans, P., and Schubert, E. 2008. Relationships between expressed and felt emotions in music. *Musicae Scientiae*, 12(1):75-99.

Fehr, B., & Russell, J. A. 1984. Concept of Emotion Viewed from a Prototype Perspective. *Journal of Experimental Psychology: General*, 113(3):464-486.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., and Ellsworth, P. C. 2007. The World of Emotions Is Not Two-Dimensional. *Psychological Science*, 18(12):1050-1057.

French, A., Macedo, M., Poulsen, J., Waterson, T., and Yu, A. 2008. *Multivariate Analysis of Variance (MANOVA).* [Online]. Accessed 30 October 2019. Available: http://online.sfsu.edu/efc/classes/biol710/manova/manovanewest.htm.

Frijda, N. H., Kuipers, P., and ter Schure, E. 1989. Relations Among Emotion, Appraisal, and Emotional Action Readiness. *Journal of Personality and Social Psychology*, 57(2):212-228.

Gabrielsson, A. 2001. Emotions in Strong Experiences with Music. In P. N. Juslin and J. A. Sloboda (eds), *Series in Affective Science. Music and Emotion: Theory and Research*. New York: Oxford University Press.

Gabrielsson, A., and 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.

Gabrielsson, A., and Lindström Wik, S. 2003. Strong experiences related to music: A descriptive system. *Musicae Scientiae*, 7(2):157-217.

Gagnon, L. and Peretz, I. 2003. Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition and* Emotion, 17(1):25-40.

Goudbeek, M., and Scherer, K. Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America* 128:1322-1336.

Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. 2007a. Emotions Over Time: Synchronicity and Development of Subjective, Physiological, and Facial Affective Reactions to Music. *Emotion*, 7(4):774-788.

Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. 2007b. Listening To Music As A Re-Creative Process: Physiological, Psychological, And Psychoacoustical Correlates Of Chills And Strong Emotions. *Music Perception: An Interdisciplinary Journal*, 24(3):297-314.

Haynes, B., and Cooke, P. 2001. Pitch. In D. Root (ed.), *Grove Music Online*. [Online] Accessed 25 October 2019. Available: www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000040883.

Heinichen, J. D. 1986. *Thorough-Bass Accompaniment according to Johann David Heinichen*. G. J. Buelow (ed., trans.). Michigan: UMI Research Press.

Hevner, K. 1935a. Expression in Music: A Discussion of Experimental Studies and Theories. *Psychological Review, 42*(2):186-204.

Hevner, K. 1935b. 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.

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

Hupont, I., Baldassarri, S., and Cerezo, E. 2013. Facial emotional classification: from a discrete perspective to a continuous emotional space. *Pattern Analysis and Applications*, 16(1):41-54.

Ilie, G., and Thompson, W. F. 2006. A Comparison of Acoustic Cues in Music and Speech for Three Dimensions of Affect. *Music Perception: An Interdisciplinary Journal*, 23(4):319-330.

Juslin, P. N. 1997a. Emotional Communication in Music Performance: A Functionalist Perspective and Some Data. *Music Perception: An Interdisciplinary Journal*, 14(4):383-418.

Juslin, P. N. 1997b. Perceived Emotional Expression in Synthesized Performances of a Short Melody: Capturing the Listener's Judgment Policy. *Musicae Scientiae,* 1(2):225-256.

Juslin, P. N. 1997c. Can Results from Studies of Perceived Expression in Musical Performances be Generalized across Response Formats? *Psychomusicology*, 16:77-101.

Juslin, P. N., and Laukka, P. 2003. Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? *Psychological Bulletin*, 129(5):770-814.

Julsin, P. N., and 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.

Juslin P. N., and Sloboda, J. A. 2001. Music and emotion: introduction. In P. N. Juslin and J. A. Sloboda (eds), *Music and emotion: theory and research*. New York: Oxford University Press:3-22.

Kallinen, K., and Ravaja, N. 2006. Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, 10(2):191-213.

Kantor-Martynuska, J., and Horabik, J. 2015. Granularity of Emotional Responses to Music: The Effect of Musical Expertise. *Psychology of Aesthetics, Creativity, and the Arts*, 9(3):235-247.

Kim, J. H. 2013. Shaping and Co-Shaping Forms of Vitality in Music: Beyond Cognitivist and Emotivist Approaches to Musical Expressiveness. *Empirical Musicology Review*, 8(3-4):162-173.

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

Kivy P. 1984. *Sound and Semblance: Reflections on Musical Representation*. New Jersey: Princeton University Press.

Kivy, P. 1990. *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca: Cornell University Press.

Kivy, P. 1993. *The Fine Art of Repetition: Essays in the Philosophy of Music*. Cambridge: Cambridge University Press.

Kivy, P. 2001. *New Essays on Musical Understanding*. New York: Oxford University Press.

Kivy, P. 2002. *Introduction to a Philosophy of Music*. New York: Oxford University Press.

Kleinginna, P. R., and Kleinginna, A. M. 1981. A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition. *Motivation and Emotion*, 5(4):345-379.

Konečni, V. J. 2008. Does Music Induce Emotion? A Theoretical and Methodological Analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2):115-129.

LeCoat, G. G. 1972. Comparative Aspects of the Theory of Expression in the Baroque Age. *Eighteenth-Century Studies*, 5(2):207-223.

Leman, M., Vermeulen, V., De Voogdt, L., Moelants, D., and Lesaffre, M. 2005. Prediction of Musical Affect Using a Combination of Acoustic Structural Cues. *Journal of New Music Research*, 34(1):39-67.

Lench, H. C., Flores, S. A., and Bench, S. W. 2011. Discrete Emotions Predict Changes in Cognition, Judgment, Experience, Behavior, and Physiology: A Meta-Analysis of Experimental Emotion Elicitations. *Psychological Bulletin*, 137(5):834-855.

Lindley, M. 2001. Temperaments. In D. Root (ed.), *Grove Music Online*. [Online] Accessed 25 October 2019. Available: www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000027643.

Livingstone, S. R., and Brown, A. R. 2005. Dynamic Response: Real-Time Adaptation for Music Emotion, in *2nd Australasian conference on Interactive Entertainment*. Sydney, Australia:105-111.

Livingstone, S. R., Mühlberger, R., Brown, A. R., and Loch, A. 2007. Controlling musical emotionality: an affective computational architecture for influencing musical emotions. *Digital Creativity*, 18(1):43-53.

Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., Ray, R., and Uchida, H. 2002. American-Japanese cultural differences in judgements of emotional expressions of different intensities. *Cognition and Emotion*, 16(6):721-747.

Mattheson, J. 1981. *Johann Mattheson's* Der Volkommene Capellmeister: *A Revised Translation with Critical Commentary*. E. C. Harriss (ed., trans.). Michigan: UMI Research Press.

Mehu, M., and Scherer, K. R. 2015. Emotion Categories and Dimensions in the Facial Communication of Affect: An Integrated Approach. *Emotion*, 15(6):798-811.

Mesquita, B., Frijda, H., and Scherer, K. R. 1997. Culture and Emotion. In J. W. Berry, P. R. Dasen, and T. S. Saraswathi (eds), *Handbook of Cross-Cultural Psychology, Volume 2: Basic Processes and Human Development*:255-297. Boston: Allyn and Bacon.

Mikels, J.A., Fredrickson, B. L., Larkin, G.R., Lindberg, C.M., Maglio, S.J., and Reuter-Lorenz, P.A. 2005. Emotional category data on images from the International Affective Picture System. *Behavior Research Methods*, 37(4):626-630.

Morgan, R. L., and Heise, D. 1988. Structure of Emotions. *Social Psychology Quarterly*, 51(1):19-31.

Neumann, R., and Strack, F. 2000. "Mood Contagion": The Automatic Transfer of Mood Between Persons. *Journal of Personality and Social Psychology*, 79(2):211-223.

Ortony, A., and Turner, T. J. 1990. What's Basic About Basic Emotions? *Psychological Review*, 97(3):315-331.

Palisca, C. V. 2001. Prima pratica. In D. Root (ed.) *Grove Music Online*. [Online]. Accessed 26 October 2019. Available: www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000022350.

Peretz, I. 2001. Listen to the Brain: A Biological Perspective on Musical Emotions. In P. N. Juslin and J. A. Sloboda (eds), *Series in Affective Science. Music and Emotion: Theory and Research*. New York: Oxford University Press.

Peretz, I., and Coltheart, M. 2003. Modularity of music processing. *Nature Neuroscience*, 6(7):688-691.

Quinto, L., Thompson, W. F., and 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.

Ramos, D., Bueno, J.L.O., and Bigand, E. 2010. 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. [Online]. doi:10.1590/S0100-879X2010007500148.

Reisenzein, R. 1994. Pleasure-Arousal Theory and the Intensity of Emotions. *Journal of Personality and Social Psychology*, 67(3):525-539.

Reybrouck, M., and Eerola, T. 2017. Music and Its Inductive Power: A Psychobiological and Evolutionary Approach to Musical Emotions. *Frontiers in Psychology* 8:494. [Online]. doi: 10.3389/fpsyg.2017.00494.

Richins, M. L. 1997. Measuring Emotions in the Consumption Experience. *Journal of Consumer Research*, 24(2):127-146.

Rickard, N. S. 2004. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, 32(4):371-388.

Rigg, M. G. 1940. Speed as a Determiner of Musical Mood. *Journal of Experimental Psychology*, 27(5):566-571.

Rochberg, G. 1997. Polarity in Music: Symmetry and Asymmetry and Their Consequences. *Proceedings of the American Philosophical Society*, 141(2):169-176.

Rosch, E. 1978. Principles of Categorization. In E. Margolis and S. Laurence (eds), *Concepts: Core Readings*:189-206. Cambridge: MIT Press.

Russell, J. A. 1980. A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6):1161-1178.

Russell, J. A., and Feldman Barrett, L. 1999. Core Affect, Prototypical Emotional Episodes, and Other Things Called *Emotion*: Dissecting the Elephant. *Journal of Personality and Social Psychology*, 76(5):805-819.

Scarantino, A., and Griffiths, P. 2011. Don't Give Up on Basic Emotions. *Emotion Review*, 3(4):444-454.

Schaefer, E. S., and Plutchik, R. 1966. Interrelationship of Emotions, Traits, and Diagnostic Constructs. *Psychological Reports*, 18(2):399-410.

Scherer, K. R. 1972. Acoustic Concomitants of Emotional Dimensions: Judging Affect from Synthesized Tone Sequences. Unpublished paper presented at the Eastern Psychological Association Meeting. 27-29 April, Boston.

Scherer, K. R. 1986. Vocal Affect Expression: A Review and a Model for Future Research. *Psychological Bulletin*, 99(2):143-165.

Scherer, K. R. 2000. Psychological Models of Emotion. In J. C. Borod (ed.), *The Neuropsychology of Emotion*. New York: Oxford University Press:137-162.

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.

Scherer, K. R. and Oshinsky, J. S. 1977. Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, 1(4):331-346.

Scherer, K. R., and Zentner, M. R. 2001. Emotional Effects of Music: Production Rules. In P. N. Juslin and J. A. Sloboda (eds), *Music and emotion: theory and research*. New York: Oxford University Press:361-392.

Schlosberg, H. 1952. The Description of Facial Expressions in Terms of Two Dimensions. *Journal of Experimental Psychology*, 44(4):229-237.

Schubert, E. 1999. *Measurement and Time Series Analysis of Emotion in Music*. Unpublished PhD thesis: University of New South Wales.

Schubert, E. 2004. Modeling Perceived Emotion With Continuous Musical Features. *Music Perception: An Interdisciplinary Journal*, 21(4):561-585.

Shaver, P., Schwartz, J., Kirson, D., and O'Conner, C. 1987. Emotion Knowledge: Further Exploration of a Prototype Approach. *Journal of Personality and Social Psychology*, 52(6):1061-1086.

Sloboda, J. A., and Juslin, P. N. 2001. Psychological perspectives on music and emotion. In P. N. Juslin and J. A. Sloboda (eds), *Music and emotion: theory and research*. New York: Oxford University Press:71-104.

Smith, C. A., and Ellsworth, P. C. 1985. Patterns of Cognitive Appraisal in Emotion. *Journal of Personality and Social Psychology*, 48(4):813-838.

Spencer, H. 1985. *The Principles of Psychology, Volume 1*. New York: D. Appleton and Company.

Stoothoff, R. 1985. Preface. In Descartes, R. The Passions of the Soul. In J. Cottngham, R. Stoothoff, and D. Murdoch (trans.), *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press:325.

Turing, A. M. 1950. Computing Machinery and Intelligence. *Mind*, 59(236):433-460.

van Reekum, C. M., and Scherer, K. R. 1997. Levels of Processing in Emotion-Antecedent Appraisal. *Advances in Psychology*, 124:259-300.

Vempala, N. N., and Russo, F. A. 2018. Modeling Music Emotion Judgments Using Machine Learning Methods. *Frontiers in Psychology* 8:2239. [Online]. doi: 10.3389/fpsyg.2017.02239.

Vlock, D. M. 1998. Sterne, Descartes, and the Music in Tristram Shandy. *Studies in English Literature, 1500-1900*, 38(3):517-536.

Vieillard, S., Didierjean, A., and Maquestiaux, F. 2012. Changes in the Perception and the Psychological Structure of Musical Emotions with Advancing Age. *Experimental Aging Research*, 38(4):422-441.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., and Bouchard, B. 2008. Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, 22(4):720-752.

Watson, D., and Clark, L. A. 1994. *The PANAS-X: Manual for the Positive and Negative Affect Schedule – Expanded Form*. University of Iowa: Department of Psychological & Brain Sciences Publications.

Watson, D., Clark, L. A., and 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.

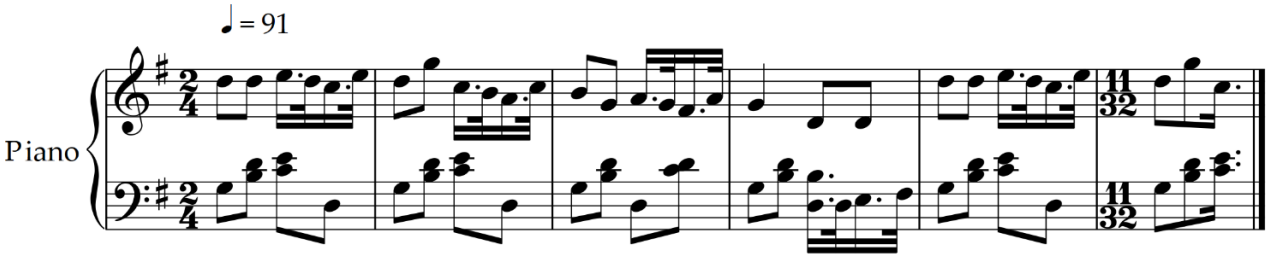
Whitfield, S. 2010. Music: Its Expressive Power and Moral Significance. *Musical Offerings*, 1(1). [Online]. doi:10.15385/jmo.2010.1.1.2.

Wilson, D. 1989. The Role of Patterning in Music. *Leonardo*, 22(1):101-106.

Wundt, W. 1897. C. H. Judd (trans.). *Outlines of Psychology*. Michigan: Scholarly Press.

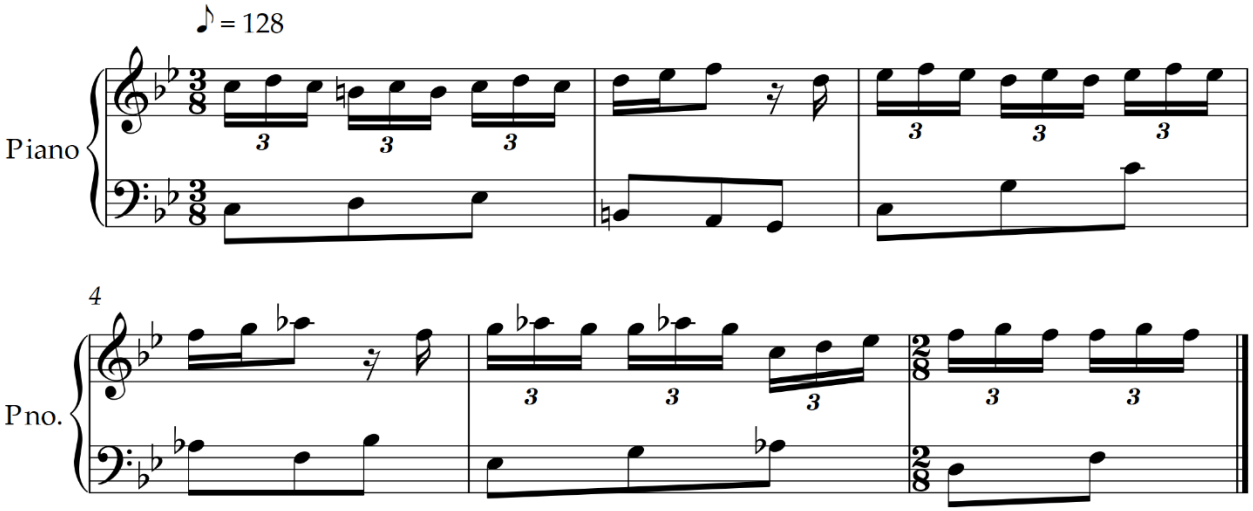
Appendix A – Musical excerpts used in the present experiment, cue levels, and Heinichen's descriptions

Excerpt 1



Identification in Vieillard *et al.*, 2008	G03
Emotion label	Happy
Time period	Modern
Mean valence score	3.97 (SD=1.10)
Mean arousal score	3.1 (SD=1.71)
BPM	91
REP	4.4
Mean pitch height	4.39
Mean interval size	4.57
Mode	Major
Texture	3

Excerpt 2



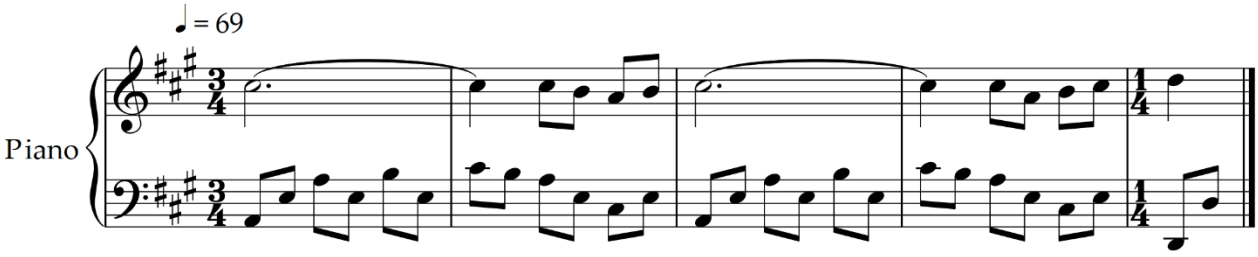
Page number in Heinichen (1986)	362-364
Emotion label	Happy
Time period	Baroque
Mean valence score	-0.13 (SD=2.06)
Mean arousal score	1.5 (SD=1.25)
BPM	128
REP	5.38
Mean pitch height	8.81
Mean interval size	2.33
Mode	minor
Texture	2

Heinichen's description, section used in categorisation in bold:

> …we could represent the result or the consequences of the search and believe that Aminta had found his love; then in this case the imagination takes the opportunity to **portray the playful looks of love**: (Heinichen, 1986:362)

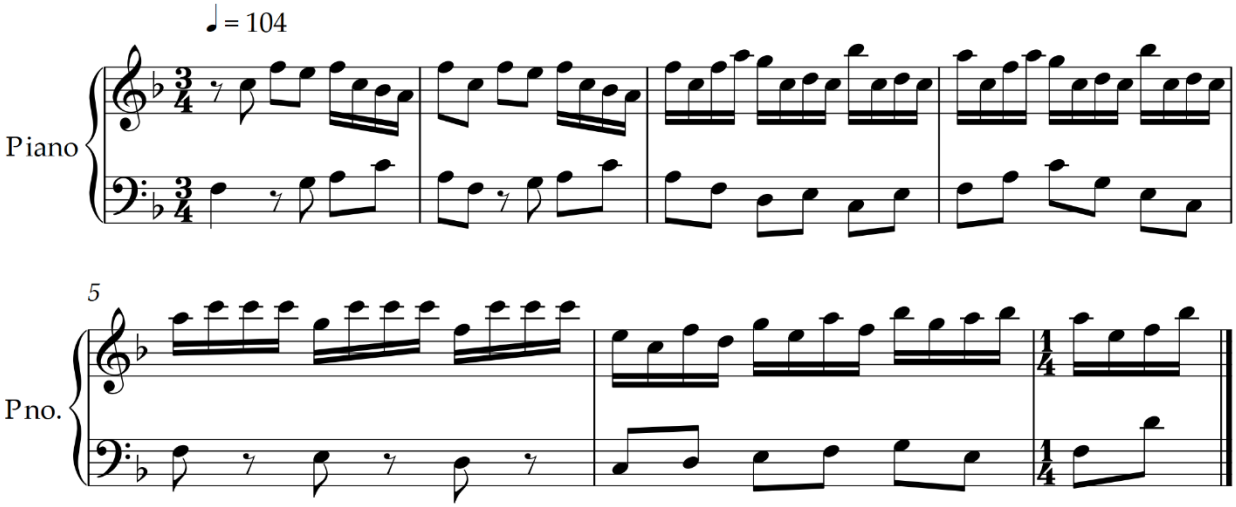
Categorisation term: *playful*
Category grouping: happiness

Excerpt 3



Identification in Vieillard *et al.*, 2008	A07
Emotion label	Peaceful
Time period	Modern
Mean valence score	0.13 (SD=2.01)
Mean arousal score	-2.3 (SD=1.27)
BPM	69
REP	2.3
Mean pitch height	-1.16
Mean interval size	4.85
Mode	Major
Texture	2

Excerpt 4



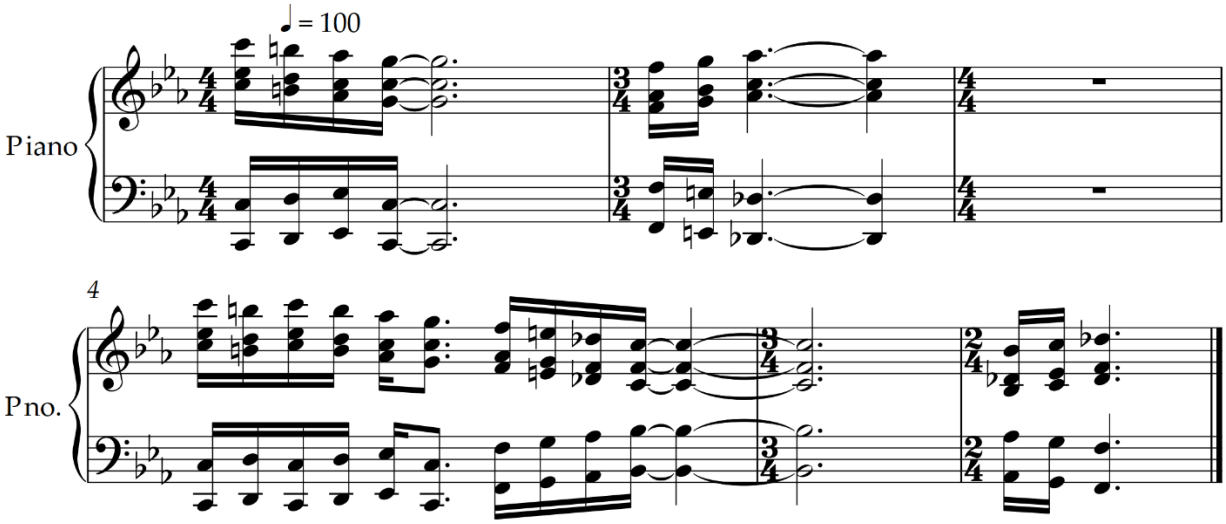
Page number in Heinichen (1986)	348-349
Emotion label	Peaceful
Time period	Baroque
Mean valence score	3.63 (SD=1.19)
Mean arousal score	3.13 (SD=1.28)
BPM	104
REP	6.18
Mean pitch height	9.69
Mean interval size	3.88
Mode	Major
Texture	2

Heinichen's description, section used in categorisation in bold:

> Should one wish to try special expressions, the words *faville, pupille, l'ardore, lo squardo* give our imagination much opportunity for pleasant and almost playful inventions. For example, one could **represent the burning fire of love** in the following invention: (Heinichen, 1986:348)

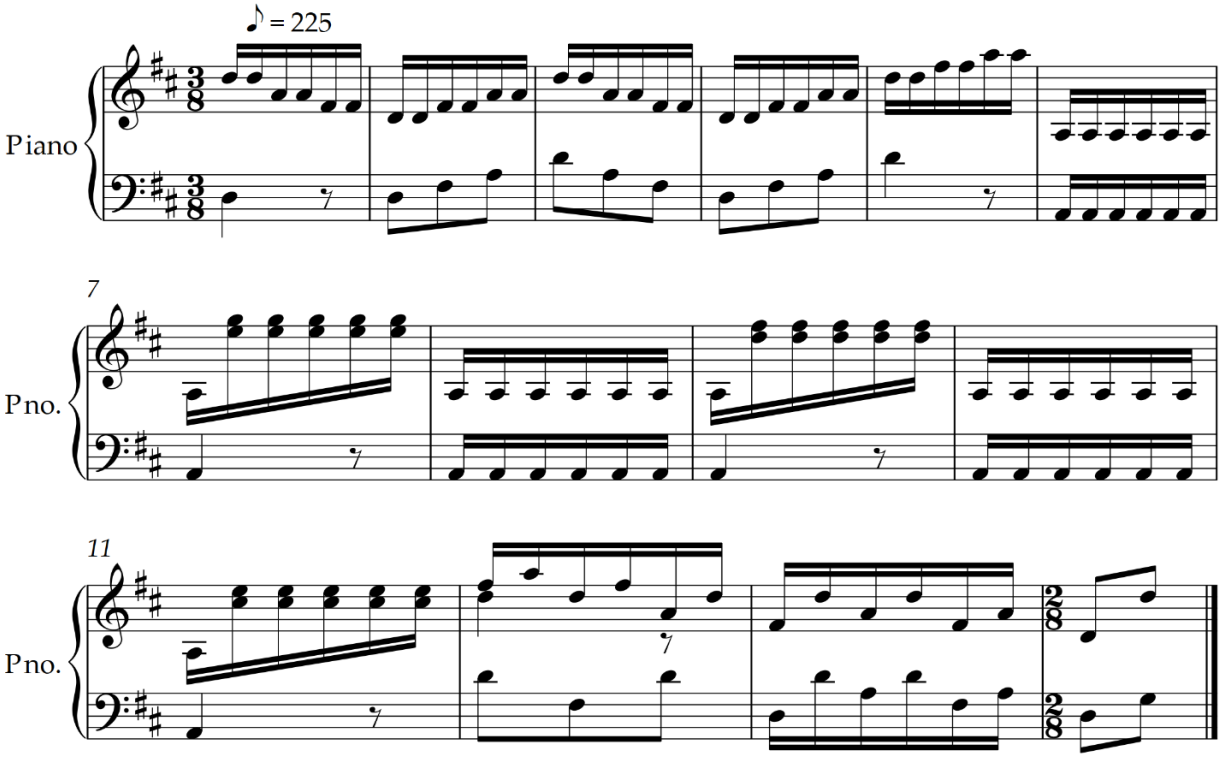
Categorisation term: *love*
Category grouping: peacefulness

Excerpt 5



Identification in Vieillard *et al.*, 2008	P07
Emotion label	Angry
Time period	Modern
Mean valence score	-1.2 (SD=1.94)
Mean arousal score	1.23 (SD=1.72)
BPM	100
REP	1.67
Mean pitch height	1.27
Mean interval size	1.92
Mode	minor
Texture	5

Excerpt 6



Page number in Heinichen (1986)	333-335
Emotion label	Angry
Time period	Baroque
Mean valence score	3 (SD=1.84)
Mean arousal score	3.43 (SD=1.81)
BPM	225
REP	7.27
Mean pitch height	1.98
Mean interval size	7.39
Mode	Major
Texture	3

Heinichen's description, section used in categorisation in bold:

> Only now the composer can derive from Metilde's intentions that this in itself dry aria can be **represented in the most furious of affections**, which should fire invention-rich composers to transform their formerly suspended thoughts into beautiful musical ideas. But should the natural fantast require still more help, one can proceed to special expressions of the recitative such as: *alti dissegni, e precipizii immensi*, and these could give something like the following expression (or ten other inventions of this type): (Heinichen, 1986:332)

Categorisation term: *fury*
Category grouping: anger

Appendix B – Full results of statistical computations

***NOTE: THROUGHOUT, "GENRE" REFERS TO TIME PERIOD**

MULTIVARIATE NORMALITY – MARDIA'S TEST

HAPPY EXCERPTS:

Test	Statistic	p value	Result
Mardia Skewness	9.18727713786162	0.0565851561656466	YES
Mardia Kurtosis	-1.49477034333358	0.134974388687436	YES
MVN	<NA>	<NA>	YES

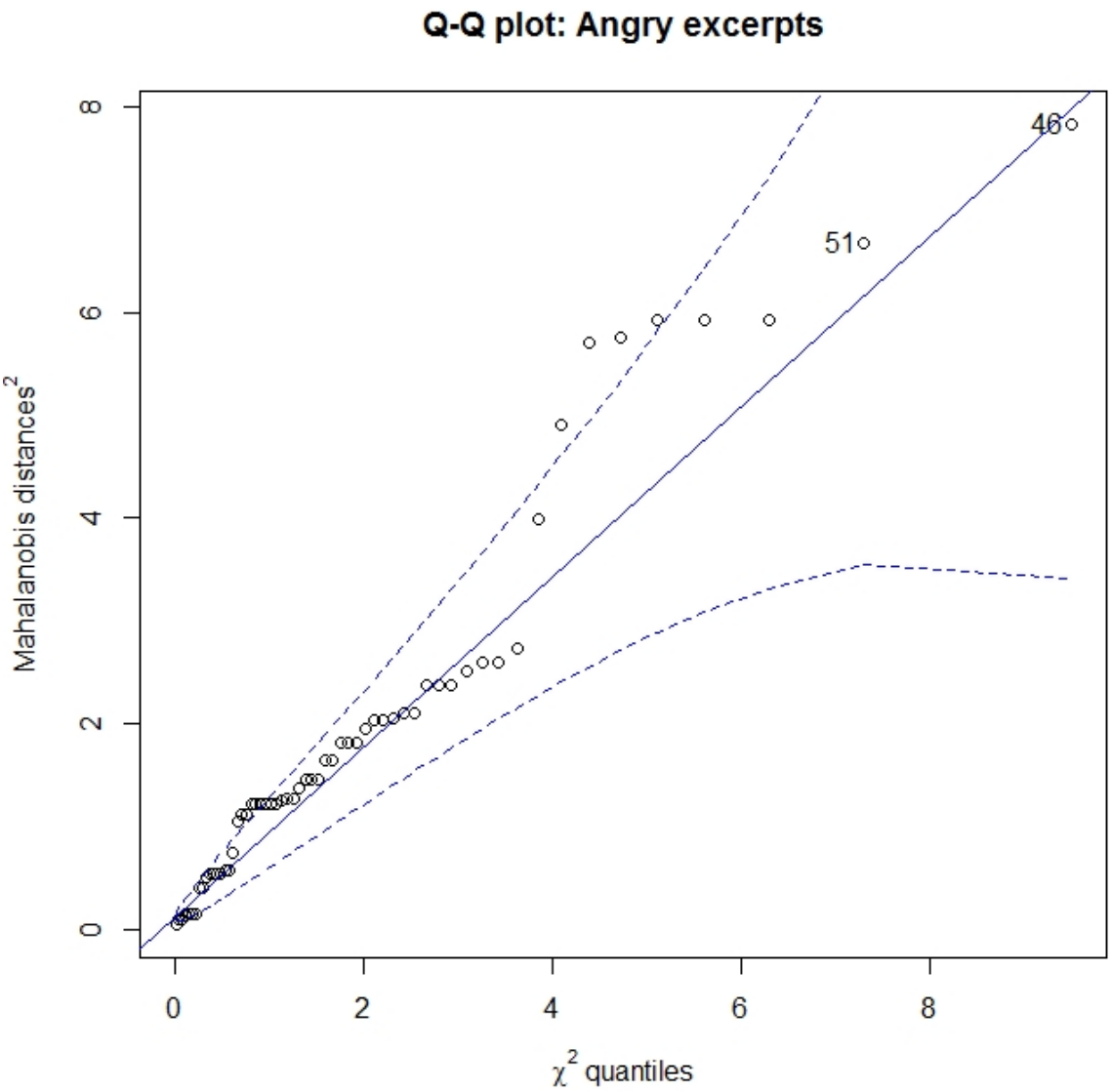
PEACEFUL EXCERPTS:

Test	Statistic	p value	Result
Mardia Skewness	8.44215883279624	0.0766599975391936	YES
Mardia Kurtosis	-1.74572872994733	0.0808581003441406	YES
MVN	<NA>	<NA>	YES

ANGRY EXCERPTS:

Test	Statistic	p value	Result
Mardia Skewness	13.2845273484154	0.00996606135729603	NO
Mardia Kurtosis	-0.547097336549365	0.584311862012778	YES
MVN	<NA>	<NA>	NO

Q-Q PLOT FOR ANGRY EXCERPTS:



PEARSON'S R

t = 8.0415, df = 178, p-value = 1.198e-13
95 percent confidence interval: 0.4001661 0.6159710
cor 0.5162155

Breusch-Pagan test

BP = 66.672, df = 11, p-value = 5.216e-10

Box's M Test – Group=Genre

Chi-Sq (approx.) = 30.957, df = 3, p-value = 8.68e-07

VIF TEST FOR MULTICOLLINEARITY – VALENCE

	GVIF	Df	GVIF^(1/(2*Df))
Genre	3	1	1.732051
Label	4	2	1.414214
Genre:Label	8	2	1.681793

VIF TEST FOR MULTICOLLINEARITY – AROUSAL

	GVIF	Df	GVIF^(1/(2*Df))
Genre	3	1	1.732051
Label	4	2	1.414214
Genre:Label	8	2	1.681793

MANOVA TESTING FOR EFFECT OF SEX

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
Sex	1	0.0018923	0.16779	2	177	0.8457
Residuals	178					

MANOVA TESTING FOR EFFECT OF PRESENTATION ORDER

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
PresOrd	1	0.020104	1.8157	2	177	0.1657
Residuals	178					

MANOVA TESTING FOR EFFECT OF AGE

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
Age	1	0.0039073	0.34715	2	177	0.7072
Residuals	178					

MAIN MANOVA

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
Genre	1	0.34302	43.596	2	167	5.831e-16 ***
Label	2	0.40206	21.135	4	336	1.467e-15 ***
Musicianship	1	0.04506	3.940	2	167	0.021282 *
Genre:Label	2	0.82101	58.495	4	336	< 2.2e-16 ***
Genre:Musicianship	1	0.00190	0.159	2	167	0.853478
Label:Musicianship	2	0.04737	2.038	4	336	0.088756 .
Genre:Label:Musicianship	2	0.10060	4.449	4	336	0.001622 **
Residuals	168					

HOTELLING'S T^2 – EFFECT OF TIME PERIOD ON HAPPY EXCERPTS

Test stat: 46.799
Numerator df: 2
Denominator df: 57
P-value: 9.432e-13

HOTELLING'S T^2 – EFFECT OF TIME PERIOD ON PEACEFUL EXCERPTS

Test stat: 151.75
Numerator df: 2
Denominator df: 57
P-value: 0

HOTELLING'S T^2 – EFFECT OF TIME PERIOD ON ANGRY EXCERPTS

Test stat: 38.163
Numerator df: 2
Denominator df: 57
P-value: 3.037e-11

FACTORIAL ANOVA – VALENCE

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Genre	1	64.8	64.8	22.169	5.2e-06 ***
Label	2	40.0	20.0	6.848	0.00138 **
Musicianship	1	4.4	4.4	1.490	0.22391
Genre:Label	2	635.7	317.9	108.740	< 2e-16 ***
Genre:Musicianship	1	0.6	0.6	0.190	0.66342
Label:Musicianship	2	4.2	2.1	0.720	0.48809
Genre:Label:Musicianship	2	23.5	11.7	4.016	0.01978 *
Residuals	168	491.1	2.9		

FACTORIAL ANOVA – AROUSAL

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Genre	1	184.0	184.02	84.163	< 2e-16 ***
Label	2	147.0	73.49	33.610	5.27e-13 ***
Musicianship	1	8.9	8.89	4.065	0.0454 *
Genre:Label	2	375.2	187.62	85.809	< 2e-16 ***
Genre:Musicianship	1	0.1	0.09	0.041	0.8405
Label:Musicianship	2	10.8	5.42	2.480	0.0868 .
Genre:Label:Musicianship	2	17.9	8.96	4.096	0.0183 *
Residuals	168	367.3	2.19		

RESULTS OF TUKEY'S HSD TEST FOR FACTORIAL ANOVA WITH VALENCE AS THE DEPENDENT VARIABLE

$Genre	diff	p adj
M-B	-1.2	5.2e-06
$Label	diff	p adj
H-A	1.01666667	0.0038755
P-A	0.98333333	0.0054578
P-H	-0.03333333	0.9937329
$Musicianship	diff	p adj
NM-M	0.3111111	0.2239131
$`Genre:Label`	diff	p adj
M:A-B:A	-4.2000000	0.0000000
B:H-B:A	-3.1333333	0.0000000
M:H-B:A	0.9666667	0.2480116
B:P-B:A	0.6333333	0.7058313
M:P-B:A	-2.8666667	0.0000000
B:H-M:A	1.0666667	0.1565332
M:H-M:A	5.1666667	0.0000000
B:P-M:A	4.8333333	0.0000000
M:P-M:A	1.3333333	0.0340998
M:H-B:H	4.1000000	0.0000000
B:P-B:H	3.7666667	0.0000000
M:P-B:H	0.2666667	0.9906255
B:P-M:H	-0.3333333	0.9744518
M:P-M:H	-3.8333333	0.0000000
M:P-B:P	-3.5000000	0.0000000
$`Genre:Musicianship`	diff	p adj
M:M-B:M	-1.0888889	0.0152963
B:NM-B:M	0.4222222	0.6456918
M:NM-B:M	-0.8888889	0.0690466
B:NM-M:M	1.5111111	0.0002585
M:NM-M:M	0.2000000	0.9451094
M:NM-B:NM	-1.3111111	0.0020559
$`Label:Musicianship`	diff	p adj
H:M-A:M	0.9000000	0.3250131
P:M-A:M	1.2333333	0.0633591

A:NM-A:M	0.4000000	0.9444980
H:NM-A:M	1.5333333	0.0084369
P:NM-A:M	1.1333333	0.1112151
P:M-H:M	0.3333333	0.9744518
A:NM-H:M	-0.5000000	0.8671286
H:NM-H:M	0.6333333	0.7058313
P:NM-H:M	0.2333333	0.9949569
A:NM-P:M	-0.8333333	0.4132792
H:NM-P:M	0.3000000	0.9839895
P:NM-P:M	-0.1000000	0.9999171
H:NM-A:NM	1.1333333	0.1112151
P:NM-A:NM	0.7333333	0.5592999
P:NM-H:NM	-0.4000000	0.9444980
$`Genre:Label:Musicianship`	Diff	p adj
M:A:M-B:A:M	-5.00000000	0.0000000
B:H:M-B:A:M	-4.13333333	0.0000000
M:H:M-B:A:M	0.93333333	0.9401099
B:P:M-B:A:M	0.40000000	0.9999652
M:P:M-B:A:M	-2.93333333	0.0003315
B:A:NM-B:A:M	-0.40000000	0.9999652
M:A:NM-B:A:M	-3.80000000	0.0000005
B:H:NM-B:A:M	-2.53333333	0.0042240
M:H:NM-B:A:M	0.60000000	0.9982699
B:P:NM-B:A:M	0.46666667	0.9998387
M:P:NM-B:A:M	-3.20000000	0.0000510
B:H:M-M:A:M	0.86666667	0.9643437
M:H:M-M:A:M	5.93333333	0.0000000
B:P:M-M:A:M	5.40000000	0.0000000
M:P:M-M:A:M	2.06666667	0.0506325
B:A:NM-M:A:M	4.60000000	0.0000000
M:A:NM-M:A:M	1.20000000	0.7437395
B:H:NM-M:A:M	2.46666667	0.0062358
M:H:NM-M:A:M	5.60000000	0.0000000
B:P:NM-M:A:M	5.46666667	0.0000000
M:P:NM-M:A:M	1.80000000	0.1566142
M:H:M-B:H:M	5.06666667	0.0000000
B:P:M-B:H:M	4.53333333	0.0000000
M:P:M-B:H:M	1.20000000	0.7437395
B:A:NM-B:H:M	3.73333333	0.0000008
M:A:NM-B:H:M	0.33333333	0.9999946
B:H:NM-B:H:M	1.60000000	0.3094494
M:H:NM-B:H:M	4.73333333	0.0000000
B:P:NM-B:H:M	4.60000000	0.0000000
M:P:NM-B:H:M	0.93333333	0.9401099
B:P:M-M:H:M	-0.53333333	0.9994178
M:P:M-M:H:M	-3.86666667	0.0000003
B:A:NM-M:H:M	-1.33333333	0.5980525
M:A:NM-M:H:M	-4.73333333	0.0000000
B:H:NM-M:H:M	-3.46666667	0.0000069
M:H:NM-M:H:M	-0.33333333	0.9999946

B:P:NM-M:H:M	-0.46666667	0.9998387
M:P:NM-M:H:M	-4.13333333	0.0000000
M:P:M-B:P:M	-3.33333333	0.0000191
B:A:NM-B:P:M	-0.80000000	0.9804163
M:A:NM-B:P:M	-4.20000000	0.0000000
B:H:NM-B:P:M	-2.93333333	0.0003315
M:H:NM-B:P:M	0.20000000	1.0000000
B:P:NM-B:P:M	0.06666667	1.0000000
M:P:NM-B:P:M	-3.60000000	0.0000025
B:A:NM-M:P:M	2.53333333	0.0042240
M:A:NM-M:P:M	-0.86666667	0.9643437
B:H:NM-M:P:M	0.40000000	0.9999652
M:H:NM-M:P:M	3.53333333	0.0000041
B:P:NM-M:P:M	3.40000000	0.0000115
M:P:NM-M:P:M	-0.26666667	0.9999995
M:A:NM-B:A:NM	-3.40000000	0.0000115
B:H:NM-B:A:NM	-2.13333333	0.0368453
M:H:NM-B:A:NM	1.00000000	0.9062244
B:P:NM-B:A:NM	0.86666667	0.9643437
M:P:NM-B:A:NM	-2.80000000	0.0008038
B:H:NM-M:A:NM	1.26666667	0.6730860
M:H:NM-M:A:NM	4.40000000	0.0000000
B:P:NM-M:A:NM	4.26666667	0.0000000
M:P:NM-M:A:NM	0.60000000	0.9982699
M:H:NM-B:H:NM	3.13333333	0.0000824
B:P:NM-B:H:NM	3.00000000	0.0002101
M:P:NM-B:H:NM	-0.66666667	0.9956068
B:P:NM-M:H:NM	-0.13333333	1.0000000
M:P:NM-M:H:NM	-3.80000000	0.0000005
M:P:NM-B:P:NM	-3.66666667	0.0000014

RESULTS OF TUKEY'S HSD TEST FOR FACTORIAL ANOVA WITH AROUSAL AS THE DEPENDENT VARIABLE

$Genre	diff	p adj
M-B	-2.022222	0
$Label	diff	p adj
H-A	-0.03333333	0.9916309
P-A	-1.93333333	0.0000000
P-H	-1.90000000	0.0000000
$Musicianship	diff	p adj
NM-M	-0.4444444	0.0453651
$`Genre:Label`	diff	p adj
M:A-B:A	-2.20000000	0.0000006
B:H-B:A	-1.93333333	0.0000158
M:H-B:A	-0.33333333	0.9524317
B:P-B:A	-0.30000000	0.9696205
M:P-B:A	-5.76666667	0.0000000
B:H-M:A	0.26666667	0.9819022
M:H-M:A	1.86666667	0.0000344
B:P-M:A	1.90000000	0.0000233
M:P-M:A	-3.56666667	0.0000000
M:H-B:H	1.60000000	0.0006313
B:P-B:H	1.63333333	0.0004474
M:P-B:H	-3.83333333	0.0000000
B:P-M:H	0.03333333	0.9999993
M:P-M:H	-5.43333333	0.0000000
M:P-B:P	-5.46666667	0.0000000
$`Genre:Musicianship`	diff	p adj
M:M-B:M	-2.0666667	0.0000000
B:NM-B:M	-0.4888889	0.3995389
M:NM-B:M	-2.4666667	0.0000000
B:NM-M:M	1.5777778	0.0000064
M:NM-M:M	-0.4000000	0.5749275
M:NM-B:NM	-1.9777778	0.0000000
$`Label:Musicianship`	diff	p adj
H:M-A:M	-0.3666667	0.9296750
P:M-A:M	-2.5333333	0.0000000
A:NM-A:M	-1.0666667	0.0633716
H:NM-A:M	-0.7666667	0.3421401
P:NM-A:M	-2.4000000	0.0000000
P:M-H:M	-2.1666667	0.0000009
A:NM-H:M	-0.7000000	0.4472170
H:NM-H:M	-0.4000000	0.9009407
P:NM-H:M	-2.0333333	0.0000047
A:NM-P:M	1.4666667	0.0023557
H:NM-P:M	1.7666667	0.0001066
P:NM-P:M	0.1333333	0.9993062
H:NM-A:NM	0.3000000	0.9696205
P:NM-A:NM	-1.3333333	0.0079283
P:NM-H:NM	-1.6333333	0.0004474

$`Genre:Label:Musicianship`	diff	p adj
M:A:M-B:A:M	-1.866667e+00	0.0326030
B:H:M-B:A:M	-2.333333e+00	0.0015458
M:H:M-B:A:M	-2.666667e-01	0.9999976
B:P:M-B:A:M	-2.666667e-01	0.9999976
M:P:M-B:A:M	-6.666667e+00	0.0000000
B:A:NM-B:A:M	-7.333333e-01	0.9696181
M:A:NM-B:A:M	-3.266667e+00	0.0000006
B:H:NM-B:A:M	-2.266667e+00	0.0024947
M:H:NM-B:A:M	-1.133333e+00	0.6241997
B:P:NM-B:A:M	-1.066667e+00	0.7091713
M:P:NM-B:A:M	-5.600000e+00	0.0000000
B:H:M-M:A:M	-4.666667e-01	0.9993504
M:H:M-M:A:M	1.600000e+00	0.1291017
B:P:M-M:A:M	1.600000e+00	0.1291017
M:P:M-M:A:M	-4.800000e+00	0.0000000
B:A:NM-M:A:M	1.133333e+00	0.6241997
M:A:NM-M:A:M	-1.400000e+00	0.2922564
B:H:NM-M:A:M	-4.000000e-01	0.9998523
M:H:NM-M:A:M	7.333333e-01	0.9696181
B:P:NM-M:A:M	8.000000e-01	0.9436524
M:P:NM-M:A:M	-3.733333e+00	0.0000000
M:H:M-B:H:M	2.066667e+00	0.0096552
B:P:M-B:H:M	2.066667e+00	0.0096552
M:P:M-B:H:M	-4.333333e+00	0.0000000
B:A:NM-B:H:M	1.600000e+00	0.1291017
M:A:NM-B:H:M	-9.333333e-01	0.8524699
B:H:NM-B:H:M	6.666667e-02	1.0000000
M:H:NM-B:H:M	1.200000e+00	0.5359135
B:P:NM-B:H:M	1.266667e+00	0.4487430
M:P:NM-B:H:M	-3.266667e+00	0.0000006
B:P:M-M:H:M	-1.776357e-15	1.0000000
M:P:M-M:H:M	-6.400000e+00	0.0000000
B:A:NM-M:H:M	-4.666667e-01	0.9993504
M:A:NM-M:H:M	-3.000000e+00	0.0000068
B:H:NM-M:H:M	-2.000000e+00	0.0147195
M:H:NM-M:H:M	-8.666667e-01	0.9050151
B:P:NM-M:H:M	-8.000000e-01	0.9436524
M:P:NM-M:H:M	-5.333333e+00	0.0000000
M:P:M-B:P:M	-6.400000e+00	0.0000000
B:A:NM-B:P:M	-4.666667e-01	0.9993504
M:A:NM-B:P:M	-3.000000e+00	0.0000068
B:H:NM-B:P:M	-2.000000e+00	0.0147195
M:H:NM-B:P:M	-8.666667e-01	0.9050151
B:P:NM-B:P:M	-8.000000e-01	0.9436524
M:P:NM-B:P:M	-5.333333e+00	0.0000000
B:A:NM-M:P:M	5.933333e+00	0.0000000
M:A:NM-M:P:M	3.400000e+00	0.0000002
B:H:NM-M:P:M	4.400000e+00	0.0000000
M:H:NM-M:P:M	5.533333e+00	0.0000000

B:P:NM-M:P:M	5.600000e+00	0.0000000
M:P:NM-M:P:M	1.066667e+00	0.7091713
M:A:NM-B:A:NM	-2.533333e+00	0.0003412
B:H:NM-B:A:NM	-1.533333e+00	0.1732902
M:H:NM-B:A:NM	-4.000000e-01	0.9998523
B:P:NM-B:A:NM	-3.333333e-01	0.9999761
M:P:NM-B:A:NM	-4.866667e+00	0.0000000
B:H:NM-M:A:NM	1.000000e+00	0.7864590
M:H:NM-M:A:NM	2.133333e+00	0.0062382
B:P:NM-M:A:NM	2.200000e+00	0.0039724
M:P:NM-M:A:NM	-2.333333e+00	0.0015458
M:H:NM-B:H:NM	1.133333e+00	0.6241997
B:P:NM-B:H:NM	1.200000e+00	0.5359135
M:P:NM-B:H:NM	-3.333333e+00	0.0000003
B:P:NM-M:H:NM	6.666667e-02	1.0000000
M:P:NM-M:H:NM	-4.466667e+00	0.0000000
M:P:NM-B:P:NM	-4.533333e+00	0.0000000

Appendix C – Data gathering materials

Questionnaire

What is your age?

……
……

What is your sex? (OPTIONAL)

……
……

Do you sing (regularly, e.g. as part of a choir) and/or play any instruments? If so, which instrument/s?

……
……

Have you received any formal training in music? If so, for how many years?

……
……

How many hours, on average, do you spend playing/practising music **per week**?

……
……

How many hours **per day**, on average, do you spend listening to music?

……
……

What genre/s of music make up the majority of your listening material?

……
……

……
……

Do you have any comments about the experiment? If so, please fill them in below.

……

……

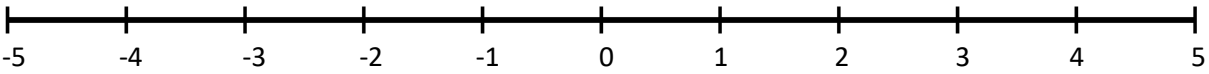
……

……

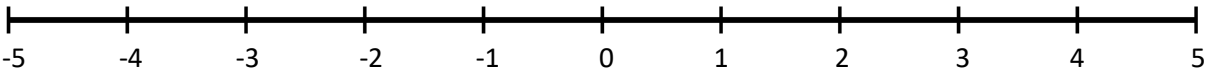
Rating sheet

EXCERPT 1

SAD NEUTRAL HAPPY

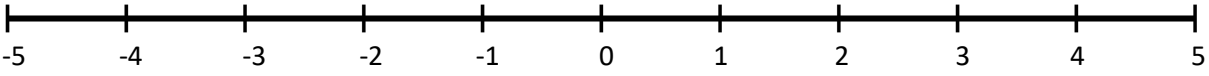


SLEEPY NEUTRAL ENERGETIC

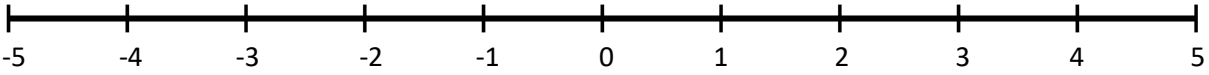


EXCERPT 2

SAD NEUTRAL HAPPY

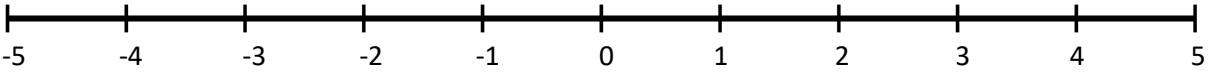


SLEEPY NEUTRAL ENERGETIC

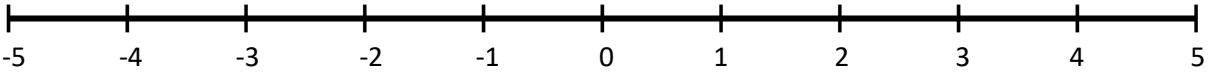


EXCERPT 3

SAD NEUTRAL HAPPY

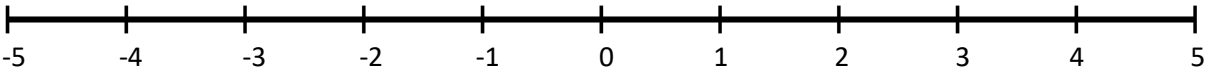


SLEEPY NEUTRAL ENERGETIC

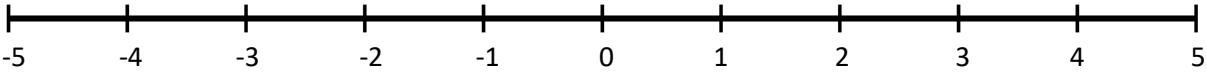


EXCERPT 4

SAD NEUTRAL HAPPY

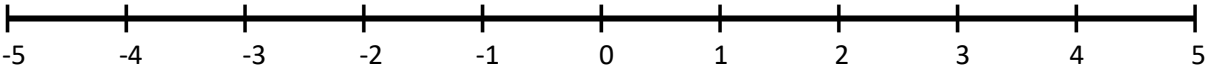


SLEEPY NEUTRAL ENERGETIC

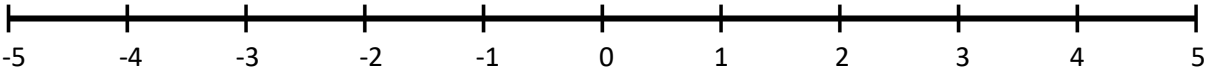


EXCERPT 5

SAD NEUTRAL HAPPY

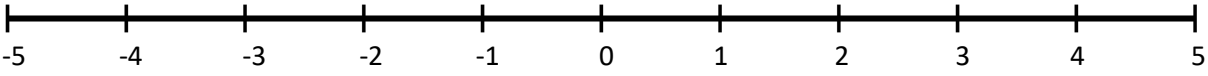


SLEEPY NEUTRAL ENERGETIC



EXCERPT 6

SAD NEUTRAL HAPPY



SLEEPY NEUTRAL ENERGETIC