

EMOTIONS, MECHANISMS, AND INDIVIDUAL DIFFERENCES IN MUSIC LISTENING: A STRATIFIED RANDOM SAMPLING APPROACH

PATRIK N. JUSLIN, LAURA S. SAKKA
Uppsala University, Uppsala, Sweden

GONÇALO T. BARRADAS
Uppsala University, Uppsala, Sweden & University
of Madeira, Funchal, Portugal

OLIVIER LARTILLOT
University of Oslo, Oslo, Norway

EMOTIONS HAVE BEEN FOUND TO PLAY A paramount role in both everyday music experiences and health applications of music, but the applicability of musical emotions depends on: 1) which emotions music can induce, 2) how it induces them, and 3) how individual differences may be explained. These questions were addressed in a listening test, where 44 participants (aged 19–66 years) reported both felt emotions and subjective impressions of emotion mechanisms (*Mec Scale*), while listening to 72 pieces of music from 12 genres, selected using a stratified random sampling procedure. The results showed that: 1) positive emotions (e.g., happiness) were more prevalent than negative emotions (e.g., anger); 2) Rhythmic entrainment was the most and Brain stem reflex the least frequent of the mechanisms featured in the BRECVEMA theory; 3) felt emotions could be accurately predicted based on self-reported mechanisms in multiple regression analyses; 4) self-reported mechanisms predicted felt emotions better than did acoustic features; and 5) individual listeners showed partly different emotion-mechanism links across stimuli, which may help to explain individual differences in emotional responses. Implications for future research and applications of musical emotions are discussed.

Received: July 6, 2021, accepted April 24, 2022.

Key words: music, emotions, mechanisms, individual differences, acoustic features

MUSICAL EMOTIONS MATTER. THEY ENRICH listeners' music experiences in profound ways (Juslin, 2019) and could also have far-reaching implications for their well-being and health (e.g., Ferreri et al., 2019; MacDonald et al., 2012). Applications of musical emotions range from pain relief and dementia care to film music and marketing (Juslin & Sloboda, 2010).

However, it might be argued that our ability to apply musical emotions effectively in society depends partly on how we answer three key questions: Which emotions does music induce? How exactly do they occur? Why do not all listeners experience the same emotion? Emotions or mechanisms that occur very rarely or that only involve a select group of people may not afford the best possibilities for efficient applications. In this article, we report data on music listening that speak directly to these issues.

EMOTION PREVALENCE: WHAT DO WE EXPERIENCE?

The term *prevalence* is used to refer to the relative frequency of occurrence of a certain phenomenon, such as emotional reactions to music, in the population of interest (Juslin et al., 2008). Prevalence data capture the phenomena that any theory of music and emotion must be able to explain, and are also important to understand the ramifications of music as an application: What, exactly, might music achieve in terms of its emotional impact?

We define an *emotion* here as a relatively brief, intense, and rapidly changing reaction to a potentially important event (a subjective challenge or opportunity) in the external or internal environment. Theories of emotion come in many forms (e.g., basic emotions, appraisal theory, psychological construction), all of which have been applied to music (see Warrenburg, 2020). Music researchers distinguish between *perception* and *induction* of emotions: We may simply perceive an emotion “expressed” in the music, or we may actually *feel* an emotion in ourselves (for reviews, see Gabrielsson, 2002, and Schubert, 2013). Most previous studies have focused on perception of emotions

(Eerola & Vuoskoski, 2013, Figure 2), yet what most people strive for is arguably to be “moved” (i.e., felt emotions).

The prevalence of specific emotions during music listening was initially mostly a matter of speculation among scholars, often based on personal experience (cf. Kivy, 1993). However, evidence is slowly accumulating from studies indicating that music could induce a fairly wide range of both “basic” (happiness, sadness, interest) and “complex” (nostalgia, pride) emotions, as well as “aesthetic” emotions (awe) (Gabrielsson, 2011; Juslin & Laukka, 2004; Juslin et al., 2011; Sloboda, 1992; Taruffi & Koelsch, 2014; Wells & Hakanen, 1991; Zentner et al., 2008). Moreover, emotions commonly regarded as “positive” in valence appear to be more prevalent than emotions commonly regarded as “negative” (Gabrielsson, 2011; Juslin et al., 2008, 2011; Juslin, Barradas et al., 2016; Sloboda et al., 2001). Positive emotions have a range of beneficial effects on physical health (Kubzansky, 2009) and subjective well-being (Fredrickson, 1998). Yet, they have been somewhat neglected in previous research on emotions (Lazarus, 1991).

Because estimates of emotion prevalence are influenced by a number of factors, such as individual characteristics, social context, and musical style, some researchers have advocated “method triangulation” to obtain representative samples of listeners, situations, and pieces of music, respectively (Juslin et al., 2010). Previous studies have sampled listeners (Juslin et al., 2011) or situations (Juslin et al., 2008). In this study, we sampled pieces of music to explore how this aspect moderates prevalence data.

MECHANISMS: HOW DO THE EMOTIONS OCCUR?

The “what” question (i.e., emotion prevalence) is closely related to the “how” question: What emotion will occur in a given musical event is largely determined by *how* the emotion was induced. This issue is often regarded as the greatest mystery of them all (cf. Dowling & Harwood, 1986; Johnson-Laird, 1992). This is because in a typical case in everyday life, an emotion is induced when an event or object is appraised as having the capacity to influence the goals of the perceiver somehow. Yet, when we listen to music, only rarely has the music per se implications for our life goals, as described by appraisal theories (cf. Ellsworth, 1994).

This could explain why many scholars have leaned towards *formalism* (e.g., Kivy, 2002), which considers music as abstract tones sequences, devoid of semantic meaning. Such a view can lead to a neglect of the “meaning-making” role of psychological processes needed for any emotion to occur, in favor of a search

for “direct” links between musical features and induced emotions (Coutinho & Cangelosi, 2011; Gomez & Danuser, 2007).

A large number of studies have attempted to predict *perceived* emotions in music based merely on acoustic features, and such attempts have admittedly been quite successful (Eerola, 2011; Juslin, 1997; Juslin & Lindström, 2010; Yang et al., 2018). However, there seems to be a “glass ceiling” with regard to the predictive accuracy that can be achieved with only acoustic features (Barthet et al., 2013). This has been referred to as a “semantic gap” between low-level acoustic features and high-level perception (Yang & Chen, 2011)—a problem that is arguably at its most acute for *induced* (i.e., felt) emotions. Daly et al. (2015) made an ambitious attempt to predict felt emotions based on acoustic features. They were able to explain less than 5% of the variance—and this was when acoustic features were augmented by EEG indices. A purely acoustic approach fared even worse. Coutinho and Cangelosi (2011) succeeded better, though they predicted changes in emotion *within* a piece of music, rather than across different pieces.

To improve predictive accuracy, researchers need to address *how* the emotional effects occur in the first place. They need to consider the underlying psychological processes that “mediate” between musical events and felt emotions. We shall refer to these processes as the *mechanism* (Juslin & Västfjäll, 2008).

Several different authors have suggested possible mechanisms for induction of emotion in music, usually involving just one or a few possibilities (Baumgartner, 1992; Berlyne, 1971; Juslin, 2001, 2013; Levinson, 1997; Meyer, 1956, 2001; Scherer & Zentner, 2001; Sloboda & Juslin, 2001). Space limitations prevent us from reviewing previous work in detail here, but the most comprehensive theory outlined so far is provided by the BRECVEMA framework (the acronym BRECVEMA derives from the first letter of each of the mechanisms featured in the theory, as listed below; for elaboration and predictions, see Juslin, 2019, Part III).

The BRECVEMA framework is consistent with a categorical approach to emotions, in the narrow sense that it assumes the existence of specific emotion categories, such as sadness and happiness. However, it does not quite adhere to any of the “traditional” emotion theories (e.g., basic emotions, appraisal, constructionism), but rather represents a novel type of theory sharing certain features with past theories. For example, it assumes that emotion mechanisms have a long evolutionary history (like basic emotion theory), it highlights the role of emotion learning and variability (like social

construction theory), and it presumes close links between cognition and emotion (like cognitive appraisal theory). However, the framework also differs from these theories in some crucial ways. For instance, we do not think that “all emotions are basic” (Ekman, 1994, p. 15); nor do we believe that emotions, just like money, “are a product of human agreement” (Barrett, 2017, p. xiii); and we do not presume that “most emotions are elicited . . . through a process of cognitive appraisal” (Scherer et al., 2002, p. 150).

Unlike these theories, BRECVEMA holds that different types of emotions are induced by different types of mechanisms at different levels of the brain. (For example, the lowest levels may involve mainly the arousal dimension and proto-emotions, such as surprise, whereas the highest levels involve more complex, and even aesthetic, emotions; Juslin & Västfjäll, 2008; Juslin, 2013, 2019). It is argued that this framework is better able to account for the complex, multi-faceted emotional responses that may occur to music, than are traditional theories. The framework postulates eight mechanisms which involve (more or less) distinct brain networks:

- *Brain stem reflex*, a hard-wired attention response to subjectively “extreme” values of basic acoustic features, such as loudness, speed, and timbre (e.g., Davis, 1984); you may become *startled* or *surprised* by the loud beginning of a piece of music (Arjmand et al., 2017; Juslin et al., 2014).
- *Rhythmic entrainment*, a gradual adjustment of an internal body rhythm such as heart rate towards an external rhythm in the music (Bason & Celler, 1972; Harrer & Harrer, 1977); you may experience excitement if your heart rate becomes gradually synchronized with a highly captivating and somewhat faster rhythm in a piece of techno music at a nightclub.
- *Evaluative conditioning*, a regular pairing of a piece of music and other positive or negative stimuli leading to a conditioned association (e.g., Blair & Shimp, 1992; Bolders et al., 2012); you may feel happy when you hear a song which has repeatedly occurred in festive contexts.
- *Contagion*, an internal “mimicry” of the perceived voice-like emotional expression of the music (e.g., Juslin, 2001); you may experience sadness when you hear a slow, quiet, low-pitched performance of a classical piece on the cello that features much vibrato and rubato (Juslin et al., 2014; see also Egermann & McAdams, 2013).
- *Visual imagery*, inner images of an emotional character conjured up by the listener through

a metaphorical mapping of the musical structure (Osborne, 1980; Taruffi & Küssner, 2019); you might become *relaxed* when you indulge in the mental images of a landscape suggested by a piece of “new-age” music while lying at home in your sofa.

- *Episodic memory*, a conscious recollection of a particular event from the listener’s past that is “triggered” by the music (see Baumgartner, 1992; Janata et al., 2007); you may experience *nostalgia* when a song evokes a vivid personal memory from the specific time you met your current partner in life (Garrido & Davidson, 2019).
- *Musical expectancy*, a response to the gradual unfolding of the syntactical structure of the music, and its expected or unexpected continuations (Huron, 2006; Meyer, 1956); you may feel *anxious* due to uncertainty created by phrases without a clear tonal center in an “avant-garde” piece (Juslin et al., 2014; see also Steinbeis et al., 2006).
- *Aesthetic judgment*, a subjective evaluation of the aesthetic value of the music, based on an individual set of weighted criteria (Juslin, 2013); you may admire the exceptional skills of a great performer at an evening concert (Juslin et al., 2021).¹

Several studies recently tested selected mechanisms in highly controlled experimental settings, using both synthesized and “real” musical excerpts (Juslin et al., 2014, 2015; Juslin, Sakka et al., 2016; Sakka & Juslin, 2018; see also Barradas et al., 2021). Target-mechanism conditions induced specific emotions in listeners largely in accordance with theoretical predictions, as shown by multiple indices (e.g., self-reported feelings, facial electromyography, psychophysiology). (For a review of neural correlates of musical emotions, see Juslin & Sakka, 2019.)

These results did not just reflect acoustic features of the music, such as tempo, sound level, or timbre. For instance, contrary to music-emotion correlations one might expect, the tempo could be faster in a piece that induced sadness than in a piece that induced happiness. Researchers concluded that this is because listeners’ responses are “driven” by mechanisms (e.g., whether a memory was evoked), rather than by acoustic features per se. In this study, we compared prediction of felt emotions based on acoustic measures with prediction based on subjective ratings of mechanisms.

¹ In addition to these mechanisms, music can also induce emotions through the default mechanism for induction of emotions: *Cognitive goal appraisal* (Scherer, 1999). You may become *annoyed* when a neighbor plays music late at night, blocking your goal to go to sleep. However, appraisal is not prevalent in music (Juslin et al., 2008).

INDIVIDUAL DIFFERENCES: WHO WILL EXPERIENCE A CERTAIN EMOTION?

The “how” question (mechanisms) is linked to the “who” question: Who will experience a certain emotion to a piece of music and who will not? Although the experimental studies cited above were designed to maximize experimental control and minimize the effects of contextual variables, the emotions induced were still not as neatly differentiated as one might hope.

Researchers have long acknowledged wide individual differences in musical emotions (e.g., Sloboda, 1996), and it has been suggested that there is greater variability for induction of emotions than for perception of emotions (Juslin, 2019). Most studies to date have tended to downplay individual differences and to focus only on means across listeners (but see, e.g., Juslin et al., 2021; Juslin, Sakka et al., 2016; Ladinig & Schellenberg, 2012; Mas-Herrero et al., 2013).

Still, it might be argued that our understanding of individual differences has important implications for the applicability of music in different contexts. Previous research has shown that personal preference and music choice play a key role for the effectiveness of music at an emotional level, both in laboratory studies (Liljeström et al., 2013) and applications (Garrido et al., 2017). Preferred music is likely to be more familiar to the listener, and familiarity itself can enable a greater number of emotion mechanisms (e.g., episodic memory) to be activated.

Why do two listeners (sometimes) respond differently to the same piece of music? Since the acoustic features of the music are the same, an explanation must clearly be sought in terms of a difference in the *emotion-mechanism* link. Specifically, individual listeners might activate different mechanisms to the same music (e.g., memory for one listener, contagion for another), thus inducing distinct emotions. In addition, for certain mechanisms (e.g., episodic memory), even the same mechanism may induce different emotions in different listeners, depending on previous personal experiences (e.g., Sakka & Saarikallio, 2020). Mechanism activation may in turn depend on factors such as attention, music training, personality traits, and individual learning history.

Because it is at the mechanistic level that individual differences in responses will tend to emerge, a mechanism focus is required to explain them. In this study, we made a first attempt to model emotion-mechanism links at an individual level, adopting a “statistical-ideographic” approach to musical emotions (Juslin, 2019; cf. Brunswik, 1956).

THE PRESENT STUDY

Most findings regarding emotions and mechanisms to date come from field studies that measure prevalence in various social and cultural contexts (e.g., Dingle et al., 2011; Juslin, Barradas et al., 2016). However, due to the lack of experimental control, it is not feasible to investigate individual differences in response to the same music. (Individual differences in response are typically confounded with differences in the music heard.)

Experimental studies, on the other hand, which have presented several listeners with the same pieces of music (Juslin et al., 2014, 2015), have tended to feature a very limited number of pieces that were selected or manipulated to target specific mechanisms. Thus, it is unclear how prevalence estimates of emotions and mechanisms may be affected if listeners are exposed to a larger and more “ecologically relevant” sample of music. The median number of musical stimuli in previous studies of music and emotion is 10, and the music selection has focused largely on classical music (Eerola & Vuoskoski, 2013), which represents a minority interest—even in the Western world (Hargreaves & North, 2010).

The overall aim of this study was thus to complement previous studies by investigating emotion-mechanism links within a controlled setting. The added value of randomly sampling *music*—as opposed to sampling *listeners* (Juslin et al., 2011) or *situations* (Juslin et al., 2008)—is that we are able to compare individual listeners, using a broad variety of musical examples. More specifically, we aimed to investigate the following questions: 1) Which emotions does music induce most frequently? 2) Which mechanisms occur most often? 3) Can emotions be predicted from self-reported mechanisms? 4) Are emotions best predicted from mechanisms or from acoustic features of the music? 5) Do different listeners show the same links among emotions and mechanisms?

These questions were addressed in a listening test using a correlational design (similar to studies in judgment analysis; cf. Cooksey, 1996). Participants listened to a large selection of musical excerpts, and were asked to rate felt emotions and subjective impressions related to various mechanisms. Whereas previous experiments have purposefully selected pieces of music so as to target particular mechanisms and induce predicted emotions, here we simply explored what emotions and mechanisms would occur spontaneously during music listening. It should be noted from the outset that such a correlational design limits the relative strength with which conclusions about causal relationships can be drawn.

We adopted a research approach characterized by five features: 1) we used a stratified random sampling procedure (e.g., Visser et al., 2000) to select the stimuli (in order to obtain a more representative sample); 2) we carried out both *nomothetic* (averaged) and *idiographic* (individual) analyses of emotion-mechanism links (an aspect of method development); 3) we focused on positive emotions (as these have been under-studied in the past); 4) we conducted extensive acoustic analyses (including a partly novel set of algorithms); and 5) we adopted a categorical—as opposed to dimensional—approach to emotion measurement (since one of our aims was to investigate the prevalence of specific emotions, and to compare our findings with those of previous studies adopting a categorical approach). We tested five hypotheses:

(H1) Based on previous studies that featured more or less representative samples of listeners (Juslin et al., 2011) or situations (Juslin et al., 2008), respectively, we hypothesized that positive emotions would be more prevalent than negative emotions.

(H2) Based on field studies that measured the prevalence of various mechanisms (Juslin et al., 2008; Juslin, Barradas et al., 2016), we hypothesized that the two mechanisms Brain stem reflex and Musical expectancy would be less prevalent than the other mechanisms.

(H3) Based on previous experimental studies (Barradas et al., 2021; Juslin et al., 2014, 2015; Sakka & Juslin, 2018), we hypothesized that the mechanisms would be reliably linked to specific emotions, as shown by accurate prediction of felt emotions (defined as $R^2 \geq .64$; a “strong” effect; Ferguson, 2009) based on mechanism ratings in multiple regression analyses.

(H4) Based on findings in Juslin et al. (2014) and theoretical arguments in Juslin (2019; see also previous section), we hypothesized that self-reported mechanism impressions would better predict felt emotions than would acoustic features.

(H5) Based on the argument that individual differences in emotional reactions to music primarily occur at the mechanistic level (Juslin, 2019, p. 396), we hypothesized that listeners would show different emotion-mechanism links in idiographic regression analyses.

In addition to testing the above hypotheses, we also correlated the ratings of emotions with measures of the Big Five personality traits (John et al., 2008), in an explorative attempt to account for individual differences. Some preliminary studies have suggested that emotion prevalence might be moderated by personality (Barrett et al., 2010; Juslin et al., 2008, 2011; Liljeström et al., 2013; McCrae, 2007; Pilgrim et al., 2017).

Method

The results reported in this paper were collected at the same time and featuring the same listeners and stimuli as the findings on how different criteria contribute to aesthetic judgments presented in Juslin, Sakka et al. (2016). There is no overlap in the data reported, however, except for the characteristics of the listener sample.

PARTICIPANTS

Forty-four listeners, 22 females and 22 males, 19–66 years old ($M = 33.57$, $SD = 14.13$) participated in the study. Their anonymous and voluntary participation was compensated with either course credits or two cinema vouchers. To obtain a broad sample not limited to students, the participants were recruited by means of posters throughout the town of Uppsala (e.g., shopping malls, the town library), as well as via a mailing list for music seminars.

Seventy-five percent of the participants played (at least) one musical instrument, while 52% stated they had received music education. Pre-screening showed that music preferences varied widely, with most listeners marking several genres of music; however, rock, classical, pop, electronica, and jazz were the most frequently reported genres. None of the participants reported a hearing problem. They were randomly divided into two groups.

MUSICAL MATERIAL

In obtaining the sample of music, we tried to achieve an optimum balance between the need to feature as many musical stimuli as possible to obtain a sufficient number of cases for the multiple regression analyses and the need to keep the experimental sessions short enough to avoid severe fatigue effects (Cooksey, 1996).

With the upper limit on the number of stimuli in mind, we decided to use the stratified random sampling procedure (Visser et al., 2000), which may help to ensure that a sufficient variety of musical genres is obtained even with a limited sample size. Each participant group listened to 40 pieces of music. (The complete list appears in Appendix A.) We selected pieces randomly, with strata corresponding to the STOMP factors for music preference proposed by Rentfrow and Gosling (2003) based on factor analysis of music preferences. (STOMP stands for Short Test of Music Preferences.) The framework is clearly the most extensive attempt so far to map the underlying structure of music preferences, and the STOMP factor structure has been replicated in several countries (Delsing et al., 2008; Gouveia et al., 2008; Langmeyer et al., 2012; Zweigenhaft, 2008). However, we modified their structure slightly to suit

Swedish conditions (e.g., dropping the Religious genre and adding Schlager/Dansband), and to obtain a similar number of genres for each factor. (The aim of the study was not to examine specific genres, but rather to obtain a sufficient variety of musical stimuli.) The four factors and their respective genres were:

1. Reflective & Complex (R&C): Classical/Opera; Jazz/Blues; Folk/Singer-songwriter
2. Intense & Rebellious (I&R): Classic rock; Heavy metal/Hard rock; Alternative/Punk
3. Upbeat & Conventional (U&C): Country; Pop; Schlager/Dansband
4. Energetic & Rhythmic (E&R): Hip-hop/Reggae; Soul/Funk; Club/House

Seventy-two pieces of music (six for each genre and 18 for each factor) were randomly sampled from the internet database Spotify. (We used the *Radio* feature to create a random playlist of songs, based on pre-specified genres. At the time the sample was drawn, we benefited from existing genre labels corresponding to those used here. The categories in Spotify have since then been changed.) The songs were randomly divided across the two listener groups with the provision that each group should have three songs per genre, leaving both groups with 36 songs apiece. In addition, four songs from each listener group (one from each factor) were randomly selected and added to the *other* group, such that eight pieces were common to the two groups—facilitating group comparison. The resulting sample may be said to be representative of what a Swedish music listener would typically encounter on the radio, in CD shops, at live concerts, or in databases on the internet at the time of the study.

Each piece was edited to a 60-second excerpt. This was done to maximize the number of pieces included without subjecting participants to an exhausting listening session. Moreover, responses to music can change quickly, so the longer the musical event, the more problematic the retrospective self-report would be. However, a previous study indicated that about 97% of the emotions experienced in daily life last longer than one minute (Frijda et al., 1991; for similar results with music, see Scherer et al., 2002). Thus, it would appear unlikely that the experienced emotion category would change much during the 60-second segment of music. The present stimuli are shorter than both the mean (135 s) and median (90 s) stimulus duration of previous studies of felt emotions in music, as reviewed by Eerola and Vuoskoski (2013). Note, however, that stimuli should not be *too* short, since different mechanisms may differ in terms of the time they need in order for the emotion-induction process

to take place (e.g., rhythmic entrainment may take longer than a brain stem reflex; Juslin, 2019, Chapter 25).

For most songs (87%), the excerpt featured the first 60 seconds of the recording. For the remaining pieces (e.g., songs with a long introduction), the recording was edited manually, so as to feature a more meaningful section (e.g., verse, chorus) of the song (see Appendix A). All excerpts ended with a quick fade and were stored in high bit-rate mp3 format.

MEASURES

Emotions

Like most emotion theorists, we assume that emotions involve multiple components, such as feeling, physiology, and expression (Scherer, 2000). In this study, we measured the *feeling* component, which is often regarded as the most crucial aspect of a musical emotion (Zentner & Eerola, 2010). Most theorists presume that self-reports of feelings have validity (Barrett, 2004). Thus, we measured the listeners' feelings using 12 rating scales: *happiness-elongation*, *sadness-melancholy*, *surprise-astonishment*, *calm-contentment*, *interest-expectancy*, *nostalgia-longing*, *anxiety-nervousness*, *pride-confidence*, *anger-irritation*, *love-tenderness*, *disgust-contempt*, and *awe-admiration*. We used two words to denote each unipolar scale to emphasize that the scale should be interpreted as referring to a *broad* emotion category that may include a variety of similar states. These scales represent a kind of compromise among the response formats currently used to measure musical emotions (for a review, see Zentner & Eerola, 2010), because the terms includes some “basic emotions” characteristic of discrete emotion theories (Izard, 1977), cover all four quadrants of the *circumplex* model in terms of *valence* and *arousal* (Russell, 1980), and feature possibly more music-related terms, such as *nostalgia*, *expectancy*, and *awe* (Juslin & Laukka, 2004). (The selected terms roughly cover the nine factors of GEMS - 9, proposed by Zentner et al., 2008, but because that scale lacks some terms that were needed in this study, e.g., surprise, we decided to use a customized list of emotion terms.) These terms include the emotions that were most prevalent in previous research, but also cover other emotions, to not pre-judge the outcome. Thus, the terms *anxiety*, *anger*, and *disgust* were mainly intended as controls. Consistent with our focus on positive emotions, the list includes a greater number of positive emotions than typical instruments that measure specific emotions (e.g., Plutchik, 1994, Chapter 5). All emotions were rated on a scale from 0 (*not at all*) to 4 (*a lot*) in response to the instruction *describe how you felt when you heard the music*. (A table of intercorrelations among emotion ratings appears in Appendix B.)

Mechanisms

We collected subjective data about the mechanisms that may have occurred, using the *MecScale* (e.g., Juslin et al., 2014). This consists of eight questions, each targeting one of the mechanisms in the BRECVEMA framework: 1) Did the music feature an event that startled you? (Brain stem reflex); 2) Did the music have a strong and captivating rhythm? (Rhythmic entrainment); 3) Did the music evoke memories of events from your life? (Episodic memory); 4) Did the music induce emotions through an association? (Evaluative conditioning); 5) Did the music evoke inner images that influenced your emotions? (Visual imagery); 6) Were you “touched” by the emotional expression of the music? (Contagion); 7) Was it difficult to guess how the music (e.g., the melody) would develop over time? (Musical expectancy); 8) Did you find the music aesthetically valuable? (Aesthetic judgment) Listeners were asked to rate each item on a scale from 0 (*not at all*) to 4 (*a lot*).

The *MecScale* does not purport to measure the induction mechanisms “directly.” Rather, it measures *subjective impressions* that are reflective of mechanisms. Some mechanisms are straightforward to report (e.g., startle reflexes, memories, and images are very distinct). Others may at least be correlated with certain impressions (e.g., a listener who becomes surprised via musical expectancy will find a song “unpredictable”). *MecScale* is intended as a cost-effective index for decision-making purposes, for which the most crucial criterion is *predictive validity* (also referred to as *concurrent validity* if predictor and criterion are indexed at the same time). The items have been reported to be predictive of both target-mechanism conditions (Juslin et al., 2014; Sakka & Juslin, 2018; see also Barradas et al., 2021) and felt emotions (Juslin et al., 2015) in previous studies. In particular, the *MecScale* has been validated in a series of highly controlled listening experiments, where musical stimuli were carefully edited so as to present or withhold the type of information required to trigger distinct mechanisms (e.g., Juslin et al., 2014). The predictive validity was quantified and the *MecScale* items showed significant and strongly positive correlations with their respective target-mechanism conditions and negative correlations with all other mechanism conditions; none of the items correlated positively with neutral control conditions. In a study by Juslin et al. (2015), the *MecScale* items could predict the right target-mechanism condition from *MecScale* ratings with an accuracy of 75% correct. A single-item measure has the obvious advantage that it minimizes participant burden, which is especially important in the case of a long listening test.

Personality

The listeners’ personality traits were measured using the Big Five Inventory (BFI; John et al., 1991; for a review, see John et al., 2008). This includes 44 items measuring the factors *Openness to experience* (O), *Conscientiousness* (C), *Extraversion* (E), *Agreeableness* (A), and *Neuroticism* (N) (OCEAN). Each factor has also been divided into personality facets that are represented by the 44 items. The domain scales of BFI have shown good internal consistency, clear factor structure, adequate convergent-discriminant validity coefficients, and substantial self-peer agreement (Fossati et al., 2011).

Participants indicated on a 5-point Likert scale the extent to which they agreed to each self-statement (from 1 = *disagree strongly* to 5 = *agree strongly*). Scores were calculated for each factor. Cronbach’s alpha reliability for the domain scales were .85 (N), .84 (E), .64 (O), .59 (A), and .80 (C), respectively.²

Acoustic Features

To be able to compare mechanism-based and acoustics-based prediction of self-reported emotions (H4), we needed a proper acoustic description of each musical stimulus, in the form of numerical values that could be entered into multiple regression analyses. Using a new set of algorithms developed by one of the authors (OL), we obtained 14 acoustic features covering six major aspects: tempo (Metroid *M*, Metroid *SD*); dynamics (Crescendo); timbre (Bass, Midrange, Treble, Roughness); harmony (Key clarity *M*, Mode *M*, Mode *SD*); rhythm (Metrical strength *M*, Metrical strength *SD*); and musical structure (Mfcc Nov, a measure of changes in timbre, KeyQuant15, a measure of tonal modulation). The goal was to include as many acoustic features as possible within the limits set by the minimum 5 to 1 ratio of cases to every predictor stated in the judgment-analysis literature (Cooksey, 1996). With 72 pieces of music, we could thus include 14 features (14 x 5 = 70 cases). A more detailed description of the features and their algorithms appears in Appendix C.

Procedure

When the participants arrived at the music laboratory, they were seated in a comfortable armchair, in front of

² Comparison with a larger Swedish sample ($N = 431$) from Zakrisson (2010) shows that the mean values were similar for *Neuroticism* (present sample: 21.84, Zakrisson: 20.07), *Extraversion* (25.00, 27.52), *Agreeableness* (34.16, 34.92), and *Conscientiousness* (31.66, 34.92). However, the present sample was higher in *Openness to experience* (40.48, 34.67), which could reflect a greater interest in music among the participants.

a computer monitor. A brief introduction and instructions concerning the experiment were shown on the monitor. The participants were informed that they would listen to a variety of pieces of music, and that after each piece, they should describe their experience of the music by means of rating scales on the computer screen. The instructions explained that the participants should rate their own feelings and not what the music is expressing. (Previous research has indicated that listeners can make this distinction; see Zentner et al., 2008.)

After informed consent was obtained from the participants, they were told to relax for a few minutes during silence. Then, the listening test began. After each piece, the rating scales (see Measures) appeared on the computer screen. During the rating, a 30-second segment of the musical excerpt was repeated at a lower sound level to facilitate the ratings (the segment included the beginning of the complete excerpt and was the same for all listeners). After the rating followed a 10-second pause before the next piece was played. Because the focus of the study was on the listener responses rather than on the musical stimuli, the stimulus order was randomized and kept constant across listeners so as to enable individual comparisons (as in a psychometric test). Had we used a unique stimulus order for each participant, any difference between listeners might simply reflect order effects (which appear to be strong in music; see Flores & Ginsburgh, 1996). All participants were tested individually. The only other people present during testing were the experimenters (LS, GB) who were separated from the listener by a blind. The room was dimly lit and sound attenuated. Stimuli were played through a pair of high-quality loudspeakers (*Dali Ikon 6 MK2*) at a comfortable sound level, which was kept constant across participants.

After the listening test, participants filled out a background questionnaire and the BFI. Sessions lasted between 100 and 180 min, depending on the participant. As noted above, we also collected other data during the sessions that are reported separately (Juslin, Sakka et al., 2016).

Results

DATA CHECKS

With 44 listeners, 20 rated variables (12 emotions and 8 mechanisms), and 40 excerpts per listener, we obtained a total of 35,200 data points to analyze in depth. Data quality checks indicated that the ratings of each variable used the full range of the scale (0-4); that there was no

straightlining³ of responses; and that there were few missing values: < 4 % for emotions; < 2.5 % for mechanisms. Examination of the distributions of the variables showed no skewness values outside the -2 to +2 range, and no (absolute) kurtosis values outside the -7 to +7 range (e.g., Byrne, 2010; Hair et al., 2010; West et al., 1995),⁴ with the only exception of the ratings of *disgust-contempt* (which featured a very large number of zeros). However, we retained this variable in the first analyses, because a recent Monte Carlo simulation study, featuring 10,000 replications of 1,308 conditions, concluded that ANOVAs are highly robust against deviations from normality (Blanca et al., 2017).

EMOTIONS

Prevalence of emotions was measured by means of the mean ratings of the 12 emotion scales. To explore how emotion prevalence varied as a function of emotion category, we conducted a two-way mixed ANOVA, with Emotion as within-subjects factor (12 levels) and Group as between-subjects factor (2 levels). The results showed a significant main effect of Emotion, $F(11, 462) = 66.955$, $p < .001$, partial $\eta^2 = 0.615$, approaching a “strong” effect in terms of Ferguson’s (2009) guidelines. By contrast, there was no significant effect of Group, $F(1, 42) = 1.713$, $p = .198$, partial $\eta^2 = 0.039$; nor was there a significant interaction between Emotion and Group, $F(11, 462) = 1.200$, $p = .284$, partial $\eta^2 = 0.028$. As illustrated in Figure 1, upper panel), the overall trends were strikingly similar for the two groups, who listened to the same musical genres, albeit (mostly) different musical excerpts.

The impression that the two listener groups reacted similarly was reinforced by further analyses. Recall that eight musical excerpts were heard by both groups (Method section). This enabled us to compare the responses of the groups to the same pieces of music. We conducted a two-way mixed ANOVA, with Piece as within-groups factor (8 levels / pieces) and Group

³ Straightlining, also referred to as “non-differentiation in ratings,” occurs when a respondent gives identical (or nearly identical) answers to a series of questions using the same response scale, which reduces data quality. This may happen if the respondent is losing motivation (e.g., because he or she is bored).

⁴ Hair et al. (2010) and Byrne (2010) argue that data is considered to be normal if skewness is between -2 to +2 and kurtosis (proper) is between -7 to +7. Notably, common statistics software (e.g., SPSS, Statistica, SAS) use a correction factor (-3) to set the kurtosis measure to zero for a normal distribution. This is referred to as excess kurtosis, as opposed to absolute kurtosis or kurtosis proper, which for a normal distribution corresponds to 3.

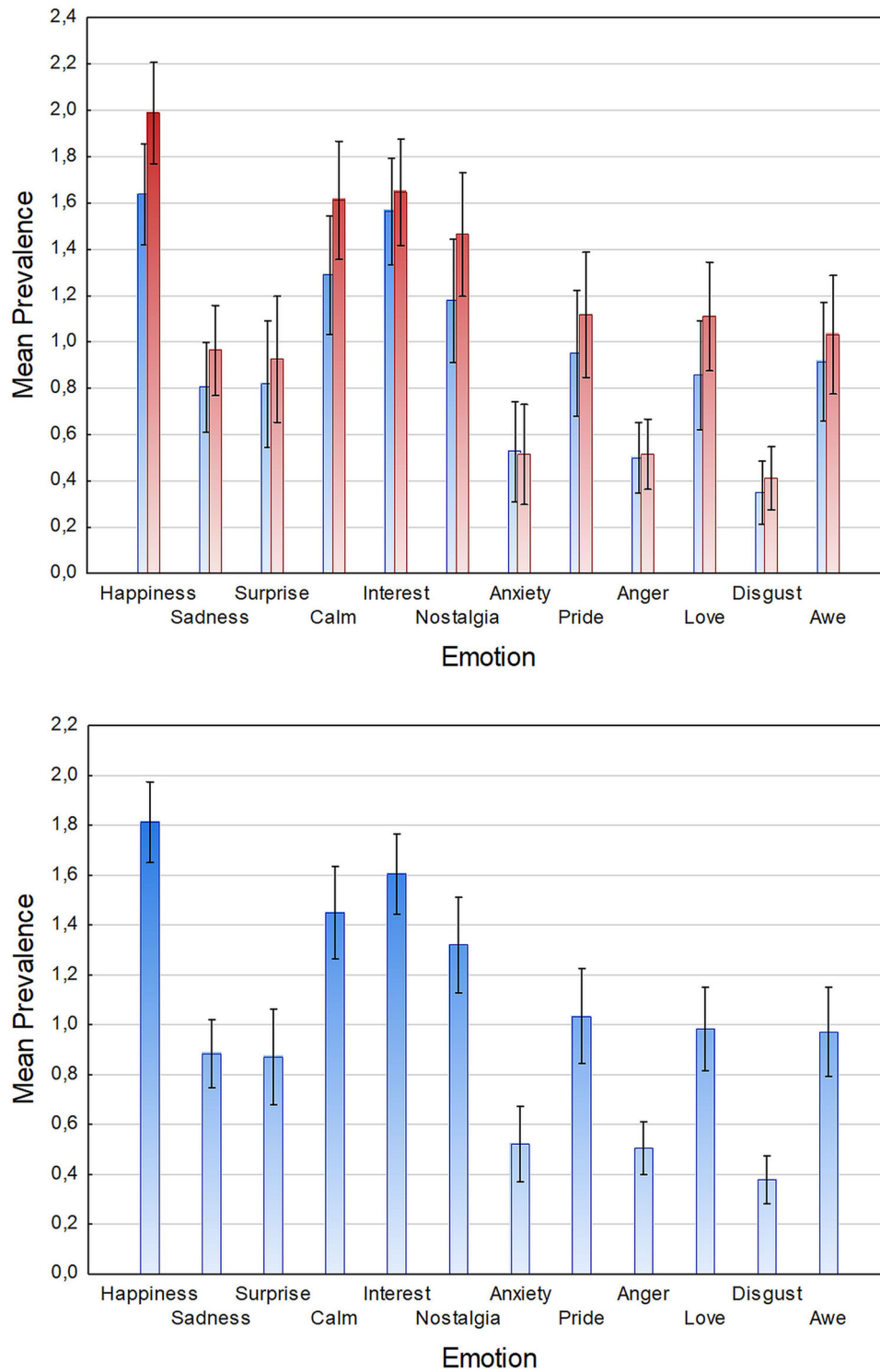


FIGURE 1. Upper panel: Mean prevalence of emotion categories as a function of listener group (blue bars = Group 1, red bars = Group 2). Lower panel: Mean prevalence of emotion categories across listener groups.

TABLE 1. Post Hoc Tests (Tukey's HSD) of Contrasts Between Emotion Categories in Estimated Prevalence

Emotion	1	2	3	4	5	6	7	8	9	10	11
1. Happiness	-										
2. Sadness	< .001	-									
3. Surprise	< .001	.999	-								
4. Calm	< .001	< .001	< .001	-							
5. Interest	.235	< .001	< .001	.703	-						
6. Nostalgia	< .001	< .001	< .001	.883	.014	-					
7. Anxiety	< .001	< .001	< .001	< .001	< .001	< .001	-				
8. Pride	< .001	.737	.625	< .001	< .001	.012	< .001	-			
9. Anger	< .001	< .001	< .001	< .001	< .001	< .001	.999	< .001	-		
10. Love	< .001	.983	.958	< .001	< .001	< .001	< .001	.999	< .001	-	
11. Disgust	< .001	< .001	< .001	< .001	< .001	< .001	.812	< .001	.897	< .001	-
12. Awe	< .001	.993	.98	< .001	< .001	< .001	< .001	.999	< .001	.999	< .001

Note: Values indicate *p* values.

as between-groups factor, for each emotion scale. After Bonferroni adjustment for multiple tests ($n = 36$) from $\alpha = .05$ to $\alpha = .0014$, results revealed a significant effect of Piece on seven of the emotion scales (*happiness-elation*, *surprise-astonishment*, *nostalgia-longing*, *irritation-anger*, *love-tenderness*, *disgust-contempt*, and *awe-admiration*). By contrast, there was no significant main effect of Group, nor any significant interaction involving Group. These data suggest that the two groups experienced the common excerpts in a largely similar manner, at a nomothetic level of analysis.

Given the absence of both main and interaction effects of Group, we collapsed the data across groups. Figure 1 (lower panel) shows mean values and 95% confidence intervals, and Table 1 presents post hoc tests of all differences (in the form of Tukey's HSD). *Happiness-elation* was significantly more prevalent than all other emotions, except *interest-expectancy*, which, in turn, was more prevalent than all the remaining emotions except *calm-contentment*. *Calm-contentment* was more prevalent than all remaining emotions except *nostalgia-longing*. The above emotions were all significantly more prevalent than a set of slightly less frequent, but still fairly common states (*sadness-melancholy*, *surprise-astonishment*, *pride-confidence*, *love-tenderness*, and *awe-admiration*) that did not differ significantly from one another. The least common states were *anxiety-nervousness*, *anger-irritation*, and *disgust-contempt*, which were significantly less common than the other emotions, but did not differ significantly from one another. Notably, the descriptive statistics presented in Figure 1 conceal some quite large individual differences in overall prevalence of specific emotions (i.e., across pieces of music). Even in the most prevalent emotion category (i.e., *happiness-elation*), there was considerable variability, as indicated by the lowest (0.675) and highest (2.800) individual mean values.

To test whether positive emotions were more prevalent than negative emotions (H1), we computed the individual mean values for both types of emotions and carried out a within-subjects *t*-test of the difference. Based on previous emotion research (Plutchik, 1994; Russell, 1980), we categorized five emotions as positively valenced (*happiness-elation*, *calm-contentment*, *interest-expectancy*, *pride-confidence*, and *love-tenderness*), and four emotions as negatively valenced (*sadness-melancholy*, *anxiety-nervousness*, *irritation-anger*, *disgust-contempt*). Three emotions (*surprise-astonishment*, *nostalgia-longing*, *awe-admiration*) were difficult to categorize with respect to valence and were therefore left out from the analysis as a precaution.⁵ The results showed that mean ratings of emotion prevalence were significantly higher for positive emotions ($M = 1.377$, $SD = 0.494$) than for negative emotions ($M = 0.608$, $SD = 0.418$), $t_{43} = 12.142$, $p < .001$, $d = 1.287$, a "large" effect (Cohen, 1988). This tendency for positive emotions to be more prevalent than negative emotions occurred for all individual listeners except two, which corresponds to 95% of the sample.

As an exploratory follow-up analysis, we also looked at how the prevalence of specific emotions varied depending on genre. Figure 2 presents means and 95% confidence intervals for the 12 emotion categories, as a function of the STOMP factors (see Method section). As seen, the prevalence patterns were mostly similar across the four factors. However, note that the R&C category evoked significantly more *calm-contentment* and *awe-admiration* than the I&R category

⁵ *Surprise* is usually regarded as a neutral (proto) emotion (Simons, 1996); *nostalgia-longing* is usually regarded as a "bittersweet" or "mixed" emotion, which involves both positive and negative affect (Wildschut et al., 2006); and *awe* is also regarded as incorporating elements of both positive and negative affect (Haidt & Seder, 2009).

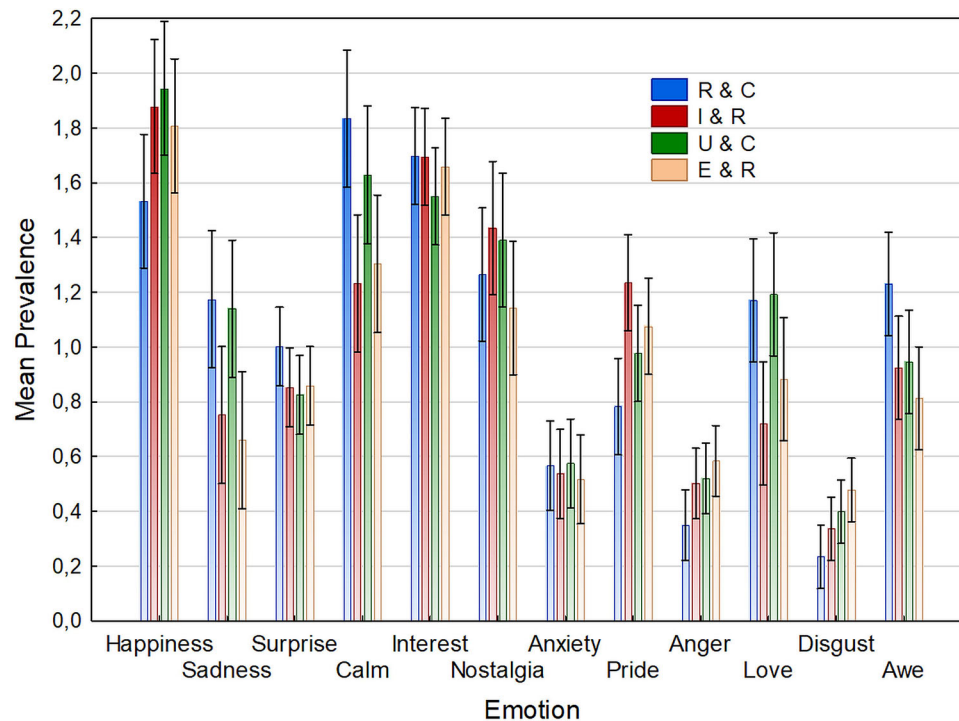


FIGURE 2. Mean prevalence of emotion categories as a function of music genre factor (STOMP). R & C = Reflective & Complex, I & R = Intense & Rebellious, U & C = Upbeat & Conventional, E & R = Energetic & Rhythmic.

(both p 's < .05, as shown by post hoc tests in the form of Tukey's HSD).

In the following emotion analyses, we will leave out the three emotions with the lowest prevalence (*anxiety-nervousness*, *irritation-anger*, and *disgust-contempt*), consistent with our focus on positive emotions, and because the low prevalence ($M < 0.60$) meant that there were few cases.

MECHANISMS

To examine how mechanism prevalence varied as a function of mechanism type, we conducted a two-way mixed ANOVA with Mechanism as within-subjects factor (8 levels) and Group as between-subjects factor (2 levels). The results revealed a significant main effect of Mechanism, $F(7, 294) = 34.808$, $p < .001$, partial $\eta^2 = 0.453$, a "moderate" effect (Ferguson, 2009). In contrast, there was no significant effect of Group, $F(1, 42) = 2.589$, $p = .115$, partial $\eta^2 = 0.058$. Nor was there a significant interaction between Mechanism and Group, $F(7, 294) = 1.131$, $p = .343$, partial $\eta^2 = 0.026$. As illustrated in Figure 3 (upper panel), the mean trends were quite similar for the two listener groups (except that Group 2 showed marginally higher ratings for all mechanisms except Brain stem reflex and Musical expectancy).

We also compared the listener groups regarding the eight pieces of music that were rated by both groups. We conducted a two-way mixed ANOVA, with Piece as within-groups factor (8 levels/pieces) and Group as between-groups factor for each mechanism scale. After Bonferroni correction for multiple tests ($n = 24$) from $\alpha = .05$ to $\alpha = .0021$, the results showed a significant effect of Piece for each mechanism. Conversely, there were no significant main effects of Group and only a single significant interaction involving Group. (Group 1 rated the Musical expectancy mechanism higher than Group 2 for the Sepultura excerpt; Appendix A.)

Given the overall absence of significant group effects, we collapsed data across groups. Figure 3 (lower panel) shows mean values and 95% confidence intervals, and Table 2 shows post hoc significance tests of all differences (in the form of Tukey's HSD). As can be seen in Figure 3, the most frequent mechanisms were Rhythmic entrainment, Aesthetic judgment, Evaluative conditioning, Visual imagery, and Episodic memory. Rhythmic entrainment was rated as significantly more common than all other mechanisms, except Aesthetic judgment, which was significantly more common than all mechanisms, except Evaluative conditioning and Visual imagery. These mechanisms in turn did not differ

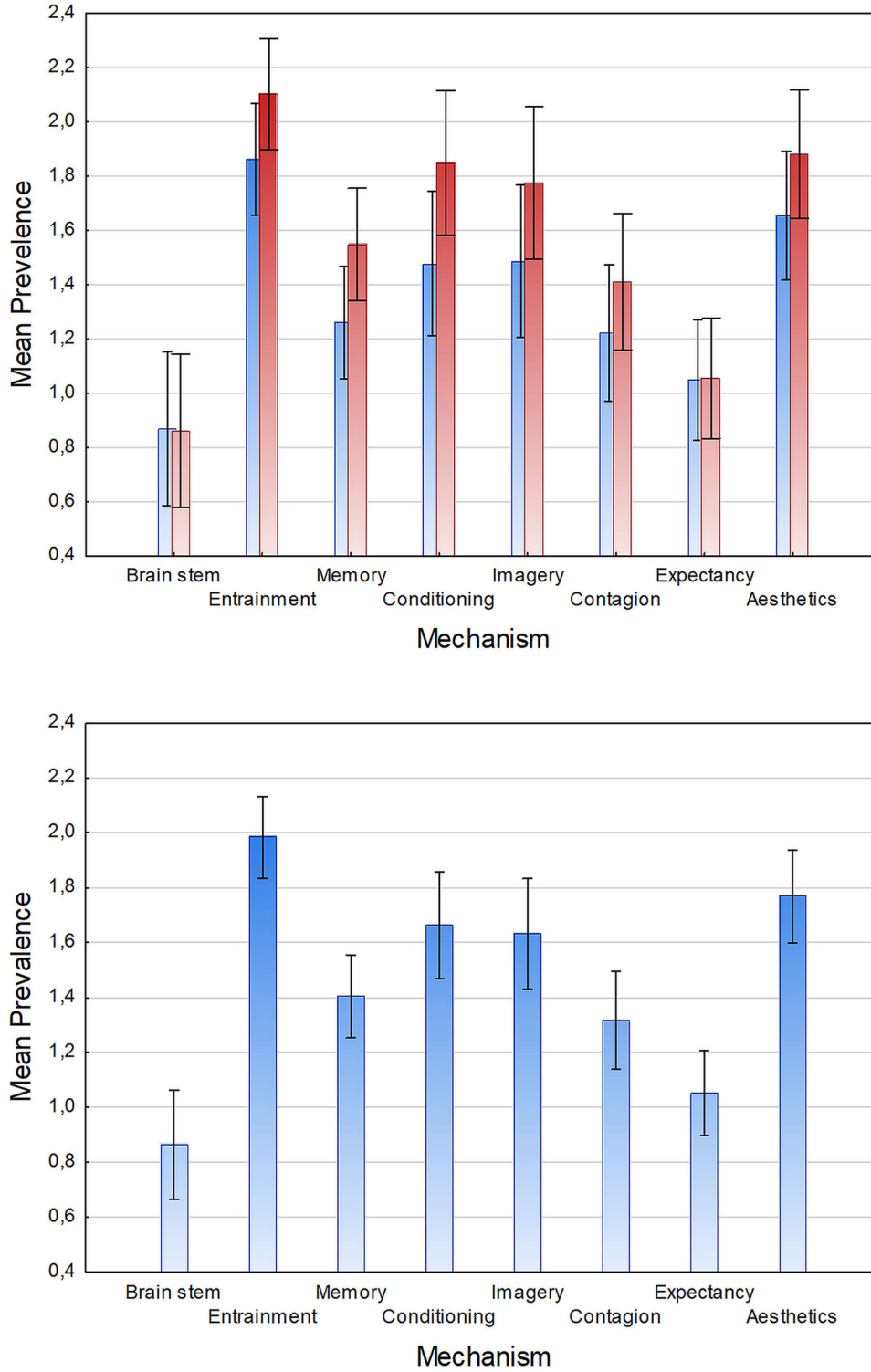


FIGURE 3. Upper panel: Mean prevalence of emotion mechanisms as a function of listener group (blue bars = Group 1, red bars = Group 2). Lower panel: Mean prevalence of emotion mechanisms across listener groups.

TABLE 2. Post Hoc Tests (Tukey's HSD) of Contrasts Between Induction Mechanisms in Estimated Prevalence

Mechanism	Brain stem	Entrainment	Conditioning	Contagion	Imagery	Memory	Expectancy
Entrainment	< .001	–					
Conditioning	< .001	.009	–				
Contagion	< .001	< .001	.003	–			
Imagery	< .001	.002	.999	.012	–		
Memory	< .001	< .001	.078	.979	.184	–	
Expectancy	.430	< .001	< .001	.063	< .001	.002	–
Aesthetic	< .001	.249	.939	< .001	.795	.001	< .001

Note: Values indicate *p* values.

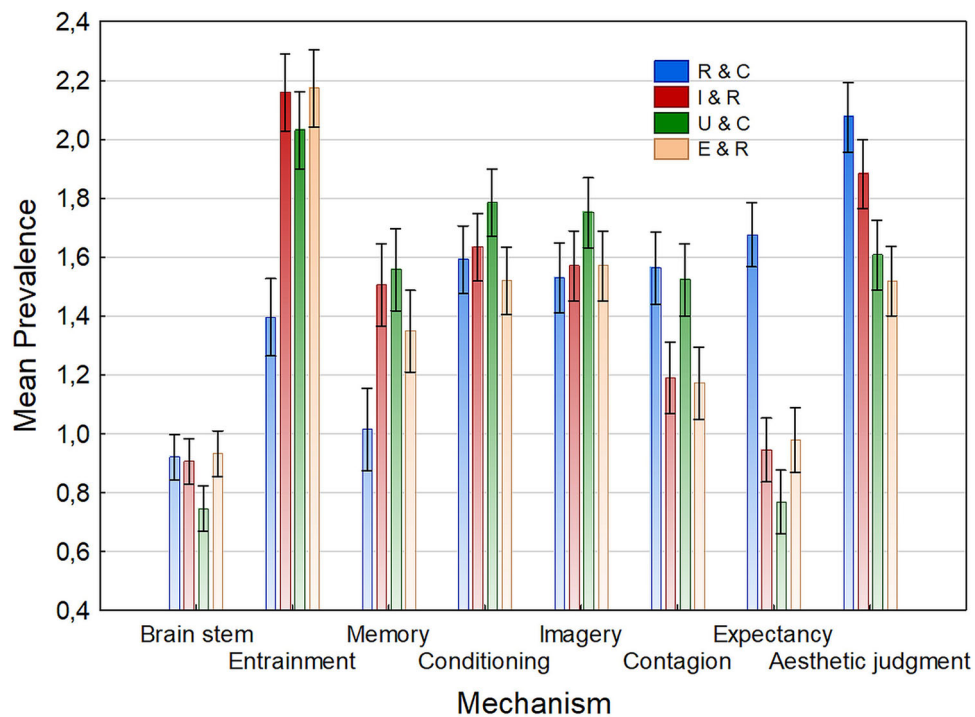


FIGURE 4. Mean prevalence of emotion mechanisms as a function of music genre factor (STOMP). R & C = Reflective & Complex, I & R = Intense & Rebellious, U & C = Upbeat & Conventional, E & R = Energetic & Rhythmic.

significantly from one another, but were both more frequent than all other mechanisms except Episodic memory, which did not differ from Contagion. In addition, as hypothesized (H2), Brain stem reflex and Musical expectancy received significantly lower mean ratings than all other mechanisms, except that the contrast between Musical expectancy and Contagion did not reach significance ($p = .06$).

Similar to our analyses of the emotion categories, we also plotted the prevalence of emotion mechanisms as a function of STOMP category. Figure 4 shows the means and 95% confidence intervals. As can be seen, mechanism prevalence did not vary much across genre

categories. However, note that the R&C genre category produced lower ratings of Rhythmic entrainment and higher ratings of Musical expectancy than the other categories (all p 's < .05, Tukey's HSD).

EMOTION-MECHANISM LINKS: NOMOTHETIC ANALYSES

To investigate how well listeners' ratings of mechanism impressions (*MecScale*) could predict ratings of specific emotions at a nomothetic level (H3), we conducted a series of multiple regression analyses, based on the mean values across listeners. More specifically, the mean rating on the respective emotion scale was the dependent variable, and the mean ratings on the *MecScale*

TABLE 3. Summary of Nomothetic Regression Models of Emotion-Mechanism Links

Dependent / Emotion	R	F	Predictor / Mechanism item (β)							
			Brainstem	Entrainment	Condit.	Contagion	Imagery	Memory	Expectancy	Aesthetic
Happiness-elation	.869	24.223	.126	.642	-.001	.006	.286	.022	-.040	.152
Sadness- melancholy	.864*	22.757	-.080	-.480	.489	.472	-.180	-.370	-.230	.153
Surprise- astonishment	.758*	10.493	.476	-.070	.048	-.140	.163	-.020	.466	-.130
Calm-contentment	.754*	10.228	-.240	-.060	.191	.316	.258	-.470	.033	.305
Interest-expectancy	.871*	24.416	.255	.589	-.460	.814	.032	.175	.381	.017
Nostalgia-longing	.930	50.535	-.100	-.110	.056	.161	.033	.471	-.170	.385
Pride-confidence	.864	23.118	.204	.491	.014	-.290	.191	.046	-.280	.382
Love-tenderness	.854	21.230	.026	-.310	.287	.555	.364	-.550	-.290	.019
Awe-admiration	.863	23.009	.074	-.100	.147	.237	.191	-.180	-.050	.542

Note: All multiple correlations (R) are statistically significant at $p < .0056$, with Bonferroni correction for multiple tests, $n = 9$. Standardized beta weights (β) that are both statistically significant ($p < .05$) and positive in direction are marked in bold. $N = 72$, except for * where $N = 71$, after one outlier ($> \pm 3$ standard residuals) has been removed.

items were the independent variables (all variables coded continuously). The unit of analysis was piece of music; that is, each piece was a case, with a mean value for each variable. Because there were such small differences in ratings between the two listener groups and the analysis proceeded at a mean level, we decided to include all 72 pieces in the analyses. For the eight excerpts that were rated by both groups, we simply used the mean across groups. We adopted a “simultaneous” regression approach, which should be used whenever there is no theoretically defensible order of the predictors (e.g., Cohen & Cohen, 1983). Consideration of Tolerance values (the inverse of the Variation Inflation Factor) showed that multi-collinearity among the predictors was not a problem (Tabachnick & Fidell, 2001).

Table 3 shows a summary of the results. After Bonferroni correction for multiple tests ($n = 9$), from $\alpha = .05$ to $\alpha = .0056$, all multiple correlations remained statistically significant. As can be seen, the multiple correlations varied depending on the emotion scale—largest for *nostalgia-longing*, and smallest for *calm-contentment*—but did not fall below $R = .754$ ($M = .847$). The mean variance accounted for ($R^2 = 72\%$, 95% CI [.649, .793], $SD = 0.093$) shows that listeners’ felt emotions could be predicted rather well based on mechanism ratings.

Table 3 also presents the beta weights (β) for the predictors (i.e., the mechanism items) of each model, which (in accordance with how the regression models were computed) should be interpreted in a row-wise fashion. The beta weights that are both statistically significant ($p < .05$) and positive in direction have been marked in red.

As seen in Table 3, each emotion category was linked with a few key mechanisms (i.e., predictors); for instance, *happiness-elation* was predicted by Rhythmic entrainment; *sadness-melancholy* was predicted by Evaluative conditioning and Contagion; *surprise-astonishment* was predicted by Brain stem reflex and Musical expectancy; *calm-contentment* was predicted by Contagion and Aesthetic judgment; *nostalgia-longing* was predicted by Episodic memory, as well as Aesthetic judgment; *pride-confidence* was predicted primarily by Entrainment and Aesthetic judgment; *love-tenderness* was mainly predicted by Contagion and Visual imagery; and *awe-admiration* was predicted by Aesthetic judgment, and also to some extent Contagion. *Interest-expectancy* stood out by being associated with several mechanisms (see Table 3).

COMPARISON OF PREDICTIVE ACCURACY: MECHANISMS VS. ACOUSTICS

To compare our prediction of felt emotions based on the ratings of mechanism items with one based on acoustic features (H4), we conducted a series of multiple regression analyses for the various emotions, using the data obtained with the computational algorithms (see Method section). The mean rating on the emotion scale was the dependent variable, and the acoustic features were the independent variables (all variables coded continuously). The analyses were in other regards similar to the ones conducted based on mechanism ratings. As acoustic features are more or less correlated, possible multicollinearity among predictors was a concern. However, checks revealed no Tolerance values < 0.25 .

TABLE 4. Comparison of Predictive Accuracy for Models Based on Mechanisms vs. Acoustic Measures

Dependent / Emotion	Type of Model				Comparison	
	Mechanisms		Acoustics		R^2 diff.	p^2
	R	F	R	F		
Happiness-elation	.869	24.223*	.598	2.263	0.397	< .001
Sadness-melancholy	.864 ¹	22.757*	.696 ¹	3.750*	0.262	< .001
Surprise-astonishment	.758 ¹	10.493*	.500	1.360	0.325	.002
Calm-contentment	.754 ¹	10.228*	.708 ¹	4.014*	0.068	.217
Interest-expectancy	.871 ¹	24.416*	.543	1.704	0.464	< .001
Nostalgia-longing	.930	50.535*	.674 ¹	3.326*	0.411	< .001
Pride-confidence	.864	23.118*	.589	2.168	0.399	< .001
Love-tenderness	.854	21.230*	.681	3.520*	0.265	< .001
Awe-admiration	.863	23.009*	.612	2.436	0.370	< .001

Note: Multiple correlations (R) marked with * are statistically significant at $p < .0056$, with Bonferroni correction for multiple tests, $n = 9$. For comparison, model data for mechanisms from Table 3 have been included also. $N = 72$.

¹ $N = 71$ after one outlier ($> \pm 3$ standard residuals) has been removed.

² Tested by means of Steiger's (1980) Z test of the difference between two dependent correlations with one variable in common (one-tailed test; see Lee & Preacher, 2013).

Hence, although there was some moderate collinearity, levels were clearly within acceptance (see Tabachnick & Fidell, 2001).

Table 4 presents the main results from the analyses in terms of multiple correlations and F values. Also included are the results from our mechanism-based prediction, thus enabling a comparison of the variance accounted for by the two types of models (the right-most column). As can be seen, the multiple correlations (R) for the models based on acoustic features ranged from .500 to .708 ($M = .622$, $SD = .072$). Only four of the models were statistically significant, but the mean variance accounted for ($R^2 = 39\%$, 95% CI [.324, .460], $SD = 0.088$) shows that felt emotions could be predicted to a moderate extent based on acoustic features. (The precise patterns of features are beyond the focus of this study, but a summary of correlations between emotions and features is provided in Appendix D.)

However, it may also be seen in Table 4 that mechanism-based models were better able to predict felt emotions. Note that the multiple correlation (R) was significantly larger for the mechanism-based model than for the acoustics-based model for all emotion categories except *calm-contentment* (tested by means of Steiger's, 1980, Z test for dependent correlations with one variable in common, one-tailed tests). The mechanism-based models accounted for 72% of the variance in emotion ratings, on average, as compared with only 39% for the acoustics-based models.

EMOTION-MECHANISM LINKS: IDIOGRAPHIC ANALYSES

To investigate the extent to which individual listeners display the same emotion-mechanism links (H5), we conducted a series of idiographic multiple regression

analyses. One disadvantage of idiographic analyses is that they tend to yield large amounts of data. Thus, we restricted our analyses to a subset of 10 participants (five males and five females), which were randomly selected using the *create a random sample* function in the *Statistica* software. Since the aim was to compare individual listeners, we sampled listeners from just one of the listener groups (Group 2 - also randomly selected), who heard the same musical excerpts. We focused on the four most prevalent emotion categories: *happiness-elation*, *calm-contentment*, *interest-expectancy*, and *nostalgia-longing*. (Notably, this included both the *most* successfully and the *least* successfully predicted of the emotions in the nomothetic analyses.) The number of cases ($N = 40$) is relatively small for a multiple regression analysis, but exceeds the minimum 5 to 1 ratio of cases to every predictor (cf. Cooksey, 1996). The aim was not to generalize individual patterns of results to the general population, but to investigate the nature of the links between emotions and mechanisms for the specific individuals sampled.

Table 5 shows a summary of the results in terms of the individual multiple correlations and semi-partial correlations with each mechanism item (*MecScale*). Rather than performing significance tests on these small samples, we report effect sizes along with interpretations of these. All semi-partial correlations $\geq .20$ are marked with an asterisk (*). This corresponds to the recommended minimum effect size representing a practically significant effect for social science data (RMPE), as proposed by Ferguson (2009).

Careful inspection of Table 5 suggests both some general tendencies, similar to those observed in the nomothetic analyses earlier, and some idiosyncratic

TABLE 5. Summary of Idiographic Regression Models of Emotion-Mechanism Links

Emotion	R	Mechanism (r_{sp})							
		Brainstem	Entrainment	Condit.	Contagion	Imagery	Memory	Expectancy	Aesthetic
<i>Happiness-elation</i>									
Listener 1.	.750	-.092	.143	.283	.163	-.132	-.254	-.031	-.176
2.	.909	.106	.271	.151	-.159	.125	.067	-.106	.154
3.	.681	.162	.244	-.088	.204	.136	.086	-.109	-.167
4.	.733	.193	.424	.039	-.140	.053	.128	-.119	.188
5.	.740	-.011	.388	-.111	-.104	.194	.042	.187	.158
6.	.810	-.093	.328	-.146	.186	-.062	.224	-.008	.075
7.	.724	-.230	.355	-.043	-.134	.069	.109	.229	.272
8.	.751	.162	.221	-.090	.223	.093	.240	.059	.038
9.	.841	-.060	.274	.044	.151	-.005	.315	.141	.129
10.	.806	-.120	.417	.196	-.018	.052	.203	.088	.146
<i>Calm-contentment</i>									
Listener 1.	.755	-.308	.047	.189	.283	-.068	-.404	.249	.181
2.	.450	.181	.017	.043	.018	-.001	.028	-.157	.195
3.	.638	-.033	.051	-.252	.259	-.011	-.064	-.072	.026
4.	.829	-.011	-.027	.102	.211	-.036	.075	-.137	.499
5.	.641	.053	-.331	-.033	.164	.026	-.070	-.070	-.424
6.	.656	-.306	.181	-.250	.369	.224	.064	.029	-.071
7.	.839	-.299	-.074	.071	.094	-.031	-.291	.036	.437
8.	.631	.144	-.284	.003	.303	-.251	.127	.013	.110
9.	.656	.070	-.243	.130	.292	-.024	.003	-.044	-.084
10.	.739	-.402	-.027	.105	-.169	.080	-.070	.379	.468
<i>Interest-expectancy</i>									
Listener 1.	.831	-.001	.127	-.068	.069	.194	-.015	-.031	.019
2.	.901	.233	.144	.017	.043	.101	-.061	-.115	.384
3.	.926	.212	-.040	-.107	.173	.086	-.099	.274	.022
4.	.727	.243	.223	.004	.067	.041	.069	.172	.120
5.	.860	.327	.003	.280	.040	-.081	.103	.189	.111
6.	.787	-.088	.247	-.165	.303	.094	.068	.187	.063
7.	.849	-.104	-.078	-.118	.075	.175	.102	.364	.212
8.	.792	.076	.237	.001	.259	.048	-.063	.256	.151
9.	.694	.065	.105	.057	-.066	-.195	.171	.508	.170
10.	.783	.072	.129	-.081	.159	.195	.220	.191	.075
<i>Nostalgia-longing</i>									
Listener 1.	.691	-.060	-.152	.173	.085	-.120	.086	-.179	.176
2.	.947	-.048	.088	-.068	.106	.090	.203	.032	.027
3.	.807	-.240	.265	.073	.359	-.106	.319	-.008	-.053
4.	.768	.082	-.055	.155	.083	-.020	.292	-.176	.238
5.	.781	-.218	-.173	.150	.243	.027	.143	-.050	-.186
6.	.832	.075	-.002	-.046	.128	-.091	.231	-.203	.323
7.	.765	-.259	-.175	-.063	.171	.221	.040	.066	.240
8.	.713	.137	-.215	.041	.133	.247	.179	-.149	.058
9.	.863	-.149	.039	.109	.019	.028	.354	.008	.141
10.	.844	-.057	-.176	.059	-.112	.279	.405	-.221	.343

Note: R = multiple correlations, r_{sp} = semi-partial correlations (those marked in bold are both positive and $\geq .20$ = RMPE). $N = 39$ – 40 , due to a few missing cases (no outliers removed).

relationships unique to individual listeners. For instance, note that *happiness-elation* was typically predicted by the Rhythmic entrainment mechanism, which met the RMPE criterion for 9 out of 10 listeners. However, for a subset of the listeners (40%), *happiness-elation* was also (and in some cases even more strongly) linked to Episodic memory; there were further some

truly idiosyncratic emotion-mechanism links, for example that the ratings of *happiness-elation* were predicted mainly by Evaluative conditioning for Listener 1, but by Musical expectancy and Aesthetic judgment for Listener 7.

For the *calm-contentment* category, the Contagion mechanism was the most frequent strong predictor (as

TABLE 6. Correlations Between Big Five Personality Traits and Individual Prevalence of Emotions and Mechanisms Respectively

	Personality Traits				
	O	C	E	A	N
<i>Emotions</i>					
Happiness-elation	-.174	-.180	.115	-.051	.007
Sadness-melancholy	.070	-.161	.072	-.098	.051
Surprise-astonishment	.034	-.323*	.148	-.211*	.197
Calm-contentment	-.059	-.279*	.137	-.132	.017
Interest-expectancy	-.178	-.232*	.083	-.124	.091
Nostalgia-longing	-.071	-.221*	.107	-.028	.077
Pride-confidence	.060	-.335*	.121	-.239*	.153
Love-tenderness	-.049	-.213*	.074	-.149	.113
Awe-admiration	.067	-.232*	.086	-.107	.025
<i>Mechanisms</i>					
Brain stem reflex	.012	-.191	.006	-.113	.164
Rhythmic entrainment	-.136	-.199	.220*	-.031	-.036
Evaluative conditioning	-.019	-.173	.094	-.092	.233*
Contagion	-.020	-.149	.029	-.008	.074
Visual imagery	.129	-.213*	.011	.110	.133
Episodic memory	-.013	-.157	.195	.040	.254*
Musical expectancy	.005	-.096	-.084	.141	-.166
Aesthetic judgment	-.159	-.008	-.078	-.040	.027

Note: Values show Pearson correlations. O = Openness to Experience; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism. All correlations $\geq r = .20$ are marked with an asterisk. This corresponds to “the recommended minimum effect size representing a “practically” significant effect for social science data” (RMPE; see Ferguson, 2009). $N = 44$.

implied by the nomothetic regression model also), reaching the RMPE criterion for 60% of the listeners. Moreover, *calm-contentment* was linked to the Aesthetic judgment mechanism for three listeners. Note further that for Listener 6, *calm-contentment* was well predicted by the Visual imagery mechanism.

For the *interest-expectancy* category, the most frequent predictive relationships occurred for the Musical expectancy, Brain stem reflex, and—to a lesser extent—Rhythmic entrainment mechanisms. However, many different mechanisms were linked to this emotion depending on the listener (e.g., Evaluative conditioning; Contagion; Aesthetic judgment; see Table 5).

As regards *nostalgia-longing*, Episodic memory was (unsurprisingly) the mechanism most frequently implicated (for at least 60% of the listeners). Note, however, that Aesthetic judgment was also a good predictor of this emotion, for at least four of the listeners. Again, there were some idiosyncratic tendencies (e.g., links with Contagion for two listeners).

LINKS TO PERSONALITY TRAITS

In view of the wide individual differences in emotion prevalence and mechanisms, we explored whether the differences were linked to various personality traits, as indexed by the BFI. Table 6 presents the Pearson correlations (r) between individual BFI test scores for the Big

Five traits and individual prevalence of emotions and mechanisms, respectively.

Since these analyses were primarily exploratory in nature, we decided to refrain from performing significance tests and to just report the effect sizes along with interpretations of these. All correlations $\geq r = .20$ (RMPE, Ferguson, 2009) are marked with an asterisk (*) in Table 6. Only 13 out of 85 correlations (15%) met this criterion. All of these effects can be categorized as “small” to “moderate” (Cohen, 1988).

As seen in Table 6, Neuroticism was positively correlated with prevalence of *surprise-astonishment* and the mechanisms Brain stem reflex, Evaluative conditioning, and Episodic memory. Extraversion was positively correlated with the mechanism Rhythmic entrainment. Agreeableness was negatively correlated with the emotions *surprise-astonishment* and *pride-confidence*. Conscientiousness was negatively correlated with seven of the emotions, as well as with the mechanisms Rhythmic entrainment and Visual imagery.

Discussion

WHICH EMOTIONS DOES MUSIC EVOKE MOST FREQUENTLY?

The results showed that prevalence varied markedly depending on the emotion category (a “large” effect; Ferguson, 2009), and replicated a number of previous

findings. First of all, as hypothesized (H1), positively valenced emotions were significantly more common in reaction to music, than negatively valenced emotions (a “large” effect; Cohen, 1988; for similar results in previous studies, see Juslin et al., 2008, 2011; and Sloboda et al., 2001). Second, the music induced both “basic” and “complex” emotions (see Gabrielsson, 2011; Juslin & Laukka, 2004; Juslin, Barradas et al., 2016; Sloboda, 1992). Third, certain emotion categories were significantly more frequent than others: *happiness-elation*, *interest-expectancy*, *calm-contentment*, and *nostalgia-longing* were the most prevalent emotions, whereas *anxiety-fear*, *irritation-anger*, and *disgust-contempt* were the least.

The main trends obtained in this study, which featured a random sample of music, were thus similar to those found in previous research, which featured random samples of situations (Juslin et al., 2008) and listeners (Juslin et al., 2011). This suggests some degree of stability of the prevalence estimates across different methods. Indeed, some recent cross-cultural findings indicate that the broad patterns of prevalence of emotions may even be similar across cultures (Juslin, Barradas et al., 2016; see also Cowen et al., 2020).

However, we also observed some differences; for instance, the emotion *awe-admiration* was more frequent here than in previous studies, which sampled music episodes more broadly in everyday contexts. This could reflect the particular circumstances of the present study (e.g., attentive listening, no concurrent activity, no social interaction), which may have benefited an “aesthetic attitude” towards the music (Juslin, 2013).

Follow-up analyses indicated that emotion prevalence varied to some extent depending on the broad genre category (i.e., the STOMP factor); for instance, the R&C category evoked more *calm-contentment* and *awe-admiration* than the I&R category. Furthermore, there were wide individual differences in prevalence of emotion, which should be investigated further in future research. Exploratory analyses of links to the Big Five traits revealed few correlations that could explain individual differences in emotion prevalence, except perhaps that listeners scoring high in Conscientiousness showed lower prevalence for all emotions.

WHICH MECHANISMS OCCUR MOST OFTEN?

Differences in mechanism prevalence were smaller, overall, than differences in emotion prevalence (a “moderate” effect) – suggesting that all mechanisms contributed to some extent. The most prevalent mechanisms were Rhythmic entrainment, Aesthetic judgment, Evaluative conditioning, Visual imagery, and

Episodic memory. In particular, Rhythmic entrainment was significantly more prevalent than all other mechanisms in this study. This might partly reflect that popular music, as opposed to classical music, dominated the sample, and that the listeners did not choose the musical excerpts themselves. (Personal choice would perhaps have yielded a higher prevalence of episodic memories and evaluative conditioning; e.g., Juslin et al., 2008, 2011). As expected (H2), the two mechanisms Brain stem reflex and Musical expectancy were less frequent than the other mechanisms, except that Musical expectancy was not significantly different from Contagion.

Similarly to what was the case with the emotions, some mechanisms were more frequent here than in previous studies that sampled situations or listeners. For example, visual imagery was clearly more frequent here than in previous field studies (Juslin et al., 2008). Again, these differences may reflect the circumstances of the listening test (e.g., focused music listening in a dimly lit laboratory room with limited visual stimulation).

Mechanism prevalence also varied to some degree depending on the STOMP category (musical genre). For example, the R&C genre category produced lower ratings of Rhythmic entrainment and higher ratings of Musical expectancy, than the other genre categories. This could be because different genres provide different “affordances” in terms of the information needed for various mechanisms to be activated. In fact, it seems plausible that differences in prevalence of emotions between genres were largely mediated by differences in mechanism activation. Hence, it may not be a coincidence that, for example, the R&C category—which induced more *calm-contentment* than the I&R category—also displayed a higher prevalence of the Contagion mechanism, which was predictive of the same emotion in the (nomothetic) regression analyses (Figures 2 and 4).

We also obtained a few relatively weak links between the Big Five personality traits and particular mechanisms—for example between the trait Neuroticism and the mechanisms Brain stem reflex, Evaluative conditioning and Episodic memory; and between Extraversion and the mechanism Rhythmic entrainment.

CAN EMOTIONS BE PREDICTED FROM SELF-REPORTED MECHANISMS?

We investigated the links between emotions and mechanisms by conducting regression analyses that aimed to predict emotions based on self-reported mechanism impressions in the form of the *MecScale*. As expected (H3), results indicated that listener’s emotional responses could be predicted quite well based

on ratings of mechanisms. A mean multiple correlation of .847 meant that ca. 72% of the variance in the listeners' felt emotions could be accounted for by the mechanisms in the BRECVEMA framework. This is remarkable considering that none of the excerpts were selected to “trigger” specific mechanisms or evoke specific emotions: we simply observed what emotions and mechanisms would occur spontaneously during listening to a representative sample of music. Yet emotions were systematically related to mechanisms—and in ways consistent with the theoretical framework (Juslin, 2019, Part III).

Each emotion category was associated with a few key predictors, which were largely as expected: For example, *happiness-elation* was predicted by Rhythmic entrainment; *sadness-melancholy* was predicted by Evaluative conditioning and Contagion; *surprise-astonishment* was predicted by Brain stem reflex and Musical expectancy; *nostalgia-longing* was predicted by Episodic memory; and *awe-admiration* was predicted by Aesthetic judgment. Conversely, *interest-expectancy* was associated with a wide range of mechanisms.

Some of the links may seem “counterintuitive” but could simply reflect that different mechanisms do not occur in isolation in “real” pieces of music—as opposed to in systematic experiments where mechanisms have been isolated through digital editing (e.g., Juslin et al., 2014). Music often includes information relevant for several mechanisms. For example, the finding that both Episodic memory and Aesthetic judgment contributed to our prediction of *nostalgia-longing*, which is per definition strongly linked mostly to memory, could reflect a spurious correlation with respect to the stimuli—say, that those pieces that were rated highly aesthetically also happened to be more familiar and thus evoked more memories in listeners. Similarly, the result that both Aesthetic judgment and Contagion contributed to prediction of *awe-admiration* could simply reflect that the pieces that were highly emotionally expressive and caused contagion also were perceived as very beautiful, and thus activated the Aesthetic judgment mechanism. (Several authors have considered possible overlap in aesthetic impact between expressivity and beauty; see e.g., Bicknell, 2009; Clynes, 1977; Juslin, 2013).

ARE EMOTIONS BEST PREDICTED FROM MECHANISMS OR ACOUSTIC FEATURES?

We also attempted to predict felt emotions based on a wide range of acoustic features, covering tempo, dynamics, timbre, harmony, rhythm, and structure. Models based on these features could indeed predict emotions to some extent, but the mean variance

accounted for (39%) was quite modest. Hence, as expected (H4), the prediction of felt emotions was more accurate based on mechanisms (72%), than based on acoustic measures. Acoustically based models accounted for merely 54% of the variance in felt emotion that was accounted for by models based on mechanisms, on average. This test may be considered “conservative” since the acoustically based analyses included nearly twice as many predictors as the mechanism based analyses, which (all other things being equal) should favor the former type of model.

Moreover, previous studies have shown that the effects of acoustic features are mainly *additive* in nature (Juslin & Lindström, 2010), whereas it has been argued that the effects of mechanisms may be more interactive in nature (e.g., Juslin, 2019, Chapter 32). If so, the present comparison—which considered only additive effects—would arguably favor an acoustically based regression model.

It is noteworthy in this context that the emotion category that was best predicted by a mechanism-based model (*nostalgia-longing*) was *worst* predicted by an acoustically based model—perhaps because nostalgia is a prime example of where the underlying mechanism makes all the difference, as opposed to the acoustic patterns. For other emotion categories, this distinction may be more blurred.

Thus, for instance, the emotion *calm-contentment* was worst predicted by a mechanism-based model, though best predicted by an acoustically based model. This emotion is perhaps a case where purely acoustic patterns go a long way towards explaining induced emotions, as mediated by low-level mechanisms (though *calm-contentment* was also induced when music met listeners' criteria for aesthetic value; Aesthetic judgment). Another possible explanation is that *calm-contentment* was evoked by slow, low-arousal rhythmic entrainment, but that the scale item for this mechanism was formulated in such a way that it tended to emphasize only the high-arousal aspects of entrainment (see Method). A mechanism-based model could then have failed to “catch” calm instances of entrainment, thus leading to worse prediction for this particular emotion.

DO DIFFERENT LISTENERS SHOW THE SAME LINKS AMONG EMOTIONS AND MECHANISMS?

Idiographic regression analyses of a subset of randomly selected listeners revealed that felt emotions could be predicted well based on mechanism ratings also at an individual level. Most importantly, results indicated as expected (H5) that the emotion-mechanism links were different for different listeners; that is, partly different

mechanisms predicted their emotional reactions. We do not wish to suggest that the precise idiographic patterns can be generalized beyond the present cases; however, mere observation that the emotion-mechanism links vary in these cases means that we can already rule out the possibility that they are uniform across listeners in the general population.

Some of the findings from the nomothetic regression analyses may be re-interpreted in light of the idiographic data. For example, whereas a nomothetic model may seem to suggest that listeners' experiences of *happiness-elongation* were caused almost exclusively by Rhythmic entrainment, idiographic models showed that for 40% of the listeners, *happiness-elongation* was also—and in some cases even more strongly—linked to Episodic memory. Such idiosyncratic tendencies are “hidden” in nomothetic models, which may prevent a full understanding of the emotion-induction process. Having said that, however, the individual differences found here in the idiographic models were still smaller than those observed with regard to the weighting of criteria in aesthetic judgments (Juslin et al., 2021; Juslin, Sakka et al., 2016).

LIMITATIONS OF THE PRESENT STUDY

There are several limitations that should be taken into consideration, when interpreting the results. First, most of the data analyzed in this study were based on self-reports. Note that participants only report what they can or are willing to report, and that their responses can be affected by social desirability and demand characteristics (e.g., Visser et al., 2000). However, the validity of self-report depends on the type of questions asked. Clearly, listeners should be able to report their feelings more or less accurately. It can, of course, be argued that listeners' responses are constrained by the use of a list of emotion labels, but similar patterns of results to those reported here have been obtained in a previous study which measured the prevalence of emotions using an open-ended response format (Juslin et al., 2011, pp. 190-191).

As regards mechanisms, it can be difficult for listeners to monitor the precise processes. However, as explained above, the *MecScale* does not aim to measure mechanisms “directly.” Rather, it measures *subjective impressions*, which are reflective of mechanisms. Some of the mechanisms are straightforward to report because of their salience in conscious experiences; listeners will hardly be mistaken about a startle reflex, a conscious recollection of a memory, or a visual mental image.

More implicit mechanisms may at least be highly correlated with distinct impressions. Thus, the contagion item (Were you “touched” by the emotional expression of the music?) is based on the notion that the expression

of the music will be salient in cases where emotional contagion occurs. (In Gabrielsson's, 2011, study, “the emotional expression” was the single most frequently reported cause by the listeners.) One might perhaps be tempted to think that the relatively low prevalence of the expectancy mechanism is because this is a more implicit mechanism than, say, episodic memory, and that it is therefore less accurately reported. The workings of the expectancy mechanism may indeed be more subtle, and we are aware of the theoretical claim that the brain is constantly involved in prediction during listening to music (Vuust & Kringelbach, 2010). However, how often such computations actually *induce a felt emotion* is a different matter. The types of emotions one would theoretically expect from the expectancy mechanism (Huron, 2006; Meyer, 1956) are not very prevalent (Juslin, 2019).

The expectancy item in the *MecScale* is, essentially, an expectancy rating. We presume that a listener who becomes surprised via the musical expectancy mechanism will experience the music as “unpredictable.” (Studies in Cognitive Science, Empirical Aesthetics, and Music Cognition often use ratings of (un)expectness, and it is commonly assumed that those ratings are valid; Margulis, 2005.) In accordance with modern theories of expectancy in the emotion field, the expectancy item is based on the assumption that the strongest emotions are induced when a listener's expectations are disrupted, rather than when they are confirmed (see Miceli & Castelfranchi, 2015), and mainly when these violations are quite large. It appears plausible that such events are detected by a listener. The *MecScale* item simply taps into the subjective impression associated with such events, and seems to do it rather well, based on our previous experimental studies, which manipulated musical expectancies with highly controlled stimuli (e.g., Juslin et al., 2014).

The high predictive validity of the *MecScale* items in the present study, and the way that mechanisms and emotions showed links consistent with the BRECVEMA framework, clearly supports the notion that listeners are able to report conscious impressions reflecting emotional processes. Even so, self-report indices of causal mechanisms need to be interpreted with some caution. (This is all the more so considering that *MecScale* is a preliminary instrument for the measurement of mechanisms, which may be subject to revisions based on empirical findings.) A multi-item version of *MecScale* is currently being developed, but in this study, we relied on one item per mechanism, to minimize participant burden in the long listening test. Reliability in terms of *internal consistency* (Cronbach's alpha) is not a relevant criterion for single-item scales, and the notion of *test-re-test reliability* is problematic when a rated attribute does not remain

constant over time. Both emotions and mechanisms are in constant flux. For instance, Gabrielsson's (2011) study of strong experiences with music found plenty of cases where the emotional response to the same piece varied over time, sometimes quite drastically so (e.g., p. 72, p. 270), even just 15 minutes later (p. 340). In fact, results showed that the majority of the listeners had not had the same response when they encountered the same music again (p. 400).

This should not be surprising, since there are numerous variables in the listener and the situation that influence a response, and these may have a differential impact depending on the mechanism. Indeed, the BRECVEMA model predicts that the same piece of music can induce different emotions at different occasions; that two listeners may experience the same emotion though due to different mechanisms; and that two listeners may experience different emotions despite activation of the same mechanism (e.g., memories; Juslin, 2019, p. 382, p. 490; Sakka & Saarikallio, 2020). All of this conspires to make it difficult to estimate test-retest reliability for the *MecScale* items. However, since an unreliable measure cannot form a relationship that yields high predictive validity, a single-item scale can be regarded as sufficiently reliable if it shows high predictive validity (see Gorsuch & McFarland, 1972). High predictive validity for *MecScale* was observed in previous experimental studies featuring stimuli that were carefully manipulated and edited so as to isolate particular mechanisms (e.g., Juslin et al., 2014).

There is a popular notion in the psychometric literature that multiple-item measures are more valid than single-item measures, for *all* types of constructs (e.g., Nunnally & Bernstein, 1994), but this assumption has not been borne out by studies. The advantage of multiple-item measures applies mainly to abstract constructs that are multi-constituent—such as personality traits which feature various “facets.” If a rated attribute can be conceptualized as concrete and singular (such that it can be easily imagined), there is no real difference in predictive validity between single-item and multiple-item measures (e.g., Bergkvist & Rossiter, 2007). Although further research is clearly required on the measurement of various mechanisms, it is plausible that the mechanisms may to all intents and purposes be conceived of as concrete and singular (e.g., “did the music evoke a memory?”).

One obvious limitation of this study, however, is the modest sample of listeners, which calls for replication. However, the present study, which sampled 72 pieces of *music* randomly, should be considered in the context of previous studies involved in our “method triangulation”

strategy which sampled 762 *listeners* (Juslin et al., 2011) or 573 listening *situations* randomly (Juslin et al., 2008). Although it would have been desirable to sample all three factors—music, listeners, situations—randomly in the same study, this was not practically feasible.

Still, the moderate sample (which was due to practical difficulties in recruiting a larger number of listeners) may perhaps go some way towards explaining the slightly disappointing results regarding the Big Five personality traits. However, links to personality have tended to be rather weak and mostly inconsistent also in previous studies, featuring considerably larger listener samples (see e.g., Juslin, et al., 2011; Juslin, Barradas et al., 2016). Indeed, another possible explanation of this absence of strong links between personality traits and the prevalence of emotions may be that personality traits primarily affect how listeners *use* music in everyday life (cf. Chamorro-Premuzic & Furnham, 2007; Vella & Mills, 2017). Such differences in use, which will shape which music listening events individuals seek out and get exposed to in daily life, will clearly have an effect on emotion prevalence in “real life.” By contrast, when such context effects are removed (like in this study where all listeners heard the same music in the same context), the effect of personality on emotion prevalence may not be apparent in the same way.

This line of reasoning highlights a final (important) limitation of this study: The lack of context. This is, of course, the inevitable price to be paid for achieving complete control over the musical stimuli, such that listeners heard the same music and could be directly compared. However, leaving out the context can give the impression that the task of predicting reactions to music is more straightforward than it really is. Several aspects of the context could help to determine which mechanisms are actually activated by influencing factors such as the music choice, the listener's attention and the *functions* of the music during a specific activity (North & Hargreaves, 2008). Familiarity with the music, which was unfortunately not indexed in this study, may enhance a listener's liking (North & Hargreaves, 1995), and also enables a greater number of memory-based mechanisms (e.g., evaluative conditioning, episodic memory) to be activated. In addition, extrinsic factors, such as the status of the performer and program notes, may also have effects on certain mechanisms (Kroger & Margulis, 2017). Thus, prediction of musical emotions could be more complex if contextual variables are taken into consideration. It is perhaps promising in this regard that contextual variables *themselves* appear to be able to predict felt emotions during music listening to some extent (Juslin et al., 2011, pp. 193-195).

IMPLICATIONS FOR FUTURE RESEARCH

Whatever the limitations of the present study, we argue that the results have three major implications for future research and application. First of all, they indicate that the mechanism Rhythmic entrainment plays an important role in emotional reactions to music—at least as far as popular music is concerned (see Juslin, Barradas et al., 2016, for similar findings in a cross-cultural sample). Although this mechanism has received some serious attention (Clayton, 2012; Trost et al., 2017), experimental evidence proving that rhythmic entrainment during music listening is able to induce emotions is still limited (Juslin, 2019) and mixed (cf. Bason & Celler, 1972; Mütze et al., 2020). Thus, more research is clearly needed to elucidate this seemingly crucial mechanism.

Second, the present findings provide further evidence that the mechanisms constitute the nexus where *control* over emotional responses is possible, and where music interventions need to be focused for maximum efficacy. Recent advances in psychological theories of how music achieves its emotional effects on listeners have not yet been translated into the clinical sphere, although researchers in the music and health domain increasingly call for mechanistic studies—as opposed to mere efficacy trials—because mechanistic studies may play a key role in optimizing interventions (Bradt, 2018; Juslin, 2011). Depending on the emotion one seeks to induce in a client, specific mechanisms may need to be selected and manipulated. We hope that the present study may set the stage for additional, more large-scale work examining how specific mechanisms might differentially influence listeners with particular personal histories.

Finally, our findings suggest that it is at the mechanistic level that individual differences will tend to occur. Though there are plenty of individual, demographic, and contextual factors that may affect emotions, the moderating effects of those factors are most likely “mediated” by mechanisms. (For instance, recent research suggests that the BRECVEMA mechanisms might play a role in explaining the effects of personality

traits on emotional reactions; see Larwood & Dingle, 2021). Individual differences are increasingly recognized as a genuine challenge in the emotion field. Davidson (2012) refers to them as “the most salient characteristic” of emotions (p. 7). It may be argued that much of the apparent diversity in both musical interventions and psychological experiments aimed at emotional reactions to music is not due to the inadequacy of the emotion theories applied (if any), but rather due to the fact that these need to be applied on an individual level. Understanding how individual variability arises within the confines of general mechanisms may clearly benefit future applications. The present results suggest that a full understanding of how music evokes emotions needs to involve idiographic analyses. This reflects that although musical activities are social, through and through (Juslin, 2021), our *emotional experiences* of music are ultimately and irrevocably individual in nature.

Author Note

Patrik N. Juslin, Department of Psychology, Uppsala University; Laura S. Sakka, Department of Psychology, Uppsala University; Gonçalo T. Barradas, Madeira Interactive Technologies Institute (M-ITI), University of Madeira; Olivier Lartillot, RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo.

This research was supported by a grant from the Swedish Research Council (Project No. 2004-2045), as well as by grants from the Research Council of Norway through its Centers of Excellence scheme (Project No. 262762) and the MIRAGE project (Grant No. 287152).

Correspondence concerning this article should be addressed to Dr. Patrik N. Juslin, Department of Psychology, Uppsala University, Box 1225, SE - 751 42 Uppsala, SWEDEN. E-mail: patrik.juslin@psyk.uu.se

References

- ARJMAND, H. A., HOHAGEN, J., PATON, B., & RICKARD, N. S. (2017). Emotional responses to music: Shifts in frontal brain asymmetry mark periods of musical change. *Frontiers in Psychology*, 8, 2044.
- BARRADAS, G. T., JUSLIN, P. N., & BADIA, S. B. (2021). Emotional reactions to music in dementia patients and healthy controls: Differential responding depends on the mechanism. *Music and Science*. <https://doi.org/10.1177/20592043211010152>
- BARRETT, F. S., GRIMM, K. J., ROBINS, R. W., WILDSCHUT, T., SEDIKIDES, C., & JANATA, P. (2010). Music-evoked nostalgia: Affect, memory, and personality. *Emotion*, 10, 390–403.
- BARRETT, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, 87, 266–281.
- BARRETT, L. F. (2017). *How emotions are made*. Houghton-Mifflin Harcourt.

- BARTHET, M., FAZEKAS, G., & SANDLER, M. (2013). Music emotion recognition: From content- to context-based models. In M. Aramaki, M. Barthelet, R. Kronland-Martinet, & S. Ystad (Eds.), *From sounds to music and emotions. Lecture Notes in Computer Science* (Vol. 2900, pp. 228–252). Springer.
- BASON, P. T., & CELLER, B. G. (1972). Control of the heart rate by external stimuli. *Nature*, 273, 279–280.
- BAUMGARTNER, H. (1992). Remembrance of things past: Music, autobiographical memory, and emotion. *Advances in Consumer Research*, 19, 613–620.
- BERGKVIST, L., & ROSSITER, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, XLIV, 175–184.
- BERLYNE, D. E. (1971). *Aesthetics and psychobiology*. Appleton Century Crofts.
- BICKNELL, J. (2009). *Why music moves us*. Palgrave-Macmillan.
- BLAIR, M. E., & SHIMP, T. A. (1992). Consequences of an unpleasant experience with music: A second-order negative conditioning perspective. *Journal of Advertising*, 21, 35–43.
- BLANCA, M. J., ALARCÓN, R., ARNAU, J., BONO, R., & BENDAYAN, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29, 552–557.
- BOLDERS, A. C., BAND, G. P. H., & STALLEN, P. J. (2012). Evaluative conditioning induces changes in sound valence. *Frontiers in Psychology: Emotion Science*, 3, 106.
- BRADT, J. (2018). *Music therapy in health care: Challenges and opportunities*. Paper presented at the Nordic Music Therapy Congress, Stockholm, Sweden.
- BRUNSWIK, E. (1956). *Perception and the representative design of psychological experiments*. University of California Press.
- BYRNE, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.
- CHAMORRO-PREMUZIC, T., & FURNHAM, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology*, 98, 175–185.
- CLAYTON, M. (2012). What is entrainment? Definition and applications in musical research. *Empirical Musicology Review*, 7, 49–56.
- CLYNES, M. (1977). *Sentics: The touch of emotions*. Doubleday.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic Press.
- COHEN, J., & COHEN, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- COOKSEY, R. W. (1996). *Judgment analysis*. Academic Press.
- COUTINHO, E., & CANGELOSI, A. (2011). Musical emotions: Predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion*, 11, 921–937.
- COWEN, A. S., FANG, X., SAUTER, D., & KELTNER, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 1924–1934.
- DALY, I., WILLIAMS, D., HALLOWELL, J., HWANG, F., KIRKE, A., MALIK, A., et al. (2015). Music-induced emotions can be predicted from a combination of brain activity and acoustic features. *Brain and Cognition*, 101, 1–11.
- DAVIDSON, R. J. (2012). *The emotional life of your brain*. Penguin Books.
- DAVIS, M. (1984). The mammalian startle response. In R. C. Eaton (Ed.), *Neural mechanisms of startle behavior* (pp. 287–342). Plenum Press.
- DELSING, M. J. M. H., TER BOGT, T. F. M., ENGELS, R. C. M. E., & MEEUS, W. H. J. (2008). Adolescents' music preferences and personality characteristics. *European Journal of Personality*, 22, 109–130.
- DINGLE, G. A., SAVILL, S., FRASER, C., & VIETH, H. (2011). *Music that moves us: In search of the mechanisms linking music listening with emotional response*. Poster presented at the meeting of the International Society for Research on Emotions, Kyoto, Japan.
- DOWLING, W. J., & HARWOOD, D. L. (1986). *Music cognition*. Academic Press.
- EGERMANN, H., & McADAMS, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception*, 31, 139–156.
- EEROLA, T. (2011). Are the emotions expressed in music genre-specific? An audio-based evaluation of datasets spanning classical, film, pop and mixed genres. *Journal of New Music Research*, 40, 349–366.
- EEROLA, T., & VUOSKOSKI, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, 30, 307–340.
- EKMAN, P. (1994). All emotions are basic. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 15–19). Oxford University Press.
- ELLSWORTH, P. C. (1994). Levels of thought and levels of emotion. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 192–196). Oxford University Press.
- FERGUSON, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40, 532–538.
- FERRERI, L., MOUSSARD, A., BIGAND, E., & TILLMANN, B. (2019). Music and the aging brain. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 623–644). Oxford University Press.
- FLORES, R. G., & GINSBURGH, V. A. (1996). The Queen Elisabeth musical competition: How fair is the final ranking? *The Statistician*, 45, 97–104.

- FOSSATI, A., BORRONI, S., MARCHIONE, D., & MAFFEI, C. (2011). The Big Five Inventory (BFI). *European Journal of Psychological Assessment*, 27, 50–58.
- FREDRICKSON, B. L. (1998). What good are positive emotions? *Review of General Psychology*, 2, 300–319.
- FRIJDA, N. H., MESQUITA, B., SONNEMANS, J., & VAN GOOZEN, S. (1991). The duration of affective phenomena, or emotions, sentiments and passions. In K. Strongman (Ed.), *International review of emotion and motivation* (pp. 187–225). Wiley.
- GABRIELSSON, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special Issue 2001–2002*, 123–147.
- GABRIELSSON, A. (2011). *Strong experiences with music: Music is much more than just music*. Oxford University Press.
- GARRIDO, S., & DAVIDSON, J. W. (2019). *Music, nostalgia, and memory: Historical and psychological perspectives*. Palgrave MacMillan.
- GARRIDO, S., DUNNE, L., CHANG, E., PERZ, J., STEVENS, C. J., & HAERTSCH, M. (2017). The use of music playlists for people with dementia: A critical synthesis. *Journal of Alzheimer's Disease*, 60, 1129–1142.
- GOMEZ, P., & DANUSER, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, 7, 377–387.
- GORSUCH, R. L., & MCFARLAND, S. G. (1972). Single versus multiple-item scales for measuring religious values. *Journal for the Scientific Study of Religion*, 11, 53–64.
- GOUVEIA, V., PIMENTEL, C., SANTANA, N., CHAVES, W., & PARAÍBA, C. (2008). [Short test of music preference (STOMP): Evidences of its factor validity and reliability]. *Psico*, 39, 201–210.
- HADIT, J., & SEDER, P. (2009). Admiration and awe. In D. Sander & K. R. Scherer (Eds.), *The Oxford companion to emotion and the affective sciences* (pp. 4–5). Oxford University Press.
- HAIR, J., BLACK, W. C., BABIN, B. J., & ANDERSON, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- HARGREAVES, D. J., & NORTH, A. C. (2010). Experimental aesthetics and liking for music. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 515–546). Oxford University Press.
- HARRER, G., & HARRER, H. (1977). Music, emotion, and autonomic function. In M. Critchley & R. A. Henson (Eds.), *Music and the brain. Studies in the neurology of music* (pp. 202–216). William Heinemann Medical Books.
- HURON, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.
- IZARD, C. E. (1977). *The emotions*. Plenum Press.
- JANATA, P., TOMIC, S. T., & RAKOWSKI, S. K. (2007). Characterization of music-evoked autobiographical memories. *Memory*, 15, 845–860.
- JOHN, O. P., DONAHUE, E. M., & KENTLE, R. L. (1991). *The Big Five Inventory - Versions 4a and 54*. University of California, Berkeley.
- JOHN, O. P., NAUMANN, L. P., & SOTO, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). Guilford Press.
- JOHNSON-LAIRD, P. N. (1992). Introduction: What is communication? In D. H. Mellor (Ed.), *Ways of communicating* (pp. 1–13). Cambridge University Press.
- JUSLIN, P. N. (1997). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae*, 1, 225–256.
- JUSLIN, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309–337). Oxford University Press.
- JUSLIN, P. N. (2011). Music and emotion: Seven questions, seven answers. In I. Deliège & J. Davidson (Eds.), *Music and the mind: Essays in honour of John Sloboda* (pp. 113–135). Oxford University Press.
- JUSLIN, P. N. (2013). From everyday emotions to aesthetic emotions: Toward a unified theory of musical emotions. *Physics of Life Reviews*, 10, 235–266.
- JUSLIN, P. N. (2019). *Musical emotions explained: Unlocking the secrets of musical affect*. Oxford University Press.
- JUSLIN, P. N. (2021). Mind the gap: The mediating role of emotion mechanisms in social bonding through musical activities. *Behavioral and Brain Sciences*, 44, e80.
- JUSLIN, P. N., BARRADAS, G., & EEROLA, T. (2015). From sound to significance: Exploring the mechanisms underlying emotional reactions to music. *American Journal of Psychology*, 128, 281–304.
- JUSLIN, P. N., BARRADAS, G. T., OVSIANNIKOW, M., LIMMO, J., & THOMPSON, W. F. (2016). Prevalence of emotions, mechanisms, and motives in music listening: A comparison of individualist and collectivist cultures. *Psychomusicology*, 26, 293–326.
- JUSLIN, P. N., HARMAT, L., & EEROLA, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, 42, 599–623.
- JUSLIN, P. N., INGMAR, E., & DANIELSSON, J. (2021). Aesthetic judgments of music: Reliability, consistency, criteria, self insight, and expertise. *Psychology of Aesthetics, Creativity, and the Arts*. Advance on-line publication: <https://doi.org/10.1037/aca0000403>
- JUSLIN, P. N., & LAUKKA, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, 33, 217–238.

- JUSLIN, P. N., LILJESTRÖM, S., LAUKKA, P., VÄSTFJÄLL, D., & LUNDQVIST, L.-O. (2011). Emotional reactions to music in a nationally representative sample of Swedish adults: Prevalence and causal influences. *Musicae Scientiae*, 15, 174–207.
- JUSLIN, P. N., LILJESTRÖM, S., VÄSTFJÄLL, D., BARRADAS, G., & SILVA, A. (2008). An experience sampling study of emotional reactions to music: listener, music, and situation. *Emotion*, 8, 668–683.
- JUSLIN, P. N., LILJESTRÖM, S., VÄSTFJÄLL, D., & LUNDQVIST, L.-O. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford University Press.
- JUSLIN, P. N., & LINDSTRÖM, E. (2010). Musical expression of emotions: Modeling listeners' judgments of composed and performed features. *Music Analysis*, 29, 334–364.
- JUSLIN, P. N., & SAKKA, L. S. (2019). Neural correlates of music and emotion. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 285–332). Oxford University Press.
- JUSLIN, P. N., SAKKA, L. S., BARRADAS, G. T., & LILJESTRÖM, S. (2016). No accounting for taste? Idiographic models of aesthetic judgment in music. *Psychology of Aesthetics, Creativity, and the Arts*, 10, 157–170.
- JUSLIN, P. N., & SLOBODA, J. A. (Eds.). (2010). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.
- JUSLIN, P. N., & VÄSTFJÄLL, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31, 559–621.
- KIVY, P. (1993). Auditor's emotions: Contention, concession and compromise. *Journal of Aesthetics and Art Criticism*, 51, 1–12.
- KIVY, P. (2002). *Introduction to a philosophy of music*. Clarendon Press.
- KROGER, C., & MARGULIS, E. H. (2017). "But they told me it was professional": Extrinsic factors in the evaluation of musical performance. *Psychology of Music*, 45, 49–64.
- KUBZANSKY, L. D. (2009). Health and emotion. In D. Sander & K. R. Scherer (Eds.), *The Oxford companion to emotion and the affective sciences* (pp. 204–205). Oxford University Press.
- LADINIG, O., & SCHELLENBERG, E. G. (2012). Liking unfamiliar music: Effects of felt emotion and individual differences. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 146–154.
- LANGMEYER, A., GUGLHÖR-RUDAN, A., & TAMAI, C. (2012). What do music preferences reveal about personality? A cross-cultural replication using self-ratings and ratings of music sample. *Journal of Individual Differences*, 33, 119–130.
- LARWOOD, J. L., & DINGLE, G. A. (2021). The effects of emotionally congruent sad music listening in young adults high in rumination. *Psychology of Music*. Advance online publication. <https://doi.org/10.1177/0305735620988793>
- LAZARUS, R. S. (1991). *Emotion and adaptation*. Oxford University Press.
- LEE, I. A., & PREACHER, K. J. (2013, September). *Calculation for the test of the difference between two dependent correlations with one variable in common* [Computer software]. Available from <http://quantpsy.org>
- LEVINSON, J. (1997). Emotion in response to art. In M. Hjort & S. Laver (Eds.), *Emotion and the arts* (pp. 20–34). Oxford University Press.
- LILJESTRÖM, S., JUSLIN, P. N., & VÄSTFJÄLL, D. (2013). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychology of Music*, 41, 577–597.
- MACDONALD, R., KREUTZ, G., & MITCHELL, L. (Eds.). (2012). *Music, health, and well-being*. Oxford University Press.
- MARGULIS, E. H. (2005). A model of melodic expectation. *Music Perception*, 22, 663–714.
- MAS-HERRERO, E., MARCO-PALLARES, J., LORENZO-SEVA, U., ZATORRE, R. J., & RODRIGUEZ-FORNELLS, A. (2013). Individual differences in music reward experiences. *Music Perception*, 31, 118–138.
- MCCRAE, R. R. (2007). Aesthetic chills as a universal marker of openness to experience. *Motivation and Emotion*, 31, 5–11.
- MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago University Press.
- MEYER, L. B. (2001). Music and emotion: Distinctions and uncertainties. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 341–360). Oxford University Press.
- MICELI, M., & CASTELFRANCHI, C. (2015). *Expectancy and emotion*. Oxford University Press.
- MÜTZE, H., KOPIEZ, R., & WOLF, A. (2020). The effect of a rhythmic pulse on the heart rate: Little evidence for rhythmic 'entrainment' and 'synchronization'. *Musicae Scientiae*, 24, 377–400.
- NORTH, A. C., & HARGREAVES, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, 14, 77–93.
- NORTH, A. C., & HARGREAVES, D. J. (2008). *The social and applied psychology of music*. Oxford University Press.
- NUNNALLY, J. C., & BERNSTEIN, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- OSBORNE, J. W. (1980). The mapping of thoughts, emotions, sensations, and images as responses to music. *Journal of Mental Imagery*, 5, 133–136.
- PILGRIM, L., NORRIS, J., & HACKATHORN, J. (2017). Music is awesome: Influences of emotion, personality, and preference on experienced awe. *Journal of Consumer Behaviour*, 16, 442–451.
- PLUTCHIK, R. (1994). *The psychology and biology of emotion*. Harper-Collins.

- RENTFROW, P. J., & GOSLING, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84, 1236–1256.
- RUSSELL, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.
- SAKKA, L. S., & JUSLIN, P. N. (2018). Emotional reactions to music in depressed individuals. *Psychology of Music*, 46, 862–880.
- SAKKA, L. S., & SAARIKALLIO, S. (2020). Spontaneous music-evoked autobiographical memories in individuals experiencing depression. *Music and Science*, 3, 1–15. DOI: 10.1177/2059204320960575
- SCHERER, K. R. (1999). Appraisal theories. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). Wiley.
- SCHERER, K. R. (2000). Psychological models of emotion. In J. Borod (Ed.), *The neuropsychology of emotion* (pp. 137–162). Oxford University Press.
- SCHERER, K. R., & ZENTNER, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361–392). Oxford University Press.
- SCHERER, K. R., ZENTNER, M. R., & SCHACHT, A. (2002). Emotional states generated by music: An exploratory study of music experts. *Musicae Scientiae, Special Issue 2001-2002*, 149–171.
- SCHUBERT E. (2013). Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives. *Frontiers in Psychology*, 4, 837.
- SIMONS, R. C. (1996). *Boo! Culture, experience, and the startle reflex*. Oxford University Press.
- SLOBODA, J. A. (1992). Empirical studies of emotional response to music. In M. Riess-Jones & S. Holleran (Eds.), *Cognitive bases of musical communication* (pp. 33–46). American Psychological Association.
- SLOBODA, J. A. (1996). Emotional responses to music: A review. In K. Riederer & T. Lahti (Eds.), *Proceedings of the Nordic acoustical meeting* (pp. 385–392). The Acoustical Society of Finland.
- SLOBODA, J. A., IVALDI, A., & O'NEILL, S. A. (2001). Functions of music in everyday life: An exploratory study using the experience sampling methodology. *Musicae Scientiae*, 5, 9–32.
- SLOBODA, J. A., & JUSLIN, P. N. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 71–104). Oxford University Press.
- STEIGER, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- STEINBEIS, N., KOELSCH, S., & SLOBODA, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18, 1380–1393.
- TABACHNICK, B. G., & FIDELL, L. S. (2001). *Using multivariate statistics* (4th ed.). Allyn and Bacon.
- TARUFFI L., & KOELSCH S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS ONE*, 9, e110490.
- TARUFFI, L., & KÜSSNER, M. B. (2019). A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology*, 29, 62–74.
- TROST, W., LABBÉ, C., & GRANDJEAN, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia*, 96, 96–110.
- VELLA, E. J., & MILLS, G. (2017). Personality, uses of music, and music preference: The influence of openness to experience and extraversion. *Psychology of Music*, 45, 338–354.
- VISSER, P. S., KROSCHICK, J. A., & LAVRAKAS, P. J. (2000). Survey research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 223–252). Cambridge University Press.
- VUUST, P., & KRINGELBACH, M. L. (2010). The pleasure of music. In M. L. Kringelbach & K. C. Berridge (Eds.), *Pleasures of the brain* (pp. 255–269). Oxford University Press.
- WARRENBURG, L. A. (2020). Comparing musical and psychological emotion theories. *Psychomusicology*, 30, 1–19.
- WELLS, A., & HAKANEN, E. A. (1991). The emotional uses of popular music by adolescents. *Journalism Quarterly*, 68, 445–454.
- WEST, S. G., FINCH, J. F., & CURRAN, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Sage.
- WILDSCHUT, T., SEDIKIDES, C., ARNDT, J., & ROUTLEDGE, C. (2006). Nostalgia: Content, triggers, functions. *Journal of Personality and Social Psychology*, 91, 975–993.
- YANG, Y.-H., & CHEN, H. H. (2011). *Music emotion recognition*. CRC Press.
- YANG, X., DONG, Y., & LI, J. (2018). Review of data features-based music emotion recognition methods. *Multimedia Systems*, 24, 365–389.
- ZAKRISSON, I. (2010). *Big Five Inventory (BFI): Utvärdering för svenska förhållanden. Samhällsvetenskapliga rapporter från Mittuniversitetet, 2010:3*. Mid Sweden University.
- ZENTNER, M. R., & EEROLA, T. (2010). Self-report measures and models. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 187–221). Oxford University Press.
- ZENTNER, M. R., GRANDJEAN, D., & SCHERER, K. R. (2008). Emotions evoked by the sound of music: Differentiation, classification, and measurement. *Emotion*, 8, 494–521.
- ZWEIGENHAFT, R. (2008). A do re mi encore: A closer look at the personality correlates of music preferences. *Journal of Individual Differences*, 29, 45–55.

APPENDIX A Stratified Random Sampling of Musical Stimuli

Factor	Genre	Artist	Piece	Order	
	--- Group 1 ---				
R&C	Classical/Opera	Bach (Yo Yo Ma)	Cello suite no. 4, courante	40	
		Stanley Myers (John Williams)	Cavatina	3	
	Jazz/Blues	Michael Nyman	The heart asks pleasure	28	
		Muddy Waters	I'm a king bee	20	
		Michael Bublé	Feeling good ^e	35	
		Stevie Ray Vaughn	Telephone song*	29	
	Folk/Singer-songwriter	John Denver	Leaving on a jet plane	5	
		Bob Dylan	The times they are a changing*	1	
		Paolo Nutini	Last request	17	
		Joni Mitchell	A case of you	11	
		The White Stripes	Seven nation army*	39	
I&R	Classic rock	Jimi Hendrix	Purple haze	26	
		Rolling Stones	Start me up	7	
		Kings of Leon	Use somebody ^e	6	
	Heavy metal/Hard rock	Alice in Chains	Would?	34	
		Sepultura	Roots bloody roots*	36	
		Iron Maiden	2 minutes to midnight ^e	30	
		Green Day	21 guns ^e	38	
	Alternative/Punk	Muse	Undisclosed desires	16	
		Nirvana	Come as you are	22	
	U&C	Country	Robert Plant & Alison Kraus	Gone gone gone (done moved on)	23
			Harry Nilsson	Everybody's talking	24
Willie Nelson			Always on my mind	33	
Pop		Jessie J	Domino	14	
		Coldplay	Paradise ^e	37	
Schlager/Dansband		Bruno Mars	The lazy song	4	
		Abba	Waterloo*	31	
		Anita Lindblom	Sânt är livet	8	
		Loreen	Euphoria	2	
E&R		Hip-hop/Reggae	Carola	Fångad av en stormvind*	10
			Eminem	Superman	32
	Talib Kweli		Get by	25	
	Soul/Funk	The Heptones	Cool rasta	13	
		Kapten Röd	Ju mer dom spottar*	15	
		Jordan Sparks	No air (featuring Chris Brown)	21	
		Whitney Houston	Will always love you	19	
	Club/House	Earth, Wind & Fire	Let's groove	18	
		Britney Spears	Till the world ends*	12	
		Flo Rida	Wild ones (featuring SIA)	27	
		Chris Brown	Beautiful people	9	
	--- Group 2 ---				
R&C	Classical/Opera	Rachmaninov (Garrick Ohlsson)	Piano concerto no. 3, finale	1	
		Debussy (Carolyn Jones)	Deux arabesque no. 1, E major	15	
		Wright/Forrest/ Borodin (Sarah Brightman)	Stranger in paradise ^e	7	
	Jazz/Blues	John Coltrane	Giant steps	40	
		Weather Report	Birdland	38	
		Wayne Shorter	Footprints	30	
		Stevie Ray Vaughn	Telephone song*	28	
	Folk/Singer-songwriter	Marissa Nadler	Your heart is a twisted vine	21	
		Bob Dylan	The times they are a changing*	14	
		Princess Music	White wave	13	

(continued)

APPENDIX A (continued)

Factor	Genre	Artist	Piece	Order
I&R	Classic rock	The White Stripes	Seven nation army*	31
		U2	Beautiful day	6
		Blur	Song 2	8
	Heavy metal/Hard rock	Thin Lizzy	Boys are back in town	12
		Sepultura	Roots bloody roots*	2
		Metallica	Enter sandman ^e	20
		Guns n' Roses	Paradise city	25
		Cheyenne Mize	It lingers	22
	Alternative/Punk	Alpine	Gasoline	39
		Ramones	Now I want to sniff some glue	23
Kenny Rogers		The gambler ^e	17	
U&C	Country	The Steeldrivers	If it hadn't been for love	16
		First Aid Kit	Emmylou	37
		Beyonce	Sweet dreams	32
	Pop	Owl City	Fireflies	9
		Pink	Please don't leave me	26
		Abba	Waterloo*	4
		Sten & Stanley	Jag vill vara din, Margareta	11
	Schlager/Dansband	Tommy Körberg	Stad i ljus	5
		Carola	Fångad av en stormvind*	24
		Jay Z & Alicia Keys	Empire state of mind	10
E&R	Hip-hop/Reggae	Kapten Röd	Ju mer dom spottar*	34
		Snoop Dogg	Drop it like it's hot	35
		Bobby Womack	Across the 110th street ^e	18
	Soul/Funk	Donell Jones	U know what's up	27
		Ray Charles	Hallelujah I love her so	36
		Avicii	Levels (radio edit) ^e	29
	Club/House	Britney Spears	Til the world ends*	19
		Madeon	Icarus	3
		Nicki Minaj	Starships	33

Note: R&C = reflective and complex, I&R = intense and rebellious, U&C = upbeat and conventional, E&R = energetic and rhythmic. *Songs that occur in both groups 1 and 2. (Stimulus order in each group is included to facilitate replication). All excerpts included the first 60s of the recording (except those marked^e, which had an intro edited out).

APPENDIX B Intercorrelations Among Emotion Ratings

Emotion	1	2	3	4	5	6	7	8	9	10	11
1. Happiness	–										
2. Sadness	–.02	–									
3. Surprise	.27	.15	–								
4. Calm	.39	.21	.17	–							
5. Interest	.55	.13	.45	.41	–						
6. Nostalgia	.51	.32	.20	.39	.39	–					
7. Anxiety	–.16	.31	.23	–.17	–.04	–.01	–				
8. Pride	.59	–.04	.24	.34	.46	.45	–.09	–			
9. Anger	–.34	.01	.03	–.31	–.31	–.25	.47	–.18	–		
10. Love	.44	.38	.21	.53	.39	.56	–.06	.40	–.27	–	
11. Disgust	–.29	.06	.02	–.25	–.29	–.19	.41	–.14	.76	–.19	–
12. Awe	.46	.27	.32	.51	.54	.53	–.03	.49	–.24	.60	–.20

Note: Values show Pearson correlations (r).

Appendix C: Description of Acoustic Features

Each musical excerpt was described by 14 numerical values, where each value corresponds to a specific audio or musical characteristic, or *feature*. Below, we provide a general description of each feature, followed by technical details. All features can be extracted using MIRtoolbox (a Matlab toolbox freely available).

TEMPO AND RHYTHM

Metroid

This feature is related to tempo, which is expressed in beats per minute (bpm). The higher the tempo, the faster the underlying beat. Because it may vary over time, tempo is estimated on successive, short excerpts, thus producing a tempo curve. The resulting features are the mean (Metroid M) and standard deviation (Metroid SD) of the overall tempo.

Musical meter is hierarchical: if there is a tempo related to eighth notes at 90 bpm in a 4/4 meter, 45 bpm could be another acceptable tempo, related to quarter notes, as well as 180 bpm, related to sixteenth notes, if these metrical levels are prevalent in the music. A score can be assigned to each tempo candidate based on its perceptual salience. One common strategy consists of choosing the candidate with the highest score. (This is the default strategy for *mirtempo* in MIRtoolbox.) However, ignored metrical levels with lower scores may play an important role in the appreciation of the general speed of the music. The proposed solution, called metrical centroid, or *Metroid*, is to take the centroid of all the tempo candidates (Lartillot & Grandjean, 2019); an increase of the Metroid value over time indicates a general increase of the overall tempo and/or an increase of the prevalence of faster metrical levels, such as sixteenth notes. Although these relate to two different musical aspects, they both contribute to a general appreciation of increase of speed in the music. The estimation of the tempo follows the method presented in Lartillot and Grandjean (2019) with some improvements described in Lartillot (in preparation).

Metrical strength

From the metrical analysis, one can also derive a *Metrical strength* measure, obtained by summing the scores related to the metrical levels defining the metrical structure (Lartillot, in preparation). A high level of metrical strength indicates that there is a clear pulsation on multiple metrical levels. For instance, for a 4/4 meter, a clear emphasis of the quarter notes, eighth notes, half notes, first beat of the bar, and sixteenth notes will yield a high metrical strength. The measure is calculated on each

successive excerpt. The resulting features are the mean (Metrical strength M) and standard deviation (Metrical strength SD).

DYNAMICS

Crescendo

Based on a dynamics curve, the *Crescendo* feature indicates the progressive increase of dynamics over time. Crescendo sequences are detected, and a general Crescendo value is given, based on the relative length of the overall crescendo region and the ambitus in dynamics (Lartillot, in preparation). A higher value indicates a longer and/or stronger overall crescendo sequence.

Estimating perceived dynamics from audio recordings is more challenging than it might seem, because it depends on several factors, such as the way music has been recorded, mixed, and mastered; how it is played back to listeners; timbre characteristics of acoustic instruments and voices possibly indicating performance dynamics, etc. Thus, instead of assessing a global dynamics level, we measured the temporal profile of the dynamics along each music excerpt. We developed a new method which met the requirements that a sudden increase of dynamics should be properly represented, and that micro-silences (shorter than one second) should not have an impact. We also attempted to model saturation effects seemingly taking place within separate frequency registers (Lartillot, in preparation).

The dynamics curve is computed as follows. First, the energy in log scale of the audio signal is decomposed along Mel frequencies, leading to a Mel spectrogram with fairly high temporal resolution (20 ms frames and 10% hop factor). Frequencies from 0 to 5000 Hz are selected, corresponding to 35 Mel bands. An envelope follower, applied to each Mel band separately, immediately reacts to high rises, while decreasing slowly. These envelopes are then summed along the Mel bands.

To detect the crescendo sequence, a second dynamics curve is computed, increasing as quickly as the dynamics curve, but decreasing even more slowly, representing an adaptation to the signal. Only the sections where this curve rises are treated as actual amplitude-increase phases; the remaining ones are considered simply as local modulations not contributing to the perception of a general increase of dynamics. These phases are combined to form crescendo episodes. Phases with low slopes are ignored, and an episode is terminated when the temporal distance between the end of its final (non-ignored) phase and the next (non-ignored) phase exceeds a given threshold. Crescendo episodes that are

too small or with an overall slope that is too low are filtered out. For each crescendo episode, a value is obtained by multiplying the duration with the difference of amplitude in dB (capped to 20 dB in maximum). These values are summed together and divided by the length of the overall excerpt and divided by 20 dB.

TIMBRE

Bass, Midrange, Treble

Tonal balance (i.e., distribution of sound energy across the spectrum) plays a crucial role in the appreciation of timbre. Similarly to how the spectrum is usually decomposed in audio engineering, we define the *Bass*, *Midrange* and *Treble* features as the amount of energy in the corresponding regions. To improve these measures, we also compute as a reference the average tonal balance of the whole music corpus. By expressing the energy for each region relative to the mean energy across the corpus, these normalized measures enable us to distinguish excerpts with significantly low or high energy along the different regions. A negative Bass value indicates that the excerpt has relatively little bass; a Midrange value of zero corresponds to an excerpt with average midrange; and a positive Treble value indicates higher than average treble (Lartillot, in preparation).

More precisely, a spectrogram decomposed into Mel-bands is computed from the audio recording, with 50 ms frame and half-overlapping. The spectrum distribution is then summed over time, leading to a single spectrum distribution for the excerpt. Each amplitude related to each separate Mel band is then expressed as a Z-score with respect to the corresponding data across the music corpus. The treble and midrange values are the sum of the Z-scores on Mel bands 24 to 56, and 3 to 22, respectively. For the bass value, because the sub-bass range lies below the first Mel band, the computation is carried out on the spectrogram without Mel-band decomposition. The 20 first bins of the spectrum, corresponding to the frequencies below 200 Hz, are represented in terms of Z scores. A weighted curve is added to emphasize the low frequencies. As before, the weighted energy of those bins are summed together. (Previous approaches, based on either the ratio of energy above and below a given threshold [*brightness*; Juslin, 2000] or the computation of a spectral centroid, fail to describe the spectrum distribution in sufficient detail.)

Roughness

Auditory roughness, or sensory dissonance, relates to a sensation of sound fluctuation, provoked by a beating phenomenon between partials that are close in frequency. This measure is computed on short time frames

and the results are summed across frames. A high *Roughness* value suggests that a music excerpt frequently includes a significant amount of beating partials.

In a model by Sethares (2005), each pair of partials additively contributes to the overall roughness in the form of a roughness term that (1) is an estimation of the beating phenomenon due to the frequency difference between the two partials (based on Plomp and Levelt's (1965) model), and (2) is proportional to the minimum amplitude within the pair of partials. The roughness terms are summed across all possible pairs of partials. The sum is normalized in amplitude by dividing by the same sum of minimum amplitudes within pairs of partials, this time without the Plomp-Levelt contribution (Lartillot & Weisser, 2021; Lartillot, in preparation).

HARMONY

Key Clarity

A tonal descriptor that is rather simple but sufficient for our needs is provided by the feature Key clarity *M*. The higher the value of this feature, the closer a musical excerpt conforms to a standard conception of tonality, as described by the key profiles given by Krumhansl (1990).

A measure of key clarity is obtained by estimating a pitch class profile of the excerpt, showing the prevalence of each of the 12 pitch classes, and further comparing this profile to profiles corresponding to the 24 possible major and minor keys (Krumhansl, 1990). The key with the highest similarity (or highest key strength) will be considered as the most probable key. The highest key strength is referred to as *Key clarity*. By choosing the duration of the time window moving throughout the excerpt, one can control the temporal granularity of the analysis, from single chords to over-all keys. The chosen granularity used in our analysis is 8 second frames, moving at a pace of 0.8 second per frame. The mean across frames yields the feature Key clarity *M*.

Mode

Another feature that can be derived from the above representation is *Mode*. A large positive value suggests that the music is clearly major, a large negative value that the music is clearly minor. A value around zero indicates some ambiguity between major and minor mode. The Mode is simply defined as the difference between the highest major and minor key strengths (i.e., the difference between the scores of the most probable major and minor keys, respectively). From this

estimate, we derive the mean across frames (Mode *M*) and the standard deviation (Mode *SD*).

The pitch class profile should ideally present a statistical description of the pitches of the played notes. However, automated transcription of audio recordings is a challenging and error-prone task. To simplify, the pitch class profile is computed as a *chromagram*: a simple spectrogram is computed and for each frequency (e.g., 440 Hz), the corresponding spectrum magnitude is assigned to the related pitch class (A for 440 Hz) and the magnitudes for each pitch class are summed together. This is a rather coarse approximation, since energy at 440 Hz could also relate for instance to the second harmonic of a pitch of fundamental frequency 146.7 Hz (D). But this approximation may be somewhat compensated by incorporating the harmonic series in the pitch class profiles of the 24 different major and minor keys (Gómez, 2006). However, because the resulting pitch class profiles lead to a bias toward minor modes, we used a slightly modified profile for the minor keys that ensured a better balance between major and minor keys (Lartillot, in preparation).

STRUCTURE

Timbral Changes

The feature *Mfcc Nov* is a measure of the extent to which a music excerpt includes a clear succession of parts with contrastive timbres (a high value) or not (a low value).

The feature is based on Mel-Frequency Cepstrum Coefficients (*Mfcc*), a descriptor of timbre often used in computational audio and music analysis research because it is a practical way to place sounds into a timbral space to establish timbral distances. *Mfcc* can be understood as a description of the “shape” of a sound’s spectrum (computed along Mel bands). The description is carried out by computing a spectrum of the spectrum (corresponding to the concept of cepstrum). However, instead of using Fourier transform, it is computed by means of Discrete Cosine Transform, which offers a more compact description.

Mfcc’s do not offer a clear perceptual interpretation, and cannot be used directly to characterize individual sounds. Thus, the *Mfcc Nov* feature was obtained in the following way: First, *Mfcc*, from ranks 2 to 13, are computed on short frames of duration 7 seconds and moving at a pace of 1.4 second per frame. All frames are compared one to one another using the cosine distance, yielding a similarity matrix (Foote & Cooper, 2003). This matrix reveals timbral transitions in the form of a generally irregular checkerboard pattern, which can be numerically assessed by computing a cross-correlation with a checkerboard kernel along the main diagonal of the matrix. Transitions between parts with contrastive timbres appear as peaks in the resulting novelty curve. The higher the peaks, the longer and more contrastive the successive parts are (Foote & Cooper, 2003).

We present a few improvements to the computation of the novelty curve (Lartillot, in preparation). In particular, the novelty curve is decomposed into two parts: a “back” novelty detection when a homogeneous segment ends, and a “forth” novelty detection when a novel homogeneous segment starts. The *Mfcc Nov* feature is obtained by extracting the 3 highest peaks in the novelty curves (both back and forth curves) and summing together their squared amplitudes.

Tonal Modulation

The feature *KeyQuant15* assesses the tonal homogeneity of an excerpt: a low value indicates that some keys or chords are very distant from each other, whereas a high value indicates a tonally congruent sequence of keys or chords.

This feature is obtained by first measuring the key strength for each of the 24 major and minor keys on time frames of 8 seconds, moving every 160 ms. These key-strength series are then compared between each pair of frames, successive or not, using the cosine distance. The resulting feature *KeyQuant15* corresponds to the 15% quantile within this series of distances.

References

- FOOTE, J., & COOPER, M. L. (2003). Media segmentation using self-similarity decomposition. *SPIE Storage and Retrieval for Multimedia Databases*, 5021, 167–175.
- GÓMEZ, E. (2006). *Tonal description of music audio signal* [PhD thesis]. Universitat Pompeu Fabra, Barcelona, Spain.
- JUSLIN, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797–1813.
- KRUMHANSL, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press.
- LARTILLOT, O., & GRANDJEAN, D. (2019). Tempo and metrical analysis by tracking multiple metrical levels using autocorrelation. *Applied Sciences*, 9, 5121. <https://doi.org/10.3390/app9235121>

LARTILLOT, O., & WEISSER, S. (2021). Roughness, crackliness, buzzingness . . . : Characterizations of sonic unsteadiness and application to the analysis of traditional music from Ethiopia, Kenya, Morocco and India. In R. Timmers et al. (Eds.), *Proceedings of the Sixteenth International Conference on Music Perception and Cognition*. University of Sheffield, Sheffield, UK.

PLOMP, R., & LEVELT, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 37, 548–560.

SETHARES, W. A. (2005). *Tuning, timbre, spectrum, scale* (2nd ed.). Springer.

APPENDIX D Summary of the Relationships Between Felt Emotions and Acoustic Features

Group	Feature	Emotion								
		Happiness	Sadness	Surprise	Calm	Interest	Nostalgia	Pride	Love	Awe
(Tempo)	Metroid (<i>M</i>)	.291	-.076	-.043	.014	.097	.198	.187	-.034	.093
	Metroid (<i>SD</i>)	-.207	.083	.055	-.024	.008	-.088	-.093	.075	-.004
(Dynamics)	Crescendo	.098	-.058	-.040	.013	-.042	.114	.071	.042	.040
(Timbre)	Bass	.005	-.052	-.129	.048	-.004	-.119	.055	.019	.050
	Midrange	.338	-.307	.134	-.269	.016	-.111	.273	-.255	-.149
(Harmony)	Treble	.181	-.320	-.021	-.264	.108	.060	.290	-.220	-.208
	Roughness	.125	-.173	.133	-.039	.072	-.046	.199	-.152	-.067
	Key clarity	.155	.080	-.081	.185	.278	.226	.167	.053	.151
	Mode (<i>M</i>)	.145	.035	.016	.085	-.081	.282	.029	.201	.082
(Rhythm)	Mode (<i>SD</i>)	-.133	.137	.016	-.033	.048	-.232	-.275	.207	-.063
	Metr strength (<i>M</i>)	-.033	-.045	.046	-.147	.060	-.212	-.071	-.026	-.115
(Structure)	Metr strength (<i>SD</i>)	-.004	-.063	-.028	-.100	-.181	-.075	.003	-.124	-.204
	Mfcc Nov	-.026	-.027	.343	.023	.223	-.134	-.148	-.091	-.016
	KeyQuant15	-.171	-.010	-.158	-.096	-.437	-.126	-.111	-.086	-.107

Note. Values show semi-partial correlations (r_{sp}) between predicted emotions and acoustic features in regression models. The highest and lowest values for each emotion are shown in bold (see Method for details).