Exploring Appraisal Dimensions in Music: Toward the Development of a Standardised Appraisal Tool for Musical Emotions

Thomas M. Lennie
American University in Bulgaria, Blagoevgrad, Bulgaria
tlennie@aubg.edu

Abstract

The relationship between music and emotion is a central topic in music psychology, with cognitive appraisals—evaluations based on goal relevance, novelty, and coping potential—playing a key role. This study examines the multidimensional nature of musical appraisals and addresses gaps in both theory and empirical research. Using a diverse and ecologically valid set of musical stimuli, sixty-six participants evaluated 14 appraisal dimensions and rated their emotional responses (valence, arousal, and interest) using diverse, ecologically valid musical stimuli. Ten appraisal items proved statistically robust, and exploratory factor analysis revealed three latent dimensions—Goal-congruence, Novelty, and Complexity—that underlie appraisals during music listening. Overall, this study provides the first proof of concept for a multidimensional appraisal framework in music, advancing the field by offering a systematic method for assessing music-induced emotions and laying the groundwork for a standardised appraisal tool.

Keywords: music, emotion, appraisal, goal-directed

Introduction

The interplay between music and emotion has intrigued researchers across disciplines for a long time, prompting investigations into how musical experiences evoke emotional responses. Central to these investigations is the concept of cognitive appraisal, whereby individuals evaluate the significance of stimuli—based on factors such as goal relevance, novelty, and coping potential—thereby shaping their emotional experiences. Appraisal theory, rooted in utilitarian emotion psychology (Moors & Scherer, 2013; Scherer, 2013), is also a key component in many theories of musically induced emotions (Céspedes-Guevara, 2023; Eerola, 2017; Juslin, 2019; Lennie & Eerola, 2022; Scherer & Coutinho, 2013). Although these theories differ regarding whether appraisal processes serve as mediators or central mechanisms, they all agree that appraisals contribute to individual and contextual differences in emotional responses to music.

Musical appraisals

Empirical research on music-induced emotions underscores the importance of appraisal processes. Studies on familiarity (Bosch et al., 2013; Pereira et al., 2011) have found that familiar music elicits stronger emotional responses by enabling listeners to predict musical events, with cultural background further shaping perception. In terms of complexity, North and Hargreaves (1998) demonstrated that liking for pop music follows an inverted-U relationship with subjective complexity, in contrast to a simpler linear relationship with familiarity. Research by Cheung et al. (2019) revealed that when music meets listeners' expectations, pleasure and satisfaction increase; conversely, violation of expectations can lead to mixed responses. Moreover, goal relevance—how well music aligns with an individual's current emotional state or desired outcomes—influences both music selection and mood regulation (Greb et al., 2019; Randall & Rickard, 2017).

These studies clearly illustrate the impact of individual appraisal dimensions. However, most have examined them in isolation, leaving the dynamic interplay between multiple appraisals in emotional episodes largely unexplored.

Interaction in appraisal dimensions

Research in broader affective sciences has shown that the interaction between appraisal dimensions can shape emotional responses significantly. For example, work by Silvia (2005b) and Turner and Silvia (2006) indicates that novel stimuli appraised as manageable tend to evoke interest and curiosity, whereas unmanageable novelty may lead to anxiety or avoidance. In addition, Laukka and Elfenbein (2012) demonstrated that listeners can reliably infer multiple appraisal dimensions from vocal expressions, highlighting their role in emotional

communication—a finding that supports similar investigations in musical contexts. There are many appraisal interactions that could be noted, but a full list lies beyond the scope of this paper; for an overview of appraisal dimensions see the first meta-analysis published on the topic (Yeo & Ong, 2024). It suffices to note that exploring such ideas in a musical context is both beneficial and timely.

However, emotional responses to music generally arise from the interplay between factors such as goal relevance, familiarity, and complexity. For instance, familiarity can moderate other appraisals by shaping expectations; Gold et al. (2019) found that predictability and uncertainty influence both pleasure and engagement. When listeners' expectations are met, they experience satisfaction, while unexpected musical events can evoke surprise or tension, which can also be emotionally impactful. Recognising these interactions is essential for developing a comprehensive model of musical emotions (Lennie & Eerola, 2022), that explains why individuals may respond so differently to the same musical piece.

Aims

Further research is required to bridge both theoretical and empirical gaps in our understanding of musical appraisals, particularly given the current lack of an appropriate assessment tool in this context. This study marks the first exploratory evaluation of multiple appraisal dimensions in music and takes an initial step towards developing and validating a standardised tool. In doing so, it also extends previous work by emphasising the

Table 1. Selected appraisal terms and their relationship to the theoretical structure of appraisal categories.

Appraisal group	Sub-category	Questions format	Scale
Relevance	Novelty	How predictable is the music? How familiar is the music? How complex is the music?	Unpredictable / Predictable Unfamiliar / Familiar Simple / Complex
	Intrinsic Pleasantness	How pleasant is the music?	Unpleasant / Pleasant
	Goal-relevance	<How expressive is the music?> <Did the music have a strong effect on you?>	Unexpressive / Expressive No Effect / Strong Effect
Implications / Consequences	Expectations	Did you expect to hear music like this?	Did not expect / Expected
	Goal-congruence	Did the music produce a desirable outcome? * <How satisfying is the music? >	Undesirable / Desirable Unsatisfying / Satisfying
Coping Potential	Control	* How comprehensible is the music?	Incomprehensible / Comprehensible
Norm-compatibility	Internal standards	* Do you find the music aesthetically beautiful? * Do you find the music aesthetically challenging? Does the music represent your personal ideas of good music?	Un-beautiful / Beautiful Unchallenging / Challenging Poor / Good
	External standards	Does the music conform to socially accepted ideas of good music?	Violates / Conforms

Note:

[a] Multiple appraisal formats are suggested if there is not sufficient evidence to conclude that one may offer a better interpretation for musical appraisal.

[b] In line with other appraisal theories, this musical interpretation suggests that some appraisal formats may be representative of an appraisal group instead of a single sub-category. These appraisals are marked with an asterisk (*).

[c] Some appraisal formats are implicit measures of appraisals, less available to consciousness. These appraisals are marked with angle brackets (<>).

predictive relationships between appraisal factors and core affective states such as valence, arousal, and aesthetic interest. Specifically, the study aims to:

1. Investigate the validity and reliability of appraisal dimensions as they pertain to musical stimuli.
2. Explore the predictive power of appraisal dimensions (e.g., goal congruence, familiarity, and complexity) on core affective and aesthetic responses to music.
3. Lay the groundwork for the development of a standardised appraisal tool for music-induced emotions, informed both by empirical data and theoretical advances.

Method

Measures

Following the utilitarian appraisal framework (Moors et al., 2013), appraisals were assessed across four key categories: relevance, implications, coping potential, and norm-compatibility. This structure is true for the most popular theoretical adaption of appraisal for music, too (Scherer & Coutinho, 2013). Specific appraisal dimensions were developed from literature sources (Céspedes-Guevara, 2023; Lennie & Eerola, 2022; Robinson, 2005, 2009; Scherer, 2013; Scherer & Coutinho, 2013) and refined by the author and two music psychology experts to suit an online listening study (Table 1). For example, utilitarian appraisal dimensions such as 'adjustment' (the ability to change one's listening situation) and 'causal attribution' (identifying the source of the music) were excluded as they did not fit this study design. In addition to the appraisal dimensions, participants were asked to rate three further items. Arousal was described as "how sleepy/ awake do you feel the music to be". Valence was described as "how positive / negative do you feel the music to be". Interest was described as "how interesting / uninteresting do you feel the music to be". All scales were rated on a 7-point Likert scale.

Stimuli were sourced from the DEAM dataset (Alajanki et al., 2016), which comprises 1,800 royalty-free 45-second excerpts across 10 genres, all sampled at 44100 Hz. Sixty excerpts were selected to represent five commonly studied Western genres (classical, electric, jazz, world/international, pop/rock; Rentfrow et al., 2011). For each genre, stimuli were chosen based on valence and arousal variability: high (SD > 2.0) and low standard deviation (SD < 1.0) for both valence and arousal, with three excerpts selected per variance category. All tracks featured no lyrics, indiscernible lyrics, or lyrics in a language unlikely to be understood by UK participants. A full list of tracks is provided in Appendix A, and the collected valence and arousal scores (Appendix B) confirm a diverse affective range.

Participants and procedure

A total of 66 participants (65% female; age range 18-69, M_{age} = 34.5, SD = 11.3) were recruited using Prolific. Eligibility criteria required participants to be UK nationals that were born and are currently residing in the UK. Participants first read detailed instructions, including definitions of valence, arousal, and interest, and completed a volume check. They were asked to maintain a constant volume throughout the study. Each participant was randomly presented with 20 out of the total of 60 tracks, ensuring an even representation of high and low standard deviation tracks for both valence and arousal across the five genres. Participants could rate the appraisal terms as they listened to the 45s excerpt and could listen to the track multiple times.

Following the listening task, participants completed the G-MSI musical training questionnaire (Müllensiefen et al., 2014) and self-rated their musicality on a scale from non-musician to professional. They also indicated their music genre preferences (e.g. rock/heavy, pop/electro, classical/ethnic, other) and provided demographic information, including age and gender. The largest proportion of the participants represented music-loving non-musicians (n = 34) and amateur musicians (n = 17) with a G-MSI musical training score (M = 20.47, SD = 11), slightly lower than the British average (Müllensiefen et al., 2014). Genre preferences indicated the most popular genre was pop/electro (35%), followed by other (22%), rock/heavy (18%), and classical (3%).

Results

All data were analysed in RStudio, with G-MSI scores computed following Müllensiefen et al. (2014); the R scripts and data are available from the author.

Internal consistency

Cronbach's alpha was used to assess the internal consistency of each appraisal dimension. Nearly all terms demonstrated good reliability ($\alpha > 0.70$), except for 'Effect' (goal-relevance sub-category)

and 'Challenge' (internal-standards sub-category), which showed very low alpha values. These two dimensions were therefore excluded from further analyses. Confidence intervals indicated reasonable homogeneity across rating scales, supporting the robustness of the data for examining individual differences (e.g., age, genre preferences, and genre type) in musical appraisals. The confidence interval and Cronbach' alphas for each scale can be seen in Appendix C.

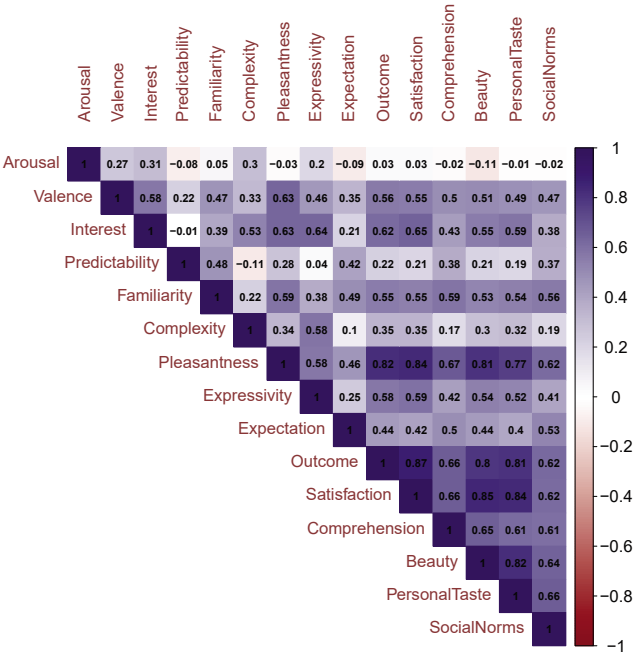

Figure 1. Correlation table for individual appraisal dimensions with arousal, valence, and interest using Pearson correlation coefficient.

Correlations

After removing the two invalid appraisal dimensions, correlations between the remaining appraisal dimensions and the dependent variables were examined. Several appraisal terms were used to represent the same dimension because it was unclear if one interpretation was theoretically superior, and there may be significant overlap in the variance they explain. Moreover, examining these correlations is informative given previous work on the overlap between core-affect and interest with appraisals.

Pearson correlation coefficients were computed to examine relationships among appraisal dimensions (Figure 1). The results revealed clear correlations between 'Valence' and 'Interest' with several appraisal dimensions. In contrast, 'Arousal' showed almost no correlation with the appraisal dimensions, aside from weak associations with 'complexity' and 'expressivity'. Individual appraisal dimensions varied in their correlations, ranging from weak to strong. Notably, very high correlations (>.80) were found between terms representing the same underlying appraisal dimension. For instance, 'Satisfaction' and 'Outcome' correlated at .87 and were theoretically representative of goal-congruence; hence, they were averaged to form a composite scale named 'Goal-congruence'. Similarly, 'Beauty' and 'Personal Taste' correlated at .82 and were averaged to form a composite scale 'Internal Standards'.

Although four scales showed similarly high correlations (> .80), they represent different theoretical levels within the appraisal model and were therefore retained separately to preserve the hierarchical structure. Much evidence supports this utilitarian structure (Davidson et al., 2009). It is possible that these high correlations partly reflect characteristics of the DEAM dataset or the selected stimuli.

Exploratory factor analysis

The new set of ten appraisal dimensions underwent parallel analysis to estimate the number of components and factors. The analysis suggested a two-component and three-factor solution (Appendix D). Subsequently, an exploratory factor analysis (EFA) was then conducted (Watkins, 2018) assuming three factors, utilising a principal factor solution with an oblimin rotation, which allows for correlated latent factors.

The three-factor solution showed that all items had sufficient loadings (|factor loadings| > 0.30). Although Comprehension and Social Norms displayed small cross-loadings, the differences (>.12) were substantial enough not to compromise factor interpretation or independence. The Tucker Lewis Index (TLI = 0.972) indicated an acceptable fit (> 0.9; Bentler & Bonett, 1980), and the Root Mean Square Error of Approximation (RMSEA = 0.0649, 90% CI: 0.0587–0.0712) was acceptable to good (Hu & Bentler, 1999). In addition, all sums of squared loadings were greater than 1 (Kaiser's rule), confirming the meaningfulness of the factors. The

full table of factor loading coefficients is available in Appendix E.

Based on loading strength and theoretical considerations, Factor 1 was named 'Goal-congruence' (accounting for 36% of the variance), Factor 2 'Novelty' (13% of the variance), and Factor 3 'Complexity' (18% of the variance). Individual item loadings and inter-factor correlations are presented in Figure 2.

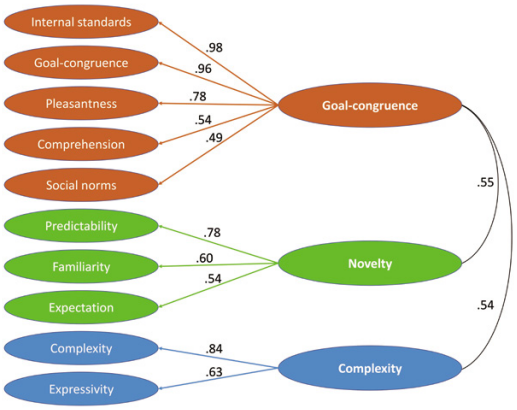

Figure 2. Factor analysis with component loadings and factor correlation coefficients.

Discussion

This discussion focuses on the validity and latent structure of the appraisal terms. For clarity, it differentiates between three levels of terminology: appraisal groups and latent factors (denoted in capitals within single quotation marks, e.g. 'Goal-congruence'), individual appraisal dimensions or sub-categories (in lowercase within single quotation marks, e.g. 'novelty'), and the terms used in the study (in double quotation marks, e.g. "predictability").

The validity of appraisal dimensions for music

The study originally proposed 14 appraisal dimensions, of which 12 showed good internal consistency across participants. In several instances, multiple terms were used to represent a single appraisal dimension—for example, both "outcome" and "satisfaction" were used to capture the appraisal dimension 'goal-congruence', and "beauty" alongside "personal taste" represented 'internal standards'. Strong correlations (>.80) between such terms justified the formation of composite measures, thereby reducing the final set to ten appraisal items representing all sub-categories and the four wider appraisal groups.

A notable finding was that some appraisal dimensions, such as 'novelty', were represented by multiple terms (e.g. "predictability", "familiarity", and "complexity") that did not correlate strongly enough to justify merging them into a composite measure. This outcome suggests that these dimensions are multifaceted constructs, perhaps reflecting different linguistic or cognitive interpretations of the same underlying process. In addition, some terms were originally hypothesised to represent broader appraisal groups rather than isolated sub-categories. For instance, "satisfaction" was originally intended to indicate the wider 'Implications and Consequences' group, while "beauty" was considered to be a proxy for the 'Norm-compatibility' group. The present results underscore that individual appraisal dimensions can be meaningfully distinguished, and that theories treating appraisal groups as unitary constructs may benefit from a more nuanced understanding of these terms.

An additional consideration is that several appraisal terms are designed to measure processes that may be less accessible to conscious introspection. For example, "expressivity" and "effect" were used to capture aspects of the 'goal-relevance' appraisal dimension—processes that Scherer et al. (2006) have suggested are implicit in nature. While the low internal reliability of "effect" ultimately led to its exclusion, the overall pattern indicates that implicit measures can be seen as indicative of the underlying mechanisms. Thus, the final set of appraisal measures provides a valid framework for the cognitive assessment of music-induced emotions.

The structure of musical appraisal

To assess the latent structure of the appraisal dimensions, a parallel analysis was conducted, which initially indicated a two-component solution. This aligns with the valence–arousal structure inherent in the stimuli (Alajanki et al., 2016), suggesting that the cumulative variance captured by the appraisal dimensions is largely explained by two broad dimensions. However, when an exploratory factor analysis (EFA) was performed using a principal factor solution with an oblimin rotation (to allow for correlated factors, as expected in appraisal research), a three-factor solution emerged as the best representation of the data.

This three-factor solution diverges from the traditional four-factor model of appraisal—typically comprising 'Relevance', 'Implications', 'Coping

Potential', and 'Norm-compatibility' (Ellsworth & Scherer, 2003). Instead, the analysis indicates that components from multiple wider appraisal groups are integrated within the latent factors.

Factor 1 ('Goal-congruence'). This factor includes items from all four appraisal groups. Notably, the composite scales for 'internal standards' (with a loading of .98) and 'goal-congruence' (with a loading of .96) dominated this factor. In line with Smith and Lazarus (1991), this finding suggests that appraisals related to social and normative compatibility are subsumed under a broader goal-congruence construct. There is of course the possibility that the distinction between goal-congruence (in the 'Implications/Consequences' appraisal group) and 'Normative Significance' is not captured in this set of stimuli (the stimuli did not vary significantly in these dimensions from those in goal-congruence). Similarly, the study participants (British nationals) may not capture meaningful distinctions between what people see as socially and individually normative as opposed to goal-congruent without a comparison group. Regardless of the additional inclusion of 'internal standards', these results suggest that goal-congruence plays a significant role in the cognitive evaluation of music.

Factor 2 ('Novelty'). This factor comprised of items representing both the 'novelty' and the 'expectations' dimensions, predominantly reflecting utilitarian interpretations of appraisal. The strongest loadings—"predictability" (.78) and "familiarity" (.60)—support the conceptualisation of novelty as a key aspect of musical appraisal.

Factor 3 ('Complexity'). This factor is characterised by the loadings of "complexity" (.84) and "expressivity" (.63). Although "complexity" was initially associated with 'novelty' (Table 1) based on Silvia's (2005a, 2005b; 2006) work on aesthetic appraisal, its strong and distinct loading here indicates that it may capture a different aspect of musical processing. It is worth noting that Silvia used an implicit experimental manipulation of more and less complex visual images as representative of complexity; his design did not use the term "complexity" explicitly, as was done here. There is an ongoing debate in the literature over the role of "complexity"—whether as a marker of novelty, 'Norm-compatibility' (Céspedes-Guevara, 2023), or emotional complexity (Robinson, 2005)—that remains unresolved. Robinson's (2005) interpretation appears closest to the results generated here. Her philosophical account of aesthetic appraisal interprets the term "complexity" as referring to the emotional complexity of what is being expressed by music; similar to the implicit measure of goal-relevance ("expressivity") presented in this adaption of appraisal terms. While the term "Complexity" was selected to reflect the distinct nature of this factor in the present study, it should be interpreted with caution, given the many interpretations noted in previous literature. What is clear from this analysis is the apparent link with the goal-relevance of the music.

Limitations and future directions

The chosen stimuli covered a broad affective space (valence and arousal) but did not systematically manipulate all appraisal dimensions. A collection of stimuli that varies across all appraisal dimensions would be incredibly large, time-consuming for participants, and would likely require a validation task, as no such dataset currently exists. Future studies would benefit from such implicit manipulations of appraisal dimensions in experimental stimuli (e.g., musical stimuli that vary in complexity or goal-relevance). However, achieving such a mammoth task is likely to be built on incremental studies, testing single appraisal dimensions at a time.

Participants may have reported hypothetical or imagined emotional responses - an established critique in appraisal research (Robinson & Clore, 2001). In addition, the short (45-second) musical excerpts may not always induce full emotional episodes, despite prior evidence suggesting this duration is often sufficient (Eerola & Vuoskoski, 2012). However, this alone is not a sufficient criterion for an emotional episode to occur (i.e., if 45 seconds of music is heard, an emotional episode will follow). Regardless of whether a full-blown emotional episode was experienced, this research provides ample support for the conclusion that appraisal is a meaningful component of the emotional experience of music.

The study's sample size limits the robustness of conclusions about individual appraisal dimensions and their latent structure. Factor analysis requires approximately ten participants per variable for stability, and a larger sample would allow for confirmatory analysis, an important next step.

Appraisals likely fluctuate over time rather than being static judgments (Scherer, 2009). Future work could explore continuous appraisal responses, but this is best approached incrementally, by investigating individual appraisals in real-time.

The relationship between appraisal and core-

affect (valence and arousal) is a key component of utilitarian theories (Russell, 2003, 2009) and music-emotion theories (Céspedes-Guevara, 2023; Eerola, 2017; Lennie & Eerola, 2022; Scherer & Coutinho, 2013). The strong correlations between appraisal dimensions and valence highlight the potential for advanced modelling of appraisal contributions to core affect, which would offer a novel addition to the field.

The appraisal terms used were adapted from existing theoretical models. Future research could refine these through empirical analysis of commonly used descriptors to enhance validity and refine dimensions. This would provide a robust grounding for appraisal in music and improve the internal validity of (or offer alternatives to) the dimensions proposed.

Conclusions

This study demonstrated the validity of ten appraisal mechanisms in musically induced emotional episodes across an ecologically valid stimulus set. Exploratory factor analysis grouped these appraisal dimensions into three latent factors: 'Goal-congruence', 'Complexity', and 'Novelty'. This study provides the first proof of concept for applying the broader appraisal framework to music and represents the first analysis of multiple appraisals in music.

References

Alajanki, A., Yang, Y. H., & Soleymani, M. (2016). Benchmarking music emotion recognition systems. *PLOS ONE, 11*(3), e0151936. https://doi.org/10.1371/journal.pone.0151936

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bosch, I. v. d., Salimpoor, V. N., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience, 7*, 534. https://doi.org/10.3389/fnhum.2013.00534

Céspedes-Guevara, J. (2023). A constructionist approach to emotional experiences with music. *Advances in Cognitive Psychology, 19*(4), 46–62. https://doi.org/10.5709/acp-0402-4

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J., & Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. *Current Biology, 29*(23), 4084–4092.e4. https://doi.org/10.1016/j.cub.2019.09.067

Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (Eds.). (2009). *Handbook of affective sciences*. Oxford University Press.

Eerola, T. (2017). Music and emotions. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 539–554). Springer. https://doi.org/10.1007/978-3-662-55004-5_29

Eerola, T., & Vuoskoski, J. K. (2012). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception, 30*(3), 307–340. https://doi.org/10.1525/mp.2012.30.3.307

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 572–595). Oxford University Press.

Gold, B. P., Pearce, M. T., Mas-Herrero, E., Dagher, A., & Zatorre, R. J. (2019). Predictability and uncertainty in the pleasure of music: A reward for learning? *The Journal of Neuroscience, 39*(47), 9397–9409. https://doi.org/10.1523/JNEUROSCI.0428-19.2019

Greb, F., Steffens, J., & Schlotz, W. (2019). Modeling music-selection behavior in everyday life: A multilevel statistical learning approach and mediation analysis of experience sampling data. *Frontiers in Psychology, 10*, 390. https://doi.org/10.3389/fpsyg.2019.00390

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

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

Laukka, P., & Elfenbein, H. A. (2012). Emotion appraisal dimensions can be inferred from vocal expressions. *Social Psychological and Personality Science, 3*(5), 529–536. https://doi.org/10.1177/1948550611428011

Lennie, T. M., & Eerola, T. (2022). The CODA model: A review and skeptical extension of the constructionist model of emotional episodes induced by music. *Frontiers in Psychology, 13*, 822264. https://doi.org/10.3389/fpsyg.2022.822264

Moors, A., & Scherer, K. R. (2013). The role of appraisal in emotion. In M. Robinson, E. Watkins, & E. Harmon-Jones (Eds.), *Handbook of cognition and emotion* (pp. 135–155). Guilford Press.

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review, 5*(2), 119–124. https://doi.org/10.1177/1754073912468165

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLOS ONE, 9*(2), e89642. https://doi.org/10.1371/journal.pone.0089642

North, A. C., & Hargreaves, D. J. (1998). Complexity, prototypicality, familiarity, and the perception of musical quality. *Psychomusicology: Music, Mind, and Brain, 17*(1–2), 77–80. https://doi.org/10.1037/h0094058

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLOS ONE, 6*(11), e27241. https://doi.org/10.1371/journal.pone.0027241

Randall, W. M., & Rickard, N. S. (2017). Personal music listening: A model of emotional outcomes developed through mobile experience sampling. *Music Perception, 34*(5), 501–514. https://doi.org/10.1525/mp.2017.34.5.501

Rentfrow, P. J., Goldberg, L. R., & Zilca, R. (2011). Listening, watching, and reading: The structure and correlates of entertainment preferences. *Journal of Personality, 79*(2), 223–258. https://doi.org/10.1111/j.1467-6494.2010.00662.x

Robinson, J. (2005). *Deeper than reason: Emotion and its role in literature, music, and art*. Oxford University Press.

Robinson, J. (2009). Emotional responses to music: What are they? How do they work? And are they relevant to aesthetic appreciation? In P. Goldie (Ed.), *The Oxford handbook of philosophy of emotion* (pp. 651–680). Oxford University Press.

Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin, 27*(11), 1520–1532. https://doi.org/10.1177/01461672012711012

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion, 23*(7), 1259–1283. https://doi.org/10.1080/02699930902809375

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion, 23*(7), 1307–1351. https://doi.org/10.1080/02699930902928969

Scherer, K. R. (2013). Measuring the meaning of emotion words: A domain-specific componential approach. In J. J. R. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 7–30). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199592746.003.0002

Scherer, K. R., & Coutinho, E. (2013). How music creates emotion: A multifactorial process approach. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 121–145). Oxford University Press.

Scherer, K., Dan, E., & Flykt, A. (2006). What determines a feeling's position in affective space? A case for appraisal. *Cognition & Emotion, 20*(1), 92–113. https://doi.org/10.1080/02699930500305016

Silvia, P. J. (2005a). Cognitive appraisals and interest in visual art: Exploring an appraisal theory of aesthetic emotions. *Empirical Studies of the Arts, 23*(2), 119–133. https://doi.org/10.2190/12AV-AH2P-MCEH-289E

Silvia, P. J. (2005b). What is interesting? Exploring the appraisal structure of interest. *Emotion, 5*(1), 89–102. https://doi.org/10.1037/1528-3542.5.1.89

Silvia, P. J. (2006). Artistic training and interest in visual art: Applying the appraisal model of aesthetic emotions. *Empirical Studies of the Arts, 24*(2), 139–161. https://doi.org/10.2190/DX8K-6WEA-6WPA-FM84

Smith, C. A., & Lazarus, R. S. (1991). Emotion and adaptation. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 609–637). Guilford Press.

Turner, S. A., & Silvia, P. J. (2006). Must interesting things be pleasant? A test of competing appraisal structures. *Emotion, 6*(4), 670–674. https://doi.org/10.1037/1528-3542.6.4.670

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Yeo, G. C., & Ong, D. C. (2024). Associations between cognitive appraisals and emotions: A meta-analytic review. *Psychological Bulletin, 150*(12), 1440–1471. https://doi.org/10.1037/bul0000452

https://doi.org/10.17234/9789533793085.22

Appendix

Appendix A

Table 1. Selected track numbers and high/low sd grouping from the DEAM dataset (Alajanki et al., 2016).

DEAM - track number	Genre	Core-affect	sd
138	Classical	Valence	High
159	Classical	Valence	High
246	Classical	Valence	High
456	Electronic	Valence	High
423	Electronic	Valence	High
454	Electronic	Valence	High
707	Jazz	Valence	High
676	Jazz	Valence	High
634	Jazz	Valence	High
35	World	Valence	High
1638	World	Valence	High
1611	World	Valence	High
952	Pop	Valence	High
797	Pop	Valence	High
781	Pop	Valence	High
158	Classical	Valence	Low
171	Classical	Valence	Low
198	Classical	Valence	Low
425	Electronic	Valence	Low
459	Electronic	Valence	Low
482	Electronic	Valence	Low
1757	Jazz	Valence	Low
1792	Jazz	Valence	Low
657	Jazz	Valence	Low
1615	World	Valence	Low
1683	World	Valence	Low
1606	World	Valence	Low
851	Pop	Valence	Low
902	Pop	Valence	Low
938	Pop	Valence	Low

(Cont.)

DEAM - track number	Genre	Core-affect	sd
207	Classical	Arousal	High
150	Classical	Arousal	High
172	Classical	Arousal	High
387	Electronic	Arousal	High
422	Electronic	Arousal	High
423	Electronic	Arousal	High
631	Jazz	Arousal	High
676	Jazz	Arousal	High
696	Jazz	Arousal	High
1684	World	Arousal	High
1673	World	Arousal	High
1675	World	Arousal	High
776	Pop	Arousal	High
858	Pop	Arousal	High
781	Pop	Arousal	High
1118	Classical	Arousal	Low
1150	Classical	Arousal	Low
1132	Classical	Arousal	Low
390	Electronic	Arousal	Low
427	Electronic	Arousal	Low
1332	Electronic	Arousal	Low
722	Jazz	Arousal	Low
1775	Jazz	Arousal	Low
649	Jazz	Arousal	Low
1654	World	Arousal	Low
1693	World	Arousal	Low
1632	World	Arousal	Low
756	Pop	Arousal	Low
767	Pop	Arousal	Low
804	Pop	Arousal	Low

Appendix B

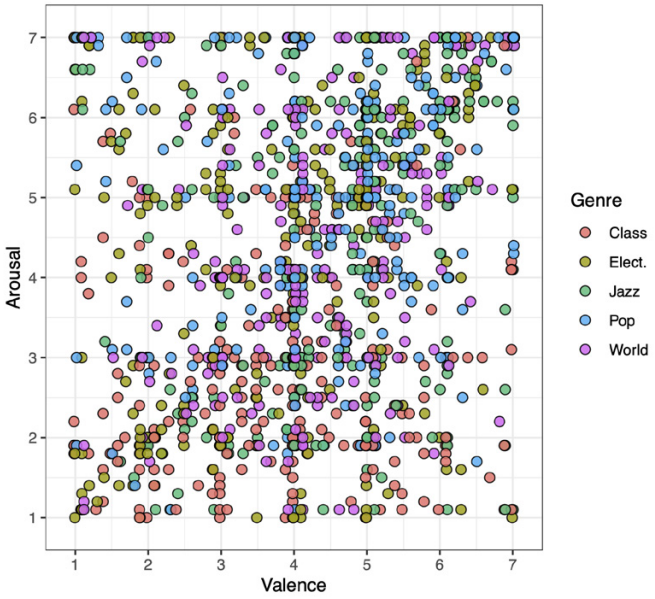

Figure 1. Valence and arousal ratings for selected stimuli as rated in the DEAM dataset (Alajanki et al., 2016) and grouped by genre. This demonstrates selected tracks have a wide distribution across the affective space.

Appendix C

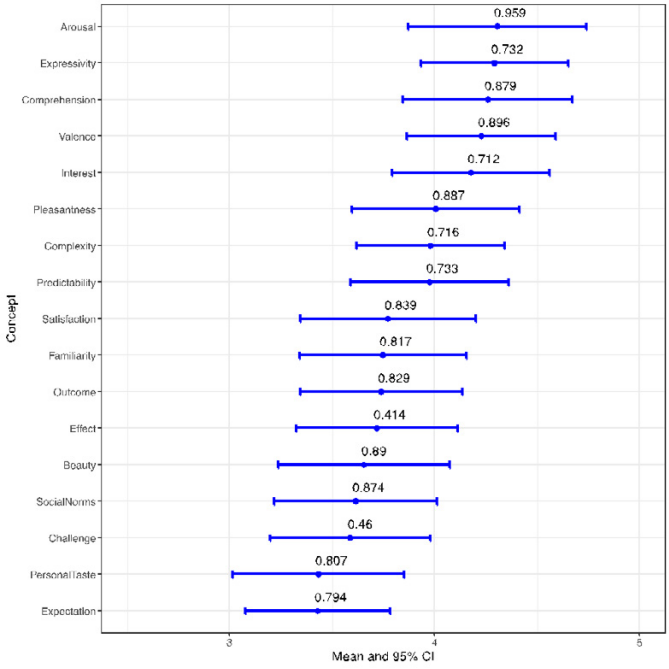

Figure 2. Confidence interval and Cronbach's alphas scores for individual for each scale of measurement.

Appendix D

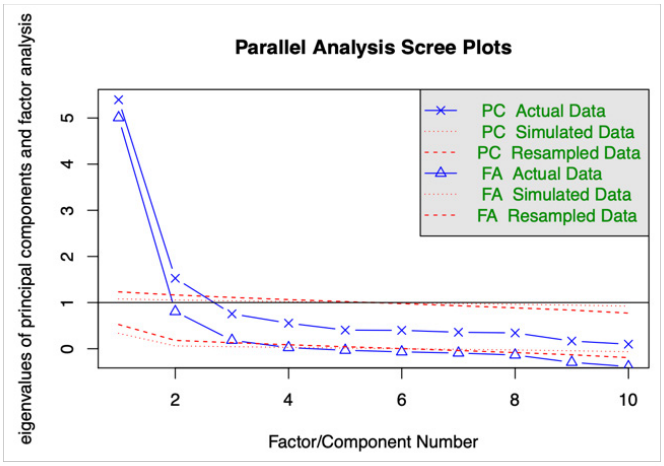

Figure 3. Parallel analysis scree plot

Appendix E

Table 2. Individual loadings for exploratory factor analysis.

	PA1	PA2	PA3
Internal Standards	0.9776101	-0.0457095	-0.0270843
Goal congruence	0.9571385	-0.0436784	0.0357015
Pleasantness	0.7811119	0.1077825	0.0866645
Comprehension	0.5383979	0.3559365	-0.0057234
Social Norms	0.4905038	0.3733337	0.0190381
Predictability	-0.1198802	0.7784378	-0.0898834
Familiarity	0.1820440	0.6048153	0.1821497
Expectation	0.1820440	0.5428005	0.0282370
Complexity	-0.0955535	-0.0219592	0.8392279
Expressivity	0.2614558	0.0225413	0.6280607