Musical Aesthetic Sensitivity

Ana Clemente, Marcus T. Pearce, and Marcos Nadal

Author Note

Ana Clemente, Human Evolution and Cognition Research Group (EvoCog), University of the Balearic Islands, Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), Associated Unit to CSIC, Palma, Spain; Marcus T. Pearce, School of Electronic Engineering & Computer Science, Queen Mary University of London, UK, and Centre for Music in the Brain, Department of Clinical Medicine, Aarhus University, Denmark; Marcos Nadal, Human Evolution and Cognition Research Group (EvoCog), University of the Balearic Islands, Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), Associated Unit to CSIC, Palma, Spain.

The project leading to these results has received funding from “La Caixa” Foundation (ID 100010434) under agreement LCF/BQ/ES17/11600021, and from the Spanish Ministerio de Economía, Industria y Competitividad with grant PSI2016- 77327-P. All authors approved the final version of this manuscript, and thank E. G. Chuquichambi and C. Rey for help in collecting the data.

Correspondence concerning this article should be addressed to Ana Clemente, Department of Psychology, University of the Balearic Islands, Crta Valldemossa km 7.5 Palma de Mallorca 07122 Spain. E-mail: ana.c.magan@gmail.com

Abstract

Empirical aesthetics has mainly focused on general and simple relations between stimulus features and aesthetic appreciation. Consequently, to explain why people differ so much in what they like and prefer continues to be a challenge for the field. One possible reason is that people differ in their aesthetic sensitivity, i.e., the extent to which they weigh certain stimulus features. Studies have shown that people vary substantially in their aesthetic sensitivities to visual balance, contour, symmetry, and complexity, and that this variation explains why people like different things. Our goal here was to extend this line of research to music and examine aesthetic sensitivity to musical balance, contour, symmetry, and complexity. Forty-eight non-musicians rated their liking for 96 4-second Western tonal musical motifs, arranged in four subsets varying in balance, contour, symmetry, or complexity. We used linear mixed-effects models to estimate individual differences in the extent to which each musical attribute determined their liking. The results showed that participants differed remarkably in the extent to which their liking was explained by musical balance, contour, symmetry, and complexity. Furthermore, a retest after two weeks showed that this measure of aesthetic sensitivity is reliable, and suggests that aesthetic sensitivity is a stable personal trait. Finally, cluster analyses revealed that participants divided into two groups with different aesthetic sensitivity profiles, which were also largely stable over time. These results shed light on aesthetic sensitivity to musical content and are discussed in relation to comparable existing research in empirical aesthetics.

Keywords: aesthetic sensitivity, aesthetics, liking, music, sensory valuation

Musical Aesthetic Sensitivity

What are the laws that govern the relations between the physical and the mental? Fechner (1860) was convinced that this question could be answered by probing the quantitative relations between stimulus magnitude and sensation magnitude. He believed, however, that sensation could not be measured directly, so he developed indirect measures of the stimulus values necessary to produce differences in sensation. Although sensation itself could not be measured, differences in sensation could: People could notice whether a sensation was present or absent, or that one sensation was greater than, equal to, or smaller than another (Boring, 1950). Differential sensitivity was, thus, central to psychophysics.

Empirical aesthetics was, in its origin and essence, applied psychophysics. Fechner used empirical aesthetics to tackle the problems of aesthetics in the same way he had used psychophysics to tackle the mind–body problem (Murphy, 1929): to identify the lawful manner in which the mind translates stimulus properties into appreciation. The sensations of beauty and pleasantness could not be measured directly, so he devised methods to quantify how changes in the magnitude of stimulation produced changes in the magnitude of beauty and pleasantness. In the early days of empirical aesthetics, researchers assembled diverse sets of materials and developed new paradigms to explore how variations in certain aspects of stimuli lead to variations in appreciation. Differences in line orientation, length, curvature, thickness (Martin, 1906), proportion (Angier, 1903; Davis, 1933; Haines & Davies, 1904), polygon complexity (Beebe-Center & Pratt, 1937), level of curvature (Lundholm, 1921), symmetry (Pierce, 1894), or uniformity of figure and arrangement (Otis, 1918) led to differences in perceived beauty or pleasantness.

Over a century of research in empirical aesthetics confirms that people generally prefer symmetry to asymmetry (Jacobsen & Höfel, 2001; Gartus & Leder, 2013; Pecchinenda, Bertamini, Makin, & Ruta, 2014), complexity to simplicity (Nadal, Munar, Marty, & Cela-Conde, 2010; Machado et al., 2015), and curved to angular contours (Palumbo, Ruta, & Bertamini, 2015; Bertamini, Palumbo, Gheorghes, & Galatsidas, 2016; Gómez-Puerto et al., 2018; Corradi, Chuquichambi, Barrada, Clemente, & Nadal, 2020). Most of these preferences seem to transcend cultural boundaries (Che, Sun, Gallardo, & Nadal, 2018), and even species boundaries (Munar, Gómez-Puerto, Call, & Nadal, 2015), but they also seem to be modulated by personality, familiarity, expertise, and experimental task (Marin & Leder, 2018; Palumbo & Bertamini, 2016; Cotter, Silvia, Bertamini, Palumbo, & Vartanian, 2017; Weichselbaum, Leder, & Ansorge, 2018; Leder et al., 2018; Pecchinenda et al., 2014; Vartanian et al., 2019).

It was noted early on, however, that these general relations between stimulus features and aesthetic responses coexisted with important individual differences. Clark, Quackenbush, and Washburn (1913) used the concept *affective sensitiveness* to distinguish between people who strongly tended to like and dislike materials of different sorts, including tones, colors, and speech sounds, from people who were relatively indifferent to those materials (Babbitt, Woods, & Washburn, 1915). Washburn, Hat, and Holt (1923) showed that poets were more affectively sensitive than science students, meaning that affective sensitiveness was related to experience and expertise in art and aesthetics. Clark and colleagues' (1913) concept of affective sensitiveness captures differences in the magnitude of people's response to visual and auditory stimuli, but their materials were not designed to include increments along a specific dimension. Thus, affective sensitiveness does not relate the increase in response to the increase in stimulation. It is a measure of how responsive people are to certain materials, but not a measure of how responsive they are to variations in specific features of those materials.

Corradi et al. (2019, 2020) have recently proposed a conceptualization of *aesthetic sensitivity* intended to capture differences among people in the extent to which aesthetic appreciation depends on, and is explained by, variations in specific stimulus features. It is a measure of the degree to which variations in a given feature influence someone's aesthetic appreciation. In this sense, someone is aesthetically sensitive to complexity, for instance, if their aesthetic appreciation of an object varies as a function of its complexity: e.g., they like complex designs more than simple ones, or vice versa. Conversely, someone is aesthetically insensitive to complexity if this feature is irrelevant to their aesthetic appreciation: their liking is indifferent to complexity.

In this regard, aesthetic sensitivity is not equivalent to perceptual sensitivity: it does not gauge whether participants can discriminate fine variations in complexity, for instance. It is also not a measure of receptiveness to artistry—to artful execution or artistic excellence. Aesthetics and art are, to some extent, overlapping fields, although not identical (Brown & Dissanayake, 2009; Pearce et al., 2016). According to Corradi and colleagues (2019, 2020), aesthetic sensitivity is the extent to which variations in a particular stimulus property lead to variations in an individual's liking for the stimulus.

Corradi and colleagues (2019, 2020) were not the first to put forward a definition of aesthetic sensitivity. For instance, Meier (1928) defined aesthetic sensitivity as "the ability to recognize compositional excellence in representative art-situations, or the ability to 'sense' quality (beauty?) in an aesthetic organization" (Meier, 1928, p. 185). Eysenck conceived aesthetic sensitivity as a distinct ability that (i) enabled some people to appreciate objective beauty better than others ("[this ability], independently of intelligence and personality, determines the degree of good or bad taste;" Eysenck, 1983, p. 231); (ii) explained performance on virtually all measures of artistic ability ("it covers a large number of, probably all, pictorial tests;" Eysenck, 1940, p. 100); and (iii) was immutable because it was biologically determined, innate ("[it] presumably [has] a genetic foundation in the structure of the nervous system;" Götz, Borisy, Lynn, & Eysenck, 1979, p. 801), and unalterable through experience ("[it] is independent of teaching, tradition, and other irrelevant associations;" Eysenck, 1940, p. 102). Parker (1978) defined musical aesthetic sensitivity as a person's biologically-based competence of making value judgments in agreement with a consensus of musical sophisticates on the appropriateness to society's aesthetic values. According to this notion, to demonstrate good taste, an individual must prefer what others had judged to be more beautiful.

Corradi and colleagues' (2019, 2020) conception of aesthetic sensitivity, which we apply to the music domain in this study, differs in several regards from previous conceptions of aesthetic sensitivity (Eysenck, 1940; Meier, 1928; Myszkowski, Çelik, & Storme 2018). First, unlike Eysenck's (1983) or Meier's (1928) notion, it does not rely on the assumption that aesthetic value is an attribute of objects: under our conception of aesthetic sensitivity, aesthetic value is a quality of the experience of objects. Second, unlike Götz et al.'s (1979) or Parker's (1978) notion, there is no external normative standard set by any authority: aesthetic sensitivity is the extent to which sensory features influence someone's valuation. Third, unlike Eysenck's (1940) or Meier's (1928) conception, aesthetic sensitivity need not be a unitary construct: people might be sensitive to some features but not others (Clark et al., 1913). Fourth, unlike Götz et al.'s (1979) or Parker's (1978) notion, aesthetic sensitivity need not be immutable: people's aesthetic sensitivity might be influenced by context, experience, expertise, and maybe even fatigue (Robbins, Smith, & Washburn, 1915).

Corradi and colleagues (2020) mapped out the variation inherent to their conception of aesthetic sensitivity defined in the previous paragraphs regarding balance, contour–curvature, symmetry, and complexity in the visual modality. Although, in general, balance was preferred to unbalance, curvature to angularity, symmetry to asymmetry, and complexity to simplicity, people differed considerably from each other in the extent to which they were aesthetically sensitive to each of those attributes. Whereas some people were insensitive to complexity, others consistently preferred complex designs, and others consistently preferred simple ones. The same was true for symmetry, balance, and contour–curvature. Additionally, Corradi and colleagues (2020) did not find strong evidence of relations among aesthetic sensitivities to these four attributes. This supports the notion of aesthetic sensitivity as a multidimensional construct: someone's liking can be strongly determined by one attribute but not another.

These findings raise the question of whether people also differ to such a great extent in their aesthetic sensitivity to attributes in other sensory modalities. As noted by Clemente et al. (2020), the aesthetic appreciation or valuation of music depends on many factors, such as cultural background, familiarity, experience, perceived complexity, or predictability (Brattico & Pearce, 2013; Koelsch, Vuust, & Friston, 2018; Pereira et al., 2011; Van den Bosch, Salimpoor, & Zatorre, 2013). People are not passive responders to music. Musical experiences are actively *constructed* by each individual relying on perceptual, cognitive, and affective processes that depend on knowledge, past experience, personal and cultural meaning, motivations, goals, and other individual and contextual circumstances.

Thus, the appreciation of music is a complex phenomenon that can, and must, be studied from a variety of perspectives using a variety of approaches. The question we ask here is whether people construct different preferences because, among many other things, they take into account different musical aspects to different extents: Might two people differ in their preference for a musical piece because, in constructing their preferences, one takes complexity into account and the other does not? If so, do they do so consistently? There is evidence suggesting that this is the case. For example, dissonance contributes to the perceived complexity of Western music, but people differ considerably in the extent to which they dislike dissonance (e.g., Plomp & Levelt, 1965; McDermott, Lehr, and Oxenham, 2010). There are also remarkable differences across cultures in the extent to which dissonance is disliked (e.g., McDermott, Schultz, Undurraga, & Godoy, 2016, McPherson et al., 2019). The musical context in which the stimuli are presented and the degree of Western tonal-functional enculturation seem to be key factors explaining variations in individual preference for consonance.

The present study had three main goals: First, we wished to apply Corradi and colleagues' (2020) conception of aesthetic sensitivity to music. Specifically, we wished to characterize musical aesthetic sensitivity to four attributes that figure prominently in the literature on visual aesthetics: balance, contour, symmetry, and complexity. There is some evidence for common effects of complexity on the appreciation of visual and musical materials (e.g., Marin, Lampatz, Wandl, & Leder, 2016; Marin & Leder, 2013) and there is also some evidence for individual differences in aesthetic sensitivity to complexity (Güçlütürk, Jacobs, and van Lier, 2016; Güçlütürk & van Lier, 2019; Marin & Leder, 2018). In the present paper, we aim to corroborate and generalize this work to balance, symmetry, and contour. Thus, we assessed aesthetic sensitivity to these four attributes through sets of stimulus features that define them. Based on Corradi et al.'s (2019, 2020) results, we hypothesized (i) significant effects of these attributes on liking, and (ii) substantial variation in the extent to which these attributes influence individuals' aesthetic valuation. Second, we examined the temporal stability of musical aesthetic sensitivities. Considering Corradi et al.'s (2020) findings in the visual domain, we hypothesized (iii) that people's aesthetic sensitivity to musical balance, contour, symmetry, and complexity are also stable in time. Third, we analyzed the relations among aesthetic sensitivities to probe whether the individual magnitude and direction of sensitivity are common across attributes and if such personal sensitivities converge into sensitivity profiles. Corradi and colleagues (2020) found no strong relations among visual aesthetic sensitivities to balance, contour, symmetry, and complexity. However, the extant literature does not allow us to form a particular hypothesis regarding the clustering of participants based on the pattern of their aesthetic sensitivities. Therefore, this analysis was conducted on an exploratory basis.

Method

Participants

Forty-eight self-reported non-musicians (39 women and nine men) aged 18–44 years (*M* = 21.560, *SD* = 5.845) and recruited at the university campus took part in the study. No participant had received musical education at a university level, and the mean duration of their formal education in music was 5.354 years (*SD* = 4.111). Before participation, all gave informed consent and reported normal or corrected-to-normal vision and hearing and no cognitive impairments. They were unaware of the study's purpose and treated under the local ethical guidelines and the Declaration of Helsinki. The study received approval from the Committee for Ethics in Research of the Balearic Islands (IB 3573/17 PI).

Materials

Clemente et al.'s (2020) MUST set of stimuli consists of 4-s monophonic piano-like motifs in C-Major that systematically vary in musical balance, contour, symmetry, and complexity. They were composed expressly for empirical studies and designed to combine experimental control and musical appeal. The design of the MUST stimuli is schematically depicted in Table A1 (Appendix A). In the present study, we used the MUST abridged stimulus set, which includes 24 motifs (plus four examples) for each of the four attributes. The abridged Balance subset includes 12 *balanced* and 12 *unbalanced* motifs; the abridged Contour subset includes 12 *smooth* and 12 *jagged* motifs; the abridged Symmetry subset includes 12 *symmetric* and 12 *asymmetric* motifs; the abridged Complexity subset includes 12 *simple* and 12 *complex* motifs (Figure 1). The stimuli were presented in WAV format using Open Sesame (Mathôt, Schreij, & Theeuwes, 2012).

The MUST (Clemente et al., 2020) also includes composite computational measures specific for the structural features characteristic of each subset: Balance and symmetry were defined by single composite measures of balance (BC1) and symmetry (SC1), respectively. Two components quantified the structural parameters of contour: one for melodic (pitch-related) contour (CC1) and the other for rhythmic contour (CC2). Likewise, two components quantified complexity: a measure of melodic complexity KC1 (event density and pitch-related entropy) and a measure of rhythmic complexity KC2 (duration entropy). Higher values correspond to greater unbalance, jaggedness, asymmetry, and complexity, respectively. We include a summary of the computational measures in Table A2 (Appendix A). The computational assessment showed that stimuli in each of the attributes' poles differ substantially in the corresponding defining features, and the behavioral assessment showed that people rate them as belonging to two opposite extremes (Clemente et al., 2020). The MUST set and computational measures are available at https://osf.io/bfxz7/.

Figure 1. Sample scores of auditory stimuli in each subset, all to be played in ♩ = 120 (i.e., quarter note at 120 bpm).

Procedure

Participants undertook the experimental tasks in the laboratory. They were first welcomed and briefed about the entire procedure. Each participant was then asked to enter one of the individual sound-attenuated testing cabins, all of which have the same computers, software, adequate light conditions, and headphone sets. After providing verbal and onscreen instructions, each subset was presented in a separate block, which consisted of four practice trials (two illustrative of each pole) and the 24 task stimuli. All stimuli were presented through headphones. The order of the blocks was counterbalanced between participants, and the order of presentation within each block was randomized for each participant. During the practice trials, participants adjusted their headsets to personal comfort levels, which remained unmodified for the whole experiment. After the experimenter had made sure participants understood the task and all doubts had been resolved, participants listened to and rated the task stimuli.

Participants rated how much they liked each of the 24 musical motifs in each subset twice: in the test and retest experimental sessions two weeks apart. They rated each motif using a keyboard on a 5-point Likert scale anchored by *not at all* (1) to *very much* (5). They were explicitly requested to base their responses on the subjective internal feelings of pleasure, interest, enjoyment, or desirability evoked, inspired, or provoked by the music. They were allowed to take breaks between blocks, and to replay a stimulus before rating it if they so desired. A brief questionnaire (included as Appendix B) followed the fourth block in the test phase and asked about demographics (i.e., age, sex, and education) and formal musical education (i.e., highest degree attained, onset, and duration). Finally, participants were debriefed and thanked for their time and participation. Test and retest sessions lasted about 30 and 15 min, respectively.

Data Analysis

All analyses were performed within the R environment for statistical computing, v. 3.6.0 (R Core Team, 2018). In the course of conducting the analyses, we found that two pairs of stimuli belonging to the Symmetry and Complexity abridged subsets were duplicated, that is to say, the same stimulus had erroneously been included in the symmetry and complexity subsets: S4 = K8 and S5 = K9. We, therefore, decided to exclude them from the analyses, leaving us with 12 *balanced* – 12 *unbalanced,* 12 *smooth* – 12 *jagged*, 10 *symmetric* – 12 *asymmetric*, and 10 *simple* – 12 *complex*.

Musical aesthetic sensitivity.

We performed four linear mixed-effects analyses (Hox, Moerbeek, & van de Schoot, 2010; Snijders & Bosker, 2012) to assess the effect of the main predictors on participants' liking judgments in the test phase. This method accounts simultaneously for the between-subjects and within-subjects effects of the independent variables (Baayen, Davidson, & Bates, 2008), and models random error at all levels of analysis simultaneously, relying on maximum-likelihood procedures to estimate coefficients. Therefore, it provides the most accurate analysis of hierarchically structured data in which, as is the case here, responses to stimuli are dependent on, or nested within, individual participants (Nezlek, 2001). Linear mixed-effects models provide other additional advantages, such as meaningful estimates of subject- and group-level variance components, unbiased handling of outliers, and ability to handle incomplete and unbalanced data and to accommodate continuous and categorical predictors (Judd, Westfall, & Kenny, 2012). Importantly, they allow deriving conclusions that generalize to other participants besides those providing the data (Judd, Westfall, & Kenny, 2017; Nezlek, 2001). They are especially well suited to the purposes of the current study because they provide estimates for both group-level effects, which can be compared with those of previous studies, and participant-level effects, which constitute our measure of individual aesthetic sensitivity (as in Corradi et al., 2020).

The models were designed to reflect the effect of the features varied on participants' liking. Thus, we modeled participants' responses as a function of Clemente et al.'s (2020) MUST composite measures: liking for musical balance predicted by BC1; liking for musical contour with predictors CC1, CC2, and their interaction (CC1*CC2); liking for musical symmetry predicted by SC1; and liking for musical complexity with predictors KC1, KC2, and their interaction (KC1*KC2). All predictors were mean-centered. The models included the respective composite measures as fixed effects. The models of liking for contour and complexity also included the interaction between melodic (CC1) and rhythmic (CC2) contour, and between melodic (KC1) and rhythmic (KC2) complexity, respectively, as fixed effects. The four resulting models also included the slope for each feature and their interactions (when appropriate) as random effects within participants, and random intercepts within stimuli, following Barr, Levy, Scheepers, and Tily's (2013) recommendation to model the maximal random-effects structure justified by the experimental design. To assess the effects of familiarity, we also ran the models including repeated listening as a predictor.

Although the mixed-effects models produce group estimates, the main goal of this study was to understand individual differences in the extent to which these four attributes influence people's liking. In linear mixed-effects models, this corresponds to the individual slopes. Thus, we defined participants' aesthetic sensitivity to each composite measure as the individual slope estimated from the models' random-effect structure. Therefore, after running each model, we extracted each participant's slopes. We used these values to describe individual aesthetic sensitivities to musical balance, contour, symmetry, and complexity, and to explore relations among them. We investigated the distribution of slopes for each predictor and used Shapiro–Wilk tests to assess their normality.

We performed these analyses using the *lmer()* function of the 'lme4' package (Bates, Maechler, Bolker, Walker, 2015) fitted with REML estimation. The 'lmerTest' package (Kuznetsova, Brockho, & Christensen, 2012) was used to estimate the *p*-values for the *t*-tests based on the Satterthwaite approximation for degrees of freedom, which has been shown to produce acceptable Type-I error rates (Luke, 2017).

Test–retest differences.

The estimation of participants' aesthetic sensitivity to the four attributes was done exactly as described above for the test and retest data. Thus, for each participant, we had two measures of aesthetic sensitivity for each of the four attributes taken two weeks apart. We were, therefore, able to determine the test–retest reliability of aesthetic sensitivity to each feature. The analysis was based on Bland and Altman's (1986) method and the smallest real difference (SRD; Vaz, Falkmer, Passmore, Parsons, & Andreou, 2013). Like other methods to estimate test-retest reliability, Bland and Altman's (2003) method quantifies variation between repeated measurements. The advantages of their graphical method are that it is robust to the data variability and can detect systematic biases in the differences between two repeated measurements. This method establishes statistical boundaries for detecting a test–retest difference, namely the threshold for change or minimal detectable true change (Vaz et al., 2013). The limits of agreement are set at 1.96 times the standard deviation above and below this difference, defining the smallest real difference (SRD; Vaz et al., 2013). When this interval contains the value 0, the test–retest difference can be attributed to error (Beckerman et al., 2001). Otherwise, it can be attributed to some form of systematic bias. Bland and Altman's (1986) graphs plot the test–retest differences against the average and, thus, allow identifying cases where changes indicate a shift in aesthetic sensitivity. We used the R package 'BlandAltmanLeh' (Lehnert, 2015).

Relations among aesthetic sensitivities.

To investigate how aesthetic sensitivities were related within individuals, we first inspected the correlations between individual slopes. Second, we wished to know whether combinations of sensitivities characterized the liking distributions, and if such combinations were finite and followed any pattern, so we performed a cluster analysis.

Cluster analysis or clustering is a common procedure in exploratory data mining and a standard for statistical data analysis. It is used in many fields, including machine learning, pattern recognition, image analysis, or music information retrieval. Clustering consists in grouping a set of objects in such a way that objects in the same group or cluster are more similar in a particular aspect to each other than to those in other groups or clusters. Therefore, it is an iterative process of knowledge discovery or interactive multi-objective optimization. Cluster analysis can be achieved by various algorithms that differ significantly in their definition of clusters and how to find them efficiently. The appropriate clustering algorithm and parameter settings (e.g., distance function to use, density threshold, number of expected clusters) depend on the particular data set and intended use of the results. The most prominent examples of clustering algorithms include hierarchical clustering, centroid-based clustering (such as the popular *k*-means, in which the number of clusters is predetermined to *k*), distribution-based clustering (e.g., Gaussian mixture models), density-based clustering, or grid-based clustering.

We applied Gaussian finite-mixture models fitted via the expectation-maximization (EM) algorithm for model-based clustering, classification, and density estimation, including Bayesian regularization, dimension reduction for visualization, and resampling-based inference. We chose it over partitioning methods because the data points were not necessarily assumed to belong to only one cluster, and this method allows the number of clusters to emerge from the data (Melnykov & Maitra, 2010). This analysis was applied to both test and retest data to ascertain whether the clustering structure held over time. We used the R package 'mclust' (Scrucca, Fop, Murphy, Brendan, & Raftery, 2016).

Impact of Demographics.

As a matter of routine, we examined the correlations between aesthetic sensitivities and demographic variables. To test whether, and to what extent, these traits predicted the clustering, we also performed a multiple linear regression analysis. As we did not have any specific hypothesis nor expect these variables would exert any effect on liking or on the configuration of the clustering, these analyses were deemed exploratory.

Results

Musical Aesthetic Sensitivity

We modeled liking judgments for stimuli in each subset as a function of the corresponding parameters of variation, as assessed by the MUST composite measures. This made a total of four linear mixed-effects models in the test phase and four more in the retest phase. In this section, we report the results of the analyses corresponding to the test phase. For each feature, we first report the group-level trends and then descriptive statistics that characterize the distributions of individual aesthetic sensitivities.

Balance.

The analysis of the balance model showed that, overall, participants found the stimuli appealing (intercept: ß = 3.240, $t_{(42.827)} = 28.537$, $p < .001$), and that they liked the balanced motifs more than the unbalanced ones: ß = -0.276, $t_{(23.845)} = -3.057$, $p = .005$ (Figure 2A). The individual slopes of liking for balance ranged from -0.482 to -0.126, indicating different degrees in the extent to which participants liked the balanced motifs, with $M = -0.276$, $SD = 0.074$. The Shapiro–Wilk normality test showed that the slopes of liking for balance were normally distributed ($W = 0.985$, $p = 0.774$) (Figure 3A).

Contour.

The analysis of the contour model revealed that, overall, liking judgments were positive (intercept: ß = 3.450, $t_{(52.084)}$ = 34.649, p < .001) and predicted by the two composite contour measures separately but not by their interaction: ß = -0.023, $t_{(23.825)}$ = -0.409, p = .686. In general, participants liked more melodic jaggedness (CC1; ß = 0.224, $t_{(36.306)}$ = 2.698, p = .011; Figure 2B) and rhythmic smoothness (CC2; ß = -0.185, $t_{(21.050)}$ = -2.173, p = 0.041; Figure 2C). The individual slopes of liking for CC1 ranged from -0.392, indicating a greater liking for melodic smoothness, to 0.829, indicating a greater liking for melodic jaggedness, with M = 0.224, SD = 0.283. The individual slopes of liking for CC2 ranged from -0.365 to -0.059, indicating a greater liking for rhythmic smoothness, with M = -0.185, SD = 0.077. The Shapiro–Wilk normality test indicated that the individual sensitivities to contour were normally distributed for CC1 (W = 0.957, p = 0.079; Figure 3B) and CC2 (W = 0.959, p = 0.092; Figure 3C).

Symmetry.

The analysis of the symmetry model showed that, overall, participants rated the stimuli positively (intercept: ß = 3.549, $t_{(44.777)}$ = 33.115, p < .001) and liked the asymmetric more than the symmetric motifs: ß = 0.198, $t_{(20.316)}$ = 2.506, p = .021 (Figure 2D). The individual slopes of liking for symmetry ranged from 0.160 to 0.233, indicating a greater liking for asymmetric motifs, with M = 0.198, SD = 0.018. The liking slopes for symmetry were normally distributed according to the Shapiro–Wilk normality test (W = 0.967, p = .187; Figure 3D).

Complexity.

The analysis of the model of liking for complexity unveiled that, overall, participants' liking was positive (intercept: ß = 3.211, $t_{(51.427)}$ = 36.436, p < .001) and increased with melodic complexity (KC1), the only significant effect: ß = 0.466, $t_{(45.415)}$ = 6.236, p < .001

(Figure 2E). The effect of rhythmic complexity (KC2) was not significant: ß = -0.080, $t_{(21.427)}$ = -1.376, p = .183 (Figure 2F). The effect of the interaction between KC1 and KC2 was also not significant: ß = -0.022, $t_{(19.877)}$ = -0.494, p = .627. The estimated slopes for KC1 ranged from -0.771, indicating a greater liking for melodically simple motifs, to 0.981, indicating a greater liking for melodically complex motifs, with M = 0.466, SD = 0.343. According to the Shapiro–Wilk normality test, the slopes of liking for melodic complexity (KC1) were not normally distributed (W = 0.905, p < .001), with skew = -1.150, and kurtosis = 1.615 (Figure 3E). The slopes of liking for KC2 ranged from -0.315, indicating a greater liking for rhythmically simple motifs, to 0.134, indicating a greater liking for rhythmically complex motifs, with M = -0.080, SD = 0.099. The slopes of liking for KC2 were normally distributed according to the Shapiro–Wilk normality test (W = 0.977, p = .457; Figure 3F).

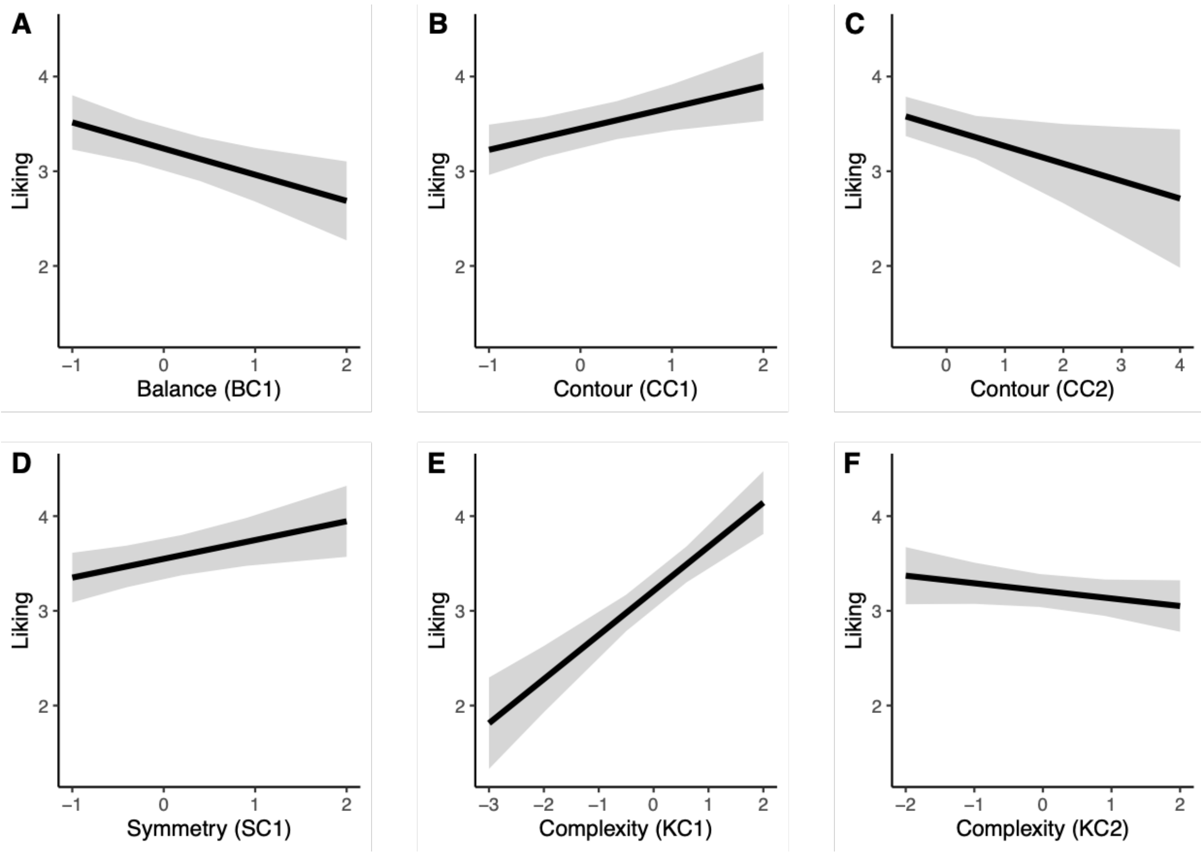

Figure 2. Main effects of balance (A), contour (B & C), symmetry (D), and complexity (E & F) on participants' liking. High values on the computational measures mean unbalanced (BC1), melodically (CC1) and rhythmically (CC2) jagged, asymmetric (SC1), and melodically (KC1) and rhythmically (KC2) complex, respectively.

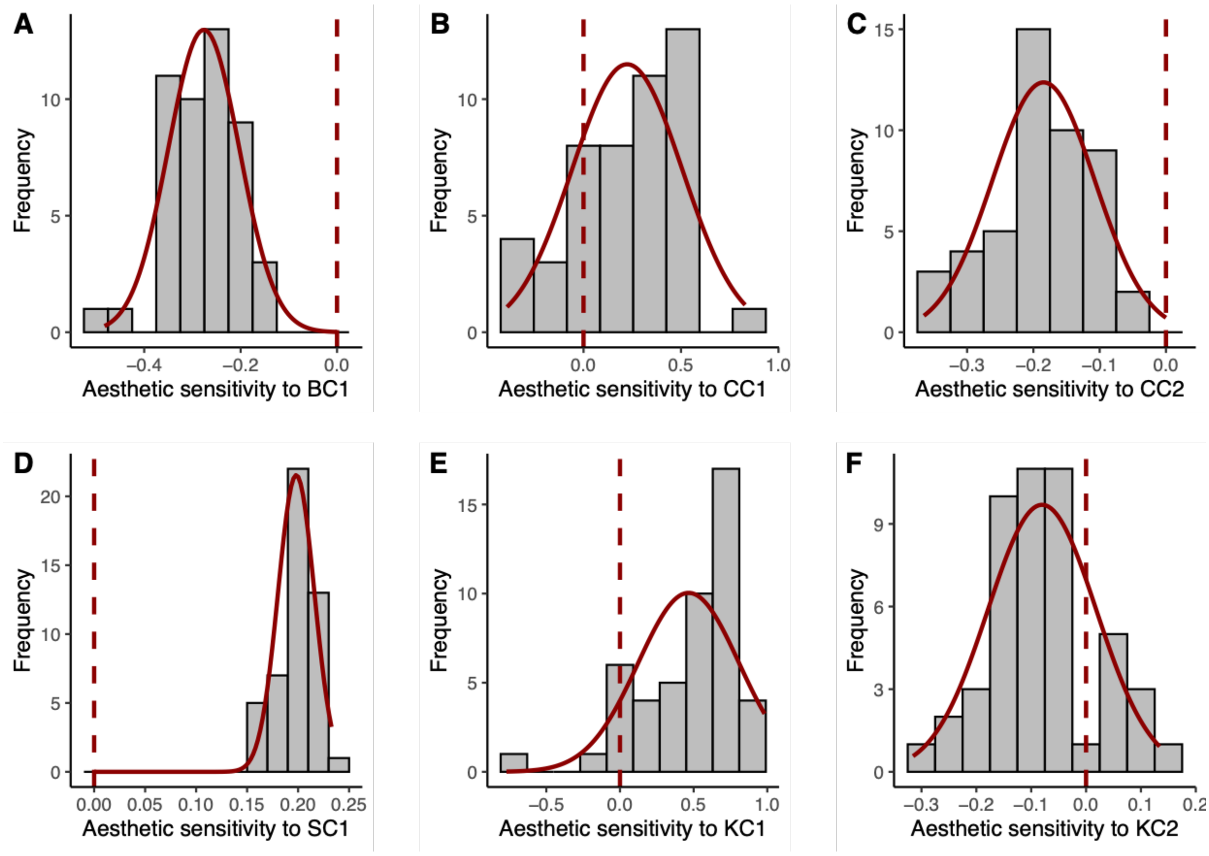

Figure 3. Individual aesthetic sensitivity to musical attributes: Histograms of individual slopes of liking for balance (A), contour (B & C), symmetry (D), and complexity (E & F). Vertical dashed lines correspond to a slope of 0, meaning absolute indifference or insensitivity toward each attribute concerning liking judgments. Positive slopes indicate a higher liking for unbalanced (BC1), melodically (CC1) and rhythmically (CC2) jagged, asymmetric (SC1), and melodically (KC1) and rhythmically (KC2) complex motifs. Negative slopes indicate a higher liking for balanced (BC1), melodically (CC1) and rhythmically (CC2) smooth, symmetric (SC1), and melodically (KC1) and rhythmically (KC2) simple motifs. Normal curves are overlaid in dark red. Only the individual slopes of liking for symmetry (SC1) and melodic complexity (KC1) were not normally distributed.

Test–retest Differences

We analyzed the retest data following the same procedure as reported for the test phase. Then, we examined test–retest changes in individual participants' aesthetic sensitivity to each feature applying Bland–Altman's graphic method and the smallest real difference (SRD).

Table 1 shows the results of the analysis based on the smallest real difference (SRD), and Figure 4 displays the corresponding Bland–Altman graphs. These analyses revealed that whereas the test–retest differences in the assessment of aesthetic sensitivity to melodic contour (CC1) and melodic (KC1) and rhythmic (KC2) complexity can be attributed to random error, this is not the case for aesthetic sensitivity to musical balance (BC1), symmetry (SC1), and rhythmic contour (CC2), where there is a systematic bias in the differences. Participants were more sensitive to rhythmic contour (CC2) and symmetry (SC1) in the retest phase. In the case of balance (BC1), participants were less sensitive in the retest phase.

The Bland-Altman analyses showed that these systematic biases owed to changes in the aesthetic sensitivity of few participants. In total, only 17 out of 288 individual sensitivities to the four aesthetic attributes (6%) exceeded the SRD. In the case of musical balance (BC1), three participants (6%) showed lower sensitivity in the retest. Regarding rhythmic contour (CC2), five participants exceeded the SRD: four (8%) were more sensitive in the retest, and one (2%) in the test phase. As for symmetry (SC1), two participants (4%) were more and one was less (2%) sensitive in the retest phase. One participant exceeded the SRD for three of the features, and another participant, for two of them. No participant exceeded the SRD for more than three features.

Table 1. *Test–retest Differences: Bland–Altman Analysis*

Feature	**Mean Retest – Test Difference**	**95% CI**		**Smallest Real Difference (SRD)**
		Lower	Upper	
BC1	-0.045	-0.073	-0.017	0.189
CC1	-0.018	-0.082	0.045	0.426
CC2	0.044	0.019	0.069	0.167
SC1	0.026	0.006	0.046	0.135
KC1	-0.030	-0.104	0.043	0.496
KC2	-0.009	-0.073	0.017	0.190

Note. Mean difference and smallest real difference: measures of test–retest reliability for aesthetic sensitivity to musical balance (BC1), melodic (CC1), and rhythmic (CC2) contour, musical symmetry (SC1), and melodic (KC1) and rhythmic (KC2) complexity.

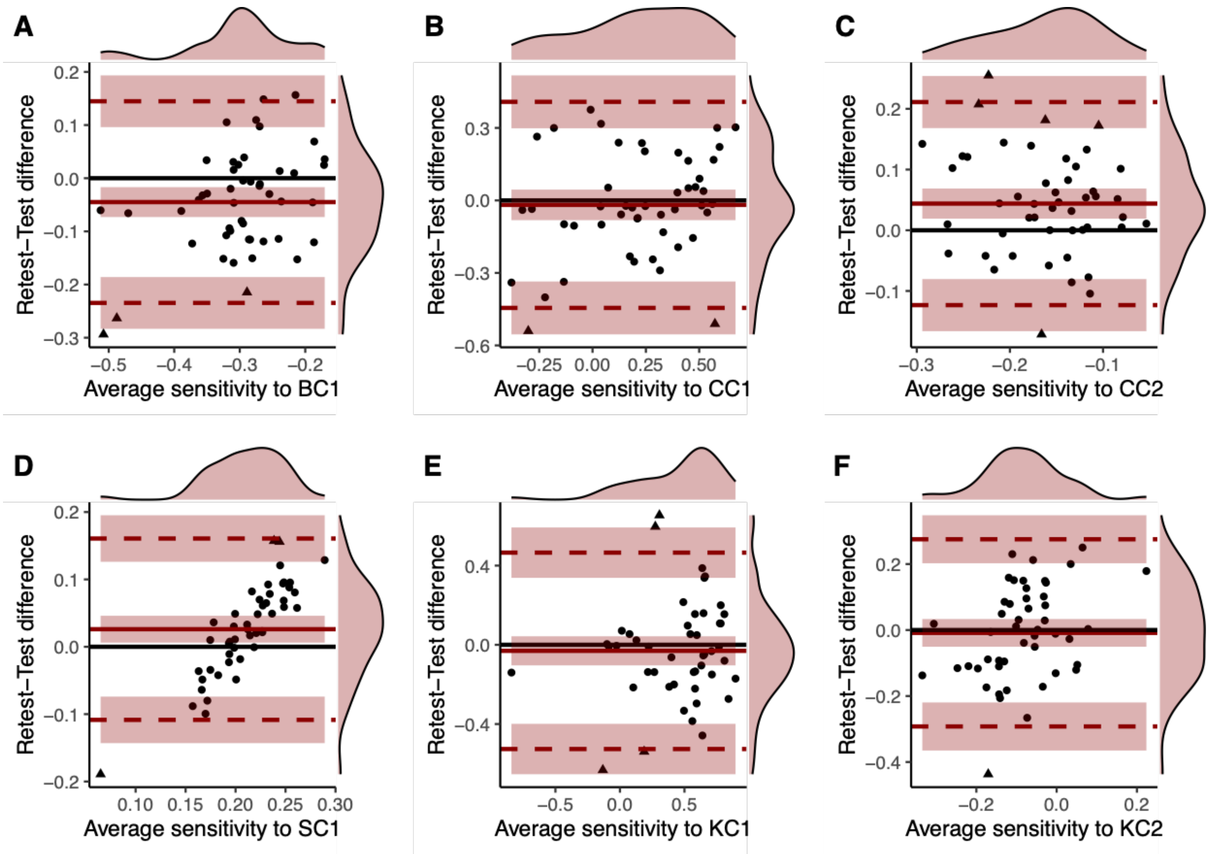

Figure 4. Bland–Altman graphs for the test–retest reliability of aesthetic sensitivity to balance (BC1; A), melodic contour (CC1; B) rhythmic contour (CC2; C), symmetry (SC1; D), melodic complexity (KC1; E) and rhythmic complexity (KC2; F). Horizontal black lines represent no test–retest change. Horizontal continuous red lines indicate the mean test–retest difference. Horizontal dashed lines mark the lower and higher limits of agreement. Horizontal ribbons comprise 95% CI. Circles correspond to participants whose test–retest difference is smaller than the smallest real difference (SRD). Triangles correspond to participants whose test–retest difference is larger than the SRD.

Relations Between Aesthetic Sensitivities

As a first approach to the relationships between sensitivities within participants, and given that not all distributions of aesthetic sensitivities were normal according to the Shapiro–Wilk tests, we calculated Spearman correlations between aesthetic sensitivities in the test phase (Table 2). These indicate that people who like melodically jagged contours also tend to prefer rhythmically jagged, less balanced, and more asymmetric and complex motifs;

and people who like more melodically complex music also tend to like more asymmetric and jagged motifs. Aesthetic sensitivities to melodic and rhythmic contour show particularly strong correlations, suggesting that people who like more jagged contours do so for both pitch-related and rhythmic aspects of musical contour. Also, aesthetic sensitivity to melodic contour is moderately correlated with that for balance, suggesting that people who like more jagged melodies also tend to like less balanced ones. Interestingly, preference for either form of musical jaggedness shows moderate to strong correlations with melodic complexity, whereas aesthetic sensitivity to these three structural properties present weaker (moderate) correlations with aesthetic sensitivity to symmetry.

Table 2. *Spearman Correlations among Aesthetic Sensitivities in the Test Phase*

	BC1	**CC1**	**CC2**	**SC1**	**KC1**
CC1	0.553****				
CC2	0.442**	0.828****			
SC1	0.157	0.382**	0.299*		
KC1	0.232	0.613****	0.641****	0.333*	
KC2	0.175	0.033	0.086	-0.057	-0.181

Note. Pairwise Spearman correlations among aesthetic sensitivities to musical balance (BC1), melodic contour (CC1) rhythmic contour (CC2), musical symmetry (SC1), melodic complexity (KC1) and rhythmic complexity (KC2). Significance codes: **** $p < .0001$; *** $p < .001$; ** $p < .01$; * $p < .05$

In addition to pairwise correlations, clustering provides a comprehensive picture of the multiple relationships of individual aesthetic sensitivities within and between individuals. In the test phase, the Gaussian finite-mixture model fitted by the EM algorithm revealed the existence of two clusters (log-likelihood $_{(48)}$ = -361.341, BIC = -780.750, ICL = -785.724).

Cluster 1T (for *test*) included 21 participants who, overall, liked more balanced, smooth, symmetric, and simple motifs. Cluster 2T included 27 individuals who generally liked more unbalanced, jagged, asymmetric, and complex motifs (Table 3).

Table 3. *Model-based Clustering of Individual Slopes of Liking Ratings in the Test Phase*

	BC1	**CC1**	**CC2**	**SC1**	**KC1**	**KC2**
Cluster 1T (21)	-0.631	-0.858	-0.770	-0.460	-0.618	-0.096
Cluster 2T (27)	0.516	0.702	0.630	0.376	0.505	0.079

Note. Estimates of aesthetic sensitivity for each cluster. Positive values indicate a greater liking for unbalanced (BC1), melodically (CC1) and rhythmically (CC2) jagged, asymmetric (SC1), and melodically (KC1) and rhythmically (KC2) complex motifs. Negative values indicate a greater liking for balanced (BC1), melodically (CC1) and rhythmically (CC2) smooth, symmetric (SC1), and melodically (KC1) and rhythmically (KC2) simple motifs.

In the retest phase, the Gaussian finite-mixture model fitted by the EM algorithm revealed the existence of three clusters (log-likelihood $_{(48)}$ = -355.645, BIC = -800.328, ICL = -807.542). Cluster 1R (for *retest*) was made up of 16 participants who, overall, liked more balanced, smooth, asymmetric, and melodically simple motifs. Cluster 2R included 19 individuals who generally liked more unbalanced, jagged, symmetric, melodically complex, and rhythmically simple motifs. Cluster 3R was made up of 13 participants who tended to like more balanced, jagged, asymmetric, and complex motifs (Table 4). Cluster 1R corresponds for the most part to Cluster 1T, whereas Clusters 2R and 3R suggest a more detailed picture for the trends characterizing Cluster 2T. Most participants (14) belonging to Cluster 1T in the test phase integrated Cluster 1R, while five fell into Cluster 2R and two into

Cluster 3R at retest. Only two participants in Cluster 2T shifted to Cluster 1R, while the rest were distributed into Clusters 2R and 3R.

Table 4. *Model-based Clustering of Individual Slopes of Liking Ratings in the Retest Phase*

	BC1	**CC1**	**CC2**	**SC1**	**KC1**	**KC2**
Cluster 1R (16)	-0.612	-0.950	-0.864	0.173	-0.822	0.081
Cluster 2R (19)	0.621	0.362	0.182	-0.527	0.272	-0.392
Cluster 3R (13)	-0.116	0.768	0.928	0.586	0.729	0.502

Note. Estimates of aesthetic sensitivity for each cluster. Positive values indicate a higher liking for unbalanced (BC1), melodically (CC1) and rhythmically (CC2) jagged, asymmetric (SC1), and melodically (KC1) and rhythmically (KC2) complex motifs. Negative values indicate a higher liking for balanced (BC1), melodically (CC1) and rhythmically (CC2) smooth, symmetric (SC1), and melodically (KC1) and rhythmically (KC2) simple motifs.

Impact of Demographics

We examined the extent to which individual aesthetic sensitivities and cluster allocation were influenced by demographic variables. Most participants in this study had only studied music at primary and secondary school, and the mean duration of their formal education in music was five years (see Participants). We found no significant associations between liking judgments and age, sex, highest general academic degree attained, highest music degree obtained, nor onset or duration of formal education in music (Table 5).

Table 5. *Spearman Correlations among Numeric Demographic Variables and Aesthetic Sensitivities in the Test Phase*

	BC1	CC1	CC2	SC1	KC1	KC2
Age	-0.067	0.150	0.124	0.163	0.084	0.153
Education	0.154	-0.089	-0.099	-0.027	-0.190	0.234
Musical education	-0.036	-0.028	0.081	-0.069	-0.012	-0.036
Musical studies duration	-0.068	0.033	0.062	0.028	-0.012	-0.079
Musical studies onset	0.148	0.194	0.134	0.081	0.234	-0.219

Note. Pairwise Spearman correlations among demographic variables and aesthetic sensitivities to musical balance (BC1), melodic contour (CC1) rhythmic contour (CC2), musical symmetry (SC1), melodic complexity (KC1) and rhythmic complexity (KC2). Significance codes: **** $p < .0001$; *** $p < .001$; ** $p < .01$; * $p < .05$

Likewise, the multiple linear regression analysis of aesthetic sensitivities revealed no significant influence of the demographic variables on whether participants were allocated to one cluster or another (Table 6).

Table 6. *Multiple Linear Regressions for the Impact of Demographics on Cluster Allocation*

	Estimate	Std. Error	*t*-value	*Pr*(>|t|)
(Intercept)	1.001	0.420	2.385	0.022*
Age	-0.021	0.014	-1.489	0.144
Gender (woman)	-0.078	0.192	-0.408	0.685
Education	0.052	0.120	0.436	0.665
Musical education	-0.101	0.194	-0.521	0.605
Musical studies duration	-0.004	0.017	-0.227	0.821
Musical education	-0.019	0.021	-0.886	0.381

Note. Impact of demographic variables on individual loadings into clusters according to musical aesthetic sensitivities. Significance codes: **** $p < .0001$; *** $p < .001$; ** $p < .01$; * $p < .05$

Repeated Listening

Overall, only 4.2% of responses involved repeated listening (i.e., repeating the stimulus before rating it): 5.2% in the Balance subset, 3.8% in the Contour subset, 4.3% in the Symmetry subset, and 3.7% in the Complexity subset. To examine the impact of familiarity on liking, we reran the linear mixed-effect models for the test data including repeated listening as a predictor. We found no significant effects of stimulus repetition on liking ratings (all *p*s > .050). For the contour subset, the model was unusable, as it failed to converge with one negative eigenvalue (-0.17): ß = -0.348, $t_{(51.137)} = -2.009$, $p = .050$.

Discussion

Empirical aesthetics has traditionally focused on simple and general laws governing the relation between sensory features and appreciation. In this line, research shows that people generally prefer symmetry to asymmetry (Jacobsen & Höfel, 2001; Gartus & Leder, 2013; Pecchinenda et al., 2014), complexity to simplicity (Nadal et al., 2010; Machado et al., 2015), and curved to angular contours (Palumbo et al., 2015; Bertamini et al., 2016; Gómez-Puerto et al., 2018; Corradi et al., 2020).

However, general trends do not imply uniformity. Research shows, in fact, that people differ remarkably in the way they respond to symmetry (Leder et al., 2018), complexity (Chmiel & Schubert, 2018), regularity (Friedenberg, 2018), and curved contours (Corradi et al., 2019). Such differences illustrate how inadequate the notion of simple and general laws linking sensory features and hedonic valuation is (Skov & Nadal, 2020a). Aesthetic appreciation is shaped by context, cultural and personal meaning, familiarity and past experience, expertise, anticipation and expectations, as well as current mood and emotions, and bodily states (Skov & Nadal, 2020b). Previous work has shown that it is also shaped by aesthetic sensitivity: people differ in their hedonic valuation of visual objects because they consistently differ in the extent to which they rely on certain attributes (Corradi et al., 2019; 2020).

The overarching goal of the study presented here was to extend our research on aesthetic sensitivity to music, asking whether people differ in their preference for musical motifs because they take into account different attributes to different degrees. To facilitate comparison with studies in other modalities such as Corradi and colleagues' (2020), we used a set of musical motifs that enable the experimental control and computational quantification of balance, contour, symmetry, and complexity in the auditory domain while preserving musical appeal (Clemente et al., 2020). We had three specific goals: first, to characterize musical aesthetic sensitivity to balance, contour, symmetry, and complexity; second, to determine whether people's aesthetic sensitivity to these attributes is stable in time; and third, to ascertain whether there are common patterns of aesthetic sensitivities that lead people to fall into defined profiles.

The intercepts for all models were above the mid-score (i.e., ßs > 3.000), which indicates that participants found the stimuli generally appealing. The results also revealed general liking trends: overall, balanced, melodically jagged, rhythmically smooth, asymmetric, and melodically complex motifs were liked more than unbalanced, smooth, symmetric, and simple. Melodic complexity (KC1) was a much stronger predictor of perceived musical complexity than rhythmic complexity (KC2) in Clemente et al. (2020). Thus, it is not surprising that the contribution to liking judgments of melodic complexity was also greater than that of its rhythmic counterpart, which did not even reach statistical significance in the present study. Considering our results with music together with those of Corradi et al. (2020) with visual designs, liking for both music and images seems to increase with balance and complexity, whereas the trends for contour and symmetry differ between sensory domains. Further research addressing aesthetic sensitivity across modalities within participants will elucidate the implications of these similarities and differences.

Beyond these general trends, our results confirmed the hypothesized considerable differences among participants in the extent to which musical features influenced liking. These findings are in line with those of Güçlütürk et al. (2016), Güçlütürk and van Lier (2019), and Marin and Leder (2018), highlighting the importance of understanding individual differences that coexist with general trends. The estimated individual aesthetic sensitivities for musical balance (-0.482, -0.126), symmetry (0.160, 0.233), and rhythmic contour (-0.365, -0.059) varied within one pole, pointing to a consistent tendency across participants. In contrast, liking for melodic contour (-0.392, 0.829) and for melodic (0.771, 0.981) and rhythmic (-0.315, 0.134) complexity varied widely, including people either insensitive or very sensitive to these features, strongly and consistently preferring either extreme. These outcomes concur with prior findings in the visual domain (Corradi et al., 2019, 2020): also in music, a substantial proportion of the variance is accounted for by differences between individuals in the influence that such features exert upon aesthetic judgments. Hence, our study adds to mounting evidence for caution when interpreting general trends in liking and preference (Corradi et al., 2019; Güçlütürk et al., 2016; Güçlütürk & van Lier, 2019).

The results also confirmed our hypothesis that aesthetic sensitivity to musical attributes is stable in time: according to the Bland–Altman analysis, the vast majority of participants were consistent in their judgments at test and retest. The average differences in liking for balance, rhythmic contour, and symmetry were driven by a small number of participants more sensitive to rhythmic contour and symmetry, and less to balance at retest. These outcomes are comparable to those of Corradi et al. (2020) in the visual domain: they found that most participants were consistent between test and retest, and that systematic differences in aesthetic sensitivities to visual symmetry and complexity were attributed to very few participants. In both Corradi et al.'s (2020) and our study, participants showed higher sensitivity to symmetry in the retest phase, suggesting either that sensitivity to this attribute may be especially susceptible to learning, or that it is just unstable in both domains. However, the test–retest change observed for complexity differed slightly between domains: whereas sensitivity to visual complexity decreased at retest, aesthetic sensitivities to melodic and rhythmic complexity were stable.

We allowed the participants to repeat the motifs because the dimensions along which the stimuli varied might not have been graspable at first hearing, in the same way as the eyes move back and forth, reinspecting an image before assessing it. In hindsight, our precaution turned out to be unnecessary. The results suggest, first, that repetition seldom occurred, and second, that repeating the motif did not affect liking.

We found multiple significant correlations between different aesthetic sensitivities to structural features in our musical stimuli. We believe that the correlated sensitivities reflect underlying differences in participants' preference for informational predictability: whereas some people seem to prefer higher uncertainty in different forms (such as larger number of notes or interval amplitude), others seem to prefer higher predictability in different forms (such as recurrent sound patterns, and smooth profiles).

Even if preliminary, the clustering revealed that although people differ in the extent to which musical features influence their liking, there is a certain regularity: people clustered together into two groups based on their aesthetic sensitivities to the musical features we examined. Individuals falling into the first aesthetic sensitivity profile tended to like more balanced, smooth, symmetric, and simple music. Conversely, the second aesthetic sensitivity profile covered the largest number of participants and was characterized by a tendency to like more unbalanced, jagged, asymmetric, and complex music. These results resemble those of Güçlütürk and van Lier (2019) on musical complexity. The averaged strengths of these preferences vary within and between clusters: the estimated preferences of the first cluster members are slightly more extreme than those of the second one. In other words, aesthetic sensitivity appears to be somewhat higher for the first than the second profile.

The basic structure of the clustering was retained in the retest, although in a more detailed manner: Clusters 1T and 1R correspond to the first aesthetic profile, and the second profile represented by Cluster 2T is distributed into Clusters 2R and 3R. Average ratings were more extreme in Cluster 3R, showing stronger preferences for jaggedness, asymmetry, and complexity, and even reverting the tendency for balance. In contrast, Cluster 2R showed a stronger preference for balance, milder preferences for jaggedness and melodic complexity than Cluster 2T, and preferences for symmetry and rhythmic simplicity failed into this cluster instead of Cluster 1R. The shifts in the estimates are due to few participants swapping clusters from test to retest. Overall, the influence of rhythmic contour was not significant and tended to indifference at test. However, it revealed more pronounced at retest despite the relative stability of individual sensitivities, again showing how averages may conceal individual differences. Replications with larger samples are required to confirm our findings.

Our results suggest a plausible cognitive mechanism underlying the appreciation of these properties, in line with research on predictability in music (Cheung et al., 2019; Gold, Pearce, Mas-Herrero, Dagher, & Zatorre, 2019), which could transcend sensory modalities and be related to other traits. Accordingly, participants clustered together as per their preference for high (profile 1) or low (profile 2) informational predictability. This is consistent with the simple correlations found among aesthetic sensitivities, and manifested in balanced vs. unbalanced event distributions, small vs. larger and varied intervals and rhythmic figures, redundant vs. different information, and simple vs. complex motifs.

There are certain limitations to our results that are worth noting. First, our sample was musically homogeneous, mainly made up of university students with little formal musical training. Although this makes our results generalizable to non-musicians, further research with more varied samples including musicians is required to elucidate the potential influence of musical expertise, experience, ability, and sophistication. Second, our stimuli were expressly designed for research purposes, and therefore bounded in certain regards (e.g., duration, style, texture, timbre, loudness). Future research is required to determine whether our results would hold for more complex musical stimuli from a broader range of musical cultures (Jacoby, Margulis, et al., 2020). Finally, it remains to be determined how person-related factors, such as personality traits (e.g., openness to experience), socioeconomic status, and musical training and aptitude, influence aesthetic sensitivity and explain why people cluster into different aesthetic sensitivity profiles.

In conclusion, our results suggest that aesthetic experience is influenced by the balance, contour, symmetry, and complexity of music. Furthermore, the relationship between these attributes and aesthetic experience remains stable within individuals but varies between individuals. This supports a conception of aesthetic sensitivity that focuses on individual experience rather than a universal, objective aesthetic standard (Corradi et al., 2020). Individuals can show different aesthetic sensitivities to different features, leading to what we have referred to as an aesthetic sensitivity profile. The ultimate conclusion of these results questions the sense of general preference trends: if people differ so much when it comes to the way complexity (for instance) influences preference, does it make sense to say that "people generally prefer intermediate levels of complexity" when this general trend does not represent the enormous variability of ways people use (or not) complexity to determine their preference? We hope that these results contribute to creating a platform for a more sophisticated investigation of the nature of aesthetic experience in future research.

References

Angier, R. P. (1903). *The aesthetics of unequal division* (Doctoral dissertation, Harvard University).

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language, 59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Babbitt, M., Woods, M., & Washburn, M. F. (1915). Affective Sensitiveness to Colors, Tone Intervals, and Articulate Sounds. *The American Journal of Psychology, 26*(2), 289. doi:10.2307/1413259

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language, 68*(3), 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. doi:10.18637/jss.v067.i01

Beckerman, H., Roebroeck, M. E., Lankhorst, G. J., Becher, J. G., Bezemer, P. D., & Verbeek, A. L. M. (2001). Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research, 10*, 571–578. doi:10.1023/a:1013138911638

Beebe-Center, J. G., & Pratt, C. C. (1937). A Test of Birkhoff'S Aesthetic Measure. *The Journal of General Psychology, 17*(2), 339–353. doi:10.1080/00221309.1937.9918004

Bertamini, M., Palumbo, L., Gheorghes, T. N., & Galatsidas, M. (2016). Do observers like curvature or do they dislike angularity? *British Journal of Psychology, 107*(1), 154–178. doi:10.1111/bjop.12132

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307–310. doi:10.1016/s0140-6736(86)90837-8

Bland, J. M., & Altman, D. G. (2003). Applying the right statistics: analyses of measurement studies. *Ultrasound in Obstetrics and Gynecology, 22*, 85–93. doi:10.1002/uog.122

Boring, E. G. (1950). *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts.

Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts, 7*(1), 48–61. doi:10.1037/a0031624

Brown, S., & Dissanayake, E. (2009). The Arts are More than Aesthetics: Neuroaesthetics as Narrow Aesthetics. *Neuroaesthetics*, *43*, 57. doi:10.4324/9781315224091-4

Che, J., Sun, X., Gallardo, V., & Nadal, M. (2018). Cross-cultural empirical aesthetics. *The Arts and The Brain - Psychology and Physiology Beyond Pleasure*, 77–103. doi:10.1016/bs.pbr.2018.03.002

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., & Koelsch, S. (2019). Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity. *Current Biology, 29*(23), 4084-4092. doi:10.1016/j.cub.2019.09.067

Chmiel, A., & Schubert, E. (2018). Emptying rooms: When the inverted-U model of preference fails—An investigation using music with collative extremes. *Empirical Studies of the Arts, 36*, 199–221. doi:10.1177/0276237417732683

Clark, H., Quackenbush, N., & Washburn, M. F. (1913). A Suggested Coefficient of Affective Sensitiveness. *The American Journal of Psychology, 24*(4), 583. doi:10.2307/1413458

Clemente, A., Vila-Vidal, M., Pearce, M. T., Aguiló, G., Corradi, G., & Nadal, M. (2020). A Set of 200 Musical Stimuli Varying in Balance, Contour, Symmetry, and Complexity: Behavioral and Computational Assessments. *Behavior Research Methods, 52*(4), 1491–1509. doi:10.3758/s13428-019-01329-8

Corradi, G., Belman, M., Currò, T., Chuquichambi, E. G., Rey, C., & Nadal, M. (2019). Aesthetic sensitivity to curvature in real objects and abstract designs. *Acta Psychologica, 197*, 124–130. doi:10.1016/j.actpsy.2019.05.012

Corradi, G., Chuquichambi, E. G., Barrada, J. R., Clemente, A., & Nadal, M. (2020). A new conception of visual aesthetic sensitivity. *British Journal of Psychology, 111*(4), 630–658. doi:10.1111/bjop.12427

Cotter, K. N., Silvia, P. J., Bertamini, M., Palumbo, L., & Vartanian, O. (2017). Curve appeal: exploring individual differences in preference for curved versus angular objects. *i-Perception, 8*(2), 2041669517693023. doi:10.1177/2041669517693023

Davis, F. C. (1933). Aesthetic proportion. *The American Journal of Psychology*, *45*(2), 298–302. doi:10.2307/1414281

Eysenck, H. J. (1940). The general factor in aesthetic judgements 1. *British Journal of Psychology. General Section, 31*(1), 94–102. doi:10.1111/j.2044-8295.1940.tb00977.x

Eysenck, H. J. (1983). A New Measure of “Good Taste” in Visual Art. *Leonardo, 16*(3), 229. doi:10.2307/1574921

Fechner, G. (1860). 1966. Elements of psychophysics. *Elements of Psychophysics.* 1st Ed. New York, NY: Holt, Rinehart, and Winston Inc.

Friedenberg, J. (2018). Geometric regularity, symmetry and the perceived beauty of simple shapes. *Empirical Studies of the Arts, 36*, 71–89. doi:10.1177/0276237417695454

Gartus, A., & Leder, H. (2013). The small step toward asymmetry: aesthetic judgment of broken symmetries. *i-Perception*, *4*(5), 361–364. doi:10.1068/i0588sas

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?. *Journal of Neuroscience, 39*(47), 9397–9409. doi:10.1523/jneurosci.0428-19.2019

Gómez-Puerto, G., Rosselló, J., Corradi, G., Acedo-Carmona, C., Munar, E., & Nadal, M. (2018). Preference for curved contours across cultures. *Psychology of Aesthetics, Creativity, and the Arts, 12*(4), 432–439. doi:10.1037/aca0000135

Götz, K. O., Borisy, A. R., Lynn, R., & Eysenck, H. J. (1979). A new visual aesthetic sensitivity test: I Construction and psychometric properties. *Perceptual and Motor Skills, 49*, 795–802. doi:10.2466/pms.1979.49.3.795

Güçlütürk, Y., Jacobs, R. H. A. H., & Lier, R. van. (2016). Liking versus Complexity: Decomposing the Inverted U-curve. *Frontiers in Human Neuroscience, 10*. doi:10.3389/fnhum.2016.00112

Güçlütürk, Y., & van Lier, R. (2019). Decomposing Complexity Preferences for Music. *Frontiers in Psychology, 10*. doi:10.3389/fpsyg.2019.00674

Haines, T. H., & Davies, A. E. (1904). The psychology of aesthetic reaction to rectangular forms. *Psychological Review, 11*(4–5), 249–281. doi:10.1037/h0076096

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Jacobsen, T., & Höfel, L. (2001). Aesthetics Electrified: An Analysis of Descriptive Symmetry and Evaluative Aesthetic Judgment Processes Using Event-Related Brain Potentials. *Empirical Studies of the Arts, 19*(2), 177–190. doi:10.2190/p7w1-5f1f-njk9-x05b

Jacoby, N., Margulis, E. H., Clayton, M., Hannon, E., Honing, H., Iversen, J., … Wald-Fuhrmann, M. (2020). Cross-Cultural Work in Music Cognition. *Music Perception: An Interdisciplinary Journal, 37*(3), 185–195. doi:10.1525/mp.2020.37.3.185

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology, 103*, 54–69. doi:10.1037/a0028347

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with More Than One Random Factor: Designs, Analytic Models, and Statistical Power. *Annual Review of Psychology, 68*(1), 601–625. doi:10.1146/annurev-psych-122414-033702

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences, 23*(1), 63–77. doi:10.1016/j.tics.2018.10.006

Kuznetsova, A., Brockho, P. B., & Christensen, R. H. B. (2012). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). http://www.cran.r-project.org/package=lmerTest/

Leder, H., Tinio, P. P., Brieber, D., Kröner, T., Jacobsen, T., & Rosenberg, R. (2019). Symmetry is not a universal law of beauty. *Empirical Studies of the Arts, 37*(1), 104–114. doi:10.1177/0276237418777941

Lehnert, B. (2015). BlandAltmanLeh: Plots (Slightly Extended) Bland-Altman Plots. R package version 0.3.1. https://CRAN.R-project.org/package=BlandAltmanLeh

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*, 1494–1502. doi:10.3758/s13428-016-0809-y

Lundholm, H. (1921). The Affective Tone of Lines: Experimental Researches. *Psychological Review, 28*(1), 43–60. doi:10.1037/h0072647

Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta Psychologica, 160,* 43–57. doi:10.1016/j.actpsy.2015.06.005

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience, 10*, 536. doi:10.3389/fnhum.2016.00536

Marin, M. M., & Leder, H. (2013). Examining complexity across domains: relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS One, 8*(8), e72412. doi:10.1371/journal.pone.0072412

Marin, M. M., & Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Personality and Individual Differences, 125*, 80–90. doi:10.1016/j.paid.2017.12.027

Martin, L. J. (1906). An experimental study of Fechner's principles of aesthetics. *Psychological Review, 13*(3), 142–219. doi:10.1037/h0076085

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. doi:10.3758/s13428-011-0168-7

McPherson, M. J., Dolan, S. E., Ossandon, T., Valdes, J., Undurraga, E. A., Jacoby, N., … McDermott, J. (2019). Perceptual fusion of musical notes suggests universal representations of dissonance despite culture-dependent aesthetic associations. *The Journal of the Acoustical Society of America, 145*(3), 1784–1784. doi:10.1121/1.5101526

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual Differences Reveal the Basis of Consonance. *Current Biology, 20*(11), 1035–1041. doi:10.1016/j.cub.2010.04.019

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature, 535*(7613), 547–550. doi:10.1038/nature18635

Meier, N. C. (1928). A measure of art talent. *Psychological Monographs, 39*, 184–199. doi:10.1037/h0093346

Melnykov, V., & Maitra, R. (2010). Finite mixture models and model-based clustering. *Statistics Surveys, 4*(0), 80–116. doi:10.1214/09-ss053

Munar, E., Gómez-Puerto, G., Call, J., & Nadal, M. (2015). Common Visual Preference for Curved Contours in Humans and Great Apes. *PLoS One, 10*(11), e0141106. doi:10.1371/journal.pone.0141106

Murphy, G. (1929). *An historical introduction to modern psychology*. Routledge. doi:10.1037/10600-000

Myszkowski, N., Çelik, P., & Storme, M. (2018). A meta-analysis of the relationship between intelligence and visual "Taste" measures. *Psychology of Aesthetics, Creativity, and the Arts, 12*, 24–33. doi:10.1037/aca0000099

Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts, 28*(2), 173–191. doi:10.2190/em.28.2.d

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin, 27*, 771–785. doi:10.1177/0146167201277001

Otis, M. (1918). Aesthetic Unity: An Investigation into the Conditions That Favor the Apperception of a Manifold as a Unit. *The American Journal of Psychology, 29*(3), 291. doi:10.2307/1414121

Palumbo, L., & Bertamini, M. (2016). The Curvature Effect. *Empirical Studies of the Arts, 34*(1), 35–52. doi:10.1177/0276237415621185

Palumbo, L., Ruta, N., & Bertamini, M. (2015). Comparing Angular and Curved Shapes in Terms of Implicit Associations and Approach/Avoidance Responses. *PLoS One, 10*(10), e0140043. doi:10.1371/journal.pone.0140043

Parker, O. G. (1978). The relationship of musical ability, intelligence and socioeconomic status to aesthetic sensitivity. *Psychology of Music*, *6*(2), 30–35. doi:10.1177/030573567862003

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics. *Perspectives on Psychological Science, 11*(2), 265–279. doi:10.1177/1745691615621274

Pecchinenda, A., Bertamini, M., Makin, A. D. J., & Ruta, N. (2014). The Pleasantness of Visual Symmetry: Always, Never or Sometimes. *PLoS One, 9*(3), e92685. doi:10.1371/journal.pone.0092685

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. doi:10.1371/journal.pone.0027241

Pierce, E. (1894). Aesthetics of simple forms. (I) Symmetry. *Psychological Review, 1*, 483–495.

Plomp, R., & Levelt, W. J. M. (1965). Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America, 38*(4), 548–560. doi:10.1121/1.1909741

R Core Team (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Robbins, H., Smith, D., & Washburn, M. F. (1915). The Influence of Fatigue on Affective Sensitiveness to Colors. *The American Journal of Psychology, 26*(2), 291. doi:10.2307/1413260

Scrucca, L., Fop, M., Murphy, T., Brendan, & Raftery, A., E. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *The R Journal, 8*(1), 289. doi:10.32614/rj-2016-021

Skov, M., & Nadal, M. (2020a). A farewell to art: Aesthetics as a topic in psychology and neuroscience. *Perspectives on Psychological Science*, *15*, 630–642. doi:10.1177/1745691619897963

Skov, M., & Nadal, M. (2020b). The nature of perception and emotion in aesthetic appreciation: A response to Makin's challenge to Empirical Aesthetics. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. doi:10.1037/aca0000278

Snijders, T. A. B., and Bosker, R. J. (2012). *Multilevel analysis. An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE Publications.

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

Vartanian, O., Navarrete, G., Chatterjee, A., Fich, L. B., Leder, H., Modroño, C., ... & Nadal, M. (2019). Preference for curvilinear contour in interior architectural spaces: Evidence from experts and nonexperts. *Psychology of Aesthetics, Creativity, and the Arts, 13*(1), 110. doi:10.1037/aca0000150

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., & Andreou, P. (2013). The case for using the repeatability coefficient when calculating test–retest reliability. *PLoS One, 8,* e73990. doi:10.1371/journal.pone.0073990

Washburn, M. F., Hatt, E., & Holt, E. B. (1923). Affective sensitiveness in poets and in scientific students. *The American Journal of Psychology, 34*(1), 105–106. doi:10.2307/1413933

Weichselbaum, H., Leder, H., & Ansorge, U. (2018). Implicit and Explicit Evaluation of Visual Symmetry as a Function of Art Expertise. *i-Perception, 9*(2), 204166951876146. doi:10.1177/2041669518761464

Appendix A: The MUST Set and Toolbox

(Adapted from Clemente et al., 2020)

Table A1. *Summary of Parameters used to Design the Musical Stimuli in each Subset*

Attribute	Parameter	Feature	
		Balanced	***Unbalanced***
Balance	Distribution of events	Regular	Irregular
	Climax position	Centered	Skewed
	Tension	Progressive	Unprepared
		Smooth	***Jagged***
Contour	Intervals	Only small (≤ 4ths)	Large (> 4ths) & small
	Durations	Progressive, small changes	Sudden, large changes
		Symmetric	***Asymmetric***
Symmetry	Vertical mirror structure	Yes	No
		Simpler	***More complex***
Complexity	Number of events	Few	Many
	Variety of events	Low	High
	Predictability	High	Low

Table A2. *Computational Measures of the Parameters Used to Compose Musical Motifs Varying in Balance, Contour, Symmetry, and Complexity, which Constituted the Composite Measures*

Attribute	**Parameter**	**Computational measure**	**Composite measure**
Balance	Event distribution Climax position Tension	***Bisect unbalance***: Equilibrium between the two halves of a stimulus ***Center of mass offset***: Distance between center of mass and geometric center ***Event heterogeneity***: Heterogeneity in the temporal distribution of events	**BC1**
Contour	Intervals	***Average absolute interval***: Average absolute pitch interval size ***Melodic abruptness***: Average interval size of changes of direction per note ***Durational abruptness***: Proportion of the stimulus with changes of direction	**CC1**
	Durations	***Rhythmic abruptness***: Average ratio of consecutive durations	**CC2**
Symmetry	Palindromic structure	***Total asymmetry***: Direct–retrograde accumulated pitch difference ***Asymmetry index***: Proportion of the stimulus with asymmetries	**SC1**
Complexity	Event density	***Event density***: Number of note events per time unit	**KC1**
	Event variety Predictability	***Average local pitch entropy***: Average pitch entropy of .25-s sliding windows ***Pitch entropy***: Entropy of pitch distribution ***2-tuple interval entropy***: Entropy of 2-tuple interval distribution ***Weighted permutation entropy***: Permutation entropy considering the *SD* of the pitch distribution of each 3-note sequence	
		3-tuple duration entropy: Entropy of 3-tuple duration distribution	**KC2**

Appendix B: Questionnaire

Original Spanish Version

1. ¿Cuál es tu edad? (Respuesta numérica)
2. ¿Cuál es tu género? Mujer / Hombre / Otro
3. ¿Cuál es el nivel de estudios más alto que has completado hasta la fecha? Secundaria / Bachillerato o equivalente / Grado, licenciatura o equivalente / Posgrado, máster o doctorado
4. ¿Cuál es el nivel de estudios musicales más elevado alcanzado hasta el momento? Enseñanza general obligatoria (primaria y secundaria) / Enseñanza elemental de música (escuela de música o conservatorio elemental) / Enseñanza profesional de música (conservatorio profesional) / Grado en música / Posgrado, máster o doctorado en música
5. ¿Durante cuántos años has recibido educación musical formal? (Respuesta numérica)
6. ¿A qué edad comenzó tu formación musical? (Respuesta numérica)
7. ¿Te dedicas profesionalmente a la música? Sí / No
8. ¿Cuánto te han gustado los motivos musicales en general? Por favor, valora del 1 (muy poco) al 5 (mucho).
9. ¿En qué te fijas o qué consideras más importante al juzgar la música estéticamente? Dicho de otro modo, ¿en qué crees que has basado tus valoraciones? (Respuesta abierta)

Translated English Version

1. How old are you? (Numeric response)
2. What is your gender? Woman / Man / Other
3. What is the highest level of education you have ever attained? Secondary-high school or equivalent / Undergraduate / Graduate, Masters, or Ph.D.
4. What is the highest level of musical education you have ever attained? General education (primary and secondary) / Elementary musical education (music school or conservatory) / Professional musical education (music school or conservatory) / Bachelor in music / Postgraduate, Masters or Ph.D. in music

5. How many years have you received formal musical education? (Numeric response)
6. Please, specify your age at the onset of your formal musical training. (Numeric response)
7. Are you a professional musician? Yes / No
8. How much did you like the musical motifs in general? Please, rate from 1 (very little) to 5 (very much).
9. What do you take into consideration or believe most important when judging music aesthetically? In other words, what do you think you based your ratings upon? (Open response)