the british psychological society
promoting excellence in psychology

British Journal of Psychology (2020), 111, 630–658

www.wileyonlinelibrary.com

A new conception of visual aesthetic sensitivity

Guido Corradi[1], Erick G. Chuquichambi[1], Juan Ramón Barrada[2], Ana Clemente[1] and Marcos Nadal[1]*

[1]Human Evolution and Cognition Research Group (EvoCog), Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), Associated Unit to CSIC, University of the Balearic Islands, Palma, Spain
[2]Department of Psychology and Sociology, University of Zaragoza, Teruel, Spain

Aesthetic sensitivity has been defined as the ability to recognize and appreciate beauty and compositional excellence, and to judge artistic merit according to standards of aesthetic value. The Visual Aesthetic Sensitivity Test (VAST) has often been used to assess this ability, but recent research has revealed it has several psychometric problems. Such problems are not easily remedied, because they reflect flawed assumptions inherent to the concept of aesthetic sensitivity as traditionally understood, and to the VAST itself. We introduce a new conception of aesthetic sensitivity defined as the extent to which someone's aesthetic valuation is influenced by a given feature. Experiment 1 aimed to characterize aesthetic sensitivity to four prominent features in visual aesthetics: complexity, symmetry, contour, and balance. Experiment 2 aimed to replicate the findings of Experiment 1 and to assess the test–retest reliability of an instrument designed to measure aesthetic sensitivity to these features using an abridged set of stimuli. Our results reveal that people differ remarkably in the extent to which visual features influence their liking, highlighting the crucial role of individual variation when modelling aesthetic preferences. We did not find clear relations between the four measures of aesthetic sensitivity and personality, intelligence, and art interest and knowledge. Finally, our measurement instrument exhibited an adequate-to-good test–retest reliability.

One of the main goals of scientific aesthetics is to explain how people value objects, events, places, and other people. Such explanations often focus on certain sensory features, including symmetry, complexity, or prototypicality (Berlyne, 1971; Fechner, 1876; Martindale, 2001), and are intended to apply to a broad range of situations, people, and objects. They therefore rely on identifying regular response patterns and general perceptual, cognitive, and affective processes (Leder & Nadal, 2014; Pelowski, Markey, Lauring, & Leder, 2016). An example of such explanations is that people prefer symmetry because it facilitates fluent processing, which generates positive subjective feelings (Reber, Schwarz, & Winkielman, 2004).

A complementary goal of scientific aesthetics is to understand how and why some people diverge from general trends (Jacobsen, 2004; Jacobsen & Höfel, 2002). Such divergences have been attributed to the effects of personality (Chamorro-Premuzic,

*Correspondence should be addressed to Marcos Nadal, Department of Psychology, University of the Balearic Islands, Crta Valldemossa km 7.5 Palma de Mallorca 07122, Spain (email: marcos.nadal@uib.es).

DOI:10.1111/bjop.12427

Reimers, Hsu, & Ahmetoglu, 2009; Mastandrea, Bartoli, & Bove, 2009; McManus & Furnham, 2006), intelligence (Chamorro-Premuzic & Furnham, 2004; Furnham & Chamorro-Premuzic, 2004), expertise (Belke, Leder, Strobach, & Carbon, 2010; Pang, Nadal, Müller, Rosenberg, & Klein, 2013; Silvia & Barona, 2009), and other personal traits. People differ in their aesthetic valuation because they differ in interests, motivations, capabilities, knowledge, and experience. For instance, art history students prefer asymmetry more than other students, because they rely more on declarative knowledge when making deliberate valuations of visual designs (Leder *et al.*, 2019; Weichselbaum, Leder, & Ansorge, 2018).

The study of individual differences in the appreciation of art and aesthetics began as soon as psychology was applied to education at the turn of the 20th century. Efficient measures of artistic and aesthetic abilities were seen as necessary for testing achievement and for vocational guidance (Burt, 1924, 1933; Meier, 1927, 1928; Thorndike, 1916, 1917). Among such measures, aesthetic sensitivity proved to be the best option for its prognostic value and suitability for laboratory research (Meier, 1928). Meier (1927, 1928) found that aesthetic sensitivity, '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), was the most efficient and predictive measure of artistic ability.

But how exactly was aesthetic sensitivity conceived? What was thought to determine aesthetic sensitivity? Irving Child (1962, 1965) believed that individual differences in aesthetic sensitivity owed to differences in the extent to which people were familiar with, and accepted, their local tradition of aesthetic evaluation. Child (1962, 1965) argued that aesthetic sensitivity was cultivated with practice and that it was the result of general cognitive style and personality, not of a specific ability. High aesthetic sensitivity, therefore, was the manifestation of an 'actively inquiring mind, seeking out experience that may be challenging because of complexity or novelty, even alert to the potential experience offered by stimuli not already in the focus of attention, interested in understanding each experience thoroughly and for its own sake rather than contemplating it superficially and promptly filing it away in a category, and able to do all this with respect to the world inside himself as well as the world outside' (Child, 1965, p. 508).

Child's views were diametrically opposed to those of the British psychometric tradition, which regarded aesthetic sensitivity as a distinct ability that manifested itself in different tasks. According to Burt (1924, 1933, 1949), this single underlying factor explained performance on different art and aesthetics tests, covering the appreciation of relations among elements in art, among the combinations of lines and colours in painting, and among sounds and words in music and literature. Eysenck (1940, 1941c) believed that this factor, *T*, corresponded to the ability to appreciate objective beauty, that is, people's taste, or aesthetic sensitivity. In Eysenck's view, aesthetic sensitivity was a distinct, general, and stable ability. It was distinct because it was unrelated to other personal traits ('[this ability], independently of intelligence and personality, determines the degree of good or bad taste'; Eysenck, 1983, p. 231), general because it explained performance on virtually all measures of artistic ability ('it covers a large number of, probably all, pictorial tests'; Eysenck, 1940, p. 100), stable because it was biologically determined and innate ('[it] presumably [has] a genetic foundation in the structure of the nervous system'; Götz, Borisy, Lynn, & Eysenck, 1979, p. 801), and insensitive to experience ('[it] is independent of teaching, tradition, and other irrelevant associations'; Eysenck, 1940, p. 102).

Eysenck identified a second factor when the influence of *T* was minimized. This factor, *K*, was bipolar and distinguished 'those who like modern art, bright, sunny photographs, and Kolbe statues, from those who like the older masters, cloudy, foreboding photographs, and the statues of Maillol and Barlach' (Eysenck, 1941c, p. 266). Thus, the main characteristic of the *K* factor is 'one of brightness or intensity as opposed to darkness or lack of intensity' (Eysenck, 1981, p. 91).

T became for art and aesthetics what Spearman's *g* had become for intelligence (Eysenck, 1940, 1941b). If *g* could be scaled and measured, then so could *T*. In Eysenck's (1941a, 1942) view, aesthetic sensitivity scaled as the degree to which liking approximated ***true aesthetic value***. True aesthetic value could be estimated by averaging people's preference or by resorting to experts' opinion. Aesthetic sensitivity could thus be calculated by simply subtracting people's average liking ratings from group averages or from experts' judgements. Eysenck used different kinds of materials to measure this notion of aesthetic sensitivity. He first correlated liking ranks of artworks (portraits, drawings, landscapes, statues, and so on) and objects (vases, mathematical functions, flowers, clocks, etc.) with the average rankings (Eysenck, 1940). He later used simple geometric designs (Eysenck, 1972; Eysenck & Castle, 1971) taken from Birkhoff (1932) and the Barron–Welsh Figure Preference Test (Barron & Welsh, 1952). Finally, he developed the Visual Aesthetic Sensitivity Test (VAST) in collaboration with the German artist and designer Karl Otto Götz, who actually produced the stimuli (Chan, Eysenck, & Götz, 1980; Eysenck, 1983; Götz *et al.*, 1979; Iwawaki, Eysenck, & Götz, 1979).

The VAST consists, in its last version (Götz, 1981), of 50 pairs of geometric and artistic designs. Both designs in each pair are very similar, but one of them was argued to be superior to the other in terms of design: 'It is more harmonious, better balanced and better adapted in the way the elements are ordered and in the way the lines are drawn' (Götz, 1981). The task given to the participants is 'to discover which picture has been better designed' (Götz, 1981). In each of the 50 pairs, the correct response had been unanimously selected by a group of 8 painters and graphic artists (Götz, 1981; Götz *et al.*, 1979). The number of correct responses constitutes each person's aesthetic sensitivity score, and a measure of 'the degree of good or bad taste' (Eysenck, 1983, p. 231). One of Eysenck and Götz's main goals in constructing this test was to measure meaningful aesthetic judgements (Eysenck, 1983). This is the reason why it emphasized the role of composition, balance, and harmony.

The VAST was designed intending to overcome the psychometric problems common to earlier design and art judgement tests that presented participants with pairs of correct and incorrect alternatives (e.g., the Graves Design Judgment Test, Graves, 1948; or the Meier–Seashore Art Judgment Test, Meier & Seashore, 1929). The fact is, however, that like the tests it intended to surpass, the VAST exhibits low internal consistency and structural validity, and its scores are explained by intelligence, figural creativity, and personality traits such as conscientiousness, extraversion, or openness to experience (Chamorro-Premuzic & Furnham, 2004; Furnham & Chamorro-Premuzic, 2004; Myszkowski, Çelik, & Storme, 2018; Myszkowski, Storme, Zenasni, & Lubart, 2014). Contrary to Eysenck's (1941a, 1942) claims, thus, this notion of aesthetic sensitivity appears not to be a distinct ability. Rather, it seems to draw upon general cognitive processes, learning, and experience.

Given these problems with the construct of aesthetic sensitivity and the instruments used to measure it, Myszkowski and colleagues (Myszkowski & Storme, 2017; Myszkowski & Zenasni, 2016) suggested two mutually compatible ways forward. One option is to revise the VAST to produce a better instrument. Myszkowski and Storme (2017) introduced an abridged and improved version of the VAST, consisting of 25 items, with better internal consistency and structural validity. The other option is to conceive aesthetic aptitude as a complex of multiple abilities and to turn to a composite measure that includes aesthetic sensitivity (aesthetic balance recognition) together with aesthetic exploration, art expertise, sensitivity to complexity, and aesthetic empathy (Myszkowski & Zenasni, 2016).

Myszkowski and colleagues (Myszkowski & Storme, 2017; Myszkowski & Zenasni, 2016) argued that it is worth holding on to Eysenck's notion of aesthetic sensitivity – or good taste – and revise the VAST because of its usefulness in explaining phenomena (Myszkowski & Storme, 2017). Here, we argue for a different course forward. We believe that there are compelling reasons to doubt the usefulness of Eysenck's construct of aesthetic sensitivity, and the measure provided by the VAST, even in its revised form. Eysenck's construct of aesthetic sensitivity as the appreciation of objective beauty is meaningful and useful only if beauty is truly an objective value; that is to say, it resides in objects themselves, and only if such a value can be determined by averaging laypeople's scores or by expert judgements. There is, however, sufficient evidence to reject both premises.

The first premise is an expression of naïve realism. This is the belief that colour, weight, and sound – and beauty too – are attributes of objects, because through perception and cognition we receive sensory input that gets transformed into percepts and representations that accurately reflect reality (Neisser, 1967; Varela, Thompson, & Rosch, 1991). This belief is refuted by basic facts of perception and cognition. Colour, weight, and sound are not attributes of objects, and neither is beauty. They are attributes of our experience of objects. Phenomena such as colour constancy and simultaneous colour contrast – even the simplest visual illusions – demonstrate that physical properties of reflected light, such as intensity and wavelength composition, do not account for our experience of colour, and of other features (Varela *et al.*, 1991). Perception is not a passive recording of stimuli, and cognition is not about rendering an accurate representation of reality (Neisser, 1967; Singer, 2013). Perception is the active comparing of sensory features with predictions based on global configuration and context (Bar, 2004; Murray, Schrater, & Kersten, 2004; Oliva & Torralba, 2007), knowledge and experience (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Clark, 2013; Engel, Maye, Kurthen, & König, 2013), and expectations (Egner, Monti, & Summerfield, 2010). And cognition is about making meaning of the world by interacting with it based on what we know and believe about it, what we expect from it, and what we need and want from it (Bruner, 1990).

Beauty, thus, is not an attribute of objects that people are more or less apt at detecting and responding to. Beauty is an attribute of our experience of objects, an experience that is actively constructed by brain systems that seek to make meaning of those objects, their features, and their value to us (Nadal, Gallardo, & Marty, 2017). As in any domain of human experience, when it comes to liking or appreciating beauty, these systems operate on the basis of expectations and predictions (Egermann, Pearce, Wiggins, & McAdams, 2013; Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011), beliefs (Kirk, Skov, Hulme, Christensen, & Zeki, 2009; Locher, Krupinski, & Schaefer, 2015), prior experience (Kirk, Harvey, & Montague, 2011; Kirk, Skov, Christensen, & Nygaard, 2009; Pang *et al.*, 2013), currently available information (Lengger, Fischmeister, Leder, & Bauer, 2007; Mastandrea & Crano, 2019; Swami, 2013), and context (Brieber, Nadal, & Leder, 2015; Gartus & Leder, 2014; Grüner, Specker, & Leder, 2019; Pelowski, Forster, Tinio, Scholl, & Leder, 2017). The notion of aesthetic appreciation as a sort of response to object properties or configurations – a distinct human ability – lingers still in empirical aesthetics. As shown above, however, it runs against a wealth of evidence on the basic functioning of perception and cognition (Skov, 2019). Moreover, it hampers the advance of empirical aesthetics and alienates the field from developments in psychology and neuroscience (Skov & Nadal, 2018).

The second premise, an expression of the belief in immutable aesthetic value, is refuted by historical fact. Many artworks revered by experts and laypeople in their time have faded into oblivion and, conversely, many of the artworks regarded as masterpieces by experts and laypeople today were never admired – some were even rejected – in their time (Pearce *et al.*, 2016). To which experts or laypeople should we turn for the criteria to true aesthetic value? Those in the past? Those in the present? Or those in the future – for that matter? To none, of course. Aesthetic value changes with time and perspective, it is historically and culturally relative (Jacobsen, 2006), and it is, therefore, in no meaningful sense 'true' or inherent. Refuting the notion of objective beauty does not imply, however, that there are no social or cultural beauty norms.

In the absence of objective or true standards of aesthetic value and, therefore, of individual deviations from these standards, it is unclear what phenomena Eysenck's construct of aesthetic sensitivity hopes to account for, and what the VAST actually measures. In order to be meaningful and useful, the construct of aesthetic sensitivity needs to be redefined and brought in line with established psychological and neuroscientific knowledge. A meaningful and useful notion of aesthetic sensitivity should provide information about the different manners in which people construct their aesthetic experiences, and the different extents to which people respond to certain sensory features, and acknowledge the role of experience, knowledge, context, and culture (Che, Sun, Gallardo, & Nadal, 2018; Jacobsen, 2006). The only way forward, thus, is to discard the notion of aesthetic sensitivity as an innate, unalterable, and general ability to appreciate objective beauty, and to accept that the VAST only provides a measure of the ability to discriminate figures according to a specific understanding of harmony (Gear, 1986).

We define aesthetic sensitivity as the extent to which a given feature influences someone's 'aesthetic' valuation, as this regards evaluation of a stimulus using factors typically thought to connect to aesthetic interests – liking, beauty, visual pleasure (Corradi *et al.*, 2019). In this sense, someone is aesthetically sensitive to complexity, for instance, if her aesthetic valuation depends to some degree on objects' complexity: She likes complex designs more than simple ones, or vice versa. Someone is aesthetically insensitive to complexity if this feature is irrelevant to her aesthetic valuation: Her liking is indifferent to complexity. In this sense, 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 to artistic excellence. Aesthetics and art are, to some extent, overlapping fields, although not identical (Brown & Dissanayake, 2009; Pearce *et al.*, 2016). In the sense put forward here, aesthetic sensitivity is the extent to which certain variations in sensation lead to variations in someone's liking for something (Corradi *et al.*, 2019).

As noted by Corradi *et al.* (2019), this conception of aesthetic sensitivity differs in several regards from Eysenck's (Table 1 summarizes these differences), and has several advantages over Eysenck's. First, it does not rely on the unfounded premise of aesthetic value as an attribute of objects: Here, aesthetic value is an attribute of the experience of objects. Second, there is no external normative standard: Sensitivity is a measure of how responsive someone is to certain features. Third, aesthetic sensitivity is not a unitary construct: It is possible that aesthetic sensitivity is a multidimensional construct. People might be sensitive to some features but not others (Stich, Eisermann, Knäuper, & Leder, 2007). Fourth, aesthetic sensitivity is not a fixed personal trait: It can change depending on

Table 1. Differences between Eysenck's and Corradi and colleagues' conception of aesthetic sensitivity

Eysenck		Corradi *et al.*	
Objectivity	Aesthetic value is an attribute of objects	Aesthetic value is an attribute of our experience of objects	Experience
Standards	There are standards of objective aesthetic value that can be determined	There are no standards of objective aesthetic value to be determined	No standards
Ability	Humans possess the ability to detect objective aesthetic value	Humans construct their experience of objective value	Construction
Singularity	There is a single factor of aesthetic valuation	There are multiple sources for the construction of aesthetic value	Multiplicity
Autonomy	The ability to detect aesthetic value is distinct, unrelated to personality and intelligence	It is probably related to past experience, personality, intelligence, etc.	Relatedness
Context-independent	People's ability to detect aesthetic value is fixed, independently of context	People's aesthetic valuation is context-dependent	Context-dependent

context, experience, expertise, and so on (Leder *et al.*, 2019; Mastandrea & Crano, 2019). Fifth, this notion of aesthetic sensitivity agrees with the common definition of sensitivity as the quality of being receptive to sense impressions, of being responsive to external stimulation. Finally, it is in line with the methods of judgement analysis, or policy capturing (Stewart, 1988), to the domain of aesthetics (Jacobsen, 2004; Jacobsen & Höfel, 2002). These methods model and compare individuals' judgement policies, that is to say, the relations between individuals' judgements and the cues used to make those judgements (Cooksey, 1996; Hammond, Rohrbaugh, Mumpower, & Adelman, 1977; Stewart, 1988).

Our aim in this paper is to explore the construct of aesthetic sensitivity as defined by Corradi *et al.* (2019) and developed in the previous paragraphs. In Experiment 1, we characterize aesthetic sensitivity to four features: complexity, symmetry, contour, and balance. We chose to develop our new concept of aesthetic sensitivity with these four features for two pragmatic reasons. First, they have been extensively studied in empirical aesthetics (e.g., Berlyne, 1971; Bertamini, Palumbo, Gheorghes, & Galatsidas, 2016; Cotter, Silvia, Bertamini, Palumbo, & Vartanian, 2017; Gartus & Leder, 2013; Gómez-Puerto, Munar, & Nadal, 2015; Höfel & Jacobsen, 2003; Jacobsen & Höfel, 2002; Leder *et al.*, 2019; Wilson & Chatterjee, 2005). Second, researchers have developed well-tested stimulus sets to study the effects of these features on aesthetic valuation (Bertamini *et al.*, 2016; Jacobsen & Höfel, 2002; Wilson & Chatterjee, 2005). We also analyse the relations among the aesthetic sensitivities to these features, and to openness to experience, intelligence, art interest and knowledge, and desire for aesthetics, given the evidence that such variables are related to aesthetic appreciation (Chamorro-Premuzic *et al.*, 2009; Furnham & Chamorro-Premuzic, 2004; Furnham & Walker, 2001; Lundy, Schenkel, Akrie, & Walker, 2010; McManus & Furnham, 2006). We chose to analyse our data using linear mixed-effects models. As explained in greater detail below, they are a clear improvement compared to standard multiple regressions commonly used in judgement analysis (e.g., Cooksey, 1996; Stewart, 1988), as they model individual- and group-level responses in combination. In Experiment 2, we conducted a replication of Experiment 1, and studied the temporal stability of aesthetic sensitivities to complexity, symmetry, contour, and balance.

EXPERIMENT 1

Method

Participants

One hundred and sixteen adult students (76 women, M_{age} = 23.34 years, SD_{age} = 5.2 years) at the University of the Balearic Islands volunteered to participate in the experiment. All participants reported normal or corrected to normal vision. Participants were treated in accordance with the Declaration of Helsinki.

Materials

The materials included three sets of images presented on a computer screen, and three paper-and-pen questionnaires. To obtain measures of aesthetic sensitivity to visual features, we used three sets of stimuli that have been used in previous experiments. To assess aesthetic sensitivity to visual contour, we created 66 patterns following the procedure described by Bertamini *et al.* (2016; Figure 1a). Half of them had curved contours, and the other half had sharp-angled contours. To include some variety in each set, we included stimuli with 22 and 26 vertices, and stimuli with designs based on circles, ovals, and lobed ovals. Curved and sharp-angled sets included the same amount of stimuli with 22 and 26 vertices, and the same amount of stimuli designed from circles, ovals, and lobed ovals. To assess aesthetic sensitivity to visual symmetry and visual complexity, we selected 60 stimuli from Jacobsen and Höfel's (2002) set (Figure 1b). The set contains a

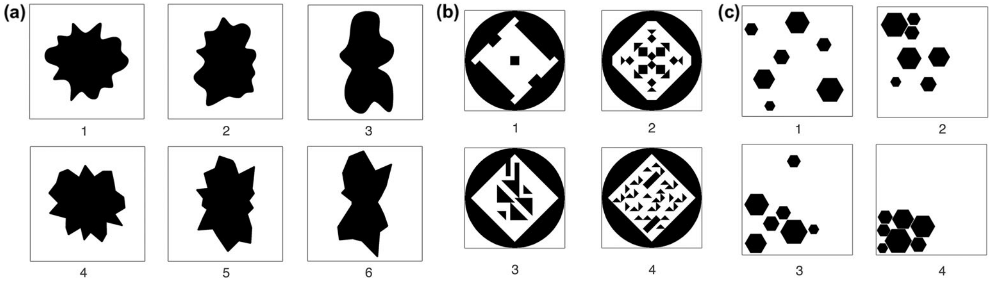

Figure 1. Examples of the stimuli included in the three sets used in Experiments 1 and 2. (a) Examples of stimuli used to assess aesthetic sensitivity to contour. They were designed following Bertamini *et al.* (2016). Stimuli on the top row (A1 to A3) have curved contours; stimuli on the bottom row are equivalent but have sharp-angled contours. Stimuli A1 and A4 were designed based on circles, A2 and A5 on ovals, and A3 and A6 on lobed ovals. (b) Examples of stimuli used to assess aesthetic sensitivity to complexity and symmetry, from Jacobsen and Höfel's (2002) set. Stimuli on the top row (B1 and B2) are symmetrical; stimuli on the bottom row (B3 and B4) are asymmetrical. Stimuli on the left (B1 and B3) are simpler than stimuli on the right (B2 and B4). (c) Examples of stimuli used to assess aesthetic sensitivity to balance, from Wilson and Chatterjee's (2005) set. Stimuli from C1 to C4 cover the range from balanced to unbalanced.

series of images of solid black circles with a centred white square containing triangles that are combined to form designs of varying complexity and symmetry. We used 30 symmetrical and 30 asymmetrical stimuli. Each of these categories included stimuli matched for different degrees of complexity, corresponding to the amount of constituting elements. To assess aesthetic sensitivity to visual balance, we used Wilson and Chatterjee's (2005) set of 65 stimuli consisting of diverse configurations of hexagons (Figure 1c). These stimuli were created to vary in balance, measured as the average of eight symmetry components over the axes of the stimuli. Each stimulus has a corresponding measure of objective balance.

All stimuli in all sets were black and white figures displayed on a medium grey background. Image sizes were 450 pixels on a 1,920 × 1,080 computer screen sized 21", and participants were placed at approximately 45 centimetres of the screen.

After completing the aesthetic sensitivity task, participants filled out three paper-and-pen questionnaires. The first was a custom experience and knowledge in visual art questionnaire adapted from Chatterjee, Widick, Sternschein, Smith, and Bromberger (2010). Five of the items asked about interest in art (1) How interested are you in art? (2) How often do you visit art museums or galleries? (3) How often do you look at art magazines or catalogues? (4) How often do you look at art on the Internet? (5) How often do you speak about art with friends or family?, and three asked about formal education in art (6) How many art history courses did you take during or after high school? (7) How many art creation courses did you take during and after high school? (8) How many hours on average do you spend creating visual art?. Participants were asked to answer each question on a 0–6 Likert scale, where 0 corresponded to *Nothing at all* (1), *Never* (2–5), or *None* (6–8), and 6 corresponded to *Very much* (1), *Once a week* (2), *Very frequently* (3–5), or *6 or more* (6–8). The second questionnaire consisted of the 12 items of the openness to experience scale of the NEO-FFI (McCrae & Costa, 2004). Finally, participants completed an abridged version of Raven's SPM (Raven, 1938; Seisdedos, 1996). We selected 26 items based on responses by a different sample of 150 respondents taken from the same population. We selected those items with at least one error in the previous experiment responses. This reduction aimed to make the whole session shorter.

Procedure

Participants undertook the experimental procedure at the psychology laboratory. They were first welcomed to the laboratory and briefed about the entire procedure. Each participant was then asked to enter one of the individual testing cabins, all of which have the same kind of computers, software, and light conditions. In the testing cabin, participants received the same standard spoken and onscreen instructions. They were told that they would be seeing images on the computer screen and that they had to rate each of them according to how much they liked them. They were instructed to use the keyboard to answer on a 1–7 Likert scale, where 1 meant 'I don't like it at all', and 7 meant 'I like it a lot'. Each stimulus was presented in the centre of the screen. Below the stimulus there was a reminder of the scale, tagged from 1 to 7. Each response was followed by a 2-s grey screen before the next trial started. The task was divided into three blocks: contour, balance, and symmetry–complexity. The order of the blocks and the order of stimuli within each block were randomized for each participant.

Data analysis

Participants' responses to stimuli in each block were analysed by means of linear mixed-effects models (Hox, 2010; Snijders & Bosker, 2012). Linear mixed-effects models account simultaneously for the between-subjects and within-subjects effects of the independent variables (Baayen, Davidson, & Bates, 2008), unlike ANOVAs. ANOVAs usually require averaging across stimuli, which can cause the empirical type I error rate to greatly exceed the nominal level, and lead to claims of significant effects that are unlikely to replicate with different samples (Judd, Westfall, & Kenny, 2012, 2017). As pointed out by Nezlek (2001), linear mixed-effects models provide the most accurate analyses of hierarchically structured data in which there is some kind of dependency, which is the case here, where responses to stimuli are dependent on, or nested within, individual participants. This is because they model random error at all levels of analysis simultaneously, relying on maximum likelihood procedures to estimate coefficients. Linear mixed-effects models have other additional advantages, even over multiple regression analyses (Hox, 2010; Snijders & Bosker, 2012): They provide meaningful estimates of subject- and group-level variance components and are able to handle incomplete and unbalanced data, to accommodate continuous and categorical predictors, unbiased handling of outliers, widespread availability, flexibility, and ease of use (Judd *et al.*, 2012). One particularly interesting feature is that they make it possible to derive conclusions that generalize to other participants besides the ones providing the data (Judd *et al.*, 2017; Nezlek, 2001). Linear mixed-effects models are, thus, well suited to analyse preference responses, given that these often vary from one person to another and also from one objet to another (Silvia, 2007). For this reason they have been used successfully in experimental aesthetics (Brieber, Nadal, Leder, & Rosenberg, 2014; Cattaneo *et al.*, 2015; Mühlenbeck, Jacobsen, Pritsch, & Liebal, 2017; Mühlenbeck, Liebal, Pritsch, & Jacobsen, 2015, 2016; Vartanian *et al.*, 2019; Wagner, Menninghaus, Hanich, & Jacobsen, 2014). They are especially well suited to the purposes of the current study, because they provide estimates for group-level effects, which can be compared with previous studies, and estimates for participant-level effects, which constitute our measure of individual aesthetic sensitivity.

In the present study, the models were set up to reflect the effect of the main predictors in each set on participants' responses. In all cases we followed Barr, Levy, Scheepers, and Tily's (2013) suggestion to model the maximal random-effects structure justified by the experimental design. This avoids the loss of power, reduces type I error, and enables the generalizability of results to other participants and stimuli. All analyses were carried out within the R environment for statistical computing, version 3.5.0. (R Core Team, 2018), using the *glmer()* functions of the 'lme4' package, version 1.1-18-1 (Bates *et al.*, 2017), fitted with REML estimation. The 'lmerTest' package, version 3.0-1 (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 produces acceptable type I error rates (Luke, 2017).

The model of liking for contour included the interaction between contour (*curved*, *sharp-angled*), shape (*circle*, *oval*, *lobed oval*), and vertices (*22*, *26*) as fixed effects. It also included the slope for each of these features and their interactions as random effects within participants. The model of liking for symmetry (*symmetricalal*, *asymmetrical*) and complexity (*number of elements*) included the interaction between both features. It also included the slope for both of these features and their interaction as random effects within participants. The model of liking for balance included balance (*objective balance index*) as a fixed effect. It also included the slope for balance as a random effect within participants. All models also included random intercepts within stimuli. In all models, categorical predictors were deviation coded. Continuous predictors were centred and, to allow comparison with categorical variables, they were scaled from −0.5 to 0.5. Reference levels for the categorical variables were: *sharp*, *lobed oval*, *22*, and *asymmetrical*.

Although the models described above produce group estimates, the main aim of this study was to understand individual differences in responsiveness to visual features driving aesthetic preference. In the linear mixed-effects models, this corresponds to the modelled individual slope for each of the four features: contour, symmetry, complexity, and balance. We thus define each participant's aesthetic sensitivity to each of these features as the individual slope estimated from the models' random-effect structure. Therefore, after running each model, we extracted each participant's slopes and used these values to describe aesthetic sensitivity to visual contour, symmetry, complexity, and balance, to explore the relations among them, and to determine whether aesthetic sensitivity to any of these features was explained by art interest, art knowledge, intelligence, or openness to experience.

Results

Contour

The results of the liking for contour model showed that overall, participants liked the curved images (m = 3.86 [3.66, 4.07]) more than the sharp-angled images (m = 2.75 [2.54, 2.96]), β = 1.11, $t_{(141,57)}$ = 9.182, p < .001 (Figure 2a). Participants also liked the figures based on lobed ovals (m = 3.42 [3.22, 3.63]) more than the circles (m = 3.20 [3.01, 3.38]), β = 0.12, $t_{(87,26)}$ = 2.552, p = .013, and the ovals (m = 3.29 [3.11, 3.48]), β = 0.11, $t_{(88,15)}$ = 2.294, p = .024. Participants' liking ratings did not differ for figures with 22 vertices (m = 3.29 [3.12, 3.47]) and for figures with 26 vertices (m = 3.32 [3.13, 3.50]), β = −0.012, $t_{(51,49)}$ = 0.441, p = .661.

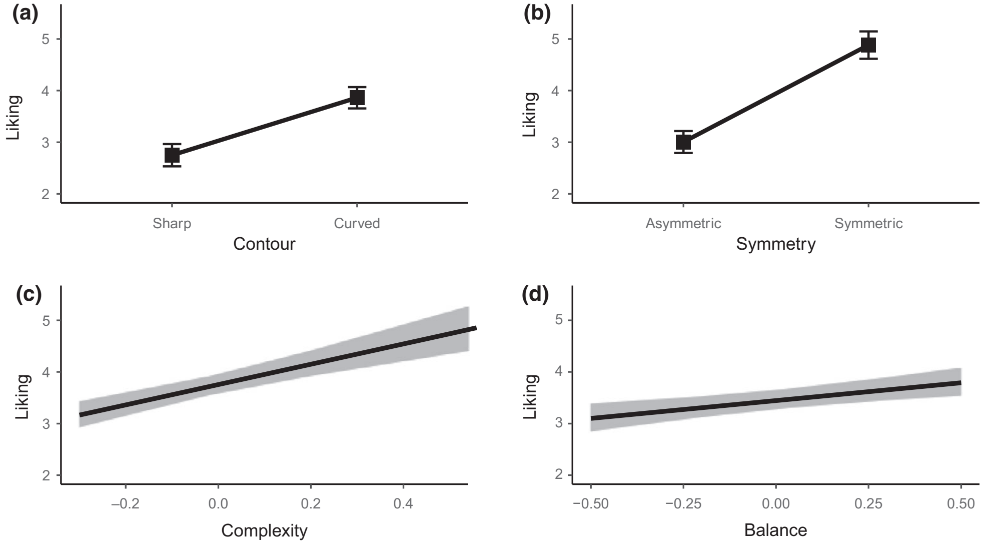

Figure 2. Main effects of contour (a), symmetry (b), complexity (c), and balance (d) on participants' liking ratings in Experiment 1.

Variation among participants in the effects of contour represented 57.47% of the model's explained variance. Removal of the random slope for contour within participants significantly reduced the model fit, $\chi^2 = 1415.8$, $df = 5$, $p < .001$. The estimated slopes for participant's liking for curved contours ranged from −1.41 (indicating higher liking for sharp-angled contours) to 4.48 (indicating higher liking for curved contours), with a mean of 1.11 and a standard deviation of 1.14 (Figure 3a). The values corresponding to the first, second, and third quartiles were 0.36, 1.05, and 1.81.

Symmetry and complexity

The model of liking for symmetry and complexity revealed that participants liked the symmetrical images ($m = 4.88$ [4.62, 5.15]) more than the asymmetrical images ($m = 3.00$ [2.79, 3.22]), $\beta = 1.88$, $t_{(130,88)} = 12.610$, $p < .001$ (Figure 2b). Participants' liking increased with complexity, $\beta = 2.13$, $t_{(78,45)} = 5.476$, $p < .001$ (Figure 2c). The interaction between complexity and symmetry was significant, indicating that the effects of complexity on liking were stronger for symmetrical stimuli than for asymmetrical stimuli, $\beta = 1.64$, $t_{(63,94)} = 2.229$, $p = .029$.

Variation among participants in the effects of symmetry represented 12.08% of the model's explained variance. Removal of the random slope for symmetry within participants significantly reduced the model fit, $\chi^2 = 885.83$, $df = 7$, $p < .001$. The estimated slopes for participant's liking for symmetry ranged from −1.36 (indicating greater liking for asymmetrical designs) to 4.07 (indicating greater liking for symmetrical designs), with a mean of 1.88 and a standard deviation of 1.02 (Figure 3b). The values corresponding to the first, second, and third quartiles were 1.18, 1.92, and 2.60.

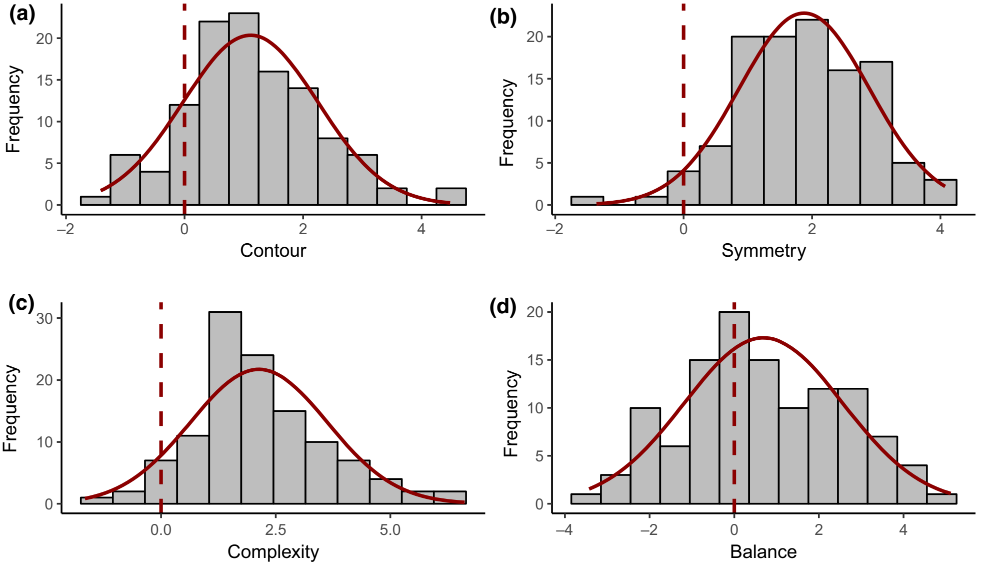

Figure 3. Histograms of individual slopes of liking for contour (a), symmetry (b), complexity (c), and balance (d) in Experiment 1. Vertical dashed lines correspond to a slope of 0, meaning absolute indifference towards each feature. Positive slopes indicate higher liking for curved, symmetrical, complex, and balanced stimuli. Negative slopes indicate higher liking for sharp-angled, asymmetrical, simple, and unbalanced stimuli. Normal curves are overlaid in dark red. [Colour figure can be viewed at wileyonlinelibrary.com]

Variation among participants in the effects of complexity represented 32.22% of the model's explained variance. Removal of the random slope for complexity within participants significantly reduced the model fit, $\chi^2 = 194.7$, $df = 7$, $p < .001$. The estimated slopes for participant's liking for complexity ranged from -1.66 (indicating greater liking for simple designs) to 6.62 (indicating greater liking for complex designs), with a mean of 2.13 and a standard deviation of 1.49 (Figure 3c). The values corresponding to the first, second, and third quartiles were 1.18, 2.01, and 2.97.

Balance

The model of liking for balance showed that participants' liking ratings increased with balance, $\beta = 0.691$, $t_{(145,52)} = 3.454, p < .001$ (Figure 2d). Variation among participants in the effects of balance represented 78.97% of the model's explained variance. Removal of the random slope for balance within participants significantly reduced the model fit, $\chi^2 = 1396.2.7$, $df = 2, p < .001$. The estimated slopes for participant's liking for balance ranged from -3.43 (indicating greater liking for unbalanced configurations) to 5.11 (indicating greater liking for balanced configurations), with a mean of 0.69 and a standard deviation of 1.87 (Figure 3d). The values corresponding to the first, second, and third quartiles were -0.52, 0.64, and 1.98.

Correlations among individual liking slopes

To determine whether there were any relations among individual liking slopes, we studied the correlations among them. The results of this analysis revealed that aesthetic sensitivity to balance was uncorrelated with aesthetic sensitivity to the rest of the features (Table 2). Aesthetic sensitivity to contour and to complexity correlated significantly, and so did aesthetic sensitivity to complexity and to symmetry. Thus, participants who liked complex stimuli also tended to like symmetrical stimuli and stimuli with curved contours.

Explaining aesthetic sensitivity

We ran one regression analysis for each feature to determine whether openness to experience, intelligence, and art interest and knowledge accounted for differences in aesthetic sensitivity among participants. Table 3 shows that art knowledge predicted aesthetic sensitivity to contour, and art interest predicted aesthetic sensitivity to symmetry. In both cases, the relation was negative, indicating that participants who declared having more knowledge of art were those who were less sensitive to contour and

Table 2. Correlations among individual slopes for contour, symmetry, complexity, and balance in Experiment 1

Feature	Contour	Symmetry	Complexity	Balance
Contour	–			
Symmetry	.17	–		
Complexity	.23*	.24**	–	
Balance	.04	.00	.07	–

Note. Spearman correlations for 116 participants.
*p < .05; **p < .01.

that participants who declared being more interested in art were those who were less sensitive to symmetry (Figure 4). Neither openness to experience nor intelligence significantly predicted aesthetic sensitivity to any of the attributes.

Discussion

The main aim of this study was to introduce a new conception of aesthetic sensitivity in the visual domain. This new conception defines aesthetic sensitivity as the degree to which a person's aesthetic valuation is influenced by a certain sensory feature. The goal of Experiment 1 was to measure and characterize aesthetic sensitivity to four features that have been studied extensively: contour, symmetry, complexity, and balance. We modelled aesthetic sensitivity as the individual slopes of the effects of each of these features on participants' liking.

At a group level, our results support previous findings on the effects of contour, symmetry, complexity, and balance on liking. People tend to like designs with curved contours that are symmetrical, complex, and balanced more than those with sharp-angled contours, and those that are asymmetrical, simple, and unbalanced (Gómez-Puerto *et al.*, 2015, 2018; Höfel & Jacobsen, 2003; Jacobsen & Höfel, 2002; Wilson &

Table 3. Regression coefficients in Experiment 1

	Openness	Intelligence	Art interest	Art knowledge
Contour	0.035	−0.038	−0.002	−0.099*
Symmetry	0.015	−0.014	−0.053**	0.033
Complexity	−0.008	−0.016	−0.034	−0.040
Balance	−0.009	0.046	0.014	−0.090

Note. Regression coefficients for each of the four predictors based on data from 116 participants. *$p < .05$; **$p < .01$.

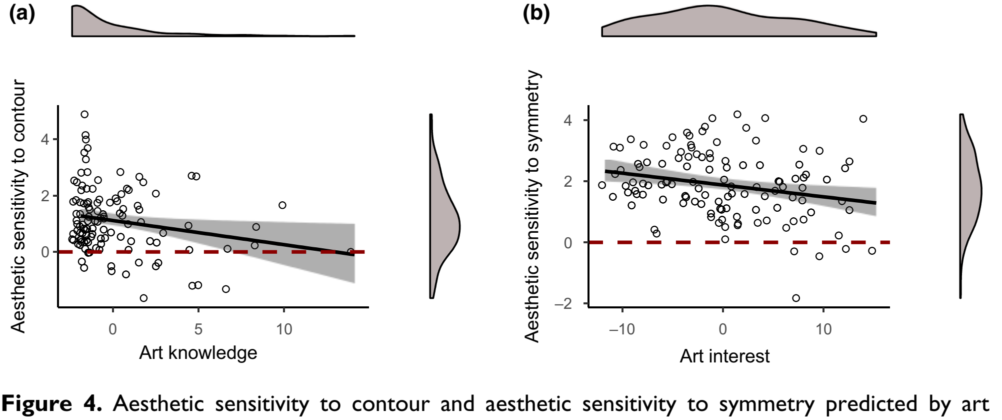

Figure 4. Aesthetic sensitivity to contour and aesthetic sensitivity to symmetry predicted by art knowledge and art interest (Experiment 1). Art knowledge predicts aesthetic sensitivity to contour (a), and art interest predicts aesthetic sensitivity to symmetry (b). The figure includes density plots (top) of art knowledge and art interest, and density plots (right) of aesthetic sensitivity to contour and symmetry. Horizontal dashed lines mark the level of aesthetic indifference to each feature. [Colour figure can be viewed at wileyonlinelibrary.com]

Chatterjee, 2005). This confirmation is, in itself, a meaningful finding. With very few exceptions (Cotter *et al.*, 2017; Gartus & Leder, 2013, 2017; Silvia, 2007), the effects of these features on liking have previously been analysed using ANOVAs or *t*-tests. We have confirmed that these effects hold when data are analysed using linear mixed-effects models, that is to say, when within- and between-participants variations are accounted for.

Our results on aesthetic sensitivity reveal that although the general trend is to like curved contours, symmetrical, complex, and balanced designs, people vary in the extent to which such features influence their liking. Differences among participants in the extent to which each of those features influenced their liking corresponded to large percentages of the variance explained by the models. In all cases, the inclusion of the random slope for the features within participants produced a much better fit to the data. It can be concluded, thus, that attending only to the general trends in liking for curved contours, symmetry, complexity, and balance overlooks a considerable variation in the extent to which such features influence individuals' liking (Jacobsen, 2004; Jacobsen & Höfel, 2002).

Given Eysenck's claim for a single factor underlying aesthetic sensitivity, we were interested in the relations among the aesthetic sensitivity scores we obtained for each of the four features. Our correlation analysis revealed that aesthetic sensitivity to the four features were either unrelated to each other or only modestly related. This suggests that as a rule, people who are highly sensitive to one feature are not necessarily sensitive to another. There were, however, modest relations between complexity and contour, and between complexity and symmetry, indicating that to some extent people who prefer complex stimuli also prefer symmetrical and curved-contour stimuli.

We were also unable to find consistent relations between aesthetic sensitivity and personality, intelligence, and art interest or knowledge. None of the measures of aesthetic sensitivity were predicted by openness to experience or intelligence. We did find a negative relation between art knowledge and aesthetic sensitivity to contour, and a negative relation between art interest and aesthetic sensitivity to symmetry, suggesting that the more knowledge and interest in art, the less people's liking is affected by these features. However, given that our sample was composed mostly of people with very little art knowledge, such conclusions need to be treated with caution.

Experiment 2 had two goals. The main goal was to ascertain the test–retest reliability of an abridged set of stimuli assessing aesthetic sensitivity to contour, symmetry, complexity, and balance. We hoped to produce an abridged version of our materials that would be less time-consuming in experiments, and still be suitable. We thus asked a new group of participants to take part in a test–retest procedure. The second goal was to replicate our findings in Experiment 1, and the test phase of Experiment 2 served this goal.

EXPERIMENT 2

Method

Participants

Participants were 91 students (M_{age} = 26.17 years, SD_{age} = 7.33 years, 45 men, all adults) attending the University of the Balearic Islands. All participants reported normal or corrected to normal vision and had not participated in the Experiment 1. The study was conducted in accordance with the Declaration of Helsinki.

Materials

Hoping to develop a more time-efficient measure of aesthetic sensitivity, we reduced the number of items participants were asked to rate. To assess aesthetic sensitivity to visual contour, we selected 24 stimuli from those used in Experiment 1. Half of these had curved contours and half had sharp angles. In each subset, we included the same number of shapes created from circles, ovals, and lobed ovals, and the same number of shapes with 22 and 26 vertices. To assess aesthetic sensitivity to complexity and symmetry, we took 20 items from our previous selection of Jacobsen and Höfel (2002) set, 10 of which were symmetrical and 10 asymmetrical. Both subsets included the same variation in complexity. To assess aesthetic sensitivity to balance, we took 22 stimuli from Wilson and Chatterjee's (2005) stimulus set, which were equally spaced in terms of balance scores. Participants completed the same art interest and activities questionnaire as in Experiment 1, the 12 items of the openness to experience scale of the NEO-FFI (McCrae & Costa, 2004), and an abridged, adapted, and translated version of the Desire for Aesthetics Scale (DAS) (Lundy *et al.*, 2010). Our adapted version of the DAS consisted of 9 items, rated on a 0 (*I completely disagree*)-to-6 (*I completely agree*) scale: (1) When I see beautiful things in daily life I rarely feel passionate about them. (2) One of the reasons I love travelling is seeing gorgeous scenery. (3) When watching a movie or series I enjoy noticing visual details (photography, framing, colours, . . .). (4) I enjoy spending time appreciating architecture. (5) I often find myself staring in awe at beautiful things. (6) I notice the details of brand logos. (7) I notice and care about design. (8) I notice and attend to the details in paintings, architecture, sculpture, and graphic work. (9) The details I notice in paintings, architecture, sculpture, and graphic art evoke emotions in me.

Procedure

The task was the same as described in Experiment 1, but it took participants less time to complete, as this abridged version contained approximately one third of the items. Participants performed the task in identical conditions as in Experiment 1, except that they performed it twice, with 14 days between the test and retest sessions. They completed the paper-and-pen questionnaires only in the test session.

Data analyses

All analyses were performed as described in Experiment 1. The exception is the new test–retest analysis. In order to examine the temporal stability of the aesthetic sensitivity measure, we conducted an analysis based on Bland and Altman's (1986) graphical method and the smallest real difference (SRD), a measure of absolute reliability (Vaz, Falkmer, Passmore, Parsons, & Andreou, 2013). Bland and Altman's (2003) graphical method has the advantage that it is unaffected by the variability in the data, as it is based upon the SRD (Vaz *et al.*, 2013), and that it can detect systematic biases in the test–retest procedure. It is based on the mean difference between each participant's scores on the test and retest phases. This method establishes the limits of agreement at 1.96 times the standard deviation above and below this difference. When this interval contains the value 0, the difference between the two measurements could be attributed to error (Beckerman *et al.*, 2001). When it does not, the difference must be attributed to some systematic bias. Bland and Altman's (1986) graphs plot the differences between the test and retest scores against the average, allowing the identification of cases where differences in the measurement are proportional to the measurement magnitude. There is no way to determine whether the limits of agreement for the difference on a given test–retest measure are wide or small. The method merely establishes the boundaries of the minimal detectable true change (Vaz *et al.*, 2013).

Results

Contour

Participants liked the curved-contour images (m = 3.80, [3.48, 4.13]) more than the sharp-angled ones (m = 2.86, [2.47, 3.16]), $\beta = 0.94$, $t_{(28,23)} = 5.11$, $p < .001$ (Figure 5a). There were no differences among participants' liking for stimuli based on lobed ovals (m = 3.44 [3.11, 3.77]), circles (m = 3.29 [2.96, 3.63]), or ovals (m = 3.26 [2.93, 3.60]) (all ps > .354). Liking did not differ for stimuli with 22 (m = 3.33 [3.03, 3.62]) and 26 (m = 3.34 [3.05, 3.63]) vertices either (p = .943).

Variation among participants in the effects of contour on liking ratings represented 50.37% of the model's explained variance. Removal of the random slope for contour within participants from the model significantly reduced the model fit, $\chi^2 = 248.23$, $df = 5$, $p < .001$. The estimated slopes for participants' liking for curved contours ranged from −2.24 (indicating greater liking for sharp-angled contours) to 3.14 (indicating greater liking for curved contours), with a mean of 0.94 and a standard deviation of 0.96 (Figure 6a). The values corresponding to the first, second, and third quartiles were 0.23, 0.87, and 1.54.

Symmetry and complexity

Participants liked the symmetrical designs (m = 4.68 [4.40, 4.96]) more than the asymmetrical ones (m = 3.70 [3.39, 4.02]), $\beta = 0.97$, $t_{(42,02)} = 6.457$, $p < .001$

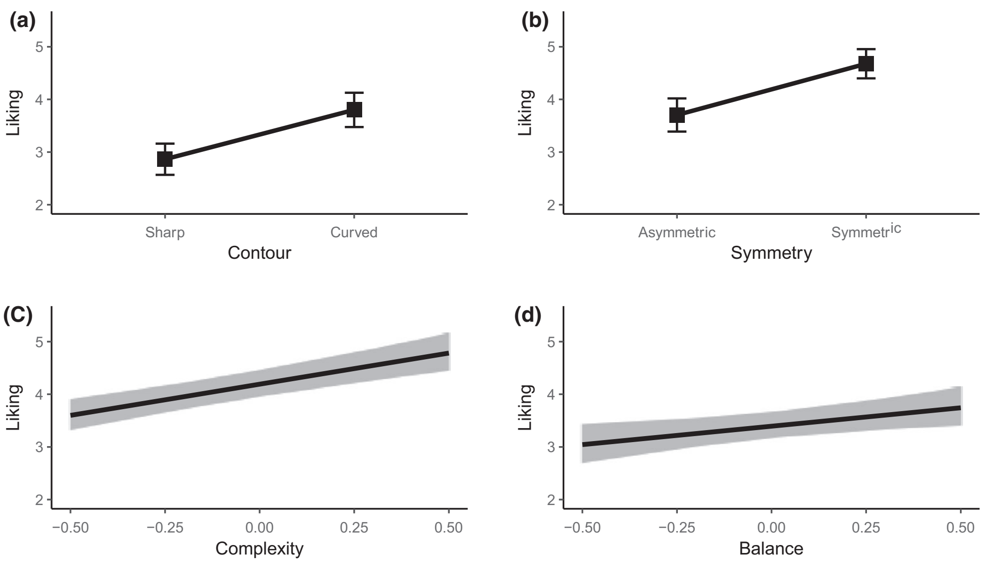

Figure 5. Main effects of contour (a), symmetry (b), complexity (c), and balance (d) on participants' liking ratings during the test phase of Experiment 2.

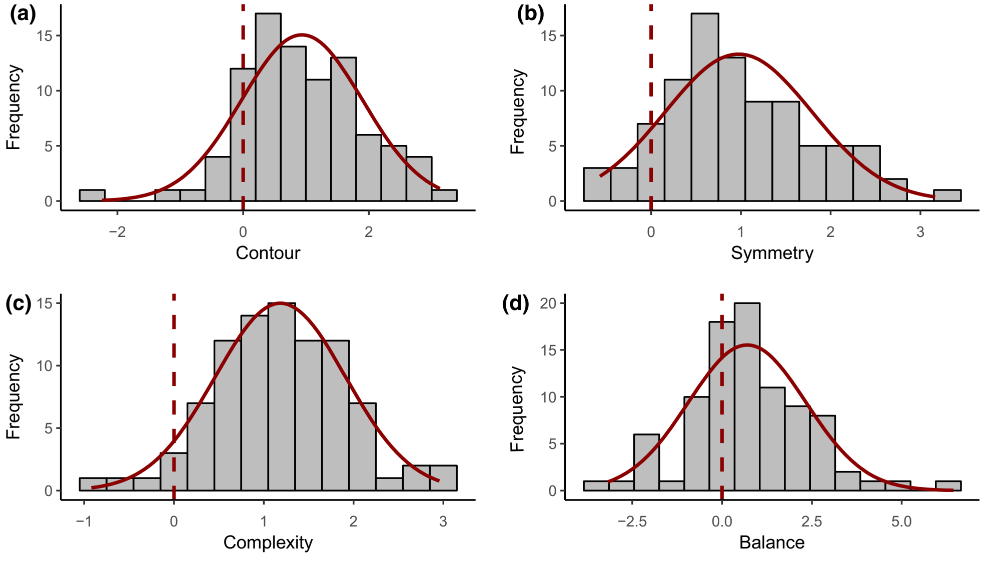

Figure 6. Histograms of individual slopes of liking for contour (a), symmetry (b), complexity (c), and balance (d) during the test phase of Experiment 2. Vertical dashed lines correspond to a slope of 0, meaning absolute indifference towards each feature. Positive slopes indicate higher liking for curved, symmetrical, complex, and balanced stimuli. Negative slopes indicate higher liking for sharp-angled, asymmetrical, simple, and unbalanced stimuli. Normal curves are overlaid in dark red. All data are from the test phase of Experiment 2. [Colour figure can be viewed at wileyonlinelibrary.com]

(Figure 5b). Participants' liking increased with complexity, $\beta = 1.185$, $t_{(26.25)} = 5.849$, $p < .001$ (Figure 5c). The interaction between complexity and symmetry was not significant, $\beta = 0.726$, $t_{(18,86)} = 1.970$, $p = .064$.

Variation among participants in the effects of symmetry on liking ratings represented 21.90% of the model's explained variance. Removal of the random slope for symmetry within participants from the model significantly reduced the model fit, $\chi^2 = 132.68$, $df = 7$, $p < .001$. The estimated slopes for participant's liking for symmetry ranged from −0.57 (indicating greater liking for asymmetrical designs) to 3.16 (indicating greater liking for symmetrical designs), with a mean of 0.97 and a standard deviation of 0.82 (Figure 6b). The values corresponding to the first, second, and third quartiles were 0.41, 0.83, and 1.46.

Variation among participants in the effects of complexity on liking ratings represented 21.93% of the model's explained variance. Removal of the random slope for complexity within participants from the model significantly reduced the model fit, $\chi^2 = 63.40$, $df = 7$, $p < .001$. The estimated slopes for participant's liking for complexity ranged from −0.92 (indicating greater liking for simple designs) to 2.96 (indicating greater liking for complex designs), with a mean of 1.19, standard deviation of 0.73 (Figure 6c). The values corresponding to the first, second, and third quartiles were 0.73, 1.17, and 1.67.

Balance

Participants' liking ratings increased with balance, $\beta = 0.70$, $t_{(57,57)} = 2.539$, $p = .014$ (Figure 5d). Variation among participants in the effects of balance on liking ratings represented 73.97% of the model's explained variance. Removal of the random slope for balance within participants from the model significantly reduced the model fit, $\chi^2 = 208.72$, $df = 2$, $p < .001$. The estimated slopes for participant's liking for balance ranged from −3.18 (indicating greater liking for unbalanced configurations) to 6.44 (indicating greater liking for balanced configurations), with a mean of 0.70 and a standard deviation of 1.64 (Figure 6d). The values corresponding to the first, second, and third quartiles were −0.24, 0.66, and 1.65.

Correlations among individual liking slopes

To determine the relations among individual liking slopes, we studied the correlations among them. The results of this analysis revealed that the only two features for which individual preference slopes correlated were contour and complexity, indicating that participants who were aesthetically sensitive to contour were also aesthetically sensitive to complexity (Table 4).

Explaining aesthetic sensitivity

We ran four regressions to determine whether openness to experience, desire for aesthetics, and art interest and knowledge explained differences among participants in aesthetic sensitivity to each of the features. These variables explained only aesthetic sensitivity to balance. Table 5 shows that art knowledge negatively predicted aesthetic sensitivity to balance ($\beta = -0.401$, $t = 2.11$, $p = .038$): Those who declared having more art knowledge were less susceptible to the effects of balance. Openness to experience, desire for aesthetics, and art interest had no significant effect on aesthetic sensitivity to any of the four attributes.

Table 4. Correlations between individual liking slopes for contour, symmetry, complexity, and balance in the test phase of Experiment 2

Feature	Contour	Symmetry	Complexity	Balance
Contour	–			
Symmetry	.07	–		
Complexity	.23*	−.07	–	
Balance	.08	.12	.08	–

Note. Spearman correlations for 91 participants.
*$p < .05$.

Table 5. Regression coefficients in Experiment 2

	Openness	Desire for aesthetics	Art interest	Art knowledge
Contour	0.007	0.012	−0.048	0.047
Symmetry	0.009	0.016	0.048	−0.116
Complexity	0.006	−0.002	0.106	−0.139
Balance	−0.052	0.052	0.280	−0.401*

Note. Regression coefficients for each of the four predictors based on data from 91 participants.
*$p < .05$.

Test–retest reliability

Table 6 shows the results of the analyses based on the smallest real difference (SRD), the absolute measure of test–retest reliability, and Figure 7 shows the corresponding Bland–Altman graphs. These analyses revealed that whereas the test–retest differences in the assessment of aesthetic sensitivity to contour and balance can be attributed to random error, this is not the case with the assessment of aesthetic sensitivity to symmetry and complexity. In both of these cases there is a systematic bias in the differences. In the case of symmetry, participants were more sensitive in the retest phase. In the case of complexity, participants were less sensitive in the retest phase. Such differences,

Table 6. Mean difference and smallest real difference measures of test–retest reliability of aesthetic sensitivity to contour, symmetry, complexity, and balance in Experiment 2

		95% CI		
Feature	Mean retest–test difference	Lower	Upper	Smallest real difference
Contour	−0.063	−0.253	0.127	1.693
Symmetry	0.237	0.071	0.402	1.474
Complexity	−0.289	−0.501	−0.076	1.898
Balance	0.144	−0.290	0.578	3.870

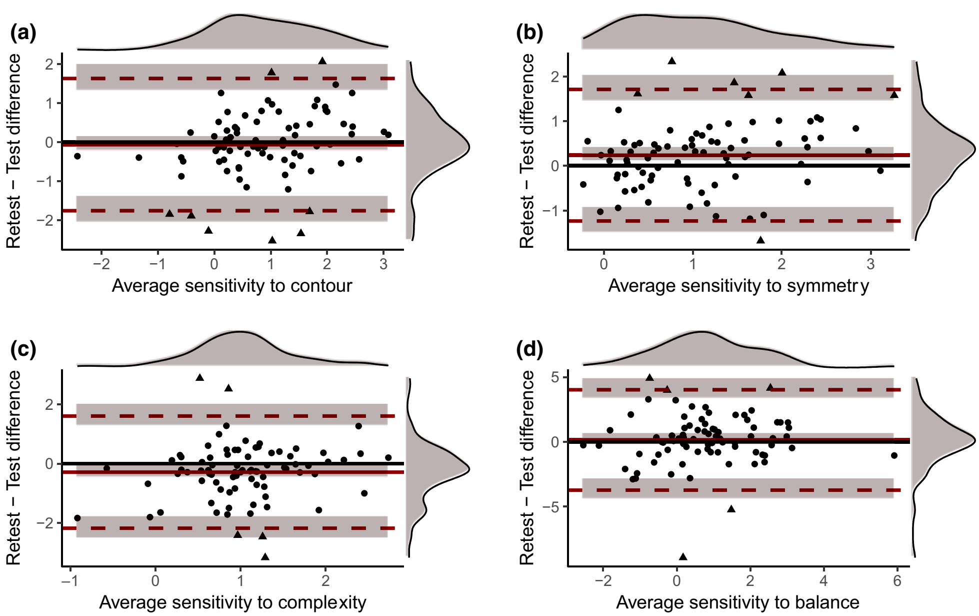

Figure 7. Bland–Altman graphs for the test–retest reliability of aesthetic sensitivity to contour (a), symmetry (b), complexity (c), and balance (d). Horizontal black lines indicate no retest–test change. Horizontal continuous red lines indicate the mean retest–test difference. Horizontal dashed lines mark the lower and higher limits of agreement. Horizontal ribbons comprise 95% CI. Circles correspond to participants whose retest–test difference is smaller than the smallest real difference (SRD). Triangles correspond to participants whose retest–test difference is larger than the SRD. [Colour figure can be viewed at wileyonlinelibrary.com]

however, can be attributed to very few participants. In the case of symmetry, seven participants exceeded the SRD: Six (6.6%) got higher scores in the retest phase and one (1.1%) in the test phase. In the case of complexity, five participants (5.5%) exceed the SRD. Three got lower scores in the retest phase, and 2 (2.2%), much higher scores in the retest phase. Only four participants exceeded the SRD for two of the features. No participant exceeded the SRD for more than two features.

Discussion

Experiment 2 had two goals. On the one hand, we wished to determine whether the results of Experiment 1 would replicate with a new sample of participants. On the other, we wished to examine the temporal stability of a computerized assessment of aesthetic sensitivity to contour, symmetry, complexity, and balance. The results of Experiment 2 replicate the results of Experiment 1, but they also suggest that our abridged assessment has an adequate test–retest reliability.

The results of Experiment 1 and Experiment 2 are remarkably similar. At the group level, participants in both experiments liked the curved-contour stimuli more than the sharp-angled ones, the symmetrical stimuli more than the asymmetrical ones, and liking increased with complexity and balance. In the case of contour and balance, the slopes of these effects were very similar. Conversely, in the case of complexity and symmetry the main effect slopes dropped almost by half in Experiment 2. At the individual level, both experiments show that there is a considerable variation among participants in the extent to which their liking is influenced by contour, symmetry, complexity, and balance. Both experiments confirm that for the four features, a substantial portion of the variance owes to differences among participants in the effects of these features and that models provided a significantly better fit for the data when including the random slopes. In both experiments, aesthetic sensitivities to the four features were barely related. The exception to this was the weak, but significant, positive correlation between aesthetic sensitivity to complexity and to contour in both experiments. In both experiments, participants who liked curved contours the most also liked complex stimuli the most. Finally, in both experiments, we found a weak influence of personality, intelligence, and education measures on aesthetic sensitivity. Art interest and art knowledge were the only scales to show some degree of influence on aesthetic sensitivity, but not in any consistent manner.

Our assessment of the test–retest stability over time of aesthetic sensitivities showed that the measures of contour and balance are stable in time. The differences in aesthetic sensitivity to both of these features measured on both occasions can be attributed to random error. Conversely, the measures of aesthetic sensitivity to symmetry and complexity were systematically biased. As measured with the abridged stimulus set, a small percentage of participants obtained higher scores for aesthetic sensitivity to symmetry in the retest phase, and lower scores for aesthetic sensitivity to complexity in the retest phase.

GENERAL DISCUSSION

Eysenck defined aesthetic sensitivity as a biologically determined ability to appreciate objective beauty. He believed this ability was distinct, in that it was independent from intelligence and personality, and general, in that it applied to many kinds of designs and artworks (Eysenck, 1940, 1941c, 1942). Aesthetic sensitivity could be measured quantitatively. It was simply the difference between someone's liking and a given norm, estimated either by averaging many laypeople's liking or by experts' judgements.

Eysenck's VAST (Götz *et al.*, 1979) was conceived to provide a valid and reliable measure of aesthetic sensitivity. Recent studies, however, revealed the VAST's psychometric weaknesses (Myszkowski & Storme, 2017; Myszkowski & Zenasni, 2016; Myszkowski *et al.*, 2014, 2018). Contrary to Eysenck's conception, aesthetic sensitivity as measured with the VAST is not a distinct ability: It is related to general intelligence, certain personality traits, and certain aspects of creativity (Myszkowski *et al.*, 2014, 2018). In addition to these measurement problems, Eysenck's notion of aesthetic sensitivity stands upon premises that have been rendered invalid with advances in neuroscience and psychology in general and empirical aesthetics in particular (Skov & Nadal, 2018). We have therefore proposed discarding Eysenck's notion of aesthetic sensitivity, and regarding the VAST as a measure of the ability to discriminate between levels of a particular notion of harmony, in line with Gear's (1986) conclusions.

In this paper, we have developed an alternative approach to aesthetic sensitivity. In line with Corradi *et al.* (2019), we have defined aesthetic sensitivity as responsiveness, as the extent to which a given feature influences someone's liking or preference. From the perspective of social judgement theory, our definition of aesthetic sensitivity corresponds to individual differences in judgement policies, that is to say, to the extent to which people's judgements depend on aesthetic cues (Cooksey, 1996; Jacobsen, 2004; Jacobsen & Höfel, 2002; Stewart, 1988). We conducted two experiments. The first aimed to introduce one possible measure of aesthetic sensitivity based on the individual slopes provided by linear mixed-effects models. We characterized aesthetic sensitivity to contour, symmetry, complexity, and balance. The second experiment aimed to replicate the results of the first using an abridged version of the task, and explore the test–retest reliability of this abridged version.

The results of both experiments confirm the general effects that have previously been reported in the literature (Gómez-Puerto *et al.*, 2015, 2018; Höfel & Jacobsen, 2003; Jacobsen & Höfel, 2002; Wilson & Chatterjee, 2005). As a group, participants liked designs with curved contours more than equivalent versions with sharp-angled contours, symmetrical designs more than asymmetrical designs, and their liking increased linearly with complexity and balance.

By applying linear mixed-effects models (Cotter *et al.*, 2017; Gartus & Leder, 2013, 2017; Silvia, 2007), both experiments also uncovered important individual variations in the impact of contour, symmetry, complexity, and balance on liking. In the four cases, individual responsiveness to these features accounted for a large proportion of variance in liking ratings. For some participants, liking was affected by variations in contour, symmetry, complexity, and balance. For other participants, liking was unaffected by such variations; they were indifferent to such features. This adds to the literature showing that group-level models conceal considerable variation among participants in the features that contribute to their liking (Jacobsen, 2004; Jacobsen & Höfel, 2002).

Both experiments also unveiled very weak correlations among aesthetic sensitivities to the four features. The only significant – although weak – correlation in both experiments was between contour and complexity. This indicates that participants who liked curved-contour designs also tended to like complex ones, that participants who liked sharp-angled contour designs tended to like simple ones, and that participants who were indifferent to one feature tended to be indifferent to the other. In sum, aesthetic sensitivity to one feature is either unrelated or only weakly related to aesthetic sensitivity to other features. People are not aesthetically sensitive in general and to all features alike. They seem to be more sensitive to some features than others. This supports the possibility of multiple relatively independent aesthetic sensitivities. Further work is required to determine the dimensions underlying aesthetic sensitivity to different features (Stich *et al.*, 2007).

In both of our experiments we found little evidence that aesthetic sensitivity to contour, symmetry, complexity, and balance is related to intelligence, openness to experience, desire for aesthetics, art interest, or art knowledge. These variables were either unrelated to aesthetic sensitivity, or only weakly and inconsistently related. Further research is also needed on this front, to better understand the relation between intelligence, personality, and experience and aesthetic sensitivity.

The results of the test–retest assessment suggest that our abridged set of stimuli has an adequate test–retest reliability regarding balance and contour, and moderate regarding symmetry and complexity. Our motivation to put together a stimulus set that is efficient for research and can be applied quickly might led us to reduce the number of stimuli excessively. Experiment 1 included between 60 and 66 items in each subset, whereas Experiment 2 included only between 20 and 24. It is possible that using between 40 and 44 items for each dimension will increase the reliability of the measures of aesthetic sensitivity to symmetry and complexity.

Our results can be seen as an extension of the application of the concepts and methods of judgement analysis, or policy capturing, to the domain of aesthetics, pioneered by Thomas Jacobsen and colleagues (Höfel & Jacobsen, 2003; Jacobsen, 2004; Jacobsen & Höfel, 2002, 2003; Jacobsen, Schubotz, Höfel, & von Cramon, 2006). One of our major steps forward, in this sense, was our use of linear mixed-effects models, which combine individual- and group-level models, a substantial advance in comparison to the common use of multiple regressions. Originally, judgement analysis was designed to quantify the relation between a person's judgement and the cues used to make that judgement (Stewart, 1988). It was intended to study experts in their natural settings making judgements about problems that are familiar to them, such as meteorologists in a laboratory forecasting the weather, or physicians in a hospital diagnosing patients (Cooksey, 1996; Stewart, 1988). When applied to situations like ours, where participants were asked to judge unfamiliar stimuli in an unfamiliar setting, it is better to conceive these as studies on policy construction, rather than on policy capturing (Brehmer & Brehmer, 1988). Because participants had not previously seen the stimuli they are asked to respond to, they did not have a developed policy; they had to develop such a policy in the course of the experimental session. The replication of the results of Experiment 1 in Experiment 2, and the reasonable temporal stability observed in the test–retest analysis, suggest that although people constructed their judgement policies in the course of the experimental sessions, they did so in a consistent manner. Our concept of aesthetic sensitivity corresponds to the kind of policy constructed by our participants. Some consistently developed a policy whereby the cues were irrelevant to judging the presented items (aesthetically insensitive). Most, however, consistently developed a policy whereby the cues were used to judge them as more or less liked or disliked.

To conclude, we have developed a new conception of aesthetic sensitivity defined as the degree to which someone's liking is influenced by a given visual feature (Corradi *et al.*, 2019). Two experiments confirm that although at a group level people like stimuli that are curved more than sharp, symmetrical more than asymmetrical, complex more than simple, and balanced more than unbalanced, there is remarkable variation among individual liking judgements. The methods and results of these experiments should encourage future researchers to examine individual differences in the extent to which object features influence aesthetic valuation. Group averages cannot continue to be treated as indicative of uniformity. We have not found compelling evidence that aesthetic sensitivity to one feature is related to aesthetic sensitivity to another, nor that aesthetic sensitivity is related to intelligence, personality, art interest, or art knowledge. But further research is definitively required to confirm this.

Variations in aesthetic sensitivity should not be treated as noise. Not everyone is cast in the same mould when it comes to aesthetic valuation. People weigh different visual features differently. Understanding why people differ in the extent to which their aesthetic valuation responds to complexity, symmetry, balance, contour, as well as other sensory features and object features (Stich *et al.*, 2007), has the potential to illuminate the process of aesthetic valuation itself. Variations in aesthetic sensitivity deserve to be studied and explained: Why are some people more aesthetically sensitive to complexity than others? Can training alter aesthetic sensitivity? Can contextual cues modulate aesthetic sensitivity? How do the different aesthetic sensitivities integrate in different people to produce an overall aesthetic value? Does aesthetic sensitivity cut across sensory domains? If people are sensitive to visual complexity, are they also sensitive to musical complexity?

Acknowledgements

The research leading to these results has received support from 'la Caixa' Foundation (ID 100010434) with fellowship code LCF/BQ/ES17/11600021, and from Grant PSI2016-77327-P, awarded by the Spanish Ministerio de Economía, Industria y Competitividad. All authors approved the final version of this manuscript. The authors are grateful to Thomas Jacobsen, Marco Bertamini, and Anjan Chatterjee for making their stimulus sets available to them.

References

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience*, *30*, 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

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*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*, 617–629. https://doi.org/10.1038/nrn1476

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*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barron, F., & Welsh, G. S. (1952). Artistic perception as a possible factor in personality style: Its measurement by a figure preference test. *The Journal of Psychology*, *33*, 199–203. https://doi.org/10.1080/00223980.1952.9712830

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Green, P. (2017). Linear Mixed-Effects Models using 'Eigen' and S4, R package version 1.1-14. Retrieved from http://cran.rproject.org/web/packages/lme4/index.html

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. https://doi.org/10.1023/A:1013138911638

Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *4*, 214–222. https://doi.org/10.1037/a0019648

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

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

Birkhoff, G. D. (1932). *Aesthetic measure*. Cambridge, MA: Harvard University Press.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307–310. https://doi.org/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. https://doi.org/10.1002/uog.122

Brehmer, A., & Brehmer, B. (1988). What have we learned about human judgment from thirty years of policy capturing. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 75–114). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S0166-4115(08)62171-8

Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica*, *154*, 36–42. https://doi.org/10.1016/j.actpsy.2014.11.004

Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PLoS ONE*, *9*(6), e99019. https://doi.org/10.1371/journal.pone.0099019

Brown, S., & Dissanayake, E. (2009). The arts are more than aesthetics: Neuroaesthetics as narrow aesthetics. In M. Skov & O. Vartanian (Eds.), *Neuroaesthetics* (pp. 43–57). Amityville, NY: Baywood.

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Burt, C. (1924). Psychological tests of educable capacity, Chapter I *Report of the Consultative Committee of the Board of Education*. London.

Burt, C. (1933). *How the mind works*. London, UK: Allen and Unwin.

Burt, C. (1949). The structure of the mind. A review of the results of factor analysis. *British Journal of Educational Psychology*, *19*, 176–199. https://doi.org/10.1111/j.2044-8279.1949.tb01621.x

Cattaneo, Z., Lega, C., Ferrari, C., Vecchi, T., Cela-Conde, C. J., Silvanto, J., & Nadal, M. (2015). The role of the lateral occipital cortex in aesthetic appreciation of representational and abstract paintings: A TMS study. *Brain and Cognition*, *95*, 44–53. https://doi.org/10.1016/j.bandc.2015.01.008

Chamorro-Premuzic, T., & Furnham, A. (2004). Art judgment: A measure related to both personality and intelligence? *Imagination, Cognition and Personality*, *24*, 3–24. https://doi.org/10.2190/U4LW-TH9X-80M3-NJ54

Chamorro-Premuzic, T., Reimers, S., Hsu, A., & Ahmetoglu, G. (2009). Who art thou? Personality predictors of artistic preferences in a large UK sample: The importance of openness. *British Journal of Psychology*, *100*, 501–516. https://doi.org/10.1348/000712608X366867

Chan, J., Eysenck, H. J., & Götz, K. O. (1980). A new visual aesthetic sensitivity test: III cross-cultural comparison between Hong Kong children and adults, and English and Japanese samples. *Perceptual and Motor Skills*, *50*, 1325–1326. https://doi.org/10.2466/pms.1980.50.3c.1325

Chatterjee, A., Widick, P., Sternschein, R., Smith, W. B. II, & Bromberger, B. (2010). The assessment of art attributes. *Empirical Studies of the Arts*, *28*, 207–222. https://doi.org/10.2190/EM.28.2.f

Che, J., Sun, X., Gallardo, V., & Nadal, M. (2018). Cross-cultural empirical aesthetics. *Progress in Brain Research*, *237*, 77–103. https://doi.org/10.1016/bs.pbr.2018.03.002

Child, I. L. (1962). Personal preferences as an expression of aesthetic sensitivity. *Journal of Personality*, *30*, 496–512. https://doi.org/10.1111/j.1467-6494.1962.tb02319.x

Child, I. L. (1965). Personality correlates of esthetic judgment in college students. *Journal of Personality*, *33*, 476–511. https://doi.org/10.1111/j.1467-6494.1965.tb01399.x

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 1–24.

Cooksey, R. W. (1996). The methodology of social judgment theory. *Thinking and Reasoning*, *2*, 141–173. https://doi.org/10.1080/135467896394483

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. https://doi.org/10.1016/j.actpsy.2019.05.012

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*, 2041669517693023. https://doi.org/10.1177/2041669517693023

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective and Behavioural Neuroscience*, *13*, 533–553. https://doi.org/10.3758/s13415-013-0161-y

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience*, *30*, 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*, 202–209. https://doi.org/10.1016/j.tics.2013.03.006

Eysenck, H. J. (1940). The 'general factor' in aesthetic judgements. *British Journal of Psychology*, *31*, 94–102.

Eysenck, H. J. (1941a). A critical and experimental study of colour preferences. *American Journal of Psychology*, *54*, 385–394. https://doi.org/10.2307/1417683

Eysenck, H. J. (1941b). The empirical determination of an aesthetic formula. *Psychological Review*, *48*, 83–92. https://doi.org/10.1037/h0062483

Eysenck, H. J. (1941c). 'Type'-Factors in aesthetic judgments. *British Journal of Psychology*, *31*, 262–270.

Eysenck, H. J. (1942). The experimental study of the 'Good Gestalt' - A new approach. *Psychological Review*, *49*, 344–363. https://doi.org/10.1037/h0057013

Eysenck, H. J. (1972). Preference judgments for polygons, designs and drawings. *Perceptual and Motor Skills*, *34*, 396–398. https://doi.org/10.2466/pms.1972.34.2.396

Eysenck, H. J. (1981). Aesthetic preferences and individual differences. In D. O'Hare (Ed.), *Psychology and the arts* (pp. 76–101). Sussex, UK: The Harvester Press.

Eysenck, H. J. (1983). A new measure of 'good taste' in visual art. *Leonardo*, *16*, 229–231. https://doi.org/10.2307/1574921

Eysenck, H. J., & Castle, M. (1971). Comparative study of artists and nonartists on the maitland graves design judgment test. *Journal of Applied Psychology*, *55*, 389–392. https://doi.org/10.1037/h0031469

Fechner, G. T. (1876). *Vorschule der Ästhetik*. Leipzig, Germany: Breitkopf und Härtel.

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality, intelligence, and art. *Personality and Individual Differences*, *36*, 705–715. https://doi.org/10.1016/S0191-8869(03)00128-4

Furnham, A., & Walker, J. (2001). The influence of personality traits, previous experience of art, and demographic variables on artistic preference. *Personality and Individual Differences*, *31*, 997–1017. https://doi.org/10.1016/S0191-8869(00)00202-6

Gartus, A., & Leder, H. (2013). The small step toward asymmetry: Aesthetic judgment of broken symmetries. *i-Perception*, *4*, 352–355.

Gartus, A., & Leder, H. (2014). The white cube of the museum versus the gray cube of the street: The role of context in aesthetic evaluations. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 311–320. https://doi.org/10.1037/a0036847

Gartus, A., & Leder, H. (2017). Predicting perceived visual complexity of abstract patterns using computational measures: The influence of mirror symmetry on complexity perception. *PLoS ONE*, *12*, e0185276. https://doi.org/10.1371/journal.pone.0185276

Gear, J. (1986). Eysenck's visual aesthetic sensitivity test (VAST) as an example of the need for explicitness and awareness of content in empirical aesthetics. *Poetics*, *15*, 555–564. https://doi.org/10.1016/0304-422X(86)90011-2

Gómez-Puerto, G., Munar, E., & Nadal, M. (2015). Preference for curvature: A historical and conceptual framework. *Frontiers in Human Neuroscience*, *9*, 712. https://doi.org/10.3389/fnhum.2015.00712

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*, 432–439.

Götz, K. O. (1981). *VAST: Visual aesthetic sensitivity test*. Düsseldorf, Germany: Concept Verlag.

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. https://doi.org/10.2466/pms.1979.49.3.795

Graves, M. (1948). *Design judgement test*. San Antonio, TX: Psychological Corporation.

Grüner, S., Specker, E., & Leder, H. (2019). Effects of context and genuineness in the experience of art. *Empirical Studies of the Arts*, *37*, 138–152. https://doi.org/10.1177/0276237418822896

Hammond, K. R., Rohrbaugh, J., Mumpower, J., & Adelman, L. (1977). Social judgment theory: Applications in policy formation. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes: Applications in problem settings* (pp. 1–29). New York, NY: Academic Press.

Höfel, L., & Jacobsen, T. (2003). Temporal stability and consistency of aesthetic judgments of beauty of formal graphic patterns. *Perceptual and Motor Skills*, *96*, 30–32. https://doi.org/10.2466/pms.2003.96.1.30

Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2nd ed.). New York, NY: Routledge. https://doi.org/10.4324/9780203852279

Iwawaki, S., Eysenck, H. J., & Götz, K. O. (1979). A new visual aesthetic sensitivity test (VAST): II. Cross-cultural comparison between England and Japan. *Perceptual and Motor Skills*, *49*, 859–862. https://doi.org/10.2466/pms.1979.49.3.859

Jacobsen, T. (2004). Individual and group modelling of aesthetic judgment strategies. *British Journal of Psychology*, *95*, 41–56. https://doi.org/10.1348/000712604322779451

Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the Psychology of Aesthetics. *Leonardo*, *39*, 155–162. https://doi.org/10.1162/leon.2006.39.2.155

Jacobsen, T., & Höfel, L. (2002). Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual and Motor Skills*, *95*, 755–766. https://doi.org/10.2466/pms.2002.95.3.755

Jacobsen, T., & Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 289–299. https://doi.org/10.3758/CABN.3.4.289

Jacobsen, T., Schubotz, R. I., Höfel, L., & von Cramon, D. Y. (2006). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, *29*, 276–285. https://doi.org/10.1016/j.neuroimage.2005.07.010

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. https://doi.org/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*, 601–625. https://doi.org/10.1146/annurev-psych-122414-033702

Kirk, U., Harvey, A., & Montague, P. R. (2011). Domain expertise insulates against judgment bias by monetary favors through a modulation of ventromedial prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 10332–10336. https://doi.org/10.1073/pnas.1019332108

Kirk, U., Skov, M., Christensen, M. S., & Nygaard, N. (2009). Brain correlates of aesthetic expertise: A parametric fMRI study. *Brain and Cognition*, *69*, 306–315. https://doi.org/10.1016/j.bandc.2008.08.004

Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *NeuroImage*, *44*, 1125–1132. https://doi.org/10.1016/j.neuroimage.2008.10.009

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). Retrieved from http://www.cran.r-project.org/package=lmerTest/

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*, 443–464. https://doi.org/10.1111/bjop.12084

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

Lengger, P. G., Fischmeister, F. P. S., Leder, H., & Bauer, H. (2007). Functional neuroanatomy of the perception of modern art: A DC-EEG study on the influence of stylistic information on aesthetic experience. *Brain Research*, *1158*, 93–102. https://doi.org/10.1016/j.brainres.2007.05.001

Locher, P., Krupinski, E., & Schaefer, A. (2015). Art and authenticity: Behavioral and eye-movement analyses. *Psychology of Aesthetics, Creativity, and the Arts*, *9*, 356–367. https://doi.org/10.1037/aca0000026

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

Lundy, D. E., Schenkel, M. B., Akrie, T. N., & Walker, A. M. (2010). How important is beauty to you? The development of the desire for aesthetics scale. *Empirical Studies of the Arts*, *28*, 73–92. https://doi.org/10.2190/EM.28.1.e

Martindale, C. (2001). How does the brain compute aesthetic preference? *The General Psychologist*, *36*, 25–35.

Mastandrea, S., Bartoli, G., & Bove, G. (2009). Preferences for ancient and modern art museums: Visitor experiences and personality characteristics. *Psychology of Aesthetics, Creativity, and the Arts*, *3*, 164–173. https://doi.org/10.1037/a0013142

Mastandrea, S., & Crano, W. D. (2019). Peripheral factors affecting the evaluation of artworks. *Empirical Studies of the Arts*, *37*, 82–91. https://doi.org/10.1177/0276237418790916

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, *36*, 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1

McManus, I. C., & Furnham, A. (2006). Aesthetic activities and aesthetic attitudes: Influences of education, background and personality on interest and involvement in the arts. *British Journal of Psychology*, *97*, 555–587. https://doi.org/10.1348/000712606X101088

Meier, N. C. (1927). Can art talent be discovered by test devices? *Western Arts Association Bulletin*, *11*, 74–79.

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

Meier, N. C., & Seashore, C. E. (1929). *The meier-seashore art judgment test*. Iowa City, IA: Bureau of Educational Research, University of Iowa.

Mühlenbeck, C., Jacobsen, T., Pritsch, C., & Liebal, K. (2017). Cultural and species differences in gazing patterns for marked and decorated objects: A comparative eye-tracking study. *Frontiers in Psychology*, *8*, 6. https://doi.org/10.3389/fpsyg.2017.00006

Mühlenbeck, C. A., Liebal, K., Pritsch, C., & Jacobsen, T. (2015). Gaze duration biases for colours in combination with dissonant and consonant sounds: A comparative eye-tracking study with orangutans. *PLoS ONE*, *10*(10), e0139894. https://doi.org/10.1371/journal.pone.0139894

Mühlenbeck, C., Liebal, K., Pritsch, C., & Jacobsen, T. (2016). Differences in the visual perception of symmetric patterns in orangutans (*Pongo pygmaeus abelii*) and two human cultural groups: A comparative eye-tracking study. *Frontiers in Psychology*, 7, 408. https://doi.org/10.3389/fpsyg.2016.00408

Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks*, *17*, 695–705. https://doi.org/10.1016/j.neunet.2004.03.010

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. https://doi.org/10.1037/aca0000099

Myszkowski, N., & Storme, M. (2017). Measuring "good taste" with the visual aesthetic sensitivity test-revised (VAST-R). *Personality and Individual Differences*, *117*, 91–100. https://doi.org/10.1016/j.paid.2017.05.041

Myszkowski, N., Storme, M., Zenasni, F., & Lubart, T. (2014). Is visual aesthetic sensitivity independent from intelligence, personality and creativity? *Personality and Individual Differences*, *59*, 16–20. https://doi.org/10.1016/j.paid.2013.10.021

Myszkowski, N., & Zenasni, F. (2016). Individual differences in aesthetic ability: The case for an aesthetic quotient. *Frontiers in psychology*, *7*, 750. https://doi.org/10.3389/fpsyg.2016.00750

Nadal, M., Gallardo, V., & Marty, G. (2017). Commentary: Neural substrates of embodied natural beauty and social endowed beauty: An fMRI study. *Frontiers in Human Neuroscience*, *11*, 596. https://doi.org/10.3389/fnhum.2017.00596

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice Hall.

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. https://doi.org/10.1177/0146167201277001

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*, 520–527. https://doi.org/10.1016/j.tics.2007.09.009

Pang, C. Y., Nadal, M., Müller, J. S., Rosenberg, R., & Klein, C. (2013). Electrophysiological correlates of looking at paintings and its association with art expertise. *Biological Psychology*, *93*, 246–254. https://doi.org/10.1016/j.biopsycho.2012.10.013

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, *11*, 265–279. https://doi.org/10.1177/1745691615621274

Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with 'real' and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, *11*, 245–264. https://doi.org/10.1037/aca0000141

Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in Human Neuroscience*, *10*, 160. https://doi.org/10.3389/fnhum.2016.00160

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

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London, UK: H.K. Lewis.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*, 257–262. https://doi.org/doi:10.1038/nn.2726

Seisdedos, N. (1996). *Raven matrices pogresivas. 2° Edición ampliada*. Madrid, Spain: TEA Ediciones.

Silvia, P. J. (2007). An introduction to multilevel modeling for research on the psychology of art and creativity. *Empirical Studies of the Arts*, *25*, 1–20. https://doi.org/10.2190/6780-361T-3J83-04L1

Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts*, *27*, 25–42. https://doi.org/10.2190/EM.27.1.b

Singer, W. (2013). Cortical dynamics revisited. *Trends in Cognitive Science*, *17*, 616–626. https://doi.org/10.1016/j.tics.2013.09.006

Skov, M. (2019). Aesthetic appreciation: The view from neuroimaging. *Empirical Studies of the Arts*, *37*, 220–248. https://doi.org/10.1177/0276237419839257

Skov, M., & Nadal, M. (2018). Art is not special: An assault on the last lines of defense against the naturalization of the human mind. *Reviews in the Neurosciences*, *29*, 699–702. https://doi.org/10.1515/revneuro-2017-0085

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

Stewart, T. R. (1988). Judgment analysis: Procedures. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 41–74). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S0166-4115(08)62170-6

Stich, C., Eisermann, J., Knäuper, B., & Leder, H. (2007). Aesthetic properties of everyday objects. *Perceptual and Motor Skills*, *104*, 1139–1168. https://doi.org/10.2466/pms.104.4.1139-1168

Swami, V. (2013). Context matters: Investigating the impact of contextual information on aesthetic appreciation of paintings by Max Ernst and Pablo Picasso. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 285–295. https://doi.org/10.1037/a0030965

Thorndike, E. L. (1916). Tests of esthetic appreciation. *The Journal of Educational Psychology*, *7*, 509–522. https://doi.org/10.1037/h0073375

Thorndike, E. L. (1917). Individual differences in judgments of the beauty of simple forms. *Psychological Review*, *24*, 147–153. https://doi.org/10.1037/h0073175

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind. Cognitive science and human experience*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/6730.001.0001

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*, 110–116. https://doi.org/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. https://doi.org/10.1371/journal.pone.0073990

Wagner, V., Menninghaus, W., Hanich, J., & Jacobsen, T. (2014). Art schema effects on affective experience: The case of disgusting images. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 120–129. https://doi.org/10.1037/a0036126

Weichselbaum, H., Leder, H., & Ansorge, U. (2018). Implicit and explicit evaluation of visual symmetry as a function of art expertise. *i-Perception*, *9*, 1–24. https://doi.org/10.1177/2041669518761464

Wilson, A., & Chatterjee, A. (2005). The assessment of preference for balance: Introducing a new test. *Empirical Studies of the Arts*, *23*, 165–180. https://doi.org/10.2190/B1LR-MVF3-F36X-XR64

Received 20 March 2019; revised version received 16 September 2019