



Global Sensory Qualities and Aesthetic Experience in Music

Pauli Brattico, Elvira Brattico* and Peter Vuust

Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark

OPEN ACCESS

Edited by:

Piotr Podlipniak,
Adam Mickiewicz University in
Poznań, Poland

Reviewed by:

L. Robert Slevc,
University of Maryland, College Park,
USA

Dan Zhang,

Tsinghua University, China

Sasa Brankovic,

Clinical Center of Serbia, Serbia

*Correspondence:

Elvira Brattico
elvira.brattico@clin.au.dk

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 30 November 2016

Accepted: 13 March 2017

Published: 05 April 2017

Citation:

Brattico P, Brattico E and Vuust P
(2017) Global Sensory Qualities and
Aesthetic Experience in Music.
Front. Neurosci. 11:159.
doi: 10.3389/fnins.2017.00159

A well-known tradition in the study of visual aesthetics holds that the experience of visual beauty is grounded in global computational or statistical properties of the stimulus, for example, scale-invariant Fourier spectrum or self-similarity. Some approaches rely on neural mechanisms, such as efficient computation, processing fluency, or the responsiveness of the cells in the primary visual cortex. These proposals are united by the fact that the contributing factors are hypothesized to be global (i.e., they concern the percept as a whole), formal or non-conceptual (i.e., they concern form instead of content), computational and/or statistical, and based on relatively low-level sensory properties. Here we consider that the study of aesthetic responses to music could benefit from the same approach. Thus, along with local features such as pitch, tuning, consonance/dissonance, harmony, timbre, or beat, also global sonic properties could be viewed as contributing toward creating an aesthetic musical experience. Several such properties are discussed and their neural implementation is reviewed in the light of recent advances in neuroaesthetics.

Keywords: music aesthetics, neuroaesthetics, musical features, naturalistic paradigm, visual aesthetics

INTRODUCTION

When the legendary music producer Phil Spector created the trademark “Wall of Sound” aesthetics during the 1960s, the point was not about music theory or song writing, or even about instrumentation, but something abstract yet firmly anchored in the world of sense: he wanted to create a saturated, dense sound that would be aesthetically appealing even when played out from the monoaural AM radio and jukebox devices of the time. Similar conclusions can be made on the basis of observations of audio and sound engineers who likewise work with abstract sonic notions that, somewhat paradoxically, refer to concrete sensory experiences. A guitar sound, for example, can be “thin” or “full”; a drum must be “singing out,” “wide-open,” “cool,” “not muffling,” “pretty tight,” to have “a little more of a smack” (Porcello, 2004, pp. 741–744).

Provided that such qualities are aesthetically important, and well-known and much used by musicians, what are they? To first coin a heuristic term, we propose to call them *global sensory qualities*. What we mean by saying that they are global is that they concern the “whole sound” distinct from any of its individual parts, instruments, harmony structure, intervals, melody, or tuning. Moreover, many or at least most of these musical qualities seem to refer to sensory qualities. For example, when a snare drum is characterized as “pretty tight,” the notion does not seem to single out a particular affective or cognitive property, let alone a property grounded in (Western or non-Western) music theory. From the context, it is clear that what is at stake is a snare drum sound not spread too wide in terms of its sensory-related acoustic dimensions (space and reverb, frequency, timbre, sustain) in order to “sit well” in the whole mix and thus to emerge distinctive

enough amongst the background of other materials. In short, the global sensory properties are both global properties, in that they concern the whole percept, but also sensory-based, since they seem to describe sensory qualities.

The premise of the present article is that global sensory qualities constitute an important yet neglected factor in a musical aesthetic experience, and could provide a fruitful avenue for research into the psychology and neurobiology of aesthetic perception. For instance, we propose that these global features are statistically extracted from the stimuli by the auditory system—or, perhaps more likely, by some subsystems (McDermott and Simoncelli, 2011; McDermott et al., 2013)—and then passed on to high-level processing, ultimately leading to the main outcomes of musical experience, namely aesthetic judgment, emotion and conscious liking, or preference (Cela-Conde et al., 2011; Brattico et al., 2013).

The idea itself is not new, especially what comes to visual aesthetics, but rarely applied to music. The notion that there are global visual sensory qualities triggering an aesthetic response has a long history, as argued for example by Bell (1914) in his theory that successful (visual) art involves a “significant form” leading to universal aesthetic experience and emotion. For Bell, the significant form, whose ultimate nature he left mysterious, consisted of “combinations and arrangements” of various visual elements such as lines, form and shapes. He wrote that “forms arranged and combined according to certain unknown and mysterious laws do move us in a particular way, and that it is the business of an artist so to combine and arrange them that they shall move us” (loc. 184).

Vision scientists have not shied away from searching for Bell’s significant formula for aesthetic experience, and, recently, a number of them have tried to locate the form in global sensory properties. Jacobs et al. (2016), for example, examined aesthetic judgments of various visual textures and argued that they correlate with global computational properties, such as the presence of lower spatial frequencies, oblique orientations, higher intensity variation, higher saturation, and overall redness. By examining industrial design and visual aesthetics, Hekkert (2006) proposed four sensory qualities that can increase the aesthetic appeal of an object: (i) maximum effect for minimum means (“economic computations are favored over more complex ones”); (ii) unity in variety (“ability to see regularities and patterns in complex observations”); (iii) most advanced, yet acceptable (“the correct balance between novelty and repetition”); (iv) and optimal match (“information from different sensory modalities should converge with each other”). Renoult et al. (2016) found out that the (algorithmically modeled) sparseness of the activity of simple cells in the primary visual cortex (V1) correlates with female face attractiveness when assessed by male participants, suggesting that there might be general, non-face recognition specific neuronal properties that factor into facial aesthetic evaluation. Spehar et al. (2015) reached similar conclusions by correlating visual sensitivity with the aesthetic properties of visual random patterns. Other candidates for global sensory properties that have been studied recently include processing fluency (Reber et al., 2004; Babel and McGuire, 2015; Forster et al., 2015), distribution of spectral frequency power (Menzel et al., 2015),

self-similarity and fractal properties (Taylor et al., 1999, 2011; Spehar et al., 2003; Hagerhall et al., 2004; Mureika et al., 2004; Graham and Field, 2007; Redies, 2007, 2015; Forsythe et al., 2011; Mallon et al., 2014).

Could similar properties play a role in determining aesthetic responses to music, and could this hypothetical causal relation be pinpointed accurately? In the following sections, we argue that this is likely the case and propose hypotheses to be tested in future research, complementing the current focus on more local factors derived from music theory. Indeed, global features constitute but one subset of auditory features relevant to music, along with others (e.g., pitch, timbre, intervals, harmony, melody, music syntax, and individual instruments), much studied both in connection with auditory processing in general (see e.g., Koelsch, 2011), but also in connection with aesthetic perception (for reviews, see Nieminen et al., 2011; Brattico and Pearce, 2013; Brattico et al., 2013; Hodges, 2016). Perhaps global sensory properties play even a special role in musical pieces of pop/rock/metal genres, in which harmony and voice leading rules are often violated but music producers follow specific professional principles toward reaching a defined aesthetic goal (Račić, 1981; Baugh, 1993; von Appen, 2007). Today almost all music is produced, recorded, reproduced and consumed electro-acoustically, and has become a ubiquitous experience in our everyday lives. Musical pieces that resemble classical music styles, such as film soundtracks (Huckvale, 1990) or computer game music (Bridgett, 2013), are today composed and produced with computers. While historically musical aesthetics has concentrated on the classical music genre, more recently also pop/rock and jazz music has received attention by aesthetic (von Appen, 2007; Juslin et al., 2016) and neuroaesthetic scholars (Limb and Braun, 2008; Janata, 2009; Berns et al., 2010; Brattico et al., 2011, 2015; Johnson et al., 2011; Montag et al., 2011; Pereira et al., 2011; Salimpoor et al., 2011, 2013; Zuckerman et al., 2012; Istok et al., 2013; Bogert et al., 2016). Indeed, even though “rock musicians never ask if a composition is aesthetically valuable,” they are still keen in evaluation “if it *sounds good*,” as observed by Račić (1981, p. 200, emphasis from the original). The study of aesthetics would be too narrowly construed if questions of what “sounds good” were ignored.

The same point can be made in the case of visual aesthetics. As pointed out by Redies (2015), the creation of visual beauty is not limited to any particular style, method, genre, or form, such as color, shape, luminance, texture, edges, or depth cues. A wide variety of materials can be used to create visually appealing objects. This suggests that the neural processes associated with aesthetic experience are not restricted to any particular feature (or corresponding neuronal circuits) or to a particular genre or style. We propose that the same might be true of music.

GLOBAL AESTHETIC SENSORY QUALITIES

We argue that global computational properties play a role in music aesthetics, and provide an overview of what we consider some of the most relevant global sensory properties to be.

We also discuss previous research in the aesthetic of music that highlights the importance of such features. This review will be limited to global sensory properties, thus for the sake of clarity we ignore properties relating to culture, history or listeners' cognitive biases that are also supposed to play a role in a musical aesthetic experience (Chapman and Williams, 1976; McPherson and Schubert, 2004; Brattico, 2009–2010). The next section is dedicated to the discussion of the possible role of global properties in brain processing. As a provisional entry to this topic, note again that it is well-known that both musicians and non-musicians do in fact use global and “holistic” notions, such as “beautiful,” “melodious,” “rhythmic,” “touching,” “harmonic,” “peaceful,” “atmospheric,” “calming,” or “versatile” when describing the personal aesthetic value of music (Jacobsen, 2004; Istok et al., 2009). Most if not all of these concepts describe abstract impressionistic and holistic properties characterizing the piece as a whole, and are not strictly dependent on (although they might interact with) music-theory based local notions, such as intervals or chords. The same point can be further appreciated by noting that aesthetic perception is in no way tied to the Western music genres, but applies equally well to non-Western music. Indeed, when we look art and aesthetics as a whole, it is true that “some kind of aesthetic activity is apparently a feature of all the 3,000 or so distinguishable cultures that are to be found on the earth's surface,” as observed by Berlyne (1971 p. 27). Hence, we believe that aesthetics or aesthetic theories should not be tied with any particular style, genre, or music-theoretical notion.

The key distinction between *global* and *local* features is best elucidated by first looking how they are used in the study of visual aesthetics, and then by extending the notion to the domain of music and auditory aesthetics. In the study of vision and visual beauty, local properties of an image constitute the individual parts of the image, such as local color patches, lines, shapes, contrast, textures, surfaces, or other visual elements. Such local elements can be either *formal*, consisting of various non-conceptual or non-representational forms, or *content-based*, consisting of elements that represent something else. Examples of the former elements are color patches, lines, and textures, of the latter faces and objects. Early processing of visual information is predominantly local, as each local point in an image is projected tonotopically to a point in a visual representation (Wurtz and Kandel, 2000). As the information processing continues, however, the local features are integrated into a whole percept, or Gestalt, that “puts each pictorial element in perceptual relation to the other elements in the artwork” (Redies, 2015, p. 6) and thus integrates the various local elements together. It is that whole Gestalt that, according to many vision researchers, is relevant to the appreciation of beauty (Ramachandran and Hirstein, 1999; Zeki, 1999). Thus, the “Global structure refers to statistical regularities in large parts of the image or in the entire image, for example the spatial frequency content of the image, the kurtosis of its luminance values, overall complexity of self-similarity” (Redies, 2015, p. 4). Hence, it is not generally possible to take a piece of art, break it into pieces and then reassemble it back in random order while automatically preserving its artistic qualities. Formulated in this way, the distinction between global and local properties becomes relative. A painting on a wall constitutes a

local feature of an even more global space, the whole wall. A modern artwork may consist of a red spot on a white background, making what in some other context would constitute a local feature a global one. These problems are kept under control by minimizing the impact of the context, for example, by framing and isolating the artwork in various ways from its natural surroundings and other objects of interest.

The global-local distinction elucidated above applies to music. In music, the local features can be best illustrated by the musical score, by separate tracks in a digital audio workstation (DAW), or by separating the performance of each band member from the rest, where each note/interval appears in isolation and is mapped to the production of physical sound with certain timbre- and rhythmic characteristics during performance and/or recording. A note carries local information concerning timbre (instrument), dynamics (loudness), pitch, pitch changes (vibrato), duration and internal change (staccato, marcato, legato). The notes are further integrated into melodies and harmonic structures and relations that can still constitute local features. In a typical multi-instrument composition, several melodic themes are weaved together to create a sense of harmonic and melodic development. A local feature can be detached from the whole musical piece simply by muting it, or by muting a whole track in a sequencer. For instance, a melody can be changed, even dramatically, by changing the pitch or duration of just one note, and this produces fast reactions in the brain (such as the mismatch negativity, MMN, and the P3a responses) reflecting both an automatic processing of the change and the reorienting of involuntary attention toward the unexpected event. In turn, we propose that global musical features involve the composition as a whole, being synthesized from individual local features as they get summed into an integrated Gestalt. One can refer to the totality of all local features as the overall “musical texture.” Although it is possible to attend to each local part selectively, this is arguably not the norm and restricted to certain artificial contexts. The idea of removing, let alone freely reassembling, some parts from a composition is quite alien to the normal production and consumption of music. Thus, as in the case of visual art, we believe that it is the totality of all such elements that determine their artistic and aesthetic value. For instance, in the production of commercial-grade music global auditory features are manipulated during the final mastering process by using limiters, compressors, equalizers, and other dynamic and spectral processors. In the same vein, listening to any of the tracks or sounds in a musical piece in isolation will typically not lead to a positive, impressive aesthetic experience; it is their combined sum that will do that. Below we provide examples of aesthetically relevant music-specific global features.

Distribution of Spectral Energy

An important aesthetic quality of music concerns the distribution and dynamics of its spectral energy. An aesthetically appealing sonic object is typically created by controlling the balance of its spectrum energy along several important dimensions such as (i) frequency, (ii) space, and (iii) time, as discussed below. “The goal” in sound engineering and mixing is “to get every aspect of the track to balance: every pitch and every noise; every

transient and every sustain; every moment in time and every region of the frequency spectrum” (Senior, 2011, loc. 4904). Orchestral and other groups of instrumentalists adhere to the same principle, explicitly, or implicitly. It is crucial that, even in a loud performance typical of rock music, for example, the instruments are balanced.

In the *frequency domain*, we propose that the crucial balance is achieved by ensuring that the musical information is distributed throughout the whole audible frequency spectrum, and that the signal-to-noise ratio for each meaningful package of musical information (i.e., instrument, singer, instrument group or, more generally, a perceived sound source) is good enough so that no lower or higher level auditory masking intervenes. Indeed, the idea that efficient coding plays a role in human perception is supported by empirical evidence. Listeners must be able to hear all instruments in a distinctive way (not as a fuzzy auditory mess) even if they focus attention only on one of them, and thus these instruments have to live inside their own “safe space” in the spatiotemporal spectrum to avoid frequency masking, even when the music is composed out of digital samples of instrument sounds. They must furthermore appear controlled and consistent. Unaesthetic dynamical changes, conflicts and overlaps are routinely cleaned up by using filters, equalizations, compressors, and other techniques. In addition, often pop/rock and jazz music thrives to fill in the whole frequency spectrum by having “bottom end” (bass, kick drum), “high end” (hihats, cymbals, high pitch sounds), and “middle range” (singers, guitars, snare drums) instruments playing simultaneously (Corozine, 2002). Systematic empirical evidence is scarce, but composers are aware that the complete lack of any of the here described components will lead into a distinctive impairment in the aesthetic quality of the overall sound.

The unaesthetic masking phenomenon referred to above might result from the biological architecture of the human auditory system. The auditory system works by decomposing the signal by several narrow cochlear filters, or critical bands, each spanning a relatively small frequency range. The number and constitution of these bands is derived from psychoacoustic masking experiments, so that they capture the upper bound on the human frequency discrimination ability (Zwicker, 1961; Moore, 2012: Ch. 3). For the most part the frequency range increases logarithmically as a function of the central frequency, and the amplitudes of the resulting filters undergo nonlinear basilar membrane compression such that they are less sensitive to higher amplitudes. Furthermore, the human ear is most sensitive to the middle frequencies around 1,500 Hz, while the sensitivity decreases for sounds with both lower and higher frequencies. The temporal resolution of the auditory system, however, surpasses that of the other senses. Indeed, temporal resolution is required in the processing of fast transients and other sound changes that occur in, e.g., natural speech (Plomp, 1964; Zatorre et al., 2002). Further processing takes place once the signal travels to the auditory cortex via several subcortical regions (Barbour and Wang, 2003). The implication is that there are limitations on how much frequency/temporal space each musical signal can occupy to be perceived distinctly and clearly by the human brain in relation to other, surrounding musical information. This is

especially relevant in the context of complex auditory signals, such as speech or music. Professional audio engineers, music producers’ and composers’ aim for distinctiveness in the sound can be interpreted as suggesting that avoidance of low- and high-level auditory masking contributes to sonic aesthetic experience. The notion is global, however, because it concerns the musical piece as a whole: how distinct various instruments and musical signals are perceived in relation to each other.

In the *space domain*, several techniques such as panning, reverbs, filtering, delays, filtering, and pre-delays are manipulated to position the musical information distinctively within the spatial field. This positioning is achieved by modeling the way the human brain encodes spatial information from the acoustic signal (Zahorik, 2002). For example, when a musical instrument is embedded within a space by using an artificial or natural reverberation, a few milliseconds of pre-delay in the reverberation can change the perceived distance of the source: a reverb with no pre-delay will position the source to the back wall of the virtual space, while 20–30 ms pre-delay will bring it closer to the listener. This models the time the reflected (reverberated) sounds will normally lag behind the direct sound. Similar manipulations are used in experiments testing the neural abilities for discriminating sound sources. Notably, these abilities, relying on the fast elaboration of differences in the incoming signal as compared with the environment at the level of the auditory cortex are very sensitive to even small variations of spatial location (Colin et al., 2002; Roeber et al., 2003; Altmann et al., 2014). But the spatial interpretation of music is global in the sense that it concerns the relative position of the listener to that of the sound source and the environment, whether these are real or virtual. The spatial dimension is also used when positioning sound sources to different locations within a virtual space in order to keep the said sources sufficiently distinct from each other.

In the *temporal domain*, the dynamical qualities of individual instruments (e.g., transients) and the whole song structure are controlled to create a sense of music development and to adjust for the inevitable sensory habituation. “In a lot of cases in commercial music,” Senior (2011) observed, “you want to have enough repetition in the arrangement that the music is easily comprehensible to the general public. But you also want to continually demand renewed attention by varying the arrangement slightly in each section” (loc. 2523). For example, to maintain listeners’ attention one is advised to “provide some new musical or arrangement diversion every 3–5 s to keep listener riveted to the radio” (loc. 2592). Thus, the balance between repetition/regularity and novelty, much discussed in the study of aesthetics and supposedly following an inverted U-shape function (Berlyne, 1971), does not concern only rhythm (Vuust and Witek, 2014; Witek et al., 2014) or melody (Green et al., 2012), but is related to a change of any kind, including changes in the global musical texture.

Musical Texture

The term “musical texture” refers to the way that local musical features, such as rhythm, melody, and harmony are integrated in a whole composition and, ultimately, into a whole Gestalt

percept in the listeners' brain (e.g., Meyer, 1956 p. 185–196). Texture is an elementary consideration in both arrangement and orchestration, processes that aim for crafting an aesthetic output from several local themes such as rhythm, melodies, counter-melodies, and harmony. The same four-way voicing, such as an arrangement for four saxophones, may have quite different textures if it is arranged in parallel compared to when the individual voices are allowed to cross one another. Music that strongly relies on music theory properties can benefit immensely from properties of the texture, as in the case of popular or film music, suggesting that texture alone can be a crucial component in determining an aesthetic response to music.

While the study of texture perception is a lively topic in the domain of vision, with by now a long tradition (e.g., Julesz, 1962), very little comparable research exists in the case of auditory modality. In one study, McDermott and Simoncelli (2011) constructed a physiologically realistic model of the auditory system, which they provided with samples of various repeating sound textures, such as rainstorms, insect swarms, river, and wind, and then used the model to extract biologically plausible time-averaged statistical properties from the textures. These statistical measures represent high-level descriptions of the sound source. They were used to synthesize the same texture sounds from white noise, and the results were compared against the natural sounds in an experiment by using human participants. Sound synthesis was either biologically realistic or unrealistic. The logic of the experiment was to use human performance as a way to benchmark the biological plausibility of the model. For example, when the synthetic sounds were indistinguishable from the natural samples by the human participants, it could be assumed that the generative model closely matched that of the human auditory system. When the participants noted marked differences with the original texture and the synthetic one, we can reason that the model did not mimic the human auditory system. A clear contrast emerged between realistic and unrealistic assumptions, suggesting that the human auditory system might indeed extract statistical properties of the sounds to encode and represent its global textural properties. For further experimental evidence that human auditory system utilizes time-averaged statistical processing to represent textures and other global features of sounds, see McDermott et al. (2013). In the latter study, the authors proposed a functional explanation for their findings, suggesting that statistical averaging is used by the auditory system to overcome memory limitations. The evidence that the auditory system uses statistical time-averages is encouraging for our hypothesis that part of the music aesthetic experience relies on global sensory properties, because it provides empirical justification for the claim that such global features could play a direct role also in auditory perception. These studies go further by proposing that there are neuronal populations within the auditory pathway that are specifically dedicated and tuned to detect global statistical properties of the auditory signal. This raises the possibility that the immediate aesthetic value in certain global sensory properties would be directly assessed by low-level modules in the brain, rather than being assembled only later when the isolated local features are merged into a whole

percept. Whatever the case, we encourage studies for testing the hypothesis that, as in the case of visual textures, musical texture would play a comparable role in the aesthetic perception of music.

Expressivity

Another relevant global quality that affects the aesthetic appeal of a sonic object is its music-emotional impact or expressivity (Robinson, 1994; Gabrielsson and Juslin, 1996). While playing synthesized chord sequences or sinewave tones in isolation and in temporally exacting sequences can indeed evoke emotions and aesthetic judgments due to their ability to represent elementary harmony relations, there is a difference between fully mechanized, synthetic version and humanly played orchestral version of the same piece such that the latter will be perceived as more aesthetic than the former (Seashore, 1929). The “humanness” in the performance of a real human being is especially relevant to the perceived emotional character of the performance. This indicates that there are global sensory features that exhibit a direct causal relationship with human emotions and the “emotional centers” of the brain (Koelsch, 2014). What these features are remains elusive, but the study of visceral affective reactions to music, such as chills, has revealed that there indeed exist prototypical sonic qualities that tend to evoke strong emotional responses in listeners. Laeng et al. (2016), for example, mentions properties such as the beginning of a piece, an entry of an instrument or human voice, melodic appoggiaturas (“extra notes or ornaments”), dynamic changes in loudness, surprising harmonic changes, and sustained high-pitch tones of instruments or voice, among other techniques (see Sloboda, 1991; Panksepp, 1995; Gabrielsson and Juslin, 1996; Rickard, 2004; Grewe et al., 2007; Gabrielsson, 2011; Branković, 2013). If we compare raw mechanical and synthesized instrumentation to that of a real human performance, a complex of dynamic and timbral differences emerge such that latter contains a continuous stream of changes in dynamics (attack, sustain, release), pitch (vibrato, true legato), timbre and spectrum, pauses (breathing, bowing), and many others.

Tempo and Mode

Researchers have shown that global properties such as tempo and mode (minor or major) influence preference and liking, possibly due to their association with basic emotions such as sadness and happiness (Hevner, 1935; Dalla Bella et al., 2001; Pallesen et al., 2003; Khalfa et al., 2005; Hunter et al., 2008; Schellenberg et al., 2008; Nieminen et al., 2012). Slow tempo and/or minor mode are associated with sadness, while fast tempo and/or major mode with happiness, the latter receiving more positive liking ratings (Husain et al., 2002). Tempo, meter and mode are global properties in the sense that they describe, not individual instruments or parts, but large segments of the compositions, or indeed the composition as a whole. Mode, for example, characterizes the underlying key (minor vs. major) upon which the composition, or a segment of the composition, is based on. It also describes the tonal center of the piece that the listener will expect the musical development to return periodically through tension and relaxation. When the mode is in major, the music

sounds happier overall than when it is in the minor key. Similarly, the meter of a song, be it, e.g., a waltz (3/4) or a march (2/4) fundamentally influences the mood of the song.

Other Properties and Experimental Expectations

In addition to the examples above, there are other global properties that are known to affect the rewarding responses to music, such as exposure or familiarity (Heingartner and Hall, 1974; Bornstein, 1989; Peretz et al., 1998; Pereira et al., 2011) and groove (Janata et al., 2012; Sioros et al., 2014; Vuust and Witek, 2014; Vuust et al., 2014; Kilchenmann and Senn, 2015; Fitch, 2016). Exposure and familiarity, in particular, affect liking in an inverted U-shape function, so that repetition will first lead to increased preference but the effect disappears if too much repetition is administered (Green et al., 2012).

In sum, alongside the more local and analytical musical features, there are several types of global sensory properties that seem to play a role in the creation of an aesthetic experience of music. One group of properties involves the distribution and dynamics of spectral energy. In aesthetically appealing music, each instrument or meaningful musical signal should occupy its own spectral space in terms of its frequency-based, spatial and dynamical dimensions in order to control auditory masking. A closely related aspect of musical aesthetics is constituted by musical texture, which refers to the overall sound that results from the combination of its local parts. Arrangement and orchestration are two ways musical texture is created, with much of the consideration having to do with distinctiveness and hence ultimately spectral dynamics. We also discussed expressivity, mode, tempo, familiarity and groove, all linked to emotions, as other possible examples of aesthetically relevant global properties.

All the global sensory properties discussed above apply equally well to Western and non-Western music and musical styles. Thus, a spectrally and spatially rich musical texture can be generated by manipulating digital instruments in a modern studio, for instance, in Western style for a pop music project, as well as by producing Balinese gamelan music in its natural surroundings. This is reasonable, since aesthetic responses are not a privilege of Western music and thus should not be explained as outcomes of only one musical genre or style.

The hypothesis linking global properties to aesthetic responses renders itself naturally to empirical experimentation. For example, the balance in spectral energy distribution can be rigorously manipulated at the stimulus level. This can be achieved by removing and/or adding sonic components at specific locations within the spectrum, irrespective of their representational or other content. If our hypothesis is correct, then such manipulation should lead into prominent changes in, e.g., aesthetic pleasure of such objects irrespective of their higher-level content (i.e., comparison between Western and non-Western music). Another relevant consideration comes from the recent naturalistic paradigm, discussed in detail in the next section, that is suited for addressing global properties

particularly well. We return to the experimental issues in the section Sensory Aesthetics as Immersion and Arousal, where we discuss the present hypothesis from a neuroaesthetics viewpoint.

THE NATURALISTIC PARADIGM FOR STUDYING GLOBAL SENSORY PROPERTIES

Some recent work toward analyzing musical stimuli in terms of their global sensory properties have been done thanks to the introduction of the naturalistic paradigm in music research. In this paradigm, the participants are required to listen attentively to a whole piece of music while their brain signal is measured. Afterwards, their brain signal is analyzed as a time-series in combination with qualities obtained by exploiting knowledge from music information retrieval (MIR), namely acoustic parameters that are relevant for identifying musical genres and extracting timbral, tonal, and rhythmic information from musical pieces. Specifically, the brain signal measured with functional magnetic resonance imaging (fMRI; Alluri et al., 2012, 2013; Burunat et al., 2016) and with electroencephalography (EEG; Poikonen et al., 2016a) has been analyzed by extracting acoustic variables from the music by using the MIR Toolbox (Lartillot and Toiviainen, 2007). This approach is based on the assumption that global computational sensory properties in naturalistic musical stimuli provide a useful window not only to technological applications but also into our appreciation of music and its neural implementation. Most of the relevant properties in these studies are spectral in nature and concern the way in which auditory energy is distributed both in frequency- and time-domains (see **Table 1**). This approach itself is a derivative of a larger research agenda of MIR that is aimed at extracting musically relevant information from whole musical pieces by using computational and statistical techniques (see Peeters, 2004; Moffat et al., 2015). MIR algorithms extract global features from the audio signal that are furthermore distantly related to the global features we claim could be relevant to aesthetics.

Particularly, in Alluri et al. (2012) the authors asked the participants to consciously listen to a musical piece (*Adios Nonino* by Astor Piazzolla) while their brain activation was simultaneously observed by fMRI scanning. The brain scans were correlated with statistical properties extracted from the song, such as overall loudness, spectral centroid, high-energy-low energy ratio, spectral entropy, spectral flux, and tonal clarity (see **Table 1**). Once these features were extracted from the whole song, the original 25 features were reduced into 9 by performing a principal component analysis (PCA) on the resulting song-wide feature vector. The remaining cluster features were global features such as Fullness, Brightness, Timbral complexity, Rhythmic complexity, Key clarity, Pulse clarity, Event synchronicity, Activity, and Dissonance, of which two (Rhythmic complexity and Event Synchronicity) were removed as they did not correlate with participants' subjective assessment in a separate behavioral experiment. Of the remaining

TABLE 1 | Acoustic features used in Alluri et al. (2012).

Feature	Description
Loudness	Root mean square energy: the square root of sum of the squares of the amplitude.
Zero crossing rate	Number of time-domain zero crossings of the signal per time unit.
Spectral centroid	Geometric center on the frequency scale of the amplitude spectrum.
High-energy—low-energy ratio	Ratio of energy content below and above 1,500 Hz.
Spectral entropy	The relative Shannon entropy, which measures peaks in the auditory spectrum.
Spectral roll-off	Frequency below which contains 85% of the total energy.
Spectral flux	Measure of the temporal changes in the spectrum.
Spectral spread	Standard deviation of the spectrum.
Spectral flatness	Wiener entropy of the spectrum, which measures as the ratio of its geometrical mean to its arithmetical mean.
Sub-Band flux	Measures the fluctuation of frequency content in 10 octave-scaled sub-bands.
Roughness	Estimates the sensory dissonance.
Mode	Strength of major or minor mode.
Key clarity	Measures tonal clarity.
Fluctuation centroid	Estimates the average frequency of rhythmic periodicities.
Fluctuation entropy	Measures the rhythmic complexity.
Pulse clarity	An estimate of the clarity of the pulse.

Of these, six clusters (Fullness, Brightness, Timbral complexity, Key clarity, Pulse clarity, Activity, and Dissonance) were created for the study by using principal component analysis (PCA). Detailed description of the features can be found from the original source and from the MIR Toolbox manual.

six global sensory properties, the authors showed that their presence and absence in the musical stimuli indeed did correlate with brain activity. For example, the timbral features (Fullness, Brightness, Timbral complexity, and Activity) were associated positively with activity in the superior temporal gyrus (BA 22) bilaterally and the cerebellum, and negatively with several regions, such as the postcentral gyrus (BA 2, 3), the left precuneus (BA 7), and the inferior parietal gyrus (BA 40). The study shows that such global statistical features do play a role in the musical experience and are indeed meaningful from the point of view of processing of music in our brains.

It remains to be seen, however, whether this approach can be applied to the study of aesthetics. Although, the statistical properties used in our previous studies (Alluri et al., 2012, 2013; Burunat et al., 2016) may be too coarse to be directly relevant for aesthetics, in particular when it comes to the masking problems, the approach is consistent with the hypothesis advanced here. Moreover, the hypothesis that any of such properties were relevant to aesthetics can be tested empirically by correlating the presence of such properties to that of listeners' subjective liking. Promising initial attempts toward that direction, namely combining the fMRI timeseries with continuous or discrete ratings have been made by Trost et al. (2015) and Alluri et al. (2015).

Several challenges must also be met when applying this naturalistic paradigm. Although it allows researchers to use realistic music stimuli, the listening conditions are less than optimal, especially in a fMRI setting, in which noise saturation, low temporal resolution and the risk of false positives in the results (Eklund et al., 2016; Liu et al., 2017) pose considerable methodological challenges to our approach. Even if replicability of brain responses to musical features using the naturalistic paradigm has been shown (mainly for timbral features; Burunat et al., 2016), the concerns for applying the current approach to fMRI data might present bottlenecks that are hard to circumvent. A promising direction would be to utilize silent neurophysiological methodologies with millisecond temporal resolution, such as magnetoencephalography (MEG) and/or electroencephalography (EEG). Two papers have obtained neural correlates of MIR features using EEG signals (Poikonen et al., 2016a,b) and we are studying the application of this approach to MEG data, which allows also a spatial resolution that is almost comparable to that of fMRI. Moreover, we do not wish to imply that this methodology be restricted to brain-imaging settings. It may be applied to behavioral experiments, and indeed many studies done in the naturalistic paradigm do involve behavioral components. In such experiments, participants are asked to evaluate naturalistic stimuli continuously, for example, by providing on-line rating or feedback of the music they are listening (Coutinho and Dibben, 2012). Also, global sensory qualities of a naturalistic stimuli can be independently manipulated in behavioral experiments in order to examine the aesthetic effects of such variables. In our view, it is possible that the optimal results are obtained by utilizing a combination of behavioral and brain-imaging methods. In such hybrid paradigms, many methodological restrictions of purely brain-imaging paradigms can be circumvented by applying behavioral methods, while the brain-imaging studies can provide detailed anatomical, physiological and time-sensitive data unavailable by using behavioral methods alone.

SENSORY AESTHETICS AS IMMERSION AND AROUSAL

In this section, we consider several possible neural explanations for the link between global sensory qualities and aesthetics. This approach is motivated by the fact that, if global sensory properties indeed are pertinent for the creation of an aesthetic experience, then there must be something in our brains, "some fundamental characteristics of the human nervous system" (Berlyne, 1971 p. 29), that explains that fact. What these fundamental characteristics are indeed constitutes a perennial problem of neuroaesthetics. The notion of global sensory properties could provide a contribution to this debate.

Historically, the search for a common aesthetic quality goes back at least to Bell's work on the aesthetics of art (Bell, 1914). Bell proposes that all art, and especially visual art, shares a universal time- and culture-independent "significant form" that is associated with aesthetic emotions. Bell thought that the form

arises from aesthetic laws pertaining to the configuration of visual features such as lines, shapes, and colors. For Bell, a crucial test for separating aesthetic art from other stimuli was its universality and time-independence: a genuine aesthetic art should be independent of time, culture, and era.

Zeki (2013) provides a modern interpretation of Bell's theory. He begins from the well-known organizational properties of our visual system, according to which the neuronal processing of visual stimuli is distributed over several quasi-independent modules in the brain, each processing its own specialized domain (movement, colors, lines, faces, direction, and such), and then proposes that each of these modules "have a certain, primitive, biologically derived combination [...] of elements for the attribute that it is specialized in processing, and that the aesthetic perception [...] is aroused when, in a composite picture, each of the specialized areas is activated preferentially" (p. 10). Aesthetic perception, according to this hypothesis, has its origin in a "preferential" activation pattern of the early sensory areas specialized in visual perception that will then lead into the activation of interest- and motivation-related brain areas and hence also to an experience of emotion, beauty, and preference (Sachs et al., 2016). This provides a neurobiological interpretation of Bell's original idea. The "significant form" would refer to the fact that some type of preferred activity occurs in various visual regions of the brain, as if each such module would have its own aesthetic principles. Artists are professionals who "create forms that activate the relevant visual areas either optimally or specifically [and] in a way that is different from that obtained by stimuli that lack the significant configuration" (Sachs et al., 2016 p. 9).

We see certain similarities to the case of music. Playing back a stereo track in mono takes away some aspect of its appeal, much in the same way as removing all reverb from a recording makes it dull and lifeless. We hypothesize that this may be because the neural systems wired to detect direction and distance of sound sources are not activated in a natural way, or they are not activated at all. If the spectral energy is further reduced by, say, removing musical material from frequencies below some threshold 500 Hz, the music becomes thinner and, again, loses part of its appeal. The neuronal systems registering lower frequencies receive no input, and therefore contribute nothing to the overall percept. The idea of avoiding too much repetition by introducing constant change derives from the same source: a dull, repeating music ceases to command our attention. In addition, if a musical piece performed by real human beings is replaced by machines mechanically playing sinewave instruments, then the performance loses some of its emotional connotations and, again, some neuronal processes linking auditory signals with emotions that would otherwise be engaged are not involved. Thus, as observed by Baugh (1993), rock music "aims at arousing and expressing feeling" (p. 23) in the listener, which we believe holds the key to sensory aesthetic experience. Music, like vision, is a composite of several qualities (direction, distance, depth, emotion) processed by semi-independent modules in our nervous system, while each such module responds to its own signature properties in the stimuli. It might be that, as in the case of vision, aesthetic appeal originates in a concerted and balanced

activation of all these modules. The global aesthetic properties in music, specifically, are aimed at optimizing the presence and balance of these qualities to keep different neural structures in the brain in a "preferential" activation and connectivity state. This brain state would, according to our hypothesis, lead to "immersion" or arousal in the listener, resulting in a rich, holistic experience (Brattico et al., 2013).

One way to refine this idea is to build on Berlyne's (1971) seminal work on aesthetics and arousal. The notion that aesthetic experience can be traced back to immersion, and especially arousal, was the cornerstone of Berlyne's work on the psychology and biology of aesthetics (e.g., Berlyne, 1971), who in turn followed much of the spirit of Fechner's (1876) pioneering work. Berlyne's main proposition was that the aesthetic experience, and aesthetic pleasure, derives from a change in organism's arousal level. The change could involve decrease (relaxing, tension reduction) or increase (excitement, expectation) of arousal level, and both could be triggered by several properties, among them novelty, surprise, complexity ambiguity for heightened arousal, and repetition, familiarity for reduced arousal levels. The global sensory qualities point toward the same direction. Thus, a sonic object evoking spatial and affective cognition, commanding the whole energy spectrum, and holding listeners' attention will lead to an immersive experience and continuous arousal: by introducing small changes and crafting a careful "building up" the artist creates a musical piece that avoids sensory habituation that would otherwise reduce its impact.

The Fechner-Berlyne approach has been subject to criticism. Their work belonged to the behaviorist-reductionist framework that sought to explain behavior in terms of bare stimulus-response principles. From such reductionist perspective, internal motivation, pleasure, or curiosity present themselves as near-paradoxical problems. An aesthetic object, in particular, is one that the organism is actively seeking to experience, and thus it presents a particularly difficult problem to explain. Berlyne's theory was an attempt to answer this problem. Many of his most strong critics, however, came from a different, humanistic-philosophical tradition involved with history, art criticism and philosophy, in which behaviorist problems played no meaningful role, and from which the whole enterprise appears unnaturally narrow (see Margolis, 1980, for an example). Today such criticism plays a much less significant role (Zeki, 2014; Bundgaard, 2015). The question of what motivates people, and makes objects desirable for them apart from their possible ecological functionality, is as relevant today as it was then. Further, the idea of explaining human behavior in terms of stimuli, brain physiology, and motoric responses cannot be substituted wholly by speculative philosophy, cultural relativism, or art history; modern neuroscience has a role in explaining human behavior. Indeed, there exists a small but active research program inside the neurosciences that can be characterized as "neuroscience of aesthetics" (for recent reviews see Jacobsen, 2006; Chatterjee, 2011; Brattico and Pearce, 2013; Orgs et al., 2013; Chatterjee and Vartanian, 2014, 2016; Pearce et al., 2016). Berlyne was in fact well-aware of the neuroscientific advances of his day, and documents such matters extensively in his work. At the same token, it is also clear that no neuroscientific or

naturalistic exploration can answer questions such as what, ultimately, is art, and what makes a piece of material constellation a genuine work of art instead of a, say, tool or random junk. This is because art is constituted by several non-appearance properties such as its history, intention, sincerity, and normativity, and not everything beautiful or appealing can be said to be art (Bundgaard, 2015). A naturalistic approach to aesthetics (Brown et al., 2011) will, therefore, pay a price in necessarily ignoring many aspects of art that we would regard as important in other contexts.

Berlyne (1971) proposed that the main categories of stimuli that can modulate arousal, and hence in his theory also involve aesthetic appreciation, fall into three distinction categories: psychophysical, ecological, and what he called “collative” (also “structural”). Psychophysical qualities refer to low-level sensory features and changes in such qualities. He mentions in this connection the fact that more intense stimuli are normally interpreted as more arousing. Ecological variables refer to stimuli that are directly associated, either innately or by means of learned association, with survival, pain, and pleasure. Finally, by the term collative or structural properties he means second-order properties that are arrived at by “summing up characteristics of several elements” (p. 69) that may be present simultaneously or could also be temporally distinct. Properties such as novelty, complexity ambiguity and surprisingness belong to this category. The global sensory properties discussed in the present work would, under this scheme, consist of a mixture of structural and psychophysiological properties: they are structural and global, in that they result from the summation of few or many individual qualities, but also sensory, in that they depend on the sensorium and are not constitutively affective or cognitive.

The immersion hypothesis, according to which aesthetic experience results from an activation of all or many brain regions specialized in the processing of the stimuli, leads to testable hypotheses, and empirical predictions. Vartanian and Goel (2004), for example, report that increased preference in the perception of visual art correlates with increased activation in the visual areas of the brain. If sensory immersion and arousal play a role in aesthetic perception, then the expected outcome is precisely that we should attest a positive correlation between aesthetic preference and the activation of the various brain areas involved in the processing of the stimuli. This prediction was also confirmed in a meta-analysis, likewise reporting an association between visual aesthetic experience and a wide-spread rather than localized brain activation (Boccia et al., 2016). In the case of music, the prediction is that the removal of relevant features, whether spatial, emotional, or spectral, should lead into a marked decrease both in the brain activation and in the aesthetic judgment. Crucially, our hypothesis predicts that this effect should not depend on local features, and should be observed entirely irrespective of musical genre, style, or (representational) content. If, in other words, the aesthetic balance in sensory qualities is achieved by means of immersion, itself based on the concurrent activation of the relevant brain regions, then what matters is the activation itself and not the particular local features present in the activating stimulus. This hypothesis could thus further be tested by invoking experimental top-down effects that

suffice to satisfy the activation condition without the presence of concrete stimulus.

However, this hypothesis predicts, if interpreted in a too simple way, that increasing the amplitude of any or all such features should always lead to increased liking. Oversaturated objects, such as overly loud music or pictures with bright colors, are not perceived as beautiful; instead, they can be perceived even as painful. Too much reverb, stereo widening or emotional expressivity makes the music incomprehensible and “wishy-washy.” This question has always puzzled those trying to understand aesthetic perception. Berlyne’s solution was to assume that stimulus levels beyond a certain moderate cutoff point would begin to activate “aversion systems” that are associated with a negative outcome (danger, unpleasantness). Zeki (2013) discusses this problem and points out that the determining factor cannot be the strength of the activation as such but, rather, there must be some quality in the original signal that prompts the positive response. He provides another interpretation of these results, according to which “it is not the strongest or maximal activity that correlates with preference but rather a specific activity that becomes optimal when stimuli of the right [aesthetic properties] are viewed” (p. 10). Hence, we are back at Bell’s mystery: there is an unknown quality in the stimulus that is preferred by the various regions in the brain specialized in processing that type of stimuli.

The case of music provides another possible interpretation. It is an established fact that the masking effect on one sound over another is amplified by the amplitude of the former. Thus, as the sound is increased in amplitude, the range of frequencies it will mask will also increase. Moreover, if we are presented with a piece of music in which one instrument is associated with overwhelming volume, our brains will attempt to adapt to the situation by attenuating the overall level. This will further reduce the perceived relative amplitudes of the rest of the musical information. Finally, the problem might not be as severe if the amplitude of all sound sources is increased in tandem, which corresponds to an increase in overall volume. Thus, we might be dealing, not with overall amplitude, but with relative amplitudes. It is possible that the reason why balanced performances instead of overly saturated ones are crucial for auditory aesthetics is because the former specifically avoids unaesthetic masking and thus keeps the musical sources distinct. This hypothesis could be tested experimentally. If the problem with amplitude concerns relative amplitudes and/or masking, then the same negative effect on aesthetic experience could be achieved by using other types of masking (noise masking) and/or also by decreasing an amplitude of a sound source relative to other sound sources.

There is another intriguing possibility. The causal relation between global sensory properties and aesthetics could be further captured in terms of the processing fluency hypothesis, as proposed in the domain of visual aesthetics (Reber et al., 2004; Babel and McGuire, 2015; Forster et al., 2015). For instance, if the crucial feature concerns the distinctiveness of each musical source in the absence of feature masking, then it is possible that the phenomenon reduces further to the notion of processing fluency (Reber et al., 2004), namely the relation between a positive aesthetic response and the ease of processing

in encoding and representing e.g., distinct sound sources. This hypothesis predicts that aiding the encoding of music and its sound sources by visual means, for example, by exhibiting the performance itself, should increase the aesthetic appeal of the piece irrespectively of whether the sound sources are overlapping or not. On the other hand, a dull, unsaturated but fluently processed sonic object might be less appealing than a complex one that incorporates the whole frequency and spatial spectrum, an obvious problem for the fluency hypothesis. This problem could be solved by combining the immersion hypothesis with the processing fluency hypothesis. Accordingly, perhaps an aesthetic appreciation requires a concerted and concurrent activation of all the relevant modules that participate in the processing of the stimulus, as assumed by the immersion hypothesis, but with the additional requirement that each module has to be able to process its input in a fluent and efficient manner, as assumed by the processing fluency theory. Under this hypothesis, the masking phenomenon linked with musical aesthetics would be interpreted as a distracting event that hinders fluent processing in any of the relevant submodules.

If we, instead, assume Zeki's hypothesis that there are "significant forms" that, by causing the various sensory submodules to enter their "preferred states," lead into aesthetic appreciation, then a rigorous definition of "significant form" is required. Zeki (2013) discusses the example of human faces in this connection. Humans have an inborn preference for perceiving, representing, and interpreting human faces, and there are specific neuronal resources dedicated to this task. These visual systems respond selectively to the properties of human faces and, moreover, some such features are perceived as more attractive than others. There are biological and evolutionary reasons why such preferences would inhabit our visual system, and the same phenomenon of "mate selection" is observed throughout the animal kingdom. The same argument can be found from several views concerning visual aesthetic, cited earlier in this paper. The idea is that the global sensory properties are shared with the biologically preferred visual images, such as natural landscapes or potential mates, which would then explain artistic preferences as a halo effect of the originally more mundane mechanism. While we do not wish to propose that all aesthetic perception derives from preferred tuning of the various sensory systems for mate selection, landscape detection, or healthy nutrition detection, this view provides a plausible argument for the *existence* of such mechanisms. Preference of certain types of mates, environments, foods, and tastes, for example, is something that our brains must be hardwired to do, although also learning and cultural exposure have an effect, while it is possible that such preferences spill over non-functionally to the perception of many types of objects, and even to abstract objects such as music. This hypothesis could be labeled as the *ecological hypothesis*. It has

REFERENCES

- Alluri, V., Toiviainen, P., Burunat, I., Bogert, B., Numminen, J., and Brattico, E. (2015). Musical expertise modulates functional connectivity of limbic regions during continuous music listening. *Psychomusicology* 25, 443–454. doi: 10.1037/pmu0000124
- Alluri, V., Toiviainen, P., Jaaskelainen, I. P., Glerean, E., Sams, M., and Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* 59, 3677–3689. doi: 10.1016/j.neuroimage.2011.11.019
- Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., et al. (2013). From Vivaldi to Beatles and back: predicting lateralized brain

been pursued in the domain of vision by examining whether global statistical sensory properties of ecological stimuli, such as natural landscapes or faces, lead into aesthetic experience when they are embedded in the context of abstract art objects or other visual stimuli. These experiments could be replicated in the case of music by extracting global statistical sensory properties from ecological sounds (wind, rain, human voice, crying, laughing) and replicating then synthetically in music or in music-type stimuli to determine if their aesthetic value can be modulated.

CONCLUSIONS

We put forward a research agenda for studying holistic qualities of musical objects that likely play an important role in creating an aesthetic response in the listener. We propose that these global features are statistically extracted from the stimuli by our auditory system—or, rather, by some subsystems (McDermott and Simoncelli, 2011; McDermott et al., 2013)—and then passed on to high-level processing, ultimately leading to the main outcomes of a musical experience, namely aesthetic judgment, emotion and conscious liking, or preference (Cela-Conde et al., 2011; Brattico et al., 2013). A shift of paradigm from conventional studies using artificial stimulation, block design, and subtraction analysis methods toward novel naturalistic paradigms with non-conventional analysis methods based on MIR combined with brain time series is called upon to accurately measure and determine the effects of global properties on brain functioning and behavior. We also discussed several possible neuronal implementations of this general hypothesis: the immersion hypothesis, processing fluency hypothesis, and the ecological hypothesis. The immersion hypothesis claims that aesthetic experience results in a concerted activation of many or all critical brain regions involved in the processing of the stimuli, irrespectively of other stimulus content; the processing fluency requires that the stimuli can be processed effortlessly by the brain; and the ecological hypothesis contents that the modules have to enter into a "preferred" neural state that is further determined by ecological conditions. Another possibility is that they all play a role.

AUTHOR CONTRIBUTIONS

PB and EB conceived the hypotheses of this paper. PB wrote most of the manuscript whereas EB wrote some parts of it. PV edited the manuscript and contributed to financing the work.

ACKNOWLEDGMENTS

This work has been funded by the Danish National Research Foundation (project number DNR117).

- responses to music. *Neuroimage* 83, 627–636. doi: 10.1016/j.neuroimage.2013.06.064
- Altmann, C. F., Terada, S., Kashino, M., Goto, K., Mima, T., Fukuyama, H., et al. (2014). Independent or integrated processing of interaural time and level differences in human auditory cortex? *Hear. Res.* 312, 121–127. doi: 10.1016/j.heares.2014.03.009
- Babel, M., and McGuire, G. (2015). Perceptual fluency and judgments of vocal aesthetics and stereotypicality. *Cogn. Sci.* 39, 766–787. doi: 10.1111/cogs.12179
- Barbour, D., and Wang, X. (2003). Contrast tuning auditory cortex. *Science* 299, 1073–1075. doi: 10.1126/science.1080425
- Baugh, B. (1993). Prolegomena to any aesthetics of rock music. *J. Aesthet. Art Crit.* 51, 23–29. doi: 10.2307/431967
- Bell, C. (1914). *Art*. London: Chatto and Windus. All citations are from the Kindle Edition.
- Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York, NY: Appleton-Century-Crofts.
- Berns, G. S., Capra, C. M., Moore, S., and Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage* 49, 2687–2696. doi: 10.1016/j.neuroimage.2009.10.070
- Boccia, M., Sarbetti, S., Piccardi, L., Guariglia, C., Ferlazzo, F., Giannini, A. M., et al. (2016). Where does brain neural activation in aesthetic responses to visual art occur? Meta-analytic evidence from neuroimaging studies. *Neurosci. Biobehav. Rev.* 60, 65–71. doi: 10.1016/j.neubiorev.2015.09.009
- Bogert, B., Numminen-Kontti, T., Gold, B., Sams, M., Numminen, J., Burunat, I., et al. (2016). Hidden sources of joy, fear, and sadness: explicit versus implicit neural processing of musical emotions. *Neuropsychologia* 89, 393–402. doi: 10.1016/j.neuropsychologia.2016.07.005
- Bornstein, R. F. (1989). Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychol. Bull.* 106, 265–289. doi: 10.1037/0033-2909.106.2.265
- Branković, S. (2013). Neuroaesthetics and growing interest in “positive affect” in psychiatry: new evidence and prospects for the theory of informational needs. *Psychiatr. Danub.* 25, 97–107.
- Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., et al. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Front. Psychol.* 2:308. doi: 10.3389/fpsyg.2011.00308
- Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., and Jacobsen, T. (2015). It's sad but I like it: the neural dissociation between musical emotions and liking in experts and laypersons. *Front. Hum. Neurosci.* 9:676. doi: 10.3389/fnhum.2015.00676
- Brattico, E., Bogert, B., and Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Front. Psychol.* 4:206. doi: 10.3389/fpsyg.2013.00206
- Brattico, E., Brattico, P., and Jacobsen, T. (2009–2010). The origins of the aesthetic enjoyment of music - A review of the literature. *Music. Sci.* Special Issue 2009–2010, 15–31. doi: 10.1177/1029864909013002031
- Brattico, E., and Pearce, M. T. (2013). The neuroaesthetics of music. *Psychol. Aesthet. Creat. Arts* 7, 48–61. doi: 10.1037/a0031624
- Bridgett, R. (2013). “Contextualizing game audio aesthetics,” in *The Oxford Handbook of New Audiovisual Aesthetics*, eds J. Richardson, C. Gorbman, and C. Vernallis (Oxford: Oxford University Press), 563–571.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., and Liotti, M. (2011). Naturalizing aesthetics: brain areas for aesthetic appraisal across sensory modalities. *Neuroimage* 58, 250–258. doi: 10.1016/j.neuroimage.2011.06.012
- Bundgaard, H. (2015). Feeling, meaning, and intentionality—a critique of the neuroaesthetics of beauty. *Phenom. Cogn. Sci.* 14, 781–801. doi: 10.1007/s11097-014-9351-5
- Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., et al. (2016). The reliability of continuous brain responses during naturalistic listening to music. *Neuroimage* 124(Pt A), 224–231. doi: 10.1016/j.neuroimage.2015.09.005
- Cela-Conde, C. J., Agnati, L., Huston, J. P., Mora, F., and Nadal, M. (2011). The neural foundations of aesthetic appreciation. *Prog. Neurobiol.* 94, 39–48. doi: 10.1016/j.pneurobio.2011.03.003
- Chatterjee, A. (2011). Neuroaesthetics: a coming of age story. *J. Cogn. Neurosci.* 23, 53–62. doi: 10.1162/jocn.2010.21457
- Chatterjee, A., and Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn. Sci.* 18, 370–375. doi: 10.1016/j.tics.2014.03.003
- Chatterjee, A., and Vartanian, O. (2016). Neuroscience of aesthetics. *Ann. N.Y. Acad. Sci.* 1369, 172–194. doi: 10.1111/nyas.13035
- Chapman, A. J., and Williams, A. R. (1976). Prestige effects and aesthetic experiences: adolescents' reactions to music. *Br. J. Soc. Clin. Psychol.* 15, 61–72. doi: 10.1111/j.2044-8260.1976.tb00007
- Colin, C., Radeau, M., Soquet, A., Dachy, B., and Deltenre, P. (2002). Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clin. Neurophysiol.* 113, 507–518. doi: 10.1016/S1388-2457(02)00028-7
- Corozine, V. (2002). *Arranging Music for the Real World: Classical and Commercial Aspects*. Pacific, MO: Mel Bay.
- Coutinho, E., and Dikken, N. (2012). Psychoacoustic cues to emotion in speech prosody and music. *Cogn. Emot.* 27, 1–27. doi: 10.1080/02699931.2012.732559
- Dalla Bella, S., Peretz, I., Rousseau, L., and Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition* 80, B1–B10. doi: 10.1016/S0010-0277(00)00136-0
- Eklund, A., Nichols, T. E., and Knutsson, H. (2016). Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7900–7905. doi: 10.1073/pnas.1602413113
- Fechner, G. T. (1876). *Vorschule der Aesthetik [Experimental Aesthetics]*. Leipzig: Breitkopf & Härtel.
- Fitch, W. T. (2016). Dance, music, meter and groove: a forgotten partnership. *Front. Hum. Neurosci.* 10:64. doi: 10.3389/fnhum.2016.00064
- Forster, M., Gerger, G., and Leder, H. (2015). Everything's relative? Relative differences in processing fluency and the effects on liking. *PLoS ONE* 10:e0135944. doi: 10.1371/journal.pone.0135944
- Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., and Sawey, M. (2011). Predicting beauty: fractal dimension and visual complexity in art. *Br. J. Psychol.* 102, 49–70. doi: 10.1348/000712610X498958
- Gabriellson, A. (2011). “Emotions in strong experiences with music,” in *Handbook of Music and Emotion: Theory, Research, Applications*, eds P. N. Juslin and J. A. Sloboda (Oxford: Oxford University Press), 547–574.
- Gabriellson, A., and Juslin, P. N. (1996). Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychol. Music* 24, 68–91. doi: 10.1177/0305735696241007
- Graham, D. J., and Field, D. J. (2007). Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. *Spat. Vis.* 21, 149–164. doi: 10.1163/156856807782753877
- Green, A. C., Baerentsen, K. B., Stodkilde-Jørgensen, H., Roepstorff, A., and Vuust, P. (2012). Listen, learn, like! Dorsolateral prefrontal cortex involved in the mere exposure effect in music. *Neurol. Res. Int.* 2012:846270. doi: 10.1155/2012/846270
- Grewe, O., Nagel, F., Kopiez, K., and Altenmüller, E. (2007). Listening to music as a re-creative process: physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Percept.* 297, 297–314. doi: 10.1525/mp.2007.24.3.297
- Hagerhall, C. M., Purcell, T., and Taylor, R. (2004). Fractal dimension of landscape silhouette – outlines as a predictor of landscape preference. *J. Environ. Psychol.* 24, 247–255. doi: 10.1016/j.jenvp.2003.12.004
- Heingartner, A., and Hall, J. V. (1974). Affective consequences in adults and children of repeated exposure to auditory stimuli. *J. Pers. Soc. Psychol.* 29, 719–723. doi: 10.1037/h0036121
- Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. *Psychol. Sci.* 48, 157–172.
- Hevner, K. (1935). The affective character of major and minor modes in music. *Am. J. Psychol.* 47, 103–118. doi: 10.2307/1416710
- Hodges, D. A. (2016). “The neuroaesthetics of music,” in *The Oxford Handbook of Music Psychology*, eds S. Hallam, I. Cross, and M. Thaut (Oxford: Oxford University Press), 247–262.
- Huckvale, D. (1990). Twins of Evil: an investigation into the aesthetics of film music. *Pop. Music* 9, 1–36. doi: 10.1017/S0261143000003718
- Hunter, P. G., Schellenberg, E. G., and Schimmack, U. (2008). Mixed affective responses to music with conflicting cues. *Cogn. Emot.* 22, 327–352. doi: 10.1080/02699930701438145
- Husain, G., Thompson, W. F., and Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Percept.* 20, 151–171. doi: 10.1525/mp.2002.20.2.151

- Istok, E., Brattico, E., Jacobsen, T., Krohn, K., Mueller, M., and Tervaniemi, M. (2009). Aesthetic responses to music: a questionnaire study. *Music. Sci.* 13, 183–206. doi: 10.1177/102986490901300201
- Istok, E., Brattico, E., Jacobsen, T., Ritter, A., and Tervaniemi, M. (2013). 'I love Rock 'n' Roll'-music genre preference modulates brain responses to music. *Biol. Psychol.* 92, 142–151. doi: 10.1016/j.biopsycho.2012.11.005
- Jacobs, R. H., Haak, K. V., Thumfart, S., Renken, R., Henson, B., and Cornelissen, F. W. (2016). Aesthetics by numbers: links between perceived texture qualities and computed visual texture properties. *Front. Hum. Neurosci.* 10:343. doi: 10.3389/fnhum.2016.00343
- Jacobsen, T. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychol. Rep.* 94, 1253–1260. doi: 10.2466/pr0.94.3.1253-1260
- Jacobsen, T. (2006). Bridging the arts and sciences: a framework for the psychology of aesthetics. *Leonardo* 39, 155–162. doi: 10.1162/leon.2006.39.2.155
- Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cereb. Cortex* 19, 2579–2594. doi: 10.1093/cercor/bhp008
- Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol.* 141, 54–75. doi: 10.1037/a0024208
- Johnson, J. K., Chang, C. C., Brambati, S. M., Migliaccio, R., Gorno-Tempini, M. L., Miller, B. L., et al. (2011). Music recognition in frontotemporal lobar degeneration and Alzheimer disease. *Cogn. Behav. Neurol.* 24, 74–84. doi: 10.1097/WNN.0b013e31821de326
- Julesz, B. (1962). Visual pattern discrimination. *IRE Trans. Inform. Theory* 8, 84–92. doi: 10.1109/TIT.1962.1057698
- Juslin, P. N., Sakka, L., Barradas, G., and Liljeström, S. (2016). No accounting for taste? Idiographic models of aesthetic judgment in music. *Psychol. Aesthet. Creat. Arts* 10, 157–170. doi: 10.1037/aca0000034
- Khalifa, S., Schon, D., Anton, J. L., and Liegeois-Chauvel, C. (2005). Brain regions involved in the recognition of happiness and sadness in music. *Neuroreport* 16, 1981–1984. doi: 10.1097/00001756-200512190-00002
- Kilchenmann, L., and Senn, O. (2015). Microtiming in Swing and Funk affects the body movement behavior of music expert listeners. *Front. Psychol.* 6:1232. doi: 10.3389/fpsyg.2015.01232
- Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Front. Psychol.* 2:110. doi: 10.3389/fpsyg.2011.00110
- Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180. doi: 10.1038/nrn3666
- Laeng, B., Eidet, L. M., Sulutvedt, U., and Panksepp, J. (2016). Music chills: the eye pupil as a mirror to music's soul. *Conscious. Cogn.* 44, 161–178. doi: 10.1016/j.concog.2016.07.009
- Lartillot, O., and Toivainen, P. (2007). "A Matlab toolbox for musical feature extraction from audio," in *Paper Presented at the 10th International Conference on Digital Audio Effects (DAFx-07)* (Bordeaux).
- Limb, C. J., and Braun, A. R. (2008). Neural substrates of spontaneous musical performance: an fMRI study of jazz improvisation. *PLoS ONE* 3:e1679. doi: 10.1371/journal.pone.0001679
- Liu, C., Abu-Jamous, B., Brattico, E., and Nandi, A. K. (2017). Towards tunable consensus clustering for studying functional brain connectivity during affective processing. *Int. J. Neural Syst.* 27:1650042. doi: 10.1142/S0129065716500428
- Mallon, B., Redies, C., and Hayn-Leichsenring, G. U. (2014). Beauty in abstract paintings: perceptual contrast and statistical properties. *Front. Hum. Neurosci.* 8:161. doi: 10.3389/fnhum.2014.00161
- Margolis, J. (1980). *Art and Philosophy*. Atlantic Highlands, NJ: Humanities Press.
- McDermott, J. H., and Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron* 71, 926–940. doi: 10.1016/j.neuron.2011.06.032
- McDermott, J. H., Schemitsch, M., and Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nat. Neurosci.* 16, 493–498. doi: 10.1038/nn.3347
- McPherson, G. E., and Schubert, E. (2004). "Measuring performance enhancement in music," in *Musical Excellence: Strategies and Techniques to Enhance Performance*, ed A. Williamon (Oxford: Oxford University Press), 61–82.
- Menzel, C., Hayn-Leichsenring, G. U., Langner, O., Wiese, H., and Redies, C. (2015). Fourier power spectrum characteristics of face photographs: attractiveness perception depends on low-level image properties. *PLoS ONE* 10:e0122801. doi: 10.1371/journal.pone.0122801
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago, IL; London: University of Chicago Press.
- Moffat, D., Ronan, D., and Reiss, J. D. (2015). "An evaluation of audio feature extraction toolboxes," in *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)* (Trondheim).
- Montag, C., Reuter, M., and Axmacher, N. (2011). How one's favorite song activates the reward circuitry of the brain: personality matters! *Behav. Brain Res.* 225, 511–514. doi: 10.1016/j.bbr.2011.08.012
- Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing, 6th Edn*. Bingley: Emerald.
- Mureika, J. R., Cupchik, G. C., and Dyer, C. C. (2004). Multifractal fingerprints in the visual arts. *Leonardo* 37, 53–56. doi: 10.1162/00240940472828139
- Nieminen, S., Istók, E., Brattico, E., and Tervaniemi, M. (2012). The development of the aesthetic experience of music: preference, emotions, and beauty. *Music. Sci.* 16, 372–391. doi: 10.1177/1029864912450454
- Nieminen, S., Istok, E., Brattico, E., Tervaniemi, M., and Huottilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex* 47, 1138–1146. doi: 10.1016/j.cortex.2011.05.008
- Orgs, G., Hagura, N., and Haggard, P. (2013). Learning to like it: aesthetic perception of bodies, movements and choreographic structure. *Conscious. Cogn.* 22, 603–612. doi: 10.1016/j.concog.2013.03.010
- Pallesen, J. K., Brattico, E., and Carlsson, S. (2003). Emotional connotations of major and minor musical chords in musically untrained listeners. *Brain Cogn.* 51, 188–190. doi: 10.1016/S0278-2626(02)00541-9
- Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Percept.* 13, 171–207. doi: 10.2307/40285693
- Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., et al. (2016). Neuroaesthetics: the cognitive neuroscience of aesthetic experience. *Perspect. Psychol. Sci.* 11, 265–279. doi: 10.1177/1745691615621274
- Peeters, G. (2004). *A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project*. Technical Report, IRCAM.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., and Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PLoS ONE* 6:e27241. doi: 10.1371/journal.pone.0027241
- Peretz, I., Gaudreau, D., and Bonnel, A.-M. (1998). Exposure effects on music preference and recognition. *Mem. Cogn.* 26, 884–902. doi: 10.3758/BF03201171
- Plomp, R. (1964). The ear as a frequency analyzer. *J. Acoust. Soc. Am.* 36, 1628–1636. doi: 10.1121/1.1919256
- Poikonen, H., Alluri, V., Brattico, E., Lartillot, O., Tervaniemi, M., and Huottilainen, M. (2016a). Event-related brain responses while listening to entire pieces of music. *Neuroscience* 312, 58–73. doi: 10.1016/j.neuroscience.2015.10.061
- Poikonen, H., Toivainen, P., and Tervaniemi, M. (2016b). Early auditory processing in musicians and dancers during a contemporary dance piece. *Sci. Rep.* 6:33056. doi: 10.1038/srep33056
- Porcello, T. (2004). Speaking of sound: language and the professionalization of sound-recording engineers. *Soc. Stud. Sci.* 34/5, 733–758. doi: 10.1177/0306312704047328
- Račić, L. (1981). On the aesthetics of rock music. *Int. Rev. Aesthet. Sociol.* 12, 199–202. doi: 10.2307/836562
- Ramachandran, V. S., and Hirstein, W. (1999). The science of art - A neurological theory of aesthetic experience. *J. Conscious. Stud.* 6, 15–51.
- Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3
- Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spat. Vis.* 21, 97–117. doi: 10.1163/156856807782753886
- Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Front. Hum. Neurosci.* 9:218. doi: 10.3389/fnhum.2015.00218
- Renoult, J. P., Bovet, J., and Raymond, M. (2016). Beauty is in the efficient coding of the beholder. *R. Soc. Open Sci.* 3:160027. doi: 10.1098/rsos.160027
- Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychol. Music* 32, 371–388. doi: 10.1177/0305735604046096

- Robinson, J. (1994). The expression and arousal of emotion in music. *J. Aesthet. Art Crit.* 52, 13–22. doi: 10.2307/431581
- Roeber, U., Widmann, A., and Schroger, E. (2003). Auditory distraction by duration and location deviants: a behavioral and event-related potential study. *Brain Res. Cogn. Brain Res.* 17, 347–357. doi: 10.1016/S0926-6410(03)00136-8
- Sachs, M. E., Ellis, R. J., Schlaug, G., and Loui, P. (2016). Brain connectivity reflects human aesthetic responses to music. *Soc. Cogn. Affect. Neurosci.* 11, 884–891. doi: 10.1093/scan/nsw009
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14, 257–262. doi: 10.1038/nn.2726
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., and Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* 340, 216–219. doi: 10.1126/science.1231059
- Schellenberg, E. G., Peretz, I., and Viellard, S. (2008). Liking for happy- and sad-sounding music: effects of exposure. *Cogn. Emot.* 22, 218–237. doi: 10.1080/02699930701350753
- Seashore, C. E. (1929). A base for the approach to quantitative studies in the aesthetics of music. *Am. J. Psychol.* 39, 141–144. doi: 10.2307/1415406
- Senior, M. (2011). *Mixing Secrets for the Small Studio*. Amsterdam: Elsevier. Kindle Edition.
- Sioros, G., Miron, M., Davies, M., Gouyon, F., and Madison, G. (2014). Syncopation creates the sensation of groove in synthesized music examples. *Front. Psychol.* 5:1036. doi: 10.3389/fpsyg.2014.01036
- Sloboda, J. A. (1991). Music structure and emotional response: some empirical findings. *Psychol. Music* 19, 110–120. doi: 10.1177/0305735691192002
- Spehar, B., Clifford, C., Newell, B., and Taylor, R. P. (2003). Universal aesthetic of fractals. *Comput. Graph.* 37, 813–820. doi: 10.1016/S0097-8493(03)00154-7
- Spehar, B., Wong, S., van de Klundert, S., Lui, J., Clifford, C. W., and Taylor, R. P. (2015). Beauty and the beholder: the role of visual sensitivity in visual preference. *Front. Hum. Neurosci.* 9:514. doi: 10.3389/fnhum.2015.00514
- Taylor, R. P., Micolich, A. P., and Jonas, D. (1999). Fractal analysis of Pollock's drip paintings. *Nature* 399, 422–422. doi: 10.1038/20833
- Taylor, R. P., Spehar, B., Van Donkelaar, P., and Hagerhall, C. M. (2011). Perceptual and physiological responses to Jackson Pollock's fractals. *Front. Hum. Neurosci.* 5:60. doi: 10.3389/fnhum.2011.00060
- Trost, W., Fruhholz, S., Cochrane, T., Cojan, Y., and Vuilleumier, P. (2015). Temporal dynamics of musical emotions examined through intersubject synchrony of brain activity. *Soc. Cogn. Affect. Neurosci.* 10, 1705–1721. doi: 10.1093/scan/nsv060
- Vartanian, O., and Goel, V. (2004). Neuroanatomical correlates of aesthetic preference for paintings. *Neuroreport* 15, 893–897. doi: 10.1097/01.wnr.0000118723.38067.d6
- von Appen, R. (2007). On the aesthetics of popular music. *Music Therapy Today* 8, 5–25.
- Vuust, P., and Witek, M. A. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Front. Psychol.* 5:1111. doi: 10.3389/fpsyg.2014.01111
- Vuust, P., Gebauer, L. K., and Witek, M. A. (2014). “Neural underpinnings of music: the polyrhythmic brain,” in *Neurobiology of Interval Timing, Advances in Experimental Medicine and Biology* 829, eds H. Merchant and V. de Lafuente (New York, NY: Springer), 339–356.
- Witek, M. A., Clarke, E. F., Wallentin, M., Kringelbach, M. L., and Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE* 9:e94446. doi: 10.1371/journal.pone.0094446
- Wurtz, R. H., and Kandel, E. R. (2000). “Central visual pathways,” in *Principles of Neural Science, 4th Edn.* eds E. R. Kandel, J. H. Schwartz, and T. M. Jessell (New York, NY: McGraw-Hill), 523–547.
- Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *J. Acoust. Soc. Am.* 111, 1832–1846. doi: 10.1121/1.1458027
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7
- Zeki, S. (1999). *Inner Vision - An Exploration of Art and the Brain*. Oxford: Oxford University Press.
- Zeki, S. (2013). Clive Bell's “Significant Form” and the neurobiology of aesthetics. *Front. Hum. Neurosci.* 7:730. doi: 10.3389/fnhum.2013.00730
- Zeki, S. (2014). Neurobiology and the humanities. *Neuron* 84, 12–14. doi: 10.1016/j.neuron.2014.09.016
- Zuckerman, M., Levy, D. A., Tibon, R., Reggev, N., and Maril, A. (2012). Does this ring a bell? Music-cued retrieval of semantic knowledge and metamemory judgments. *J. Cogn. Neurosci.* 24, 2155–2170. doi: 10.1162/jocn_a_00271
- Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands. *J. Acoust. Soc. Am.* 33, 248. doi: 10.1121/1.1908630

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Brattico, Brattico and Vuust. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.