The increasing intensive study of music by neuroscientists over the past two decades has established the neurosciences of music as a subdiscipline of cognitive neuroscience, responsible for investigating the neural basis for music perception, cognition, and emotion. In this endeavor, music perception and cognition have often been compared with language processing and understanding, while music-induced emotions are compared with emotions induced by visual stimuli. Here, we review research that is beginning to define a new field of study called neuroaesthetics of music. According to this fresh perspective, music is viewed primarily as an expressive art rather than as a cognitive domain. The goal of this emerging field is to understand the neural mechanisms and structures involved in the perceptual, affective and cognitive processes that generate the three principal aesthetic responses: emotions, judgments, and preference. Although much is known about the frontotemporal brain mechanisms underlying perceptual and cognitive musical processes, and about the limbic and paralimbic networks responsible for musical affect, there is a great deal of work to be done in understanding the neural chronometry and structures determining aesthetic responses to music. Research has only recently begun to delineate the modulatory effects of the listener, listening situation, and the properties of the music itself on a musical aesthetic experience. This article offers a review and synthesis of our current understanding of the perceptual, cognitive, and affective processes involved in an aesthetic musical experience and introduces a novel framework to coordinate future endeavors in an emerging field.

**Keywords:** music, auditory cortex, cognitive neuroscience, pitch, rhythm

---

### Overture: From Neuroscience to Neuroaesthetics of Music

People value music primarily for aesthetic reasons: for the emotions it generates, for triggering memories, and for its beauty (Juslin & Laukka, 2004; Laukka, 2007; McDonald & Stewart, 2008). Just as in other aesthetic domains such as visual art, architecture, or dance (see also Leder, Belke, Oeberst, & Augustin, 2008), listening to and performing music generates, in concert with a favorable environment and listening situation, aesthetic experiences that include specific emotions and evaluative judgments of beauty, aesthetic quality, and liking. The question of how music generates an aesthetic experience has been addressed with scientific methods since the dawn of experimental psychology. A new era of empirical work on musical perception began with Helmholtz (1863/1985) who, inspired by Hanslick (1854/1954), associated the aesthetic qualities of musical notes and scales with their psychoacoustic properties (especially frequency ratios between partials of complex tones). Wundt, a founder of experimental psychology and one-time assistant of Helmholtz in Heidelberg, developed a more psychological approach to the study of aesthetics, introspecting, for example, about how his own sensations of pleasure, tension, and excitement varied with the tempo of a metronome (see Miller & Bukhout, 1973). Elsewhere, Wundt demonstrated that physiological arousal is related to stimulus complexity and argued that aesthetic pleasure is maximal at intermediate degrees of complexity (Wundt, 1874). In his *new experimental aesthetics*, Berlyne (1971) developed this idea into an inverted U-shaped function linking the “arousal potential” of a stimulus with its “hedonic value,” such that intermediate degrees of arousal correspond to maximum pleasure and attempted to identify how stimulus properties (such as complexity, familiarity, novelty, uncertainty) influence aspects of the aesthetic experience such as arousal, pleasure, and interestingness.
Although the neurosciences of music can be now considered an autonomous subdiscipline of cognitive neuroscience (Levitin & Tirovolas, 2009; Peretz & Zatorre, 2003), the neuroaesthetics of music remains relatively undeveloped. For example, many experiments have been conducted on the neural effects of musical expertise on perceptual and cognitive skills (e.g., Brattico et al., 2009; Kraus & Chandrasekaran, 2010; Pallesen et al., 2010; Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011; Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001), but only two such studies have focused on aesthetic or affective judgments (Müller, Höfel, Brattico, & Jacobsen, 2010; Brattico et al., in preparation). By contrast, neuroscience and psychological research on the visual perception of artistic stimuli seems to have given equal weight to cognitive processing (e.g., Wiesmann & Ishai, 2011, for expertise effects in perception of Cubist paintings) and aesthetic and affective experience (e.g., Kirk, Skov, Christiansen, & Nygaard, 2009, for expertise effects on aesthetic preference for architecture). We suggest two reasons for this asymmetry: first, most musical experiences occur in nonaesthetic situations, leading to a focus on basic goal-oriented emotions (Jacobsen, 2009; Sloboda, 2010); second, interest in the relationship with language processing in the neurosciences has led to a cognitive rather than aesthetic approach to musical experience (i.e., a focus on the cognitive representations and processing of musical structure rather than the affective and aesthetic experiences that often result from this processing). Even in the psychology of music, “Psychologists have tended to avoid studying aesthetic response to music . . . [and] rather focused on more mundane aspects such as preference” (Juslin, Liljestrom, Västfjäll, & Lundqvist, 2010, p. 635). Here we view preference as an important, possibly even necessary though not sufficient, component of the aesthetic experience of music.

Research has found that beauty is an important part of most people’s understanding of musical aesthetics, for music experts and nonexperts (Istok et al., 2009), and for 6- to 9-year-old children (Nieminen et al., 2011) replicating results for visual art (Jacobsen, 2004). Drawing on proposals by Brattico, Bogert, and Jacobsen (in press); Juslin et al. (2010); Sloboda (2010); Hargreaves and North (2010); and Leder et al. (2004), we define an aesthetic experience of music as one in which the individual immerses herself in the music, dedicating her attention to perceptual, cognitive, and affective interpretation based on the formal properties of the perceptual experience. We identify here three main outcomes: first, emotion recognition (e.g., “this song is sad”) and induction (e.g., “I feel nostalgic”); second, aesthetic judgment (e.g., “this song is beautiful”); and third, liking (e.g., “I like this song”) and preference (e.g., “I love rock & roll”). As noted by Juslin et al. (2010), not all of these outcomes may be present but typically they combine to form a genuine aesthetic situation: The presence of a music-induced emotion is not sufficient for an aesthetic experience (p. 636). We examine the kinds of emotion induced by music before reviewing research on the different psychological and neural mechanisms involved in the generation of emotion by music. In doing so, we emphasize the need to take account not only of the properties of the music, but also the listener and listening situation on the aesthetic experience (Hargreaves & North, 2010).

The relevance of a scientific approach to the aesthetics of music is debated. While some philosophers of aesthetics have embraced psychological and neuroscientific research (e.g., Carroll, Moore, & Seeley, 2011; Meyer, 1956; Robinson, 2005), others argue that neuroscience has nothing to tell us about aesthetic questions (e.g., Tallis, 2008). Equally, although some scientists argue that “aesthetic philosophy is receding to a sideline ‘advisory’ role, while cognitive science takes an accustomised leadership position” (Huron, 2010, p. 151), here we try to build an interdisciplinary framework for delineating a neuroaesthetics of music that explicitly includes philosophical, psychological, neuroscientific, and evolutionary approaches.

Emotions in Musical Experience

Increasing neuroscientific interest in affective processes has spread to the neurosciences of music. However, research to date has been limited primarily to the most common emotions encountered also in everyday life, such as happiness or sadness, and their role in mood regulation. Indeed, musical emotions are either interpreted with reference to the categorical framework of “basic” emotions, supposedly universal affective states, punculturally recognized and necessary for species survival (Ekman, 1999; Peretz, 2010), or to general dimensional models of emotion (e.g., Russell, 1980; Schimmack & Grob, 2000), in which emotions are locations within a continuous 2-D or 3-D affective space. It has been suggested, however, that music generates emotions of a special kind, qualitatively different from goal-oriented, common (or utilitarian) emotions (Scherer & Zentner, 2008), although there is little neuroscientific research to date on such aesthetic emotions.

Basic Emotions

Much research on music and emotion has been inspired by work on the categorical perception of facial emotion (Ekman, 1999). This framework identifies basic (or primary) emotions such as happiness, sadness, anger, fear, and disgust, which, it is argued, are punculturally recognized and are associated with innate motor and physiological responses. Music may express and induce these basic emotions in individuals of all ages, including infants, and across many cultures (Peretz, 2010), although within a safe aesthetic environment, the negative emotions lose some of their aversive character (Juslin & Västfjäll, 2008).

The amygdala plays a central role in the processing of salient negative emotions, fear in particular, induced by aversive stimuli. The amygdala also seems to be a crucial brain structure for perception and recognition of fear in music, because patients with medial temporal resection encompassing the amygdala, and one patient with bilateral amygdala damage, confused scary music with peaceful music while showing intact perceptual consonance (Gosselin et al., 2007; 2005). The amygdala is activated by sad and dissonant music (contrasted with emotionally neutral and consonant music, respectively; Koelsch, Fritz, von Cramon, Muller, & Friederici, 2006; Mitterschiffthaler, Fu, Dalton, Andrew, & Williams, 2007; Trost, Ethofer, Zentner, & Vuilleumier, 2012) and even by single unpredictable chords (Koelsch, Fritz, & Schlaug, 2008), indicating that further research is needed to isolate the role of the amygdala in musical listening (Koelsch, 2010).

Sad emotions associated with slow minor classical piano pieces activated the left medial frontal gyrus (BA 10) and the adjacent superior frontal gyrus (BA 9) when compared with happy major
and fast pieces (Kalfia, Schon, Anton, & Liegeois-Chauvel, 2005). These regions were also activated during aesthetic judgment of pictures (Jacobsen, Schubotz, Hofel, & Cramon, 2006) and rhythms (Kornsheva, von Cramon, Jacobsen, & Schubotz, 2010), suggesting that their activation might reflect subjective preference for minor pieces rather than sadness per se. In a functional MRI (fMRI) study of emotional responses to classical music, Mitter-schiffthaler et al. (2007) found no regions activated by sad compared with happy music, but did find activation in the hippocampus/amygdala for sad compared with neutral music. Again, this study did not control for the effects of preference.

Turning now to positive emotions such as happiness, Mitter-schiffthaler et al. (2007) found that happy classical music induced activation of the ventral striatum compared with neutral music. Compared with sad music, happy music activated left superior temporal gyrus (BA 22), a region of the nonprimary auditory cortex, devoted to integrating sounds over longer time spans as opposed to the primary auditory cortex, and hence processing more abstract aspects of sounds. In a recent study, Brattico et al. (2011) confirmed these findings in music from several genres (pop, jazz, classical) and also found activation of the limbic insula, adjacent to the superior temporal gyrus. A contrast between sad music and happy music indicated recruitment of the caudate nucleus, responsible for the preparation of the chill response according to Salimpoor, Benovoy, Larcher, Dagher, and Zatorre (2011), and of the left thalamus, associated in many previous studies with processing of sad facial expressions (Fusar-Poli et al., 2009). In sum, musical emotions activate some brain structures previously associated with emotions induced by visual stimuli: this partial overlap in neural processing may facilitate crossmodal transfer of emotions, such as the modulation of neural processing of neutral faces during listening to happy music (Logeswaran & Bhattacharya, 2009).

**Dimensional Models of Emotion**

Dimensional models attempt to identify a series of dimensions capable of representing all possible emotional states. In theory, the dimensional structure should be rich enough to represent the basic emotions as points in the space. The most widely known dimensional model of emotion is the circumplex model (e.g., Russell, 1980, 2003), which distinguishes valence (pleasure-displeasure) and arousal (activating-relaxing) as two orthogonal dimensions of an emotional experience. This model has been applied to music in many behavioral and neuroscientific studies (e.g., Krumhansl, 1997; North & Hargreaves, 1997; Schmidt & Trainor, 2001; Vieillard et al., 2008). A variant on this approach (Schimmack & Grob, 2000) includes the three dimensions of valence (pleasant-unpleasant), arousal (awake-tired), and tension (tense-relaxed). In applying this model to music, Ilie and Thompson (2006) found that loudness and tempo increased judgments of arousal and tension while loudness and pitch height increased pleasantness. In a study of excerpts of film music, Eerola and Vuoskoski (2011) found that participants were able to correctly discriminate examples differing in the three dimensions, and noted that removing tension did not impair the fit of the model. Eerola and Vuoskoski also found that the categorical model of basic emotions (Krumhansl, 1997) is inferior to the dimensional model (including energy, tension/ arousal and valence as dimensions) in characterizing emotionally ambiguous examples.

In an fMRI study of emotional responses to excerpts of classical music, Trost et al. (2012) report differences in arousal and valence were reflected by changes of activation in the reward and limbic system (including the striatum, ventral tegmental area, and orbitofrontal cortex for valence and the ventromedial prefrontal cortex and the subgenual cingulate for arousal), with additional effects in brain areas related to memory, motor control, and self-reflective processes for musical arousal (Trost et al., 2012). There is further evidence for the role of physiological arousal in the aesthetic experience as proposed by Berlyne (1971). Increases in electrodermal activity, generated by the sympathetic autonomic nervous system, are greater while listening to energetic than relaxing music (Kalfia, Peretz, Blondin, & Robert, 2002) and by unexpected compared with expected chords (Steinbeis, Koelsch, & Sloboda, 2006). According to Hargreaves and North (2010), the arousal level of the autonomic nervous system predicts liking for music and leads to finer grained emotional responses related to the listener’s engagement with the music (e.g., feeling excited, bored or unsettled). However, these predictions await experimental neuroscientific research.

In view of the ongoing debate about the extent to which dimensional models of emotion can encompass categorical emotions (Eerola & Vuokskoski, 2011; Gosselin et al., 2007; Kalfia et al., 2008; Vieillard et al., 2008), the challenge for proponents of these models is to identify the smallest number of dimensions capable of representing all distinguishable emotional states. It is not clear, for example, that the circumplex model is capable of sufficiently distinguishing closely related emotions such as anger and fear. In an aesthetic context, one difficulty with existing dimensional models is that they equate valence with pleasure, which would seem to preclude the possibility of finding a frightening or sad (negative valenced) experience of music enjoyable. Schubert (1996) proposes that we enjoy negative emotions in music by inhibiting displeasure thanks to the “safe” aesthetic context that nullifies the possible dangerous consequences associated with real everyday negative emotions (see also Huron, 2006). Therefore, we suggest using valence to indicate positive and negative affective character and including pleasure or enjoyment as an extra dimension (see also the discussion of aesthetic emotions below). We also propose a finer distinction between different types of pleasure: immediate sensory pleasure and a more reflective process of enjoyment although empirical research has yet to tease apart these two different kinds of pleasure.

As a final comment, we note that dimensional models are fundamentally unable to accommodate the possibility of mixed emotions combining the extremes of a single dimension, such as feeling happy and sad at the same time (Hunter, Schellenberg, & Schimmack, 2008).

**Aesthetic Emotions**

Scherer and Zentner (2008) distinguish between aesthetic and utilitarian emotions: aesthetic emotions differ from their utilitarian counterparts by occurring in situations that do not trigger self-interest or goal-directed action. They argue that music-induced emotion reflects a multiplicative function of structural features of the music, listener features, performer features and contextual features leading to distinct kinds of emotion such as wonder, transcendence, entrainment, tension and awe. In a series of related
In music neuroscience, the aesthetic emotion of enjoyment has been investigated by focusing on the chill reaction. Possibly the most thoroughly studied aesthetic experience of music, chills correspond to physiological changes such as goose bumps and shivers down the spine, also referred to as frisson or thrills. Although not everyone experiences chills during musical listening or playing, those who experience them do so relatively frequently and reliably (Panksepp, 1995; Sloboda, 1992). In addition to being easy to record behaviorally, chills have the additional advantage of producing characteristic physiological markers including changes in heart rate, breath depth, and skin conductance (e.g., Blood & Zatorre, 2001). People who score highly on openness tend to experience chills to music (McCrae, 2007; Nusbaum & Silvia, 2011), and this effect appears to be mediated by the degree to which people listen to and value music in their everyday lives (Nusbaum & Silvia, 2011). However, age, gender, and music education appear to have no influence on the experience of chills (Grewe, Kopiez, & Altenmueller, 2009). Turning to the experience itself: chills are associated with increased subjective emotion and physiological arousal (Grewe et al., 2009) and are usually experienced as highly pleasurable (Goldstein, 1980; Panksepp, 1995; Sloboda, 1992).

In a pioneering study, Blood and Zatorre (2001) attempted to determine the neural correlates of the chill experience by asking musically trained subjects to bring their chill-evoking music to the lab and using the pieces of other subjects as the control stimulus. When correlating the metabolic brain responses measured by positron emission tomography (PET) with ratings of emotional intensity, it was found that the strongest chill responses activated the bilateral insula, the left ventral striatum (including the nucleus accumbens), the right orbitofrontal cortex (BA 14), the medial anterior cingulate and supplementary motor area (BA 6), the right thalamus and the left midbrain whereas they down-regulated the right amygdala, the left anterior hippocampus/amygdala formation, and the bilateral medial prefrontal cortex (BA 10/32). Indeed, highly pleasant, familiar music enhances connectivity between the ventral tegmental area and the nucleus accumbens, and between this latter area and the hypothalamus, pointing toward pleasure-related responses in the autonomic nervous system (Blum et al., 2010; Menon & Levitin, 2005; Salimpoor et al., 2011; Sutoo & Akiyama, 2004). Salimpoor et al. (2011) have recently shown that chills are associated with dopamine release in the ventral striatum and with activation of the nucleus accumbens while the caudate nucleus is activated during anticipation of a passage of music inducing chills.

Further evidence for the link between dopamine release and intense musical pleasure comes from the reduction of the chill reaction in music by naloxone, an opioid antagonist, whose transmission in the nucleus accumbens is associated with dopamine release in the ventral tegmental area (Goldstein, 1980). In addition, exposure to Mozart’s music in rats improves dopaminergic transmission in the neostriatum, as indicated by decreased systolic blood pressure (Sutoo & Akiyama, 2004). The ventral striatum of the basal ganglia, including the nucleus accumbens, which controls dopamine release in the ventral tegmental area, is associated with reward, pleasure and motivation derived from primary activities necessary for survival (e.g., eating, sex), and plays a central role in the transition to habitual drug use (Haber, 2009). The fact that such ancient survival-related circuitry is modulated so efficiently by an abstract stimulus such as music argues for the adaptive evolutionary status of music making (Brattico, Brattico, & Jacobsen, 2009–2010; Cross, 2003; Huron, 2003; Justus & Hutsler, 2005). It has been claimed that music has beneficial functions in cohesion, mother–infant interaction and mate choice (Wallin, Meyer, & Brown, 2000) but one cannot reject the alternative hypotheses that music appreciation is a spandrel exapted from a collection of abilities originally adapted for other reasons or else that it originated through mechanisms of biological evolution, such as genetic drift, gene flow, or nonrandom mating, not entailing any adaptive function (Brattico et al., 2009–2010).

Despite the physiological and evolutionary salience of a central aesthetic emotion such as musical pleasure, Juslin and colleagues (2010) argue against the existence of aesthetic emotions claiming: first, that defining them as emotions associated with art is not informative; and second, that defining them as lacking goals and action-oriented coupling does not make them unique to music. Similarly, Koelsch (2010; see also Koelsch, Offermanns, & Franzke, 2010) describes music-induced emotions as “real” emotions because they activate the same brain structures that are involved in everyday affective states and rejects their association with aesthetic experiences “lacking motivational components and goal relevance” (p. 131). According to Juslin and Västfjäll (2008), music may evoke emotions more frequently than other kinds of events, but the emotions themselves are indistinguishable from everyday emotions.

To address these questions, recent empirical research has compared the GEMS model with the dimensional model. In a study of emotional responses to film music, Vuoskoski and Eerola (2011) found that a dimensional model (Schimmack & Grob, 2000) showed better consistency among participants and was better able to discriminate the musical excerpts than the GEMS model. Therefore, the GEMS model, which was constructed using classical instrumental music, may not reflect the aesthetic experiences of listeners to other musical styles. In a recent fMRI study of indi-
viduals listening to classical music, Trost et al. (2012) found evidence for grouping the 9 GEMS factors into higher-order affective dimensions distinguished by arousal and valence, with some evidence for a finer categorization (e.g., vitality, unease, sublimity).

Research on the aesthetic emotions, therefore, presents a somewhat mixed picture. Following Sloboda (2010), we suggest that musical experiences in everyday contexts can induce basic emotions, such as sadness, happiness, and fear (which sometimes reach intensities comparable to those triggered by life events), while musical experiences in aesthetic situations as defined above can generate special kinds of emotion that are distinct from the other aesthetic outcomes (judgment and preference) but interact with them to produce an aesthetic experience. Key challenges for research in neuroaesthetics are to recreate the right experimental conditions for inducing a genuine aesthetic experience and to understand which minimal set of dimensions are required to represent adequately both utilitarian and aesthetic emotional responses to music.

**How Does Music Generate Emotions?**

The relationship between music and emotion is “the much-vexed question that has been at the centerpiece of musical aesthetics since . . . the late 18th century” (Kivy, 2006, p. 288). The rise of instrumental music and subsequently program music in 18th- and 19th-century Europe generated a debate between referentialists, such as Hegel and Wagner, and formalists, such as Hanslick and Stravinsky. The debate centered on the question of whether music possesses referential content such that musical structures designate nonmusical entities such as physical objects, individuals, or feelings. This prompted Hanslick (1954) to argue that the aesthetic function of music is not to induce emotion: music cannot represent definite feelings (which have objects) because it cannot represent the thoughts that support these feelings; it can represent dynamic changes in intensity of such feelings but not as properties of specific emotions since other phenomena also share such dynamic changes. Hanslick argues instead that the aesthetic effects of music are specifically musical:

The most significant factor in the mental process which accompanies the comprehending of a musical work and makes it enjoyable . . . is the mental satisfaction which the listener finds in continuously following and anticipating the composer’s designs, here to be confirmed in his expectations, there to be agreeably led astray. It goes without saying that this mental streaming . . . occurs unconsciously and with the speed of lightning. Only such music as brings about and rewards this mental pursuing . . . will provide fully artistic satisfaction. (Hanslick, 1954, p. 64)

Meyer (1956) argued that this debate had been founded on a spurious association of referentialism with expressionism: the fact that referentialists tend to be expressionists does not render expressionism incompatible with formalism. Meyer sought to develop a formalist account of emotional expression in music founded on the psychological process of expectation which creates patterns of tension and resolution, which, in turn, generate affective states such that violations of expectation are negatively valenced, indicating predictive failure (Meyer, 1956, p. 27).

In their review, Juslin and Västfjäll (2008) emphasize the importance of distinguishing emotion perception, where a listener perceives or recognizes emotions expressed in the music, and emotion induction, where music evokes an emotion in the listener. They go on to identify six psychological processes supporting the induction of emotions by music. First, brain-stem reflexes (originating from areas such as the inferior colliculus) mediate the induction of arousal by sudden, loud, dissonant or rapidly pulsing sounds. Second, in evaluative conditioning, music can induce emotion through association, as a conditioned stimulus, with an aversive or rewarding stimulus. Third, musical structures may induce emotion through emotional contagion, by mimicking other means of emotional expression such as language, posture and gait. Fourth, music may invoke emotions through the use of structures with close external referents in the sensorium thereby evoking visual imagery (e.g., a storm). Fifth, music may evoke emotion through the intermediary of an episodic memory associated with the music (“Darling, they’re playing our tune”). Finally, the generation and violation of expectations can induce experiences of tension, release, surprise, and uncertainty.

Here we focus on the psychological mechanisms most specific to musical aesthetic experience, and also most studied by neuroscientists: brainstem mechanisms (such as those producing the dissonance sensation), emotional contagion or imitation, and expectation.

**Sensory Dissonance**

Of the brainstem mechanisms discussed by Juslin and Västfjäll (2008), most attention has been paid to sensory consonance and dissonance, which have long been used by composers and musicians in Western and non-Western cultures to manipulate aesthetic responses to music. Two sounds played simultaneously are dissonant, experienced as beating amplitude modulation or roughness, when their physical distance is smaller than two-thirds of the critical bandwidth stimulating neighboring hair cells in the basilar membrane and causing neurons in the cochlear nucleus and brainstem to fire without properly resolving the two sounds (Kameoka & Kuriyagawa, 1969; Peretz, 2010; Plomp & Levelt, 1965). On reaching the primary auditory cortex, the signal causes neurons to resonate at the beat frequency (Fishman et al., 2001), generating more neuronal activity than consonant sounds (Brattico, Näätänen, Verma, Välimäki & Tervaniemi, 2000; Brattico et al., 2009; Schön, Regnault, Ystad, & Besson, 2005). These sensory responses to dissonant sounds are coupled with an affective experience of irritation, whose neural basis seems to rely on the parahippocampal gyrus, a brain region responsible for withdrawal behavior, and the amygdala, associated with salience and negative affect (Blood et al., 1999; Gosselin et al., 2006; Koelsch et al., 2006). It is interesting to note that a clear-cut lateralization of the parahippocampal gyrus during listening to affective classical music was observed by Trost et al. (2012): the left parahippocampal gyrus was recruited by highly arousing music whereas the right parahippocampal gyrus was activated by tender and nostalgic music with low arousal. In turn, the soothing sensation of consonance is usually described as an absence of dissonance, but some identify it as an active process involving reward centers in the brainstem and ventral striatum (Blood et al., 1999; Braun, 1999; Koelsch et al., 2006; Tramo, Cariani, Delgutte, & Braida, 2003). A motor circuit, including the rolandic operculum, probably related to the automatic impulse to imagine singing during pleasant music,
is also activated while nonmusicians listen to consonant music (Koelsch et al., 2006).

**Musical Reference and Imitation**

Hence . . . it becomes possible for motion in music to imitate the peculiar characteristics of motive forces in space . . . And on this, as I believe, essentially depends the power of music to picture emotion. (Helmholtz, 1985, p. 370)

By virtue of imitation (emotional contagion and visual imagery) in Juslin & Västfjäll, (2008), musical structures can denote or refer to external entities of two kinds: first, physical objects and events; second, subjective states such as emotions via their behavioral effects. The former include overt references to birdsong, thunderstorms, and the like, although this form of referential semantics is severely limited in scope to phenomena clearly identified by some auditory pattern (there is no musical structure which could unambiguously denote a castle, for example). For similar reasons, it is difficult for a musical structure to unambiguously denote an emotional state (although see the section on expectations below). However, by virtue of its temporal and spatial structure (pitch is often expressed and understood in spatial terms), a piece of music can exhibit characteristics that refer, more or less unambiguously, to behavioral expressions of emotion. Perhaps the best-known advocate of this approach is Kivy (1989), who argues that instrumental music can be expressive of specific emotions by imitating aspects of emotional speech, gesture, facial expression, gait, and so on. Thus, a slow, ponderous piece of music can be expressive of sadness, while a fast piece with high, ascending melodic contours can be expressive of lighter emotions. Other authors such as Davies (1994) and Langer (1953), also propose imitation theories of musical expression.

Clarke (2005) discusses apparent motion resulting from changes in musical structure from the perspective of Gibsonian affordances (Gibson, 1979), and several neuroimaging studies have reported activation of brain structures controlling movements or imagining actions while listening to music, especially when it is pleasurable (Alluri et al., 2012; Koelsch et al., 2006; Pereira et al., 2011). Strikingly, Koelsch, Kasper, Sammler, Schulze, Gunter, and Friederici (2004) report a semantic priming task with target words presented visually after spoken sentences or a musical excerpt, which showed comparable late negative electrophysiological responses (N400) to target words following semantically unrelated linguistic or musical contexts, demonstrating that music may imitate some qualities of objects (e.g., ascending pitch scales for staircases), contain gestural cues (e.g., a cry) or describe a concept (such as hero for a symphony by Beethoven). These findings have been replicated for single chords varying in pleasantness (Steinbeis & Koelsch, 2008) and for single sounds varying in timbre (Grieser-Painter & Koelsch, 2010).

A formalist such as Hanslick would question the specificity with which these imitative characteristics of music can relyably and exclusively express particular well-defined emotional states and Kivy himself no longer defends this particular theory of how music acquires its expressive properties (Kivy, 2002). Acknowledging that not all examples of musical expression can be explained by imitation, Kivy (1989) proposes a second mechanism based on convention. Thus, in Western tonal music the association of the minor mode with negative valence, and the activation in brain structures classically associated with negative visual emotions such as the amygdala, thalamus and brain stem (Brattico et al., 2011; Green et al., 2008; Pallesen et al., 2005), may derive from the conventional association of this mode with, for example, sad lyrics, slow tempi, and so on.

**Expectation**

Hanslick (1954), Helmholtz (1985), and Meyer (1956) have argued that musical enjoyment is linked with patterns of tension and resolution resulting from the confirmation and violation of perceptual expectations of which we are usually unconscious. These expectations might concern, for example, the pitch of the next note in a melody, the next chord in a pattern of harmonic movement, or the timing of the next note in a solo percussion performance. In each case, the preceding context of the music sets up expectations (or predictions) in the mind of the listener for what is to happen next.

Recent research suggests that expectations in music are acquired through a process of statistical learning in which listeners construct implicit probabilistic models of the next element in a musical sequence given the preceding context both at psychological (Huron, 2006; Meyer, 1957; Oram & Cuddy, 1995; Pearce & Wiggins, 2006; Tillmann, Bharucha, & Bigand, 2000) and neural levels (Loui, Wu, Wessel, & Knight, 2009; Kim, Kim, & Chung, 2011; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). Electrophysiological research has identified characteristic neural responses to violations of harmonic expectation, including an early right anterior negativity (ERAN) peaking at around 180-ms post-stimulus (Koelsch, Kilches, Steinbeis, & Schelinski, 2008; Steinbeis et al., 2006). The amplitude of this component is related to the long-term transition probability of the chord (Kim et al., 2011; Loui et al., 2009). In an electroencephalography (EEG) study of listeners to hymn melodies, Pearce et al. (2010) examined oscillatory and phase responses to high and low probability notes predicted by a complex variable-order probabilistic model of pitch expectation. Violations of expectation increased phase synchrony across a wide network of sensor locations and generated characteristic patterns of beta-band activation in the superior parietal lobe, previously associated with tasks involving auditory-motor interaction, suggesting that violations of expectation may stimulate networks linking perception with action.

Huron (2006) proposes a framework for linking expectations based on statistical learning to aesthetic responses. He distinguishes three responses to an event: first, a prediction response evaluating whether the event conforms to prior expectations; second, a reaction response, a fast, automatic, subcortical affective reaction; and third, an appraisal response, a more leisurely, cortically mediated process of consideration and assessment. Positive emotions resulting (via the prediction response) from anticipatory success are misattributed to the stimulus itself, leading to a preference for predictable events while the stress resulting from surprising events, as an indicator of maladaptive anticipatory failure, activates fast fight, flight, or freeze responses and provides negative feedback for the learning processes that generated the prediction. How is it then that surprise can be enjoyable even though it is associated with negative emotion resulting from the failure to correctly anticipate the future event? According to Huron (2006), an event that is unexpected but ultimately innocuous induces a
negative prediction response that increases, via a process of contrastive valence, the relatively positive limbic effect of the subsequent reaction or appraisal responses.

There is some empirical evidence to support the theory that the confirmation/violation of expectations is capable of leading to aesthetic experiences. In behavioral studies, probabilistic measures of stimulus predictability produce inverted U-shaped profiles of subjective pleasantness and beauty in simple tone sequences (Vitz, 1985). Expectations based on explicit knowledge of a work, and schematic expectations based on years of implicit learning through exposure (Bharucha, 1987). It appears to be impossible for us to “switch off” our schematic expectations even when we know consciously what is about to happen (just as we are unable to consciously influence our perception of certain visual illusions even when we know them to be illusions).

Huron (2006) suggests that contrastive valence may also be capable of generating pleasurable experiences of awe and laughter as well as chills although these claims await detailed neuroscientific study.

Aesthetic Judgments

It is possible to have an aesthetic experience without making an aesthetic decision or judgment about the stimulus causing that experience. When we do make an aesthetic judgment (consciously or unconsciously), such as when we decide that an object is beautiful, research in visual neuroaesthetics suggests that areas in the prefrontal regions of the brain, and specifically the dorsolateral prefrontal and orbitofrontal cortex, are activated (Nadal, Munar, Capo, Rossello, & Cela-Conde, 2008). Ventromedial prefrontal cortex is thought to be crucial for judgmental processes based on affective valence of stimuli such as in aesthetic context (Damasio, 1996; Krügelbach, 2005). Activation in orbitofrontal cortex has been found in tasks involving judging the beauty of paintings (Kawahata & Zeki, 2004), black-white abstract shapes (Jacobsen et al., 2006), and faces (Aharon et al., 2001).

Several neuroimaging studies of musical listening confirm the role of the orbitofrontal cortex in positive affective experiences associated with aesthetic judgments of preference or beauty for music (Alluri et al., 2012; Blood & Zatorre, 2001; Blood et al., 1999; Brattico et al., 2011; Pereira et al., 2011). For example, judging the beauty of a rhythmic sequence activated the ventromedial prefrontal cortex when contrasted with judging the tempo of the sequence (Kornysheva et al., 2010). In a recent fMRI study (Ishizu & Zeki, 2011), a finer localization of the area involved in beauty judgments of both musical pieces and paintings was obtained: activation in a very small region of the medial OFC, the A1 field, correlated with the intensity of beauty experienced. Expertise also modulates beauty judgments of music: Müller et al. (2010) found enhanced emotion-related neural processing for beauty judgments compared with cognitive judgments in nonmusicians but not in musicians, suggesting the latter make less use of emotion and rely on other strategies in making aesthetic judgments.

The dorsomedial midbrain nuclei, belonging to the dopaminergic reward circuit of the brain, are also activated by consonant chords judged as beautiful, irrespective of their major or minor keys, when contrasted with dissonant chords judged as ugly (Su-zuki et al., 2008). In addition, Kornysheva et al. (2010) report the activation of ventral premotor cortex and cerebellum to rhythmic patterns judged as beautiful contrasted with those judged as nonbeautiful. The involvement of these motor regions in beauty judgments of music may reflect the powerful ability of beautiful music to entrain behavioral responses such as song and dance (see Calvo-Merino and Christensen, this issue, pp. 76–88).

Preference

Another important outcome of the musical aesthetic experience is preference, which differs from enjoyment or subjective pleasure (with which it is often identified) in that it includes making a decision about the stimulus as a whole. Such a decision typically occurs after listening to an entire piece of music and may endure long after the listening episode. This decision may be based on the intensity of enjoyment, on an aesthetic judgment related to beauty or other formal properties of the stimulus (but it can also be divergent and independent from such judgment), and on other intrapersonal factors related to the listening history of an individual or her current mood or personality. Schubert (2007) demonstrated that induction of any emotion (negative or positive) by music predicts preference: the more we are moved by music, the more we like it. Another predictor of preference is the gap between the emotion expressed by the music (external locus) and emotion felt by the listener (internal locus): when we listen to a sad piece of music we may recognize the sad expression but not feel sad ourselves (a large gap) but preference tends to be higher when the gap between emotional loci is minimal (Schubert, 2007). Finally, Vuoskoski and Eerola (2011) observed that individuals characterized by high trait empathy tend to show higher preference and feel sadder than less highly empathetic individuals when listening to sad music.

There is evidence that musical preference activates lateralized brain networks. A pioneering electroencephalography study by Altenmüller, Schurmann, Lim, and Parlitz (2002) identified left-lateralized frontotemporal responses when listeners preferred classical, pop or jazz excerpts lasting 15 s, and right-lateralized anterior responses when they disliked them (neutral music generated bilateral brain responses). In a subsequent EEG and fMRI study (Flores-Gutierrez et al., 2007), preferred 30-s excerpts by Bach and Mahler similarly activated left-hemispheric regions, including Geschl’s gyrus, middle temporal gyrus and cuneus, whereas disliked
findings suggest that familiarity is one of the strongest influences on emotional and hedonic responses in the brain. Familiarity with music affects neural responses even at the level of the brainstem and auditory cortex. For instance, individuals with formal musical education show more accurate neural processing of chords that are atypical in Western tonal music, whereas discrimination in the auditory cortex of prototypical minor chords is comparable between musicians and nonmusicians (Brattico et al., 2009). Reber, Wurtz, and Zimmermann (2004) argued that the valence of aesthetic response is determined by the ease and speed with which a stimulus can be processed: the more fluent the processing, the more pleasant the experience. Given this, we might predict that if musical training improves auditory processing ability, musicians would show a preference for more complex musical styles (e.g., atonal music; cf. Brattico et al., 2009). Professional classical musicians also possess reinforced expectations for sounds following Western tonal rules, showing larger inferior frontal brain response (the ERAN discussed above) to unexpected chords in a sequence, compared with nonmusicians (Koelsch, Schmidt, & Kansok, 2002; although for discrepant findings with folk musicians, see Brattico et al., under revision). Children 5 and 6 years of age show similar reactions to unexpected chords, indicating that a few years of passive exposure to Western tonal music are sufficient to form a neural representation of harmonic syntax, although not as clearly or accurately as in older adults or in musicians (Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005). It remains to be seen whether these effects of stylistic familiarity, and ease of processing, impact on the aesthetic experience of music.

Functional neuroimaging studies have further investigated the neural structures involved in long-term memory for a particular piece. In a PET study, Satoh, Takeda, Nagata, Shimosegawa, and Kuzuhara (2006) found that familiarity judgments of piano melodies generated activation in the anterior portion of bilateral temporal lobes, posterior portion of superior temporal gyri, anterior and posterior portion of the medial frontal lobes, the bilateral cingulate gyri, the left inferior frontal gyrus, and the left superior temporal gyrus. In an fMRI study, Peretz et al. (2009) found that familiar melodies activated the bilateral superior temporal sulcus more than identical reversed melodies. Plailly, Tillmann, and Royet (2007) compared responses with familiar classical music excerpts and odors, finding activation for familiar over unfamiliar music in several left-hemisphere regions, including superior and medial frontal gyri, precentral gyrus, and superior temporal sulcus, posterior cingulate gyrus, and supramarginal gyrus. Multimodal areas activated by both familiar music and odors included superior and inferior frontal gyri, angular gyrus, precuneus and parahippocampal gyrus, suggesting that these areas are involved in neural circuits underlying the experience of familiarity during an aesthetic or hedonic experience regardless of sensory modality.

Attention

Attention is central to a musical aesthetic experience because the listener must concentrate on the music in order to appreciate the emotions and memories induced by it, judge whether it is beautiful or well performed, and decide on its aesthetic value. Several neuroimaging studies of musical listening have reported the involvement of the superior parietal lobule, the precuneus, and other parietal structures related to the ventral network for stimuli-
driven attention (e.g., Fan, McCandliss, Fossella, Flombaum, & Posner, 2005), and to the default-mode network, constantly monitoring and integrating the external environment (Raichle & Snyder, 2007). In addition, while neural processes related to the extraction of sound features, their integration into auditory objects (e.g., the pitch categories of thechromatic scale) and their maintenance in sensory memory are fast and automatic (Brattico, Terraniemi, Naatanen, & Peretz, 2006), analyzing the conformity of a chord sequence to the rules of Western tonal harmony, an important part of aesthetic appreciation, requires the allocation of attentional resources (Loui, Grent-’t-Jong, Torpey, & Wolderoff, 2005; Maidhof & Koelsch, 2011). Neuroimaging studies confirming the role of attention in a musical aesthetic experience remain to be conducted, but research in the visual domain on the importance of contemplation (Höfel et al., 2007) suggests a similar role for attentive processing in the auditory modality.

Reprise: Establishing a Neuroaesthetics of Music

We have proposed a framework for developing a neuroaesthetics of music in which music is viewed as an expressive art rather than as a cognitive domain sharing properties with language, as is often the case in the neurosciences of music. This interdisciplinary approach emphasizes the interplay between perceptual, affective and cognitive processes that generate aesthetic responses, including emotions, judgments and preference. Our approach is broadly compatible with the model of Leder et al., (2004) proposed for understanding aesthetic experiences of visual art and further developed for musical aesthetic experience by Brattico et al. (submitted). Regarding perceptual analysis, we have reviewed research on several properties of music, which are thought to be involved in aesthetic experience of music (including the effects of sensory dissonance, imitation and expectation) and how these generate emotional states. The induction of emotion by music can lead to states of variable enjoyment but these are not sufficient for an aesthetic experience per se, which typically also requires an appreciation of the properties of the music that determine the experience, an aesthetic judgment of how beautiful it is and a decision about how much we like it. The state of the listener and the listening context combine with the music to determine these judgments. Regarding implicit memory integration of these features, we have reviewed research suggesting that processing of these features is modulated by attention, familiarity and expertise. The field remains in its infancy, taking its first steps toward legitimate scientific status. As a way of assessing its current state, we recall the five tests of a healthy research paradigm suggested by Sloboda (1986, p. 199).

First, an established paradigm must have “an agreed set of central problems.” Sloboda (1986) himself identifies two central questions in music psychology: What is the nature of musical knowledge or representation? And how does music have emotional and aesthetic effects? Assuming a neuroscientific approach to music, we can view the second of these questions as the proper topic for the neuroaesthetics of music, the first being appropriate for the cognitive neuroscience of music. For more detailed questions, we propose that the set of problems distinguishing the neuroaesthetics of music from other areas of study can come from the field of neuroaesthetics, established in the late 90s by Zeki (1999) and whose research agenda is steadily maturing (e.g., Chatterjee, 2011; Livingstone & Hubel, 2002; Nadal & Pearce, 2011). The central questions of neuroaesthetics are to understand how and why (in an evolutionary sense) the human brain enables the capacity for appreciating and creating artifacts that are experienced as aesthetic. A neuroaesthetics of music would address a strong bias in the current literature on neuroaesthetics toward visual art to the virtual exclusion of other sensory modalities. In defining and delimiting the central questions of a neuroaesthetics of music, it will be profitable to examine the philosophical literature of musical aesthetics (e.g., Carroll et al., 2011; Robinson, 2005) and research in empirical aesthetics (e.g., Berlyne, 1971; Wundt, 1874).

The second condition for a healthy paradigm is, according to Sloboda, “agreed methods for working on these problems,” which, in the case of neuroaesthetics of music, are largely borrowed from neighboring disciplines, and particularly cognitive neuroscience and experimental psychology. One factor that distinguishes the field, however, is a concerted effort to focus on real-world musical examples in real-world settings so as to reliably induce aesthetic experiences. In a methodological attempt to simulate an ecologically valid situation, for example, Alluri et al. (2012) implemented a nonconventional analysis of the continuous fMRI signal correlated with the acoustic features extracted from an 8-min piece of music (Adios Nonino by A. Piazzolla). In this way, they were able to identify large-scale brain networks involved in processing of timbral, tonal, and rhythmic features during online musical listening.

The third condition proposed by Sloboda is “agreed theoretical frameworks in which to discuss [the research problems]”. We acknowledge that this is work in progress but we have aimed here and elsewhere (Brattico et al., submitted) to propose a theoretical framework for establishing a context within which future empirical results and theoretical observations within the neuroaesthetics of music may be assessed.

The fourth condition for a healthy research paradigm requires “techniques and theories which are specific to the paradigm.” We have argued above that although it inherits from music psychology, the neuroaesthetics of music is distinguished from the cognitive neurosciences of music by a focus on emotion and aesthetics rather than cognitive representation and processing. Furthermore, although it inherits from neuroaesthetics, the neuroaesthetics of music is distinguished from the neuroaesthetics of art by its subject which is a complex multidimensional, auditory signal extended in time and processed in distinct neural pathways from visual stimuli. These distinctions call, to take one example, for a specific focus in a neuroaesthetic of music on the role of time: a piece of music cannot be viewed as a static entity but rather one that unfolds in time, generating and manipulating expectations and interpretations in order to induce an aesthetic experience.

Finally, Sloboda calls for “research which is appropriate to the whole range of phenomena in the domain being studied.” As a field in its infancy, it is clear that the neuroaesthetics of music can only develop with further empirical research, which will, in turn, further clarify its status as an independent field of research and its relationships with neighboring disciplines. As a stimulus to this process of maturation, we suggest that one of the most pressing issues is to identify those topics and questions from the traditional and empirical aesthetics of music (taking a broad view of these disciplines), which can benefit from a neuroscientific approach, as
well as to formulate new questions and hypotheses that could not have been developed within existing traditional frameworks. With this in mind, we propose the following central questions for the field. Which neural regions and mechanisms are involved in the experience of aesthetic emotions to music? To what extent are they distinct from the neural mechanisms of basic emotions? What is the neural basis of imitative processes in music perception? Is there a cross-modal component to these processes and how is this reflected in neural processing? Is attentive contemplation during listening necessary to induce an aesthetic emotion? Can an understanding of underlying neural processes help to clarify how unexpected musical events can produce the pleasurable experiences of chills, awe and laughter? What are the relationships, at the cognitive and neural levels, between judgments of liking, beauty, preference and sensory pleasure: are they systematically, functionally related, or completely independent? At what stage of infancy do aesthetic emotions, preference, and aesthetic judgments appear in the maturing brain and what are their developmental trajectories? Is motor-cortex recruitment a precondition for a hedonic experience of music? Can the effects of the listening context and aesthetic attitude be identified at the level of neural processing?

We know from psychological research (e.g., Hargreaves & North, 2010) that aesthetic experience depends on the listener (e.g., expertise, internal state, mood, personality, attitude) and listening situation (e.g., social context, concurrent tasks) as well as the music (e.g., sensory dissonance, timbre, congruency with formal rules) but much work remains to be done on the underlying neuroscience of how these influences combine to generate an aesthetic experience. Perhaps most crucially, we feel it is fundamentally important for researchers in neuroaesthetics to try to replicate all facets of the aesthetic experience so as to encourage a genuine aesthetic judgment based on a fully immersed emotional experience and a focused analysis of the formal properties of a musical work. In doing so, the neuroaesthetics of music can develop an understanding that will have significant practical implications. Music is widely used to induce emotions and regulate mood (Thayer, Newman, & McClain, 1994; Särkämö et al., 2008), and music therapy is increasingly being investigated for treatment of affective and psychiatric disorders (Erkkilä et al., 2008; Erkkilä et al., 2011): a deeper knowledge of the neural mechanisms involved in the aesthetic experience of music will help put these clinical developments on a sound scientific footing.

References
Altenmüller, E., Schurmann, K., Lim, V. K., & Parlitz, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical laterisation patterns. Neuropsychologia, 40, 2242–2256. doi:10.1016/S0028-3932(02)00107-0


Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be
measured using functional magnetic resonance imaging. Trends in Cognitive Sciences, 14, 370–376. doi:10.1016/j.tics.2010.05.007
Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be
measured using functional magnetic resonance imaging. Trends in Cognitive Sciences, 14, 370–376. doi:10.1016/j.tics.2010.05.007
Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be
measured using functional magnetic resonance imaging. Trends in Cognitive Sciences, 14, 370–376. doi:10.1016/j.tics.2010.05.007
Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be
measured using functional magnetic resonance imaging. Trends in Cognitive Sciences, 14, 370–376. doi:10.1016/j.tics.2010.05.007
Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be
measured using functional magnetic resonance imaging. Trends in Cognitive Sciences, 14, 370–376. doi:10.1016/j.tics.2010.05.007


