Experiencing musical beauty: emotional subtypes and their physiological and musico-acoustic correlates

Diana Omigie[1,2], Klaus Frieler[1], Christian Bär[1], Muralikrishnan R. [1], Melanie Wald-Fuhrmann[1], Timo Fischinger[1,3,4].

1) **Music Department. Max Planck Institute for Empirical Aesthetics. Frankfurt am Main. Germany.**
2) **Psychology Department. Goldsmiths, University of London. London UK.**
3) **Freiburg Institute of Musicians' Medicine, Medical Center University of Freiburg,**
4) **University of Music Freiburg, Faculty of Medicine, University of Freiburg, Germany**
5)

Abbreviated title: Subtypes of musical beauty.

Correspondence

Diana Omigie

Music Department.

Max Planck Institute for Empirical Aesthetics.

Grüneburgweg, 14.

60322, Frankfurt am Main.

Germany.

Email: diana.omigie@aesthetics.mpg.de

Telephone: 00 49 6983 0047 9230

Abstract

A listener's aesthetic engagement with a musical piece often reaches peaks in response to passages experienced as especially beautiful. The present study examined the extent to which responses to such self-identified beautiful passages (BPs), in self-selected music, may be distinguishable in terms of their affective qualities. In an online survey, participants indicated pieces in which they considered specific passages to be outstandingly beautiful. In the lab, they listened to these pieces while physiological recordings were taken. Afterwards, they provided ratings on their experience of the BPs, where items targeted emotion response, underlying engagement mechanisms, and aesthetic evaluation. Cluster-analyses based on emotion ratings suggested three BP subtypes that we labelled low-Tension-low-Energy (LTLE), low-Tension-high-Energy (LTHE) and high-Tension-high-Energy (HTHE) BPs. LTHE and HTHE BPs induced greater interest and were more liked than LTLE BPs. Further, LTHE and HTHE clusters were associated with increases in skin-conductance, in accordance with the higher arousal reported for these BPs, while LTLE BPs resulted in the increases in smiling and respiration-rate previously associated with processing fluency and positive valence. LTLE BPs were also shown to be lower in tempo and polyphony than the other BP types. Finally, while both HTHE and LTHE BPs were associated with changes in dynamics, they nevertheless also showed distinct patterns whereby HTHE BPs were associated with increases in pitch register and LTHE BPs, with reductions in harmonic ambiguity. Thus, in line with our assumption that there is more than one kind of experience of musical beauty, our study reveals three distinct subtypes, distinguishable on a range of facets.

Key words: beautiful passages, music, emotion, physiology, aesthetic judgments.

Introduction

When individuals, with or without music expertise, articulate their positive appreciation of a piece of music, they often emphasize the value of particular moments or passages within it. These moments are commonly referred to as "beautiful passages", a phrase that goes back at least into the 18th century and which can be found in most European languages. Although a listening mode in which the listener primarily savors "beautiful passages" (rather than following the unfolding of its form) has been criticized -its most renowned critic being Theodor W. Adorno (1965)- it is a common mode of listening to music that is associated with a great variety of musical styles and genres.

Empirical work has shown that individuals invariably use the concept of beauty to express positive evaluations of works of art, as well everyday objects (Gabrielson, 2011; Istok et al., 2009; Jacobsen et al., 2004; Juslin, 2013; Juslin & Isaksson, 2014). However, in using the term beauty, everyday listeners inadvertently employ a notion that has been subject to wide-ranging discourse in philosophical aesthetics and art criticism (Sartwell, 2017; Tatarkiewicz, 1963, 1972; Levinson, 2015). Indeed, continuous debates regarding beauty's nature and how it should be defined make it a loaded concept for philosophers of aesthetics. Especially prominent in 18th and 19th century aesthetics (Sartwell, 2017), were discussions regarding whether beauty should be taken to refer to an identifiable feature of an object or rather a purely idiosyncratic response on the part of the beholder. Today, dictionaries define beauty as "a combination of qualities (...) that pleases the aesthetic senses" (Oxford dictionary) or something with "the quality of being pleasing" (Cambridge dictionary). Thus, in contrast to philosophical distinctions drawn between the beautiful and the pleasant, for instance, (Kant, 1790), perceived pleasantness may be intrinsic to everyday beauty propositions.

Despite these challenges, studying self-reports of experiences of beauty may be seen as essential for the field, given its adherence to a key principle of empirical aesthetics – namely, studying aesthetic concepts as they are understood by "ordinary people" (Fechner, 1876). Modern scholarly interpretations provide useful guidance for navigating terminological difficulties (Brattico, 2015; Stone-Davis, 2011) and thus, despite a rightful emphasis on the need for empirical aesthetics to look beyond beauty to fully understand the aesthetic experience (Consoli, 2015; Bundgaard, 2015; Conway & Rehding, 2013), the way in which listeners experience self-identified beautiful passages remains a question of central relevance.

Anecdotal evidence suggests that such self-identified beautiful passages (BPs) are regarded as highly valuable and may influence the whole experience of listening to a piece (Konrad, 2011). However, little is known about the affective qualities associated with such beauty judgments. Indeed, despite the experience of beauty being increasingly described as having an affective component (e.g. Menninghaus et al., 2018), only sparse empirical evidence for this claim exists. The present study uses a multi-modal approach to examine the extent to which BPs may be grouped into emotion subtypes that are associated with distinct patterns of engagement mechanisms, physiological signatures and musico-acoustic features.

The role of affect in beauty experiences

Research using subjective reports of beauty have not only demonstrated their importance to the aesthetic experience of non-experts (Istók et al., 2009; Jacobsen, 2010; Jacobsen, Buchta, Köhler, & Schröger, 2004; Pearce et al., 2016; Thagard, 2019) but have also been able to provide insights into the perceptual (e.g. Spehar et al., 2015) and neural underpinnings of experiences of beauty (Ishizu & Zeki, 2011; Müller, Höfel, Brattico & Jacobsen, 2010; Zeki, 2013, 2014; Starr & Starr, 2013).

For the still relatively new field of neuroaesthetics, a question of deep interest has been the neural substrates that underlie beauty experiences (Ishizu & Zeki, 2011; Müller, Höfel, Brattico,

& Jacobsen, 2010; Suzuki et al., 2008; Nieminen, Istók, Brattico, Tervaniemi, & Huotilainen, 2011; Pearce et al., 2016). Interestingly, while some studies have provided corroborating evidence of the involvement of certain brain areas across aesthetic domains, others have suggested that distinct neural substrates underlie different subtypes of beauty experiences, even within a given aesthetic domain. For instance, Zeki and colleagues have repeatedly demonstrated a role of the orbitofrontal cortex, a key area in the emotion network, in beauty judgments that are elicited by visual artworks, music and even mathematics (Ishizu & Zeki, 2011; Zeki, 2013, 2014). However, others have also shown that within music listening, major consonant chords perceived as beautiful induce different patterns of activity from those seen for minor consonant chords perceived as beautiful (Suzuki et al., 2008). Further studies using electroencephalography recordings have provided evidence that affect may be a key element underlying beauty judgments, at least for those lacking formal training (Müller, Höfel, Brattico, & Jacobsen, 2010). For instance, comparing neural responses during a correctness versus an aesthetic judgment task, Muller and colleagues (2010) demonstrated that non-experts displayed brain patterns pointing to a reliance on affective states when forming aesthetic judgments.

Investigations suggesting that different beauty experiences may have different underlying neural substrates and may highly depend on the brain's emotion network suggest that beauty experiences may be both varied on the one hand and intrinsically related to affective responses on the other. Contemporary psychologists increasingly subscribe to the dialectical idea (Kant, 1790; Shell, 2002), that the experience of beauty is an emotion at the same time as it is a judgment (Armstrong & Detweiler-Bedell, 2008; Menninghaus et al., 2018; Payne, 1980; Schindler, Hosoya, Menninghaus, Beerman, Wagner, Eid et al., 2017). In music specifically, it has been suggested that aesthetic emotions may arise from listeners forming aesthetic judgments about the music being heard (Juslin, 2013) while the aesthetic experience of music has been described as including perceptual, cognitive, and affective components (Brattico &

Pearce, 2013; Brattico, Bogert, & Jacobsen, 2013). All such accounts call for a better understanding of the affective qualities of aesthetic judgments in general, and of beauty judgments in particular.

Classical approaches from aesthetics emphasize the role of form in eliciting judgments of beauty with Hanslick famously lamenting the focus of music aesthetics on the feelings music awakens (Hanslick & Payzant, 1986; Sartwell, 2017). However, it is important to note that focusing on form would not necessarily preclude feelings being awakened by music. Indeed, a focus on form may be predicted to lead to not one, but rather multiple subtypes of beauty experiences, related to objective features of the formal structures being heard. A large body of research on the emotions experienced during music listening has examined them in the context of domain general models of emotions (Balkwill, Thompson & Matsunaga, 2004; Fritz, Jentschke, Gosselin, Sammler, Peretz & Turner, 2009; Juslin, & Sloboda, 2011). Many of those studies have used the dimensional model of emotions, either with Valence and Arousal (Wundt, 1896; Russell, 1980) or with Valence, Tension arousal and Energy arousal as dimensions (Schimmack & Grob, 2000; Schimmack & Reisenzein, 2002; see Vuoskoski & Eerola, 2011 and Eerola & Vuoskoski, 2013 for a review). In the latter cases, Energy and Tension have been measured as separate dimensions of arousal, based on claims that they are influenced by different factors to different degrees (Gold, MacLeod, Deary, & Frier, 1995; Watson, Wiese, Vaidya & Tellegen, 1999).

The dimensional model of emotion is powerful in providing a framework for even categorical models of emotion. Specifically, in order to reduce their dimensionality, the dimensional model has been used to describe not just every day emotions (e.g. happiness is seen as high in arousal with positive valence) but also music-induced categorical aesthetic emotions (e.g. see Trost, Ethofer, Zentner, & Vuilleumier, 2011). Critically, support for the usefulness of considering beauty experiences in the context of the dimensional model of emotions may be found in the philosophical and psychological literature. For instance, the existence of a specific type of beauty experience that is low in tension has been argued for in the philosophical literature. Levinson describes narrowly beautiful music as "… an immediate, unmitigated and unmixed pleasure in how the music unfolds, a pleasure in which tension and discord have at most a minor place" arguing that such music is there to charm rather than to confront or challenge (Levinson, 2012, p. 128). Levinson's account resonates with older aesthetic theories that pit feelings of beauty against those of the sublime, whereby the former is associated with feelings of tenderness and the latter with a preservation instinct (Burke, 1757/1990). Similarly, claims of the importance of arousal to the aesthetic experience are also dominant in the field of music psychology, whereby tension arousal induced by music is considered crucial to the aesthetic response (Meyer, 1956; see Lehne & Koelsch, 2015 for a review), and where moderate levels of arousal are argued to be preferred over the other extremes (Berlyne, 1971).

Thus, taken together, both the philosophical and psychological literature suggest that, if explored, subtypes of the experience of beauty may be distinguishable by at least the arousal, if not also the valence, dimensions of the three-dimensional model. The current sought to test this by examining the extent to which the experience of beauty in response to music vary meaningfully in this regard.

Potential mechanisms underlying experiences of beauty: Empathy and Interest

Perhaps just as important as identifying subtypes of musical beauty experiences may be elucidating the engagement mechanisms that drive such experiences. Two mechanisms that have been widely argued to play especially important and distinct roles across different aesthetic domains are, first, empathy, the process of embodying the emotional experience of another, and, second, cognitive stimulation, the process of engaging cognitive resources with the processing of information (Brielmann & Pelli, 2017; Gernot, Perlowski & Leder, 2017; Graf & Landwehr, 2017; Omigie, 2015).

While the word empathy is used today to encompass social feelings (see Omigie, 2015), its origins are in the 19th-century aesthetic concept of "Einfühlung", literally meaning "feeling into" and understood as the capacity to have an experience of the form and emotional content of an aesthetic object (Lipps 1903; Titchener, 1903; Vischer 1846). Empathy is held to be central to the aesthetic response even in the context of aesthetic objects that lack verbalisations (Baltes & Miu, 2014; Freedberg & Gallese, 2007; Gernot, Perlowski & Leder, 2017; Miu & Baltes, 2012) and has, in the context of music, being explained in terms of persona theory, whereby emotionally expressive music is held to make the listener imagine a person related to that emotional expression (Cone, 1974; Levinsonson, 2005; Cochrane, 2010). The feeling of "being moved", a complex emotion state characterised, amongst others, by co-activation of positive and negative affect and passive empathic observation is considered a highly related concept (Kuehnast et al., 2014; Menninghaus et al., 2015; Wassiliwizky et al., 2015). In the context of music, for instance, feelings of being moved have been shown to be predicted by the levels of trait empathy in listeners (Eerola, Vuoskoski & Kautiainen, 2016). Critically, studies examining the role of empathy mechanisms in the aesthetic response have shown that they may be able to explain the value attached to visual artworks (Gernot, Perlowski & Leder, 2017) as well as the aesthetic liking of negatively valenced music (Eerola, Vuoskoski & Kautiainen, 2016).

Similarly, studies examining the role of cognitive stimulation in the aesthetic response have shown that challenging, but largely comprehensible, objective features of aesthetic objects may lead to positive aesthetic judgments being made (Armstrong & Detweiler-Bedell, 2008; Brielmann & Pelli, 2017; Conway & Rehding, 2013; Graf & Landwehr, 2015; Silvia, 2005).

Interest, a state resulting from cognitive stimulation may be seen as an aesthetic category independent of beauty (Schlegel, 1797; Ngai, 2008; Ngai, 2012; Menninghaus et al., 2018). However, in influential psychological theories, Interest is proposed as one of two routes to aesthetic appreciation, whereby, compared to the pleasure driven by high processing fluency, aesthetic interest is held to arise from stimulus features that are relatively difficult to process (Graf & Landwehr, 2015). A route to aesthetic appreciation that is driven by interest would provide a counterpart to descriptions of those forms of beauty that are "narrow" and do not challenge (Levinson, 2012). Further, experiences varying in interest may be expected to be reflected in differing levels of subjective and physiological arousal, whereby greater interest may be associated with greater arousal. However, such potential mappings between subjective reports of emotional arousal on the one hand, and engagement mechanisms underlying beauty experiences on the other, remain to be explored.

Experiencing beauty: physiological correlates and acoustic cues?

Support for the existence of distinct subtypes of beauty experiences would come from observing differences, not only in terms of how they are reported subjectively, but also in terms of bodily responses to the BPs and the BPs' musical features. Decades of research have demonstrated that physiological responses vary with respect to the emotion dimensions. Smiling and frowning (as indexed by zygomaticus and corrugator facial muscle activity) have, for instance, been associated with positive and negative valence respectively, while increases in heart rate, respiration rate and skin conductance have, in general, tended to be associated with elevated subjective arousal (Hodges, 2009; Ogg, Sears, Marin & McAdams, 2017, see Juslin & Sloboda for a review, 2011).

Short-lived pleasurable experiences, variously referred to as chills, thrills or more generally, frisson, have particularly been associated with signatures of high physiological arousal, such as increases in skin conductance (Grewe, Nagel, Kopiez & Altenmüller, 2007b; Nagel, Kopiez, Grewe & Altenmüller, 2008; Salimpoor, Benovoy, Larecher, Dagher & Zatorre, 2011). However, since, for some listeners, the experience of chills may overlap with experiences of beauty, such physiological signatures may be expected for at least some subtypes of beauty experiences. In that case, a question that would nevertheless arise, concerns the bodily responses that may characterize those beauty experience that are not associated with high subjective and physiological arousal.

Here it is important to note that subjective arousal and valence levels reported by listeners are not only associated with certain physiological responses but are also well predicted by psychoacoustic features such as changes in loudness, pitch level and tempo (Coutinho & Cangelosi, 2011; Coutinho & Dibben, 2013; Chuen, Sears & McAdams, 2016), as well as by higher level features such as pulse or key clarity, associated with metrical and tonal regularity respectively (Alluri, Toiviainen, Jääskeläinen, Glerean, Sams & Brattico, 2012). Crucially, research linking musico-acoustic features with subjective emotion experiences parallels research on frisson that show these transient highly pleasurable experiences to be linked to changes in loudness, register, and harmonic or temporal ambiguity (Panksepp, 1995; Guhn, Hamm, & Zentner, 2007; Nagel, Kopiez & Altenmuller, 2007). Thus, taken together, one may expect that subtypes of beauty experiences, characterized by distinct emotion profiles may also be associated with distinct patterns of both physiological responses and musico-acoustic features. The current study set out to test this assumption.

The Current Study

In summary, a rich psychological and philosophical literature suggests that experiences of beauty in response to music may be expected to vary in terms of emotional dimensions (Levinson, 2012; Meyer, 1956), may be associated with different modes of engaging with features of the music (Graf & Landwehr, 2015) and may be associated with different patterns of physiological responding and musical features (Guhn et al., 2007). However, while the importance of taking a multifaceted approach to peak experiences has often been emphasized (Gabrielsson, 2003), an examination of these different elements (subjective emotional, physiological, and musical) in the context of the experiences of beauty, remains to be carried out.

Motivated by this gap in the literature, participants' physiological variables were measured while they listened to self-selected pieces containing what they considered to be BPs. Participants were required to describe their experience of these BPs with respect to the three-dimensional emotion scale, to evaluate how much empathy and interest was induced in them while listening, and to provide ratings of beauty and liking in response to all experienced BP. To address the evidence for BP subtypes in a data-driven way, we used a combination of cluster analyses methods to validate a clustering solution emerging from listeners' emotion ratings. Then, following a grouping of BPs according to the similarity of their emotional profiles, we evaluated the extent to which these BP subtypes differed in terms of engagement mechanisms (Empathy and Interest) and aesthetic evaluation (Liking and Beauty). Finally, since evidence of dissimilar patterns for each subtype may be taken as support for their distinctiveness, we examined the extent to which the different BP subtypes, suggested by the cluster analyses, resulted in distinct profiles of physiological responses and acoustic and musical features.

While a large number of empirical studies of music aesthetics have avoided using music with lyrics as stimuli, such an approach inevitably constrains the extent to which research findings can be extrapolated to the rich and various experiences of music in every-day life. Thus one key feature of the current study, similar to previous naturalistic studies (e.g., Brattico et al., 2011; Pereira et al., 2011), was that participants were encouraged to indicate BPs in pieces from all genres, including those containing lyrics.

Materials and Methods

Participants

The study was conducted in full accordance with the World Medical Association's Declaration of Helsinki and the Ethical Guidelines of the German Association of Psychologists (DGPs). All procedures were approved by the Ethics Committee of the Max Planck Society and all participants gave written informed consent.

An online survey was designed that allowed potential participants to provide information about the musical pieces that according to them contained BPs. Listeners were asked to suggest as many pieces as they could, and provide, for each piece, as much detailed meta-information as possible regarding, for instance, the name of the composer, the performer, the recording company, or when the record was published. Moreover, participants were asked to indicate, with timestamps, the exact positions of the BPs within each of their pieces.

40 participants who gave sufficient information regarding pieces containing BPs were selected for invitation to the lab. However, due to technical difficulties during testing, data from 33 participants (22 F, Age: 28.49, SD=12.7) were analyzed. (In five cases, the presentation script terminated due to issues with the file format that were difficult to track down. In another two cases, the participants were presented with the wrong stimuli due to human error). All participants were German speaking and almost all were enrolled on, or had completed, an undergraduate university degree or higher (31/33). Their musicality was assessed using the

GoldMSI (Müllensiefen, Gingras, Musil & Stewart, 2014). With regard to music performance background, more than half (56%) of participants reported having regularly practiced an instrument for more than 4 years, while only about a quarter (26%) reported never having played an instrument. With regard to music education background, 35% had undergone six or more years of lessons outside of school while 24% had undergone less than one year of music lessons outside school. Finally, almost all (94%) reported listening to more than 15 minutes of music a day.

Musical stimuli

The musical pieces proposed by listeners were examined and, where necessary, a sub-selection taken so that the total duration of self-selected musical stimuli playback for each participant summed up to approximately 15 minutes and no more than 18 minutes. The main criterion for inclusion of a piece in a participant's stimuli selection was that it both contained at least one BP (with two or more being favored to maximize data collected) and was obtainable by the authors (e.g., was present in the institute library). As can be seen in the full list of selected pieces (Table 1), most stimuli came from the pop/rock genre (51.4%) with finer analysis showing that a large proportion of constituent songs could be described as *solo ballads* while others were best described as *upbeat feel good* songs. Classical and film music were the next most prevalent classifications accounting for a joint 20% of pieces, while the remaining pieces came from a range of genres and styles including soul, world, rap/ hip-hop, jazz, and electronic dance music. Finally, up to 60% of pieces had been released within ten years of the experiment taking place.

In the vast majority of cases, the pieces were less than five minutes long (e.g., a three-minute pop song) and contained up to two to three BPs. Accordingly, such pieces were presented to the participant in their entirety. In a few cases, the audio file was edited to begin at a musically meaningful moment at least two minutes before the beginning of the first indicated BP (e.g., beginning of a musical section or phrase) and to end at another musically meaningful moment several seconds after the offset of the final indicated BP. To avoid participants requiring an adjustment of volume during the experiment, a custom written MATLAB script was used to process all audio stimuli to be presented to a given participant, such that excerpts had the same average sound intensity while keeping within-excerpt intensity variance intact. The script calibrated the multi-channel stimuli audio files for similar loudness by determining the maximum amplitude across all stimuli and normalizing them with respect to that maximum amplitude.

Preparation and recordings

Participants read a document in which all details of the experiment, including their task and what would be measured, were explained. Participants were then given the opportunity to ask any questions before signing the informed consent form. All measurements took place in a soundproof booth.

Facial electromyography (EMG) was measured using electrodes placed on the forehead and cheeks. After cleansing the skin in order to reduce electrode impedance, Ag/AgCl electrodes filled with EMG gel were placed on the left side of the face according to published guidelines (Fridlund & Cacioppo, 1986). Two electrodes were placed over the eyebrow to monitor the activity of the corrugator muscle (EMGC), while two others were placed on the cheek to monitor the activity of the zygomaticus muscle (EMGZ). The ground electrode was placed on the middle of the forehead, and impedance of all facial EMG electrodes was kept below 5 kΩ. To measure breathing activity (RESP), a custom-prepared respiration belt was wrapped around the participant's upper rib cage, on the level of the sternum. To measure galvanic skin conductance (SC) response, electrodes were attached to the middle segment of the index and the middle finger of the left hand. Finally, a plethysmograph clip was put on the ring finger of the participant's left hand, allowing the measurement of blood volume pulse, from which heart rate (HR) could be estimated. All electrodes were connected to the amplifier with data sampled at 1000 Hz. EEG data was also collected but will not be presented here. All preparation lasted between 30 and 40 minutes and participants were given a brief overview of how their signal quality was affected by their actions in order to encourage their compliance with our request that they minimize artefact-causing movements.

Procedure

Participants were seated in front of a computer monitor and all auditory stimulation was presented over loudspeakers (Neumann KH 120 A). Loudness was adjusted to individualized comfort levels using a short audio file played at the beginning of the study.

Participants were first presented with their individually selected pieces and asked to listen carefully while keeping their eyes closed and avoiding movement. In a subsequent part, participants were asked to use the space bar on the keyboard to indicate, to the best of their ability, the onset and offset of the BPs, where they experienced them in the music. Listeners were informed that they could focus on the task of annotating their BPs without any worry of signal contamination, as physiological recordings would no longer be taking place. They were informed that at the end of each so-indicated (i.e., using the space bar) BP, they would be required to indicate their agreement with the statement, "While listening to this excerpt of music, the following was triggered in me: 'Tension' ('Anspannung'), 'Energy' ('Energie'), 'Pleasantness' ('Wohlgefühl'), 'Interest' ('Interesse') and 'Empathy' ('Empathie') on a scale of 1 to 6 ("Not at all" to "A lot"). They would also be required to rate their agreement with the statements "I like this excerpt" ('Dieser Musikausschnitt gefällt mir') and "I find this excerpt beautiful" ('Diesen Ausschnitt finde ich schön') on a scale of 1 to 6 ("Not at all" to "A lot").

The generic anchor points were used rather than individual opposing bipolar scales for ease of rating across the different items. Further, a scale of 1 to 6 was used to encourage participants to give informative responses. Participants were not given definitions of any of the terms. Nor did any of the participants ask for clarification.

Data collection was part of a larger protocol in which participants also a) listened to and rated 20 40-second-long experimenter-selected excerpts (Cohrdes, Wrzus, Wald-Fuhrmann, & Riediger, 2018) and b) typed out comments after every indicated BP and experimenter-selected excerpt. However, these data, outside the scope of the present investigation, are not presented here.

Physiological data preprocessing

All signal processing was carried out in MATLAB. To obtain a continuous measure of facial muscle activity over time, the zygomaticus and corrugator facial muscle signals (henceforth, EMGZ and EMGC) were band-pass filtered between 20 and 249 Hz, and the absolute value of the Hilbert transform of the filtered signal extracted and then smoothed at 1 sec using a convolution function (http://www.fieldtriptoolbox.org/documentation). As raw respiration and blood volume pulse signals tend to show high levels of noise caused by chest and hand movement artefacts respectively, these signals were low pass filtered using a Butterworth filter and in the case of respiration signal, further smoothed using the *conv* function from MATLAB in line with previous work (e.g., Salimpoor et al., 2009). Next, peaks in the filtered signal were identified using a custom-written peak detection algorithm, with visual inspection used to confirm that peaks were correctly identified. Respiratory and heart rates (henceforth RESP and HR) were then obtained by taking a differential of the timings of maximum peaks and interpolating the resulting rate signal to a regular sampling frequency of 20 Hz. Finally, a time-series of the phasic component of the skin conductance response (henceforth, SC) for each piece was extracted from the tonic component using the Ledalab toolbox (Benedek & Kaernbach, 2010). In line with previous literature, any remaining linear trends were removed to combat slow downward drifts typical of the skin conductance signal (Salimpoor et al., 2013). No threshold was used for estimation of the skin conductance response but rather, a time integration of the continuous measure of phasic activity (average value) was taken as an index of event-related sympathetic activity in line with previous recommendations (Benedek & Kaernbach, 2010).

Statistical analysis of subjective report ratings

To provide an initial exploration of the relationship between ratings of induced emotion (Tension, Energy, Valence), engagement mechanisms (Empathy and Interest) and aesthetic evaluation (Liking and Beauty), Pearson correlations were estimated between all pairs of variables. Next, to identify potential clusters of BPs differing in terms of emotion profiles in a data-driven way, a hierarchical cluster analysis was carried out using the emotion ratings that participants gave after every indicated BP. The *hclust* function from the *stats* package, and the *Ward's* minimum variance method, which finds compact and spherical clusters, were applied to a distance matrix computed with squared Euclidean distances (R Core Team, 2014). This agglomerative procedure based on a minimum-variance approach involved each BP forming its own individual cluster in the first instance, followed by successive grouping of the most similar clusters until all BPs were merged in one cluster.

To characterize emergent clusters fully in terms of subjective experience (i.e., in terms of the emotion variables that were used to classify them in the cluster analysis, as well as with respect to engagement style and aesthetic evaluation, which were not used in that way), we adopted a 2-step approach. In the first step, we assessed the effect sizes of differences between the emergent clusters in terms of emotion ratings. This non-inferential statistical approach was taken due to the inappropriateness of inferential statistics, given that clustering algorithms specifically select cluster solutions that maximize Euclidean distances between the clusters. However, in order to ensure the soundness of our interpretation (consequent naming of the emergent clusters), and to check that our solution was not tied to our choice of clustering algorithm, we also examined alternative cluster solutions that we obtained using a subset of the BPs (hierarchical clustering with only 1 BP per piece), as well as additional clustering methods (Kmeans and PAM clustering).

In the second step, once subtypes of BPs had been validated, we estimated how the emergent subtypes differed in terms of engagement mechanisms (Interest, Empathy) and aesthetic evaluation (Liking and Beauty). As these variables were not used to obtain the emergent clusters, individual linear mixed effect models were estimated for each variable on all possible pairs of clusters using BP subtype as fixed effect and participant as a random effect with intercept modelled. Linear mixed models were estimated using the *lmer* function from the *lme4* package (Bates, Maechler, Bolker, Walker, 2015), while the *anova* function in the *car* package was used to estimate *p* values. A Bonferroni correction for multiple comparison was applied to account for multiple comparisons and standardized betas are presented allowing for effect size comparison (as per Cohen, Cohen, West, & Aiken, 2003).

Analysis of physiological data.

To examine the changes in physiological correlates associated with BP subtypes identified using cluster analysis, we compared mean responses during the BPs themselves to mean responses during the period leading up to them (up to 10 s). For both mean values (pre BP, during BP), data points falling outside of 1.5 times the interquartile range were excluded and individual linear mixed effect models, for each of the physiological measures separately, were then estimated. In these models, Period (pre BP, during BP) was modelled as the fixed effect and individual BP (a unique ID of the BP) was taken as random effect with both intercept and slope for Period modelled. Multiple comparisons were accounted for using Bonferroni correction. Here, note that although it was expected that the time period preceding indication of a BP may also include responses associated with anticipation of the BPs themselves, this was not further addressed here, given the lack of a principled way to estimate the onsets of such anticipatory effects.

Analysis of musico-acoustic features.

To explore the extent to which any identified BP types were characterized by changes in musico-acoustic features that have previously been associated with pleasurable responses to music, we investigated potential changes in register (mean pitch), dynamics (RMS or acoustic intensity), modality (major or minor mode) and harmonic and metric ambiguity (key and pulse clarity) using features automatically extracted with MIRtoolbox (Lartillot & Toiviainen, 2007). Modality was measured using "Mode", whereby a high positive value indicates "majorness", a high negative value indicates "minoress" and a value close to 0 indicates an ambiguity between the two. Harmonic and metric ambiguity were estimated using "Key Clarity", an estimate of how strongly a key is represented in musical stimuli at a given moment, and "Pulse Clarity", an estimation of the regularity of rhythmic or metrical pulsation (by extracting underlying recurring periodicities with the help of autocorrelation) in music, respectively (Lartillot & Toiviainen, 2007). Features were calculated using default window sizes as recommended by the MIRtoolbox: namely 50 ms for low-level acoustical features (RMS and Pitch) and 3 s for the higher-level features (Key Clarity, Pulse Clarity and Mode). Here as in the physiological data analysis, moments just before (up to 10 s) and over the duration of the BP were compared, whereby for each BP type, we estimated linear mixed effect models that took Period (pre, BP) as fixed effect and individual BP (a unique indicator of the BP) as the random effect.

Finally, musical features that were expected to differ between beauty emotional subtypes, but for which manual annotations have been shown to be more valid than automatic extraction (Lange & Frieler, 2018) were also estimated. Specifically, 3 expert listeners, blind to groupings of the BPs, provided manual annotations of tempo (in BPM), musical dissonance (on a scale of 1 (highly dissonant) to 5 (highly consonant)), polyphony (on a scale of 1 (highly monophonic/homophonic) to 5 (highly polyphonic)), and finally the presence of vocals in all BPs (present, not present), thus allowing comparison of these features across emergent BP types.

Results

Participants identified 420 BPs across their selected pieces but those BPs with a duration of less than 1 s were excluded (since these included false button presses) leaving 399 BPs for further investigation (Mean duration = 32.9 s, SD = 27.77).

Exploration of relationships between behavioural ratings

The results of Pearson correlation analyses, estimated between all pairs of variables exhaustively, may be seen in Table 2. With respect to how aesthetic judgment dimensions related to each other, a high correlation – the highest observed – was found between Beauty and Liking ($r = 0.7$, $p < .001$). The arousal dimensions Energy and Tension showed a moderate sized correlation with each other ($r = 0.39$, $p < .001$), although while Energy demonstrated small to moderate positive correlations with aesthetic judgments (Beauty: $r = 0.21$, $p < .001$; Liking: $r = 0.33$, $p < .001$), no such linear relationship was observed between Tension and either aesthetic judgment variable (both $p > 0.05$).

Similarly, in contrast to Energy, which was positively correlated with Valence ($r = 0.2$, $p < .001$), a small negative correlation was observed between Tension and Valence ($r = –0.12$,

$p < .05$). In turn, pointing to a tendency for Beauty and Liking judgments to be associated with positive valence (i.e., for there to be a strong association between positive valence and positive aesthetic judgments) moderate-high correlations were found between Valence and Liking ($r = 0.46$, $p < .001$) and between Valence and Beauty ratings ($r = 0.40$, $p < .001$).

With respect to how engagement mechanisms related to each other, to aesthetic evaluations and to the emotion dimensions, Interest and Empathy ratings showed a moderate correlation with each other ($r = 0.41$, $p < .001$), and were moderately correlated with aesthetic evaluations ($r = 0.36$–0.46, $p < .001$). However, while Interest showed moderate sized correlations with the arousal dimensions ($r = 0.23$–0.43, $p < .001$), only small correlations were observed between these arousal dimensions and Empathy ($r = 0.11$–0.15, $p < .05$). Finally, Valence ratings were shown to correlate to a greater extent with Interest ($r = 0.42$, $p < .001$) than with Empathy ($r = 0.25$, $p < .001$).

Clustering of BPs into emotion subtypes

The correlational analyses, displaying asymmetries in the mapping between emotion dimensions, engagement mechanisms and aesthetic evaluations, provided preliminary evidence of potential subtypes of beauty experiences, and validated the use of cluster analyses to allow identification of such BP subtypes in a data-driven way. Figure 1 shows the cluster solution obtained from clustering 399 BPs based on Likert ratings given with regard to Tension, Energy and Valence. Based on a widely used criterion, the optimal number of clusters was determined by the maximal increase in cluster height in the dendrogram (Everitt, 1974), resulting in a three-cluster-solution containing 112, 140, 147 BPs, as indicated by the red boxes in the plot.

Bar plots in Figure also 1 show the characterization of the emergent clusters in terms of emotion, engagement mechanisms and aesthetic evaluation variables. As shown there and in

Table 3, a clear pattern resulted whereby one of the emergent three clusters was best characterised by both low Tension and low Energy ratings (Cluster 1), another by low Tension and high Energy ratings (Cluster 2), and a third by both high Tension and high Energy ratings (Cluster 3). Specifically, starting with the Tension variable, large differences were observable between Clusters 1 and 3, and between Clusters 2 and 3 whereby Cluster 3 had the highest mean Tension rating, while Clusters 1 and 2 had lower tension ratings showing negligible differences from each other. In contrast, with respect to Energy, large differences were observable between Clusters 1 and 2 and between Clusters 1 and 3, where both Clusters 2 and 3 had higher levels of Energy than Cluster 1, while showing negligible differences from each other. Finally, in contrast to Tension and Energy, the differences in Valence across the three clusters were negligible. Thus, to emphasize the particularly large variations in Tension and Energy reflected in the cluster solution, we named Cluster 1 the Low-Tension/Low-Energy cluster (LTLE), Cluster 2, the Low-Tension/High-Energy cluster (LTHE), and finally, Cluster 3 the High-Tension/High-Energy (HTHE) cluster.

Figure 2 shows the results of the additional cluster analyses carried out to ensure the soundness of this interpretation namely a further hierarchical cluster analysis of a subset of the data (only 1 BP per piece), and a K-means and PAM cluster analysis of the full set of BPs. Table 4 shows the effect size (Cohen's d) of mean rating differences between BP clusters from all clustering solutions. In line with the initial hierarchical clustering analysis, and as can be seen in Table 4 and Figure 2 (where the clusters are named accordingly), all three further clustering solutions also resulted in one cluster being low in both Energy and Tension (LTLE), another low in Tension but high in Energy (LTHE), and a third high in both Tension and Energy (HTHE). Critically, the effect sizes when comparing clusters labelled as low or high in Tension or Energy were all greater than 2 standard deviations (SD), (compared to all other effect sizes which were differentiated by less than 1 SD) providing justification of the cluster labelling.

Finally, in terms of how individual BPs fell into the different cluster categories, analysis of the coincidence revealed a very high coincidence of 0.86 and 0.79, between the original hierarchical cluster solution and the PAM and Kmeans alternatives respectively. Further analysis using variants of the hierarchical solution (average, median, centroid clustering algorithms), showed qualitatively similar results, thus further validating the cluster solutions. However, examination of the compactness of the cluster using the Dunn index provided values ranging from 0.16 to 0.19, indicating that the obtained clusters show continuous rather than abrupt transitions, and may hence be better interpreted as *types* than as *classes*.

Characterization of BP subtypes in terms of engagement mechanisms and aesthetic evaluation.

Following identification of the emotional subtypes of beauty experiences, potential differences between these subtypes, in terms of engagement mechanisms and aesthetic variables, were assessed (Figure 1). Tables 5 and 6 provide descriptive statistics for these variables, for each of the three clusters. Linear mixed effect modelling revealed that the three groups did not significantly differ from each other, with respect to ratings of induced Empathy. However, these analyses showed that LTLE BPs induced less Interest than both LTHE ($B = 0.83$, $SE = 0.19$, $p < .001$, CIs [0.46, 1.18]), and HTHE BPs ($B = 0.78$, $SE = 0.2$, $p < .001$, CIs [0.4, 1.17]), while LTHE and HTHE did not, however, differ from each other in this regard ($p = 0.99$). With regard to aesthetic evaluations, no significant differences in Beauty ratings were observed across the clusters. However, LTLE displayed the lowest, LTHE the highest (significantly greater Liking than LTLE, $B = 0.42$, $SE = 0.08$, $p < .001$, CIs [0.24, 0.59]), and HTHE intermediate levels of Liking (albeit not differing from either LTLE or LTHE, both $p > 0.05$)). Random effects were found to be significant for the majority of models showing individual differences between participants.

In summary, large differences between clusters were found to be related to degree of the Tension and Energy arousal they induced in listeners, such that these dimensions could be used to parsimoniously characterize the emergent clusters. Comparisons across different clustering solutions confirmed the robustness of the categorization and interpretation, and comparison of the observed clusters showed they differed, not just in their emotional profiles but also in terms of the levels of Interest and Liking associated with them.

Physiological and acoustic correlates of the distinct BP types

Figure 3 and Tables 7 and 8 show the results of analyses comparing the mean of bodily responses to the BPs themselves, to the those measured in the preceding time-window. Respiration rate (B = 0.35, SE = 0.078, $p < .001$, CIs [0.2, 0.51]) and zygomaticus activity (B = 0.18, SE = 0.06, $p < .01$, CIs [0.05, 0.30]) were shown to be greater during LTLE BPs, compared to the moments preceding them. In contrast, an increase in skin conductance from the preceding context was seen for both LTHE and HTHE BPs (LTHE: B = 0.11, SE = 0.04, $p = .01$, CIs [0.03, 0.18]; HTHE: B = 0.13, SE = 0.04, $p = .001$, CIs [0.06, 0.2]) but not for LTLE. No other significant effects were found. Analyses using differently-sized baseline windows (up to 5 seconds and 15 seconds before BP) showed qualitatively similar results.

Next, the extent to which distinct BP types could be characterised in terms of changes in dynamics, pitch register, modality and harmonic and temporal ambiguity was examined. Figure 4 and Tables 9 and 10 summarize how these features changed in relation to the time window preceding the different BP types. These analyses showed that LTLE BPs did not differ from the preceding context in terms of dynamics, pitch register, modality or harmonic ambiguity (all $p > .05$). However, they showed that LTHE BPs were associated with increases in dynamics (RMS: B = –0.42, SE = 0.12, $p < .001$, CIs [–0.67, –0.42]), a tendency towards the major mode

(B =–0.23, SE = 0.12, p = .05, CIs [–0.46, 0]) and decreases in harmonic ambiguity (increases in Key clarity: B = –0.33, SE = 0.13, p = .011, CI [–0.56, –0.07]). In turn, HTHE BPs were associated with increases in dynamics (RMS: B = –0.63, SE = 0.13, p < .001, CIs [–0.85, –0.41]) and showed an additional association with increases in pitch register (Pitch: B = –0.31, SE = 0.12, p < .01, CIs [–0.54, –0.08]). Analysis of expert manual annotations of the BPs showed that LTLE BPs were lower in tempo than the higher arousal BP types (LTLE: M = 85.18 BPM, SD = 27.35; LTHE: M = 90.71 BPM, SD = 26.07; HTHE: M = 94.33 BPM, SD = 27.94) and that while no differences were observable across BP types in terms of expert ratings of consonance (all p > .05), less polyphony was found in LTLE BPs (M = 1.79; SD = 0.56) as compared to the other two BP types (LTHE: M=1.93, SD = 0.53; B = 0.21, SE = 0.07, p = 0.006; HTHE: M = 1.96, SD = 0.48, p < .05; B = 0.17, SE = 0.07, p = 0.05).

With regard to the style of music containing the BPs, the most BPs were found in pieces that could be described as ‘solo ballads’. Specifically, 72% of HTHE BPs, 62% of LTLE BPs, and 54% of LTHE BPs were in solo ballads. However, subtle trends were also observable across the BP types whereby LTHE BPs were very frequently in pieces best described as upbeat (33% of LTHE BPs, compared to 22 % and 17% of LTHE and HTHE respectively) while LTLE moments were very frequently in pieces from Classical and film music (16% of LTLE BPs compared to 14% and 11% of LTHE and HTHE respectively). Finally, with regard to the lyrical content of the BP subtypes, vocals/lyrics were found to be present in all BP types at high levels (79% of LTLE, 82% of LTHE and 88% of HTHE BPs).

Discussion

Although various strands of literature suggest that the experience of beauty may have an affective component, little has been done to investigate this empirically. We examined the evidence for the existence of distinct subtypes of beauty experiences in response to music, focusing specifically on responses to discrete self–identified beautiful passages within self-selected music.

We required participants to provide a list of musical pieces in which they considered particular passages to be particularly beautiful, and then, in the lab, had the participants first listen to their selected pieces while a range of physiological measures were recorded, and then provide subjective ratings on their experience of the BPs. Results of clustering analyses of the BPs based on emotional ratings were shown to be stable across different clustering methods and revealed three types of BPs (LTLE, LTHE, and HTHE) that differed primarily in terms of the patterns of Tension and Energy induced in the listener. These BP types, in turn, differed in terms of the degree of Interest and Liking reported in association with them. Critically, supporting the notion that the obtained clusters reflect different subtypes of beauty experiences to music, all three types of BPs also differed with respect to key physiological and musico-acoustic features.

Subtypes of beauty experiences in response to music

Both the philosophical and psychological literature suggested that, if explored, subtypes of the experience of beauty may be distinguishable by the arousal dimensions of the three-dimensional model (Levinson, 2012; Meyer, 1956). The results of our cluster analyses, which grouped BPs according to the three-dimensional emotion model, showed the arousal dimensions Tension and Energy to be the main discriminators of the three BP subtypes, which were otherwise similar in Valence judgments. Strong experiences in response to music have been argued to be highly multifaceted with affect (emotions and feeling) comprising an especially important component (Gabrielsson, 2003). Here, we show that the experience of beauty during music-listening, which may be considered a type of peak aesthetic experience, does indeed vary meaningfully in terms of affect.

In the context of art, positive emotional valence may be inseparably associated with positive aesthetic judgments. Thus, the finding of similarly positive Valence ratings across the three BP types likely reflects the fact that all BPs induced a largely positively valenced affective state in the listeners (Schindler et al., 2017). That in contrast to Valence, the Energy and Tension dimensions were highly discriminating of BPs, parallels previous work where dissimilarity judgments of musical excerpts using all three emotion variables were shown to be best explained by a two-dimensional space defined along the two arousal dimensions in particular (Vieillard, Peretz, Gosselin, Khalfa, Gagnon, & Bouchard, 2008).

The three-dimensional model has been widely used in music psychology studies due to its ability to capture the nuances in emotions that can be induced by music, whereby music-induced arousal may be vigorous and activated (Energetic arousal), or associated with uncertainty and anticipation (Tension arousal). In our study, a single dimension of arousal (as used in two-dimensional emotion models) would have been unable to distinguish the two high arousal BP subtypes we observed (LTHE and HTHE BPs), which not only differed in terms of subjective reports of Tension but also in terms of the musical features associated with them. Nor would a two-dimensional model have been able to distinguish clearly between the two low tension BPs (i.e. LTLE and LTHE BPs) which differed not just with respect to induced Energy but also with respect to induced Interest, musico-acoustic features and physiological variables.

By providing evidence that beauty that may be characterised by both low and high Tension, we provide some preliminary support for the notion of narrow and non-narrow beauty subtypes previously proposed in the philosophical literature (Levinson, 2012). Similarly, failing to show evidence for certain subtypes suggests a constraint on the patterns of emotional states that may accompany experiences of beauty in response to music. For instance, providing evidence for an HTHE cluster but failing to show evidence for high-tension-low-energy (HTLE) BP subtype, may indicate that induced Tension necessarily also "energises" in the context of beauty experiences. In sum, our results are valuable in demonstrating the relevance of considering musical beauty experiences in terms of the three-dimensional emotion model. We show its ability to not only clarify which dimensions of the affective state are particularly critical in distinguishing beauty experiences but also which subtypes of beauty experiences may be most prevalent for everyday listeners.

Engagement mechanisms: Greatest interest for high arousal BPs

Based on the literature suggesting different routes to aesthetic pleasure, we predicted that different BP subtypes would differ in terms of the mechanisms underlying them. Interestingly, we showed that while ratings for Interest and Empathy ratings, tended to be correlated with each other, Interest but not Empathy distinguished the different BP subtypes obtained. Specifically, in contrast to Empathy levels, which did not differ significantly across the three types of BPs, Interest levels were shown to be greatest for the two BP types that were associated with greater subjective and physiological arousal.

A possible outcome we anticipated was that greater reported Empathy would also be associated with higher arousal BP types (Jacobs, 2015). However, our finding that Empathy played a comparably sized role across BP types is plausible given that empathy is a mechanism by which all emotion states, including low arousal ones, may be induced (Omigie, 2016). In contrast, it can be argued that the cognitive stimulation from music may be more closely bound to objective features and may therefore be more likely to vary across subtypes of beauty experiences. Interestingly, that the BP subtypes observed differed in terms of induced Interest (LTHE and

HTHE greater than LTLE) is in line with Graf and Landwehr's suggestion that high Interest may underlie at least some positive aesthetic judgments (Graf & Landwehr, 2011). Indeed, while an influential theory suggests that beauty judgments are driven by processing fluency (Reber et al., 2004), and while such processing fluency is likely at play for the LTLE BPs (associated with the lowest levels of Interest), there exists strong evidence for a role of cognitive stimulation in eliciting beauty judgments. In one study in which participants were presented with artworks either with or without a distracting task, distraction was shown to reduce experiences of beauty, suggesting that an inability to fully cognitively process artworks can limit their aesthetic appeal (Brielmann & Pelli, 2017). Thus, taken together, our data provide support for the notion that while processing fluency may be one route to beauty experiences (Reber et al., 2004), high information content of arousing musical features may also spark Interest (Huron, 2006) and through that provide a route to aesthetic pleasure (Graf & Landwehr, 2011; Ngai, 2012) that results in a beauty judgments.

Aesthetic evaluations: Greatest liking for moderately arousing BPs

An outstanding question we aimed to examine with our data is the extent to which potentially different subtypes of beauty experiences may differ in their value for listeners. Here, we saw that while there were no differences in mean beauty ratings across BP subtypes, greater liking was reported for LTHE BPs, as compared to LTLE BPs. The failure of HTHE BPs to differ significantly from either LTLE or LTHE BPs precludes drawing firm conclusions regarding the nature of the relationship between arousal and preference. However, it is worthy of note here that our finding that the greatest liking ratings were assigned to the moderately-arousing LTHE BPs is in line with Berlyne's theory of aesthetic pleasure being related to an inverted U-shaped function of arousal (Berlyne, 1971).

It has been suggested that liking judgments may track a different form of aesthetic evaluation than beauty judgments (Reber et al., 2004), whereby liking correlates with an “I” related perspective and beauty with a more object centric perspective (Behne, 1986). Further, philosophical aesthetics has discussed a number of categories, other than beauty, that may lead to liking, such as the sublime, the simple or the comical (Vischer, 1846; Sibley, 1959; Ngai, 2012). That the three BPs differed in terms of the liking but not the beauty reported for them would seem to support such conceptual distinctions between beauty and liking judgments. Our results are interesting in suggesting that liking ratings may track preferences that are particularly related to the degree of emotional arousal and/or the interest afforded by particular passages of music, whereby moderate arousal and high interest levels tend to be preferred (Berlyne, 1971; Graf & Landwehr, 2011).

Here, however, it is important to consider the role that demographic factors might have played, since certain factors, such as age, have been shown to influence patterns of musical taste (Cohrdes, Wrzus, Wald-Fuhrmann, & Riediger, 2018; LeBlanc, Sims, Siivola, & Obert, 1996). Indeed, one possibility, that would need addressing in future work, is the possibility that any superior preference (highest liking) for LTHE BPs is specific to the demographic of the current sample, who were relatively young in age.

Distinctive physiological markers and acoustic features associated with BP subtypes

Finally, providing support for our assumption that emergent clusters reflect different subtypes of the beauty experience, our results showed unique patterns of both physiological and musico-acoustic signatures to be associated with these BP subtypes. LTLE moments were shown to be accompanied by increases in zygomaticus activity as well as increases in respiratory rate. Smiling as indexed by zygomaticus activity has long been associated with positively valenced emotions (Larsen, Norris, & Cacioppo, 2003) as well as the processing fluency (Reber et al., 2004) that has been ascribed to narrowly beautiful music (Levinson, 2012). Thus, the increased smiling seen here for LTLE in particular may be accounted for by the ease in processing (reflected in the lesser interest) associated with these passages compared to the other two BP types.

The increase in respiration rate seen for LTLE moments was, in contrast, less expected since an increase in respiration rate to music is typically seen as an index of arousal (Juslin & Sloboda, 2011). However, a closer examination of the literature also shows links between respiration rate and valence, whereby greater respiration rate has been associated with positively valenced music (Krumhansl, 1997), and breath length has been reported to be shortest during happiness (compared to sadness and fearfulness) induction (Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2005). More generally, rapid and deep respiration has been associated with positive compared with negative emotions (Bloch et al., 1991; Boiten, 1998; Collet et al., 1997; Wentjes, 1992) and thus, one possibility is that the increases in breathing rate shown for LTLE BPs reflects a tendency toward increasingly positive feelings relative to the preceding moments.

Further, support for the distinctiveness of the different BP subtypes may be seen in the finding that LTHE and HTHE showed an increase in a key measure of physiological arousal that was not seen for LTLE BPs. Increases in skin conductance are often linked to highly pleasurable physiological arousal in response to music (e.g., Salimpoor et al., 2009) and thus, the patterns of skin conductance change seen here suggest that high arousal BPs (HTHE and LTHE BPs) are more similar than LTLE BPs to so-called frisson, the highly arousing pleasurable responses to music that have widely been studied in previous work (Harrison & Loui, 2014).

Finally, in terms of the distinctiveness of musico-acoustic features associated with the different BP types, all three showed unique patterns in line with the subjective report and physiological variables associated with them. That the low arousal BP subtype (LTLE) was not associated with significant changes in dynamics, register, mode or harmonic ambiguity, all of which may result in increases in arousal, is in line with the idea that some forms of beauty do not confront or challenge (Levinson, 2012). The lower tempo and polyphony these BPs showed compared to the other BP types, may also have contributed to their low arousing qualities. In contrast, just as previous research suggests that change in the musical structure may lead to pleasurable music listening experiences (Nagel et al., 2007), increases in a range of musical features were observed for both LTHE and HTHE.

However, while both LTHE and HTHE were associated with a change in dynamics (an increase in mean pitch), other key features distinguished the two BP types. Specifically, HTHE musical moments were characterized by an increase in pitch register, in line with previous literature that has associated this feature with the occurrence of frisson, as well as age-old Western conventions of representing or expressing emotions via music (Hevner, 1936). An increase in mean pitch from before to during the HTHE BPs is also interpretable in view of music psychology studies showing higher pitches (both in terms of perception and production) to be more tension-inducing than lower ones (Ilie & Thompson, 2006; Lehne & Koelsch, 2015) as well as ethological studies showing a tendency for loud high-pitched sounds to be associated with alarm (Huron, 2006). Also interesting was the trend towards a change from major to minor modality (and an increase in key clarity) in HTHE BPs. Given the more negative associations listeners assign to the minor relative to the major key (Hevner, 1936), such a transition from the major to minor mode may have contributed to the feelings of this more negatively valenced arousal (tension) unique to this subtype. In contrast, the finding that LTHE was associated with an increase of tonal clarity and a tendency towards a major mode is in line with these features being associated with pleasant subjective feelings (Aljanaki & Widmer, 2018). Finally, that both LTHE and HTHE BPs were associated with greater polyphony and higher tempo than LTLE moments is coherent with the greater interest and subjective arousal that these beauty experiences were associated with.

Experience of Beautiful passages as peak experience.

Short-lived highly pleasurable experiences to music, referred to as frisson (Harrison & Loui, 2014), and first experimentally examined by Goldstein (1980) have seen a great deal of research in recent years (Panksepp, 1995; Grewe, Nagel, Kopiez & Altenmüller, 2007a; Grewe, Nagel, Kopiez & Altenmüller, 2007b; Nagel, Kopiez, Grewe & Altenmüller, 2008; Salimpoor, Benovoy, Larecher, Dagher & Zatorre, 2011). However, that body of research, in emphasizing pleasurable experiences associated with strong physiological responses, necessarily excludes consideration of other peak-like experiences in response to music that may not be associated with high physiological or subjective arousal. Our study, in taking the aesthetic judgment of beauty as opposed to physiological arousal as a starting point, provides evidence of the range of emotional qualities that may characterize positive aesthetic judgments in response to discrete passages in music.

Unfortunately, as we did not require participants to provide subjective ratings of any sections of music other than those considered BPs, it is difficult to speak to how subjective experiences of BPs in particular may compare to the experience of whole songs that a listener considers to be beautiful. In line with the idea that beautiful passages colour the whole experience of listening to music (Konrad, 2011), however, we suggest that whole pieces judged as beautiful may often be judged as such, precisely because they contain what the listeners consider to be beautiful passages. Studying beautiful passages may therefore lead to a better understanding of how holistic beauty judgments arise.

Closing statements, limitations and future outlook

The data-driven approach taken in our study revealed evidence of robustly differing subtypes of the experience of musical beauty and demonstrated that the broad concept of beauty, at least in the context of music, may need finer distinctions. Critically, our results carry the important message that empirical aesthetics investigations should, as far as possible, include emotion dimension measures that are able to differentiate these very different shades of beauty.

Unfortunately, our study is unable to speak to any potential defining features that may distinguish experiences of beauty from neutral feelings in response to music. The main aim of our study was to examine the possibility that different subtypes of beauty experiences exist and we provided support for this using a multimodal data set. However, a future study seeking to identify that which is unique to beauty experiences may profit from a design similar to that used in previous research on frisson (Salimpoor et al., 2011), whereby pairing of participants' stimuli allowed the comparison of those responses obtained when a given stimulus is considered beautiful with those when it is not considered beautiful. Another useful way would be to require participants to provide subjective report of the passages surrounding the beautiful passages. Examining the extent to which the emergent clusters that are associated with these non-beautiful passages, differ or not from those clusters found for beautiful passages, would help clarify whether there are any patterns unique to the experiences of beauty. On a related note, the current study was unable to reveal any one index of physiological responding that was common across all subtypes of beauty experiences. This is possibly because no one such index may exist. Indeed, as mapping patterns of physiological responding to even basic emotions (e.g., happiness, sadness) has proven problematic, it may be unrealistic to expect such simple mappings of physiological responses to more complex aesthetic emotions (e.g., the experience of beauty).

Given the robustness of the mere exposure effect, whereby repeated exposure leads to increasing preference (Zajonc, 2001), a question of relevance for future work is the extent to which any effects shown here are due to the listeners' high familiarity with the music heard. Previous work has shown the important influence of familiarity on experiences of music (Van Den Bosch, Salimpoor, & Zatorre, 2013) and in one study that also used a mostly pop/rock musical repertoire, emotion-related limbic and paralimbic regions as well as the neural reward circuitry were shown to be significantly more active for familiar than unfamiliar music (Pereira, Teixeira, Figueiredo, Xavier, Castro, & Brattico, 2011). Such findings suggest that familiarity may amplify emotional responses in response to BPs and, thus, while the main aim of this study was to examine responses to these highly familiar, highly valued musical excerpts, future work examining the role of familiarity in the experience of beauty is of great relevance.

Finally, one concern that may be had with regard to the current study is that by seeking out subtypes using cluster analysis, we inevitably found them. However, that our results showed a high number of qualitative differences between the emergent clusters (not just in terms of underlying mechanisms and aesthetic evaluations, but also in terms of physiological responding and musico-acoustic features), speaks to the validity of the observed subtypes. Perhaps, a more significant concern is that the current conclusions are based on data from a sample of somewhat younger participants, who proposed a limited range of musical styles. We suggest that systematically increasing the range of participant demographics and music styles may be particularly worth implementing in future work.

Acknowledgments: We are grateful to our participants and the MPIEA technicians and laboratory staff who assisted greatly in data collection. In addition, we wish to thank Winfried

Menninghaus and Christian Grüny for their valuable comments on earlier versions of this paper.

References

Adorno, T. W. (1997/1965). Schöne Stellen (1965). In R. Tiedemann (Ed.), *T. W. Adorno, Gesammelte Schriften Bd.18., Musikalische Schriften V* (pp. 695–718). Frankfurt a. M.: Suhrkamp.

Aljanaki, A., & Widmer, G. (2018). Modeling Majorness as a Perceptual Property in Music from Listener Ratings. arXiv preprint arXiv:1806.10570.

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage*, 59(4), 3677-3689.

Aristotle. (1961). *Poetics*, Butcher S. H., trans, Fergusson, F., intro. New York: Hill & Wang.

Armstrong, T., & Detweiler-Bedell, B. (2008). Beauty as an emotion: The exhilarating prospect of mastering a challenging world, *Review of General Psychology, 12*(4), 305-329.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 340–412.

Balkwill, L.L., Thompson, W. F., & Matsunaga, R. (2004.) Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research, 46*(4), 337–349.

Baltes, F. R., & Miu, A. C. (2014). Emotions during live music performance: links with individual differences in empathy, visual imagery, and mood. *Psychomusicology, Music Mind Brain 24*(1), 58–65.

Bates, D., Maechler, M., Bolker, B., & Walker S. (2014). Fitting Linear Mixed-Effects Models Using Ime4. *Journal of Statistical Software, 67*, 1–48.

Behne, K. E. (1986). *Hörertypologien: Zur Psychologie des jugendlichen Musikgeschmacks*.

[Listener typologies: On the psychology of adolescent musical taste] Regensburg: Bosse.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electro-dermal activity. *Journal of Neuroscientific Methods*, 190, 80–91.

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

Bloch, S., Lemeignan, M., & Aguilera-T, N., (1991). Specific respiratory patterns distinguish among human basic emotions. *International Journal of Psychophysiology, 11*(2), 141–154.

Boiten, F. A., Frijda, N. H., & Wientjes, C. J. (1994). Emotions and respiratory patterns: review and critical analysis. *International Journal of Psychophysiology 17*(2), 103–128.

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A Functional MRI Study of Happy and Sad Emotions in Music with and without Lyrics. *Frontiers in Psychology,* 2, 308

Brattico, E. (2015). From pleasure to liking and back: Bottom-up and top-down neural routes to the aesthetic enjoyment. Art, Aesthetics, and the Brain; Huston, J., Nadal, M., Mora, F., Agnati, L., Cela Conde, CJ, Eds, 303-318.

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology,* 4, 206.

Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. Psychology of Aesthetics, Creativity, and the Arts, 7, 48.

Brattico, E., Jacobsen, T., De Baene, W., Glerean, E., & Tervaniemi, M. (2010). Cognitive vs. affective listening modes and judgments of music–An ERP study. *Biological psychology, 85(*3), 393-409.

Brielmann, A. A., & Pelli, D. G. (2017). Beauty requires thought. *Current Biology, 27*(10), 1506-1513.

Burke, E. (1860). *A Philosophical Inquiry Into the Origin of Our Ideas of the Sublime and*

Beautiful: With an Introductory Discourse Concerning Taste. Harper and brothers.

Bundgaard, P. F. (2015). Feeling, meaning, and intentionality—a critique of the neuroaesthetics of beauty. *Phenomenology and the Cognitive Sciences, 14, 781-801*

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology 50,* 260–268.

Chuen, L., Sears, D., & McAdams, S. (2016). Psychophysiological responses to auditory change. *Psychophysiology, 53*(6), 891-904.

Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of Life Reviews, 15,* 61-88.

Cochrane, T. (2010). A simulation theory of musical expressivity. *Australasian Journal of Philosophy,* 88(2), 191-207.

Cochrane, T., Fantini, B., & Scherer, K. R. (Eds.). (2013). *The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control.* OUP Oxford.

Cohen, B. H., Davidson, R. J., Senulis, J. A., Saron, C. D., & Weisman, D. R. (1992). Muscle tension patterns during auditory attention. *Biological Psychology. 33*, 133–156

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis in the behavioral sciences (*3rd ed.). Mahwah, N J: Erlbaum.

Collet, C., Vernet-Maury, E., Delhomme, G., & Dittmar, A., (1997). Autonomic nervous system response patterns specificity to basic emotions. *Journal of Autonomic Nervous System, 62* (1– 2), 45–57.

Craig, D. G. (2005). An exploratory study of physiological changes during “chills” induced by music. *Musicae Scientiae, 9*(2), 273–287.

Cohrdes, C., Wurzus, C., Wald-Fuhrmann, M., & Riediger, M. (2018). “The sound of affect”:

Age differences in perceiving valence and arousal in music and their relation to music characteristics and momentary mood. *Musicae Scientiae* (Online First).

Cone, E. T. (1974). *The Composer's Voice.* Ernest Bloch Lectures, Berkeley: University of California Press.

Consoli, G. (2015). From beauty to knowledge: a new frame for the neuropsychological approach to aesthetics. *Frontiers in Human Neuroscience, 9,* 290.

Conway, B. R., & Rehding, A. (2013). Neuroaesthetics and the trouble with beauty. *PLoS Biology, 11(3),* e1001504.

Conway, B.R. & Rehding, A. (2013). Neuroaesthetics and the trouble with beauty. *PLoS Biol., 11*, e1001504.

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition & Emotion, 27*(4), 658-684.

Coutinho, E., & Cangelosi, A. (2011). Musical emotions: predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion, 11*(4), 921-937.

Davidson, K. R. Scherer, & Goldsmith, H. H (Eds.): *Handbook of Affective Sciences* (pp. 503–534), Oxford.

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music, 39*(1), 18–49.

Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal. 30,* 307-340

Eerola, T., Vuoskoski, J. K., & Kautiainen, H. (2016). Being moved by unfamiliar sad music is associated with high empathy. *Frontiers in psychology, 7*, 1176.

Egermann H, Pearce M, Wiggins G, & McAdams S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music.

Cogntive and Affective Behavioural Neuroscience, 13, 533–53.

Egermann, H., & McAdams, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception: An Interdisciplinary Journal, 31(*2), 139-156.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion review, 3*(4), 364-370.

Everitt, B. (1974). *Cluster Analysis. Council E for the SSR*, editor. London: Heinemann.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R.* Sage publications.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*(5), 567–589.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., & Turner, R. (2009). Universal recognition of three basic emotions in music. *Current biology, 19*(7), 573–576.

Gernot, G., Pelowski, M., & Leder, H. (2017). Empathy, Einfühlung, and aesthetic experience: The effect of emotion contagion on appreciation of representational and abstract art using fEMG and SCR. *Cognitive Processing,* 1-19.

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: Sage.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae,* Special Issue, 2001–2002: 123–147.

Gabrielsson, A., & Lindstrom, Wik S. (2003). Strong experiences related to music: A descriptive system. *Musicae Scientiae, 7*(2), 157–217.

Gabrielsson, A. & Juslin, P. N. (2003). Emotional expression in music, Oxford University Press

Goldstein A. (1980) Thrills in Response to Music and Other Stimuli. *Physiological Psychology*. 8(1),126–9.

Graf, L. K. M., & Landwehr, J. R. (2017). Aesthetic pleasure versus aesthetic interest: the two routes to aesthetic liking. *Frontiers in Psychology. 8*(15).

Graf, L. K. M., & Landwehr, J. R (2015). A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking. *Personality and Social Psychology Review, 19*(4), 395-410.

Grewe O, Nagel F, Kopiez R, Altenmüller E. (2007a). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion,* 7(4), 774–88.

Grewe O, Nagel F, Kopiez R, Altenmüller E. (2007b). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception, 24*(3), 297-314.

Grewe O, Kopiez R, Altenmüller E. (2009). The Chill Parameter: Goose Bumps and Shivers as Promising Measures in Emotion Research. *Music Perception.* 27(1), 61–74.

Greenberg, D. M, Baron-Cohen, S., Stillwell, D. J., Kosinski, M., Rentfrow, P. J. (2015) Musical Preferences are Linked to Cognitive Styles. *PLoS ONE* 10(7): e0131151.

Hanslick, E., & Payzant, G. (1986). *On the Musically Beautiful: A Contribution Towards the Revision of the Aesthetics of Music*. Hackett Publishing.

Harrison L, & Loui P. (2014). Thrills, chills, frissons, and skin orgasms: toward an integrative model of transcendent psychophysiological experiences in music. *Frontiers in Psychology,* 5, 790.

Hay, D. (1990). *Religious experiences today: Studying the facts.* London, UK: Mowbray.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, 48(2), 246-268.

Hodges, D. A. (2009). *Bodily responses to music. The Oxford handbook of music psychology,* 121-130.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: MIT Press.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception: An Interdisciplinary Journal,* 23(4), 319-330.

Ilie, G., & Thompson, W. F. (2011). Experiential and cognitive changes following seven minutes exposure to music and speech. *Music Perception: An Interdisciplinary Journal,* 28(3), 247-264.

Ishizu, T., & Zeki, S. (2011). Toward a brain-based theory of beauty. *PloS One,* 6(7), e21852.

Istók, E., Brattico, E., Jacobsen, T., Krohn, K., Müller, M., & Tervaniemi, M. (2009). Aesthetic responses to music: A questionnaire study. *Musicae Scientiae.*

Jacobs, A. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, 9, 186.

Jacobsen, T. (2010). Beauty and the brain: culture, history and individual differences in aesthetic appreciation. *Journal of Anatomy,* 216(2), 184-191.

Jacobsen, T., Buchta, K., Köhler, M., & Schröger, E. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychological Reports,* 94(3 Pt 2), 1253-1260.

Judd, C. M., Westfall, J., Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology.* 103(1):54–69.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Physics of life reviews,* 10(3), 235-266.

Juslin, P. N., & Isaksson, S. (2014). Subjective criteria for choice and aesthetic judgment of music: A comparison of psychology and music students. *Research Studies in Music Education,* 36(2), 179-198.

Juslin, P. N., & Sloboda, J. (Eds.). (2011). *Handbook of music and emotion: Theory, research, applications.* Oxford University Press.

Kreutz, G., Schubert, E., & Mitchell, L. A. (2008). Cognitive Styles of Music Listening. *Music Perception, 26,* 57–73.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(4), 336.

Krumhansl, C. L., (2002). Music: a link between cognition and emotion. *Current Directions in Psychological Science, 11*(2), 45–50.

Lartillot, O., Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. In: *International Conference on Digital Audio Effects.*

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology, 40*(5), 776-785.

LeBlanc, A., Sims, W. L., Siivola, C., & Obert, M. (1996). Music style preferences of different age listeners. *Journal of Research in Music Education, 44*(1), 49-59.

Lehne, M. & Koelsch, S. (2015). Toward a general psychological model of tension and suspense. *Frontiers in Psychology, 6*.

Levinson, J. (2005). Musical Expressiveness as Hearability-As-Expression, in Contemporary Debates in Aesthetics and the Philosophy of Art, ed. Matthew Kieran, Oxford: Blackwell.

Levinson, J. (2012). Musical Beauty. *Teorema Rev Int Filos*, 127–135.

Levinson, J. (2015). *Musical concerns: Essays in philosophy of music*. OUP Oxford.

Lipps, T. (1903). Einfühlung, innere Nachahmung und Organempfinden. *Archiv für die gesamte Psychologie 1,* 185–204

Lipps, T. (1903). *Grundlegung der Ästhetik* (Vol. 1). Voss.

Lowis M. J. (2003). Peak emotional experiences and their antecedents: A survey of staff at a British university college. *Korean Journal of Thinking & Problem Solving, 13*, 41–53.

Lundqvist, L. O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music, 37*(1), 61-90.

Madsen, C. K., Brittin, R. V., & Capperella-Sheldon, D. A. (1993). An empirical method for measuring the aesthetic experience to music. *Journal of Research in Music Education, 41*(1), 57-69.

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in human neuroscience, 10*, 536.

Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. (2013). Individual differences in music reward experiences. *Music Perception: An Interdisciplinary Journal, 31*(2), 118-138.

Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology, 24*(6), 699-704.

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PloS one, 10*(6), e0128451.

Menninghaus, W. et al (2018). What Are Aesthetic Emotions? Epub ahead of print.

Meyer L. B. (1956). *Emotion and Meaning in Music.* Chicago: Chicago University Press.

Miu, A. C., & Baltes F. R. (2012). Empathy manipulation impacts music-induced emotions: a psychophysiological study on opera. *PLoS One*. 7(1), e30618.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS one,* 9(2), e89642.

Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons—*An ERP study. International Journal of Psychophysiology*, 76(1), 40-51.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of Classification, 31*, 274-295.

Nagel, F., Kopiez, R., Grewe, O., & Altenmüller, E. (2008). Psychoacoustical correlates of musically induced chills. *Musicae Scientiae, 12*(1),101–113.

Ngai, S. (2008). Merely interesting. *Critical Inquiry*, *34*(4), 777-817.

Ngai, S. (2012). *Our aesthetic categories: Zany, cute, interesting*. Cambridge, MA: Harvard University Press.

Nieminen, S., Istók, E., Brattico, E., Tervaniemi, M., & Huotilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex,* 47(9), 1138-1146

Ogg, M., Sears, D. R., Marin, M. M., & McAdams, S. (2017). Psychophysiological Indices of Music-Evoked Emotions in Musicians. *Music Perception: An Interdisciplinary Journal, 35*(1), 38-59.

Omigie, D. (2015). Music and literature: are there shared empathy and predictive mechanisms underlying their affective impact? *Frontiers in Psychology. 6*,1250.

Omigie, D. (2016). Basic, specific, mechanistic? Conceptualizing musical emotions in the brain. *Journal of Comparative Neurology, 524*(8), 1676-1686.

Panksepp, J. (1995). The emotional sources of chills'' induced by music. *Music Perception, 13*(2), 171–207.

Panzarella, R. (1980). The phenomenology of aesthetic peak experiences. *Journal of humanistic psychology,* 20(1), 69-85.

Payne, E. (1980). Towards an understanding of musical appreciation. *Psychology of Music,* 8(2), 31-41.

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The Cognitive Neuroscience of Aesthetic Experience. *Perspectives on Psychological Science,* 11(2), 265-279.

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PloS One, 6(11), e27241*.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and social psychology review, 8*(4), 364-382.

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music, 32*(4), 371-388.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–78.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14*(2), 257–62.

Sartwell, C., (2017). Beauty, The Stanford Encyclopedia of Philosophy (2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/beauty/>.

Scherer, K., & Zentner, M. (2008). Music evoked emotions are different–more often aesthetic than utilitarian. *Behavioral and Brain Sciences, 31*(5), 595-596.

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality. 14*(4), 325–45.

Schimmack, U., & Reisenzein, R. (2002). Experiencing activation: Energetic arousal and tense arousal are not mixtures of valence and activation. *Emotion, 2*(4), 412–417.

Sibley, F. (1959). Aesthetic Concepts, *Philosophical Review,* 68, 421–450.

Silvia, P. J. (2005). Cognitive appraisals and interest in visual art: Exploring an appraisal theory of aesthetic emotions. *Empirical studies of the arts*, 23(2), 119-133.

Silvia, P. J., Fayn, K., Nusbaum, E. C., & Beaty, R. E. (2015). Openness to experience and awe in response to nature and music: Personality and profound aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 376-384.

Schlegel (1797). *Über das Studium der griechischen Poesie, Neustrelitz, Kritische Friedrich-Schlegel-Ausgabe*. Erste Abteilung: Kritische Neuausgabe, Bd. 1, München.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and non-emotional feelings. *Journal of Personality and Social Psychology. 64*, 211–220.

Stone-Davis, F. J. (2011). *Musical Beauty: Negotiating the Boundary between Subject and Object.* Wipf and Stock Publishers.

Suzuki, M., Okamura, N., Kawachi, Y., Tashiro, M., Arao, H., Hoshishiba, T., … Yanai, K. (2008). Discrete cortical regions associated with the musical beauty of major and minor chords*. Cognitive, Affective & Behavioral Neuroscience,* 8(2), 126-131.

Tatarkiewicz, W. (1972). The great theory of beauty and its decline. *Journal of aesthetics and Art Criticism, 31*(2), 165-180.

Tatarkiewicz, W. (1963). Objectivity and Subjectivity in the History of Aesthetics. *Philosophy and Phenomenological Research,* 24, 157-173.

Thagard, P. (2019). *Natural Philosophy: From Social Brains to Knowledge, Reality, Morality, and Beauty (Treatise on Mind and Society).* Oxford University Press.

Trost, W., Ethofer, T., Zentner, M., & Vuilleumier, P. (2011). Mapping aesthetic musical emotions in the brain. *Cerebral Cortex, 22*(12), 2769-2783.

Van Den Bosch, I., Salimpoor, V., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience,* 7, 534.

Vischer, F. T. (1846). Aesthetik oder Wissenschaft des Schönen. Zum Gebrauch für Vorlesungen. Reutlingen: Mäcken.

Vischer, R. (1873). *On the optical sense of form: A contribution to aesthetics*. In H. F.

Mallgrave, & E. Ikonomou (Eds., transl.) (1994), Empathy, form, and space: Problems in German aesthetics, 1873–1893. Chicago, IL: University of Chicago Press.

Vuoskoski, J. K., & Eerola T. (2011). Measuring music-induced emotion A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae, 15*, 159–173.

Vuoskoski, J. K., & Miu, A. C. (2017). The social side of music listening: Empathy and contagion in music-induced emotions. In *Music and Empathy* (pp. 124-138). Routledge.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion, 22*, 720–752.

Team RDC (2008). *R: a language and environment for statistical computing*.

Schoen, M. (1931). The Experience of Beauty in Music. *The Musical Quarterly, 17*(1), 93-109.

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., et al. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLoS One, 12*(6), e0178899.

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R. & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One, 4*(10), e7487.

Shell, S. (2002), Kant as Propagator. Reflections on “Observations on the Feeling of the Beautiful and Sublime”. *Eighteenth-Century Studies*, 35, 455–468.

Starr, G. G. (2013). *Feeling beauty: The neuroscience of aesthetic experience*. Mit Press.

Vischer, F. T. (1922). *Aesthetik oder Wissenschaft des Schönen* [1846–58]. 8 Bde. Hg. v. Robert Vischer. München, 2, 23.

Ward J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*, 236–44.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence: The structure of emotion. *Journal of personality and social psychology, 76*, 820–838.

Wientjes, C.J. (1992). Respiration in psychophysiology: methods and applications. *Biological Psychology, 34*(2–3), 179– 203.

Whaley J., Sloboda J. A., Gabrielsson A. (2009). Peak experiences in music. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 452–461). Oxford, UK: Oxford University Press.

Witvliet, C. V., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion, 21*(1), 3-25.

Wundt, W. (1896). *Grundriss der Psychologie.* Leipzig, Germany: Engelmann.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current directions in psychological science,* 10(6), 224-228.

Zangwill, N. (2003). *Beauty.* In Jerrold Levinson (ed.), Oxford Companion to Aesthetics. Oxford University Press.

Zangwill, N. (2014). Re-Centring Musicology and the Philosophy of Music. *Journal of Aesthetics and Phenomenology 1(*2), 231-240.

Zeki, S. (2013). Clive Bell's "Significant Form" and the neurobiology of aesthetics. *Frontiers in Human Neuroscience,* 7, 730.

Zeki, S. (2014). Neurobiology and the humanities. *Neuron,* 84(1), 12-14.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion, 8*(4), 494.

Figure captions

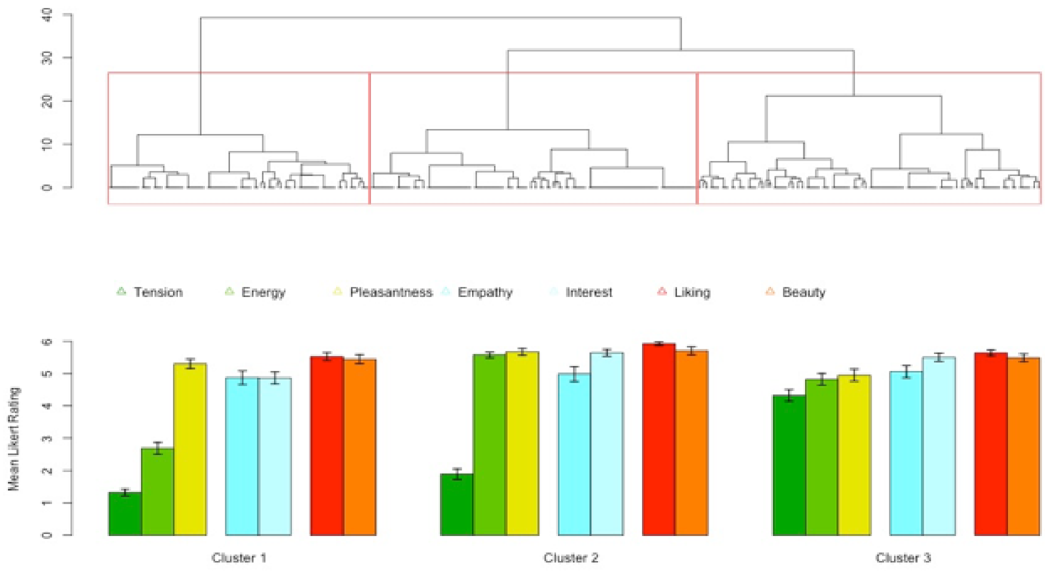

Figure 1: A cluster dendrogram showing the results of a hierarchical cluster analysis of all BPs according to emotion dimensions (top row). The red boxes indicate the three obtained clusters while bar plots show characterization of the clusters in terms of mean ratings on the Likert scales (1 to 6) (bottom row). Error bars indicate standard error of the mean.

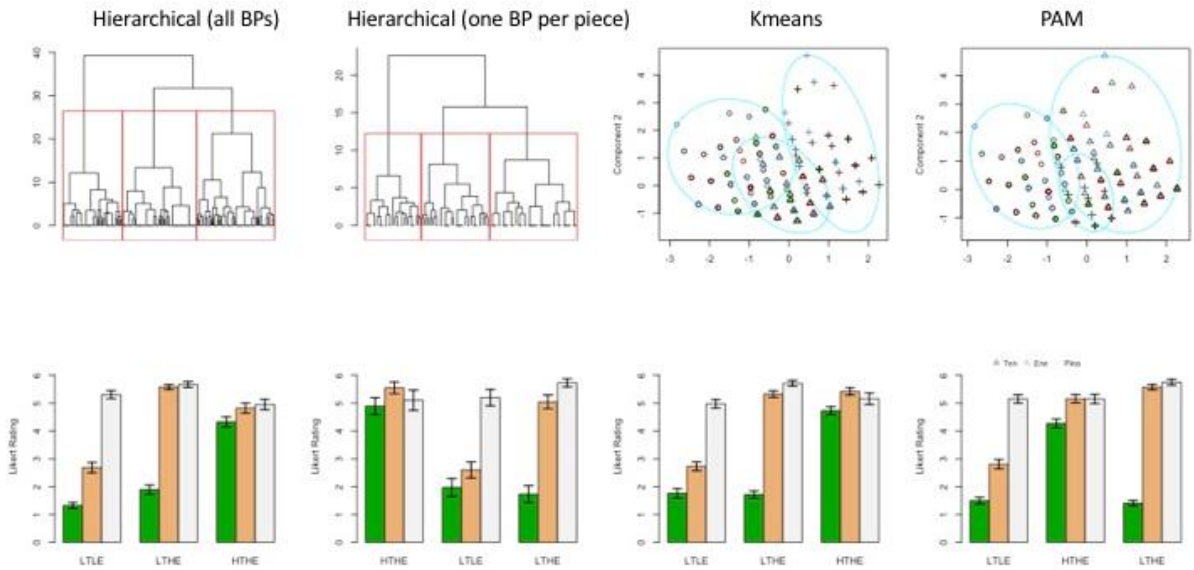

Figure 2: Illustrating the comparability of the cluster solutions obtained from clustering Beautiful passages (BPs) with hierarchical cluster analysis (Column 1) with solutions obtained from the hierarchical clustering of a subset of the BPs (Column 2), clustering using Kmeans (Column 3) and clustering using PAM Column 2). LTLE: Low Tension/Low Energy, LTHE: Low Tension/High Energy, HTHE: High Tension/High Energy. Error bars indicate standard error of the mean.

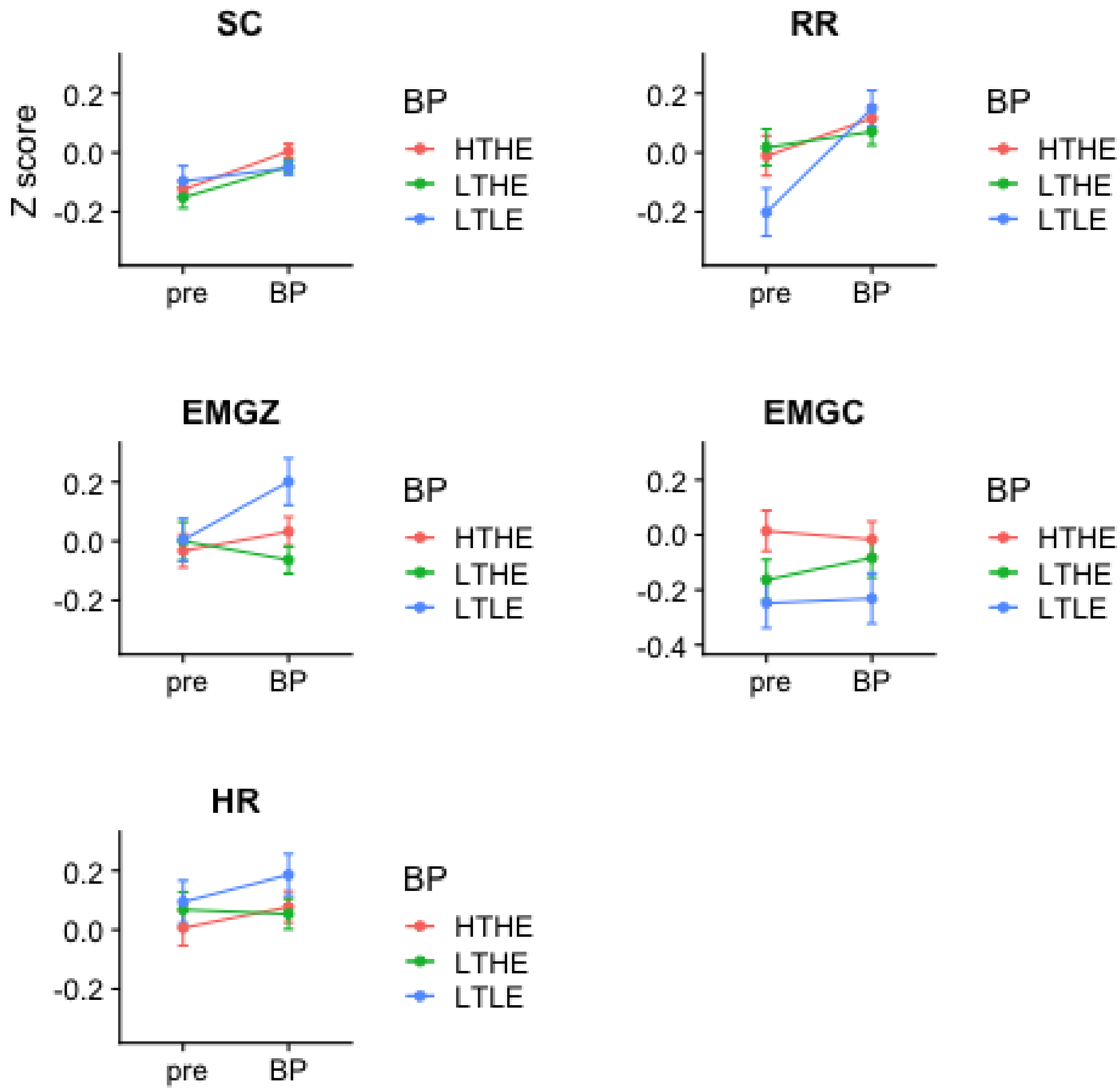

Figure 3: Plot showing how physiological responses during the three different Beautiful passages (BP) types compare to physiological responses preceding them**.** LTLE: Low Tension/Low Energy, LTHE: Low Tension/High Energy, HTHE: High Tension/High Energy. Error bars indicate standard error of the mean.

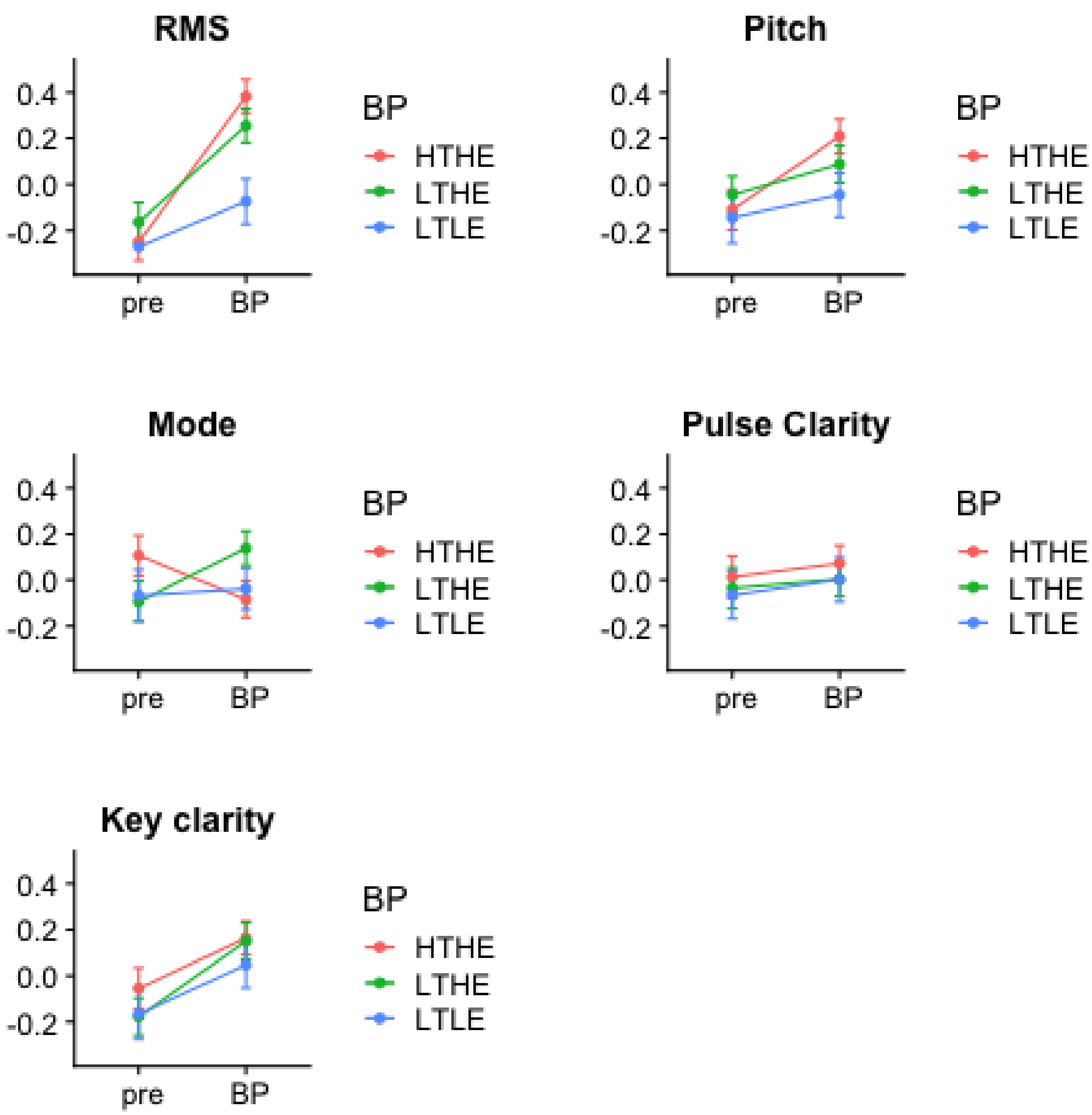

Figure 4: Plot showing how the different musico-acoustic features of the Beautiful passages (BPs) differed from the preceding musico-acoustic context. LTLE: Low Tension/Low Energy, LTHE: Low Tension/High Energy, HTHE: High Tension/High Energy. Error bars indicate standard error of the mean