The Physiology of Musical Preference: A Secondary Analysis of the *Study Forrest* Dataset

Sophia Kaltsouni Mehdizadeh & Grace Leslie
University of Colorado Boulder

The dominant research strategy within the field of music perception and cognition has typically involved new data collection and primary analysis techniques. As a result, numerous information-rich yet underexplored datasets exist in publicly accessible online repositories. In this paper we contribute two secondary analysis methodologies to overcome two common challenges in working with previously collected data: lack of participant stimulus ratings and lack of physiological baseline recordings. Specifically, we focus on methodologies that unlock previously unexplored musical preference questions. Preferred music plays important roles in our personal, social, and emotional well-being, and is capable of inducing emotions that result in psychophysiological responses. Therefore, we select the *Study Forrest* dataset "auditory perception" extension as a case study, which provides physiological and self-report demographics data for participants (N = 20) listening to clips from different musical genres. In Method 1, we quantitatively model self-report genre preferences using the MUSIC five-factor model: a tool recognized for genre-free characterization of musical preferences. In Method 2, we calculate synthetic baselines for each participant, allowing us to compare physiological responses (pulse and respiration) across individuals. With these methods, we uncover average changes in breathing rate as high as 4.8%, which correlate with musical genres in this dataset ($p < .001$). High-level musical characteristics from the MUSIC model (mellowness and intensity) further reveal a linear breathing rate trend among genres ($p < .001$). Although no causation can be inferred given the nature of the analysis, the significant results obtained demonstrate the potential for previous datasets to be more productively harnessed for novel research.

Received: July 13, 2022, accepted May 17, 2023.

Key words: musical preference, STOMPR, secondary analysis, heart rate, breathing

Music is a form of human communication that is found across cultures and eras, actively or passively permeating aspects of social interaction, relationships, memories, emotional state, self-identity, and more (Cross, 2001). Our preferences for certain types of music may begin to form relatively early on and then continue to evolve, playing non-trivial roles throughout our life (Bonneville-Roussy & Eerola, 2018; Bonneville-Roussy et al., 2017). For example, musical preference has been found to reveal some aspects of lifestyle (North & Hargreaves, 2007; Pettijohn et al., 2012), personality (Chamorro-Premuzic et al., 2010; North, 2010; Rentfrow & Gosling, 2003), identity and self-view (Hargreaves et al., 2008; Rentfrow & Gosling, 2006), and cognitive style (Greenberg et al., 2015).

Individuals also commonly use music for self-regulation of emotional state or levels of arousal (Lonsdale & North, 2011; Schäfer & Sedlmeier, 2009). Music is capable of inducing a range of emotions of various valence and arousal levels, accompanied by strong and immediate effects on the listener's physiology (Bigand et al., 2005; Juslin & Laukka, 2004; Lundqvist et al., 2009). This may be mediated by several psychological mechanisms—most comprehensively outlined in the BRECVEMA framework (Juslin et al., 2010; Juslin & Västfjäll, 2008)—such as the brain stem reflex (Davis, 1984), episodic memory (Baumgartner, 1992), musical expectancy (Steinbeis et al., 2006), and others. Two or more of these mechanisms can be recruited simultaneously, which this framework suggests might explain mixed or conflicting emotional responses to certain stimuli (Juslin, 2013). Therefore, in music perception and emotion research, especially that which focuses on measuring the neurological or physiological outputs of these mechanisms, it is critical to recognize and attempt to isolate the involved mechanisms to avoid conflicting or inconclusive results.

One direction of music emotion research investigates responses to *objective* musical characteristics: acoustic features like pitch, tempo, loudness, and timbre, or event-related features like sound onsets or section transitions. For example, acoustic features such as loudness and tempo have been shown in multiple studies to

Music Perception, volume 40, issue 5, pp. 393–407, issn 0730-7829, electronic issn 1533-8312. DOI: https://doi.org/10.1525/mp.2023.40.5.393

correspond with levels of physiological arousal, as measured through heart rate, breathing, and other cardiorespiratory metrics (Bernardi et al., 2006; Coutinho & Cangelosi, 2011; Juslin et al., 2014; Nyklíček et al., 1997; Russo et al., 2013). Musical biofeedback systems have successfully directed participants' breathing patterns using music's tempo, both with (Siwiak et al., 2009) and without (Leslie et al., 2019) their awareness.

Other studies explore responses to more *subjective* features of music: individual preference, appreciation, or familiarity. For instance, studies of affective video stimuli suggest that amusement may be accompanied by activation of the parasympathetic nervous system, with high heart rate variability (HRV) observed during approach-motivated emotions compared to avoidance-motivated emotions (Wu et al., 2019). It has also been shown that pleasurable responses to preferred music correlate with an aroused response from the autonomic nervous system (increased heart rate and respiration), as well as activation in regions of the brain associated with reward (Blood & Zatorre, 2001). Lingham and Theorell (2009) found this physiological arousal response (specifically, increased heart rate) occurred for both high and low arousal "favorite" music. However, the effects were more pronounced (and included an increase in breathing rate) for high arousal favorite music and were smaller and less reliable for the low arousal favorite music. Altogether, these studies suggest that both subjective characteristics (such as individual preference) and objective musical features jointly contribute to one's psychophysiological response to music.

A recurring challenge in studies of musical preference is the variability and subjectivity of music style categories, or genre labels. One concern is limitations that may arise from trying to restrain or conform musical stimuli to what are often broad or ill-defined categories. Individuals may choose to define certain genres differently, and these definitions may drift as genres evolve over time. This issue is particularly apparent with umbrella categories such as "rock," "pop," "metal," etc. These definitions might not only differ between participants themselves, but also with the researcher's own categorizations (Lamont & Greasley, 2008; Lex et al., 2020). Several alternative methods and models have been developed in attempts to eliminate negative effects of genre labels in research, and instead capture attributes of musical preference in ways that are "genre-free." One example is the three-dimensional model that maps various musical characteristics (such as auditory, affective, and instrumental descriptors) to the dimensions arousal, valence, and depth (Fricke et al., 2018, 2021; Greenberg et al., 2016). Rentfrow and Gosling (2003) identified four independent dimensions of musical preference through developing the Short Test of Musical Preferences (STOMP). Results from the STOMP provided a clear way to translate one's musical preferences from genre labels to quantitative, "genre-free" loadings on each of the four identified factors. The four-factor model was eventually updated following subsequent analyses, creating the revised STOMP (STOMP-R) based on a better-fitting, five-factor model. These five orthogonal factors of musical preference are (1) Mellow, (2) Unpretentious, (3) Sophisticated, (4) Intense, and (5) Contemporary ("MUSIC"). Styles of music are given standardized loadings on each of these five factors, providing a more general and quantifiable characterization (Rentfrow et al., 2011).

In this paper, we explore physiological correlates of both objective (musical characteristics) and subjective (individual preference) features of musical stimuli through the secondary analysis of the publicly available *Study Forrest* dataset (Hanke et al., 2015). Secondary analyses can be feasible, efficient, and collaborative ways of exploring multiple questions, and can provide a facilitating step towards the collection of new data. However, they may also present a unique set of challenges depending on the design of the dataset being used and the questions attempting to be asked. The primary objective of this paper is to present a feasibility report of performing these post hoc analyses on a dataset not originally designed to answer these questions, and to share the methodologies we developed to work with the given data.

From the *Study Forrest* dataset we analyze the "auditory perception" subtask, in which 20 participants listened to 25 short excerpts of researcher-selected songs from five different genres of music. Functional magnetic resonance imaging (fMRI) data, respiratory and finger photoplethysmogram (PPG) traces are provided, along with responses to a pre-task questionnaire where participants report their demographics, music training, and a list of favorite genres and musical artists. We select this dataset as a case study for the proposed secondary analysis feasibility report on account of 1) its accessibility, versatility, and documentation, which was of particular significance at the time of this work when much in-person data collection was halted due to the COVID-19 pandemic, 2) the naturalistic design of the music listening task, and 3) the types of data that were recorded during the experiment, which support our research questions into the effects of music and preference on listener physiology. It is important however to keep in mind the dataset limitation of small sample size; the Discussion section contains more details regarding limitations.

The analyses on this dataset to date have primarily explored the fMRI BOLD responses for distinct patterns between the five genre groups, individual songs, and overarching musical features (Casey, 2017; Hanke et al., 2015). In contrast, we focus on the unexplored participant self-report musical preference information, and the respiratory and pulse data recorded during the music listening task. As these data were not central to the primary analyses, we encounter two significant challenges: 1) lack of participant stimulus ratings, and 2) lack of physiological baseline recordings. In the following section we discuss in detail how we overcome these by utilizing 1) the MUSIC five-factor model, and 2) synthetic physiological baseline calculations. These methods unlock previously unexplored analyses on objective and subjective features of the musical stimuli, and their physiological correlates.

To facilitate our secondary analysis feasibility report, we develop two hypotheses with regards to the new analyses we can conduct on this data. We hypothesize that participants' degree of physiological response—absolute percent change in heart rate (HR), breathing rate (BR), and heart rate variability (HRV)—will positively correlate with their preference for each of the stimulus genres. In other words, a greater absolute change from baseline physiological readings will be seen when participants are listening to stimuli they are likely to prefer. Previous work has suggested pleasurable responses to observed stimuli correspond with increased heart rate, heart rate variability, and respiration (Blood & Zatorre, 2001; Wu et al., 2019). However, there is also evidence of possible interactions or conflicting effects of the arousal level of the stimulus (Lingham & Theorell, 2009; Russo et al., 2013). For this reason, our hypothesis focuses on the *absolute* change from baseline to account for the potential varying effects of high and low arousal music. We also hypothesize that participants' physiological response (percent change in HR, BR, and HRV) will correlate with the "mellow" and "intense" MUSIC factor correlations of the stimuli. The "mellow" and "intense" factors are more directly correlated with music loudness and tempo than the other three factors of the MUSIC model (Rentfrow et al., 2011). As reviewed above, these psychoacoustic features in particular have been found in other research to correlate with levels of physiological arousal. We therefore predict observing negative and positive changes in physiological arousal for more "mellow" and more "intense" music, respectively.

Our hypotheses are founded on results from controlled, primary studies in which stimuli and experimental conditions are designed to examine specific physiological effects. Correlational studies, such as this one, on the association between music and physiology during naturalistic music listening are less common. However, the growing availability of heterogeneous datasets continues to gain attention from members of the music perception and cognition community (Greenberg & Rentfrow, 2017; Huron, 2013). Such analyses may call for novel data wrangling or transformation techniques to uncover latent connections between auditory, behavioral, and cognitive variables (for example, Liikkanen et al., 2015). Although done on a relatively small sample size, we believe the following methodological practices, results, and feasibility discussion offer applicable suggestions for how previous datasets may be more productively harnessed for novel research.

Method

This secondary analysis uses the physiological data (cardiac and respiratory) collected in the "auditory perception" extension of the *Study Forrest* dataset (Hanke et al., 2015). All data are acquired from the *Study Forrest* website (www.studyforrest.org) where they have been made available for public reuse under the ODC Public Domain Dedication and License by the original researchers. An overview of the study and data are given in this paper before addressing the methods used in this analysis.

STUDY FORREST DATASET BACKGROUND

Participants

The participants in this study were 20 right-handed, native German-speakers (mean age: 26.6 years, 12 male). According to the pre-study questionnaire, two participants reported being musicians with formal music training (one with five years, the other with seven years). Four other participants identified themselves as musicians with at least two years of experience (mean reported: 7.25 years, maximum reported: 12 years) playing one or two instruments (but did not report any formal training). Five additional participants did not identify themselves as musicians but did report playing one or two musical instruments for at least two years (mean reported 5.5 years, maximum reported: 7 years). Of the remaining nine participants, seven reported being "music lovers" with no music training or experiences. Two participants did not respond to this question.

Task

The original task of this dataset had participants listen to the audio movie version of the film "Forrest Gump" while physiological signals and high-resolution (7 Tesla) fMRI data were recorded. For additional information on the original study, task, or participant details, readers are directed to the respective paper (Hanke et al., 2014). An extension to the original dataset was published approximately a year later with the same individuals participating in a music listening task. The data recorded from this task were the focus of this analysis.

Stimuli

Each participant listened to the same 25 naturalistic music stimuli. These were six-second-long song clips, five each from five musical genres: ambient, country, heavy metal, rock'n'roll, and symphonic. Pieces were selected by the researchers to be representative of naturalistic music listening (i.e., commercially released song recordings), with the clips taken from the middle of each file and aligned to begin on a down-beat when possible. Participant familiarity or liking for the stimuli or five genre categories was not reported (although some participants happened to include some of these genres in their self-reported favorite genres). The same stimuli had also been used in a previous study by the research team investigating representation of different musical styles' auditory features (pitch, harmony, spectrum, timbre) in the brain (Casey et al., 2012). Spectrograms for all 25 stimuli are shown in Figure 1 of Hanke et al. (2015). Country, heavy metal, and rock'n'roll pieces contained vocals, while ambient and symphonic pieces did not. The root mean square power of all stimuli was normalized.

Experiment Structure & Data Recording

There were eight repeated experimental runs, meaning each participant heard each stimulus a total of eight times. Separating the presentation of each stimulus was an alternating four, six, or eight seconds of silence. fMRI as well as PPG and respiratory trace were recorded from the participants. It is important to note that physiological signals were truncated to begin and end with the first and last stimulus of the run, respectively. No baseline physiological measurement ("resting" measurement) was included with the data. For additional information on the "auditory perception" extension to the *Study Forrest* dataset, readers are directed to Hanke et al. (2015).

CALCULATING MUSICAL PREFERENCE

In the pre-study questionnaire, participants were asked to report their top three favorite genres of music for active, passive, and live concert listening, as well as provide examples of pieces or artists they enjoy. Overall, the musical preference information was subjective and highly variable with respect to the level of detail provided by each participant. On average each participant provided four unique genres (minimum: 3, maximum: 6) and 4–5 favorite music examples (minimum: 0, maximum: 19). Five of the 20 participants did not provide any favorite music examples. Some participants used vague genre labels ("rock," "charts," "instrumental") while others referenced smaller subgenres ("triphop," "darkwave," "Irish folk"). Importantly, participants were not asked to directly rate their liking for the 25 stimuli, or the five genres these came from. Therefore, it was necessary for us to develop a method for characterizing participants' musical preferences in a generalizable and quantifiable way. This would facilitate connecting participant preferences to the stimuli they were presented with in the experiment, as well as comparing preferences between participants using a common scale.

The STOMP-R and MUSIC model were the most feasible way of achieving this. Traditional use of the STOMP-R involves individuals rating a set of musical genres each on a scale from 1–7 (1: "dislike strongly," 7: "like strongly"). Each genre has a documented correlation with each of the five MUSIC factors. Individuals receive a score on each factor based on their ratings of the genres. These scores are then a quantitative and genre-free way of characterizing the individual's musical preferences. *Study Forrest* participants did not complete the STOMP-R during the original study, nor was it possible for us to administer the test to them post hoc. However, Rentfrow et al. (2011) provides the five-factor correlations for 26 musical genres (see Table 4 of that manuscript). By coding the preference information in *Study Forrest* in terms of those genres, we are able to simulate STOMP-R style results for the *Study Forrest* participants. Stimulus and preference coding was done using MATLAB (version 2019b). More details on this method are provided below.

Defining Participant Preferences in the Five-Factor Model

For each participant, their self-reported favorite genres are coded in terms of the genres provided in Table 4 of Rentfrow et al. (2011). Any music examples reported are used to validate and inform the coding process and perform corrections and additions if necessary. Coding is further informed by the genre taxonomies in Silver et al. (2016). For example, if a participant reported enjoying rock, metal, electro, and jazz, and artists such as The White Stripes and Muse, we select the genres "alt rock," "metal," "electronica," and "traditional jazz" from Table 4 of Rentfrow et al. (2011). The five-factor correlations for each of these genres are taken from the table. The participant's overall "score," or quantitative representation of their preferences, is then calculated as the five-dimensional centroid of the genres. All participants' reported genres and music examples from the *Study Forrest* dataset are weighted equally during this process, regardless of the reported listening contexts and listening frequency. We decide to do this given the observed inconsistencies in the pre-study questionnaire responses, to overcome the heterogeneity of this behavioral data.

Defining Stimulus Genres in the Five-Factor Model

We apply a similar methodology to acquire five-factor correlations for the five *Study Forrest* stimulus genres as well. The correlations for heavy metal are taken from the "heavy metal" entry in Table 4 of Rentfrow et al. (2011). The correlations for rock'n'roll are taken from the "rock'n'roll" table entry, after verifying there was a common artist listed for this genre category in both Rentfrow et al. (2011) and Hanke et al. (2015). For the symphonic stimulus genre, we use the correlations from the "classical" category in Table 4 of Rentfrow et al. (2011).

The remaining two stimulus genres, ambient and country, are not directly listed in the reference table and have to be interpolated using similar genres. Similar genres are once again informed by Silver et al. (2016). For country, these are "mainstream country," "new country," "country rock," and "bluegrass." For ambient, we select "classical," "avant-garde classical," "electronica," and "adult contemporary." The five-factor correlations of the identified similar genres are taken from Table 4 of Rentfrow et al. (2011) and used to calculate five-dimensional centroids representing the country and ambient stimulus genres, respectively.

CALCULATING PHYSIOLOGICAL RESPONSE

We analyze eight runs of physiological data for each of the 20 participants. Three channels of physiological data were made available for each run and subject, as well as the volume acquisition trigger pulse emitted by the MRI scanner. The physiological channels were (1) the respiratory trace, (2) the cardiac trace, and (3) the oxygen saturation (Hanke et al., 2015). In this analysis, we focus on the respiratory and cardiac data.

The respiratory trace was provided by a Siemens respiratory belt and pressure sensor (Honeywell 40PC001B1A). The cardiac trace (PPG) was measured by a Nonin 8600 FO pulse oximeter. Both signals were digitized with a 12-bit converter at 200 Hz and truncated according to the MRI trigger pulse recorded in channel 1. This was the state in which the data were publicly available (Hanke et al., 2015). All preprocessing and calculations are carried out using MATLAB (version 2019b).

Pre-processing

We apply a lowpass filter to the raw respiratory signal to help eliminate any high-frequency noise. We select a second-order Butterworth filter with cutoff of 1 Hz (Leslie et al., 2019; Russo et al., 2017). For the PPG signal, we first apply a bandpass filter to help eliminate powerline interference of 50–60 Hz (Nonin Operator's Manual: Models 8600FO and 8600FOM Pulse Oxymeters, 2005) as well as signal drift (Elgendi, 2012; Elgendi et al., 2015). We select a second-order Butterworth filter with bandwidth of 0.5–15 Hz (Elgendi et al., 2014). Next we calculate the acceleration plethysmogram (APG), which is the second derivative of the PPG, using a three-point center derivative (Elgendi et al., 2015). The APG is more commonly used than the PPG and can also be more closely related to electrocardiogram (ECG) (Elgendi, 2012; Elgendi et al., 2014). Lastly, a peak detection algorithm is applied to both the filtered respiratory signal and the APG signal to identify breaths and heartbeats (represented by a-waves in the APG signal), respectively. For APG a-wave detection, the method from Elgendi et al. (2014) is followed. Outlier detection is done on the vectors of detected peaks to eliminate falsely detected or missed peaks. For the APG signal, a-a intervals shorter than 0.428 seconds and longer than 1.5 seconds (corresponding to heart rates less than 40 bpm and over 140 bpm) are excluded. Additionally, a moving average for a-wave amplitude is calculated—peaks that do not fall within +/- 30% of this mean are rejected (van Gent et al., 2018).

Alternative Baselines

The provided physiological data were truncated at the beginning and end of each run by the original researchers, meaning that no baseline measurements were made available in this dataset by which to normalize each participant's physiological response. Therefore, an alternative method for calculating baseline physiological information needs to be employed. As there is no universal solution to this problem, and it is highly dependent on the aims of the study (Quintana & Heathers, 2014), we ultimately opt to use the participant average (over all eight runs) as the baseline.

Metric Calculation

From the respiratory and APG signals, we calculate four metrics: breathing rate (BR), heart/pulse rate (HR),

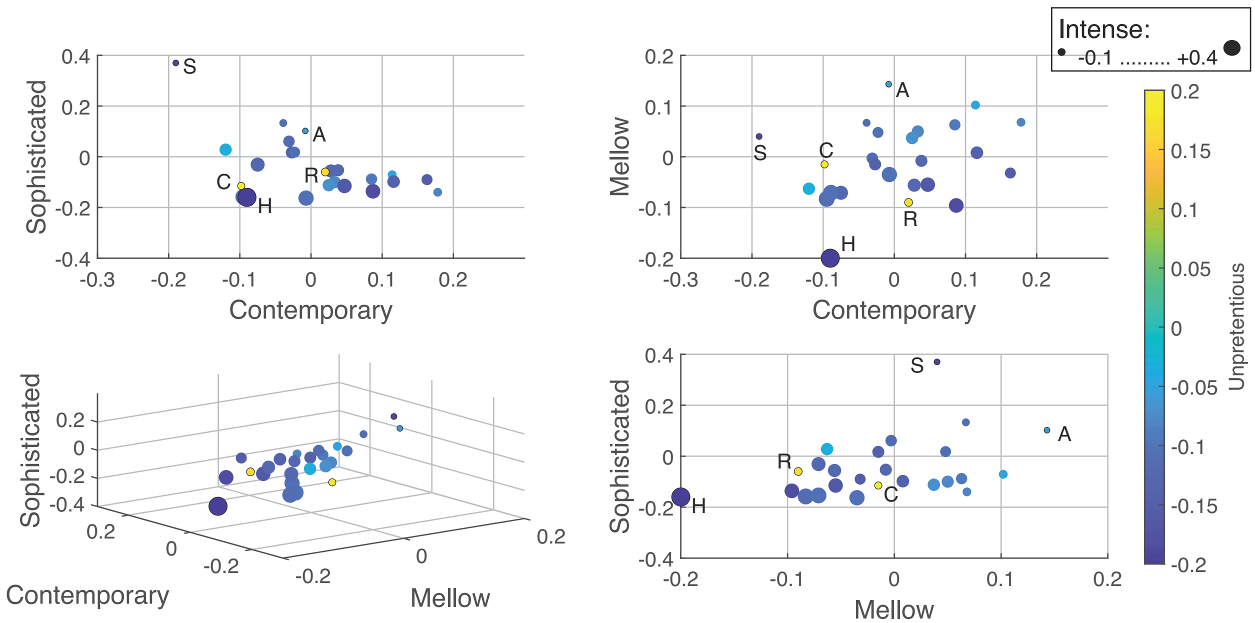

FIGURE 1. Visual representation of 5D MUSIC model space, with each factor mapped to either a spatial dimension, point size ("intense"), or color ("unpretentious"). Projections onto 2D space are also shown. Stimulus genres (labeled points by genre first letter: A- ambient; C- country; H- heavy metal; R- rock'n'roll; S- symphonic) are plotted using five-factor correlations found from the MUSIC model paper (Rentfrow et al., 2011). All *Study Forrest* participants (unlabeled points) are similarly represented in the space by averaging the MUSIC five-factor correlations for their self-report genre preferences.

standard deviation of N-N intervals (SDNN), and root mean square of successive R-R interval differences (rMSSD). BR and HR represent number of breaths, or heart beats, per minute. Because we analyze an APG signal as opposed to an ECG signal, SDNN is instead defined as the standard deviation of *a-a* intervals, and rMSSD is defined as the root mean square of successive *a-a* interval differences. Equations 1–4 show how these metrics are calculated using the inter-respiratory (IRI) and *a-a* intervals described previously (Elgendi et al., 2010).

$$BR = \frac{60}{IRI} \tag{1}$$

$$HR = \frac{60}{aa\ interval} \tag{2}$$

$$SDNN = std(aa\ intervals) \tag{3}$$

$$rMSSD = \sqrt{mean[(aa_i - aa_{i-1})^2]} \tag{4}$$

These metrics are calculated during each six-second window where there was a stimulus presented. Six seconds is considered to be a relatively short measurement epoch, particularly in the context of heart/pulse rate variability (HRV/PRV) metrics (Shaffer & Ginsberg, 2017). However, we decide to proceed with this window size as opposed to using a slightly larger ten-second window, which would include four seconds of rest in addition to the six-second stimulus. This is based on prior literature that suggests aesthetic evaluations for music are made on a very short time scale (< 3 seconds) (Belfi et al., 2018). The six-second measurement epoch limits the types of HRV measurements that can be reliably calculated, given that it falls within the lower range of ultra-short-term HRV measurements. We therefore decide to limit the derived metrics to SDNN and rMSSD time-domain metrics, as previously mentioned, and not calculate any frequency-domain metrics (Castaldo et al., 2019; Salahuddin et al., 2007; Shaffer & Ginsberg, 2017). PRV, or HRV metrics derived from PPG signal analysis, has been previously investigated as an estimate of HRV and generally found to be a suitable alternative (Pinheiro et al., 2016).

Change in BR, HR, SDNN, and rMSSD is calculated by subtracting the baseline metric from the metric during the six-second stimulus presentation. Then a percent change is calculated as shown in Equations 5–8, which

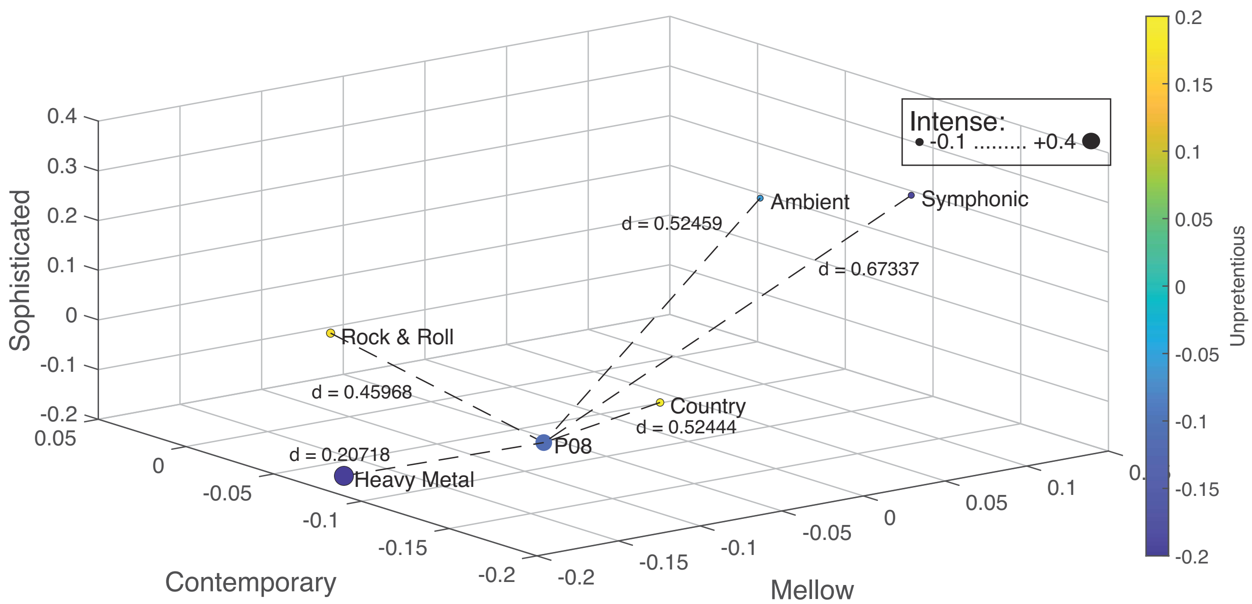

FIGURE 2. Participant 08 (center point), visualized within the 5D MUSIC model space alongside the five stimulus genres. Even though *Study Forrest* participants did not rate their preference for the five stimulus genres, translating the stimulus genres and participants' general musical preferences into the MUSIC model space allows us to quantitatively estimate this missing information. The estimate is shown via the calculated preference distances, *d*, between the participant and each stimulus genre. In this example, heavy metal has the shortest preference distance, suggesting that this participant would be most likely to prefer stimuli from this genre.

creates the normalized physiological metrics we use in our analysis.

$$\%\Delta BR = \frac{BR_{stim} - BR_{avg}}{BR_{avg}} \times 100 \tag{5}$$

$$\%\Delta HR = \frac{HR_{stim} - HR_{avg}}{HR_{avg}} \times 100 \tag{6}$$

$$\%\Delta SDNN = \frac{SDNN_{stim} - SDNN_{avg}}{SDNN_{avg}} \times 100 \tag{7}$$

$$\%\Delta rMSSD = \frac{rMSSD_{stim} - rMSSD_{avg}}{rMSSD_{avg}} \times 100 \tag{8}$$

Results

MUSIC AND MUSICAL PREFERENCE

With all 20 participants, as well as all five stimulus genres represented using the MUSIC five-factor model (Figure 1), a five-dimensional Euclidean distance (Equation 9) is calculated between each participant p, and each stimulus genre g. This results in a series of five "preference distances" per participant, where the lowest of the five distances represents the stimulus genre most likely preferred by that participant (Figure 2).

$$\begin{aligned} &Preference\ Distance\ (p, g) \\ &= \Big[(participant_{p,M} - stimulus_{g,M})^2 \\ &\quad + (participant_{p,U} - stimulus_{g,U})^2 \\ &\quad + (participant_{p,S} - stimulus_{g,S})^2 \\ &\quad + (participant_{p,I} - stimulus_{g,I})^2 \\ &\quad + (participant_{p,C} - stimulus_{g,C})^2\Big]^{1/2} \end{aligned} \tag{9}$$

The preference distance provides a quantitative estimate of the participants' liking for the stimuli they were shown in the experiment, based on their reported favorite genres in the pre-study questionnaire (Figure 3). When qualitatively compared, preference estimates mostly appear to resemble participants' questionnaire responses. However, some participants who reported a wider range of genres appear to be less well-represented by the estimate. For example, Participant 2 had reported metal as one of their favorite genres, but this is not reflected by their calculated preference distances in Figure 3. Instead of calculating a strong preference for metal, it appears

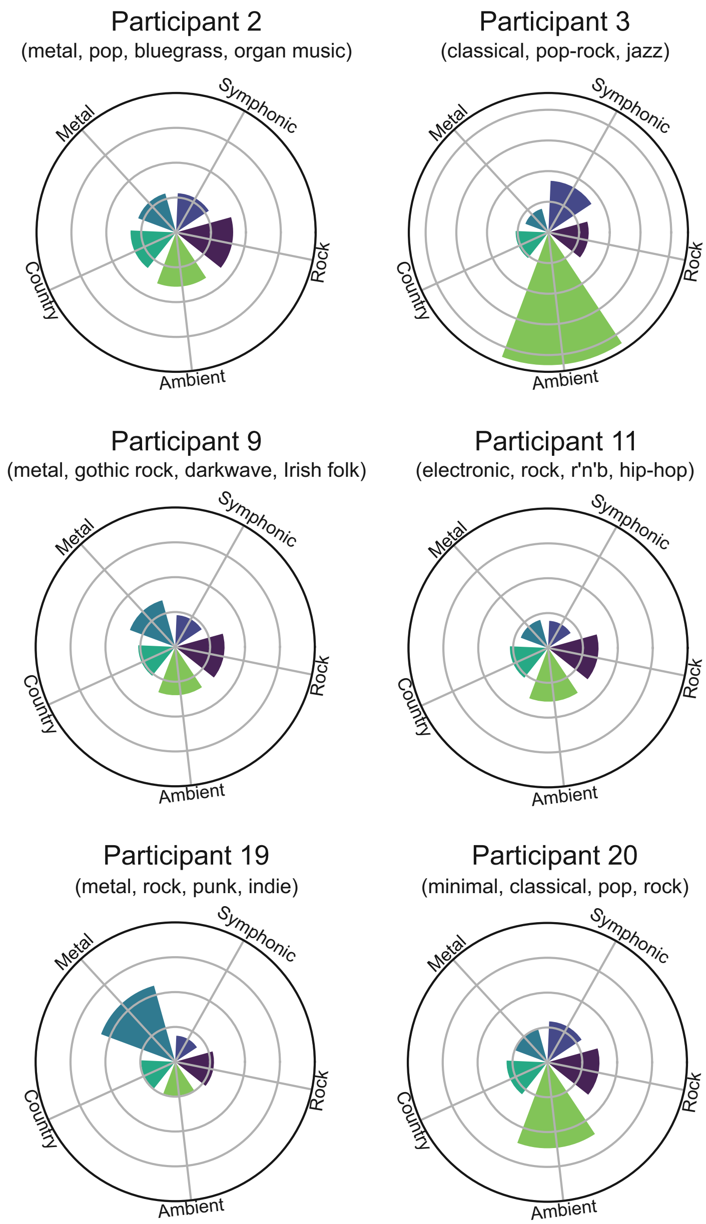

FIGURE 3. Six participants' estimated stimulus genre preferences are visualized here as an example of how different self-reported music tastes (listed above each plot) were mapped onto five the *Study Forrest* genres. Inverse preference distance is represented by bar size, such that a larger bar indicates higher estimated preference for that genre. For some participants, preference estimates appear to closely match their questionnaire responses (e.g., Participants 9, 19). However, this appears to not be the case for participants who reported a wide range of favorite genres (e.g., Participants 2, 11). Preference for ambient music appears to be frequently overestimated for participants (e.g., Participants 3, 20), perhaps due to our MUSIC interpolation of the missing ambient genre from styles including classical and electronica.

that this was diminished in the process of averaging the rest of the participant's preferences, which included pop, bluegrass, and soundtracks.

Comparing the distribution of preference distances across all participants by genre reveals some imbalances in the dataset (Figure 4). On average, participants are estimated to have higher preference for the ambient and rock stimulus genres and lower preference for the symphonic genre. This imbalance appears to mirror the pre-study questionnaire: 16 out

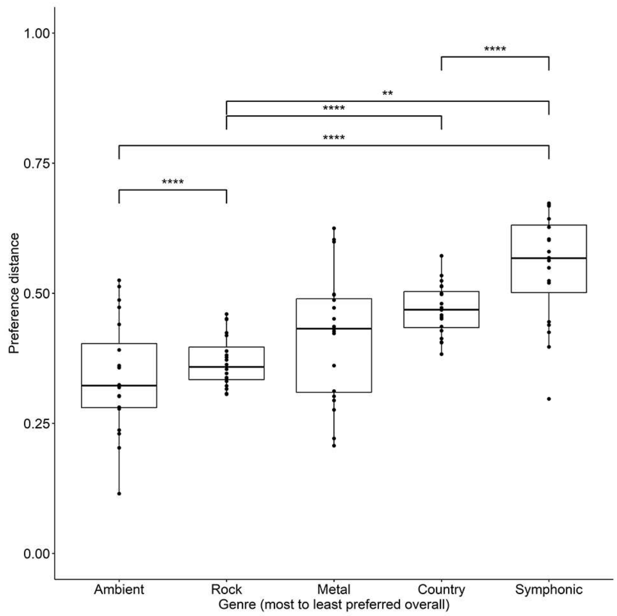

FIGURE 4. Distribution of all participants' calculated preference distances by stimulus genre. Note that values for preference distance are unitless and are derived from the five-dimensional Euclidean distance calculation shown in Equation 9. Pairwise comparisons are indicated by the horizontal bars (Bonferroni-corrected) and the significance is indicated by the asterisks (**$p < .01$, ****$p < .0001$). The distribution reveals that estimated genre preferences across participants are not balanced in this dataset, with higher estimated preferences on average for ambient and rock music, and lower estimated preference for symphonic and country. This trend appears to mostly mirror qualitative information in the pre-study questionnaire, except for the ambient genre.

of the 20 participants included rock as one of their favorite genres or music examples, while only six participants included classical music styles in their response. Although only five participants reported ambient styles in their response, this genre has the highest estimated preference across all participants. The way in which we represented the ambient stimulus genre in the MUSIC model space may be responsible for this discrepancy. Since ambient was not one of the genres listed in the MUSIC factor correlation table we were referencing, we had to represent it using the average of similar genres such as classical and electronica (see Method). Fourteen out of the 20 participants in the dataset included electronic styles as one of their favorite genres. In the absence of direct ratings of the stimuli or stimulus genres, the preference distance estimate is used to represent participants' preferences in the remainder of the analyses. All analyses are performed using SPSS (version 26) and R version 4.2 following tests of normality.

TABLE 1. *Pearson Correlation Coefficients for Fixed Effects in Preference Distance Mixed Model*

Physiological measure	$\|\%\Delta HR\|$	$\|\%\Delta SDNN\|$	$\|\%\Delta BR\|$
$\|\%\Delta HR\|$	1.000		
$\|\%\Delta SDNN\|$	.238	1.000	
$\|\%\Delta BR\|$	.292	.213	1.000

TABLE 2. *Results From Linear Mixed Effects Model for Preference Distance*

Fixed Effects	*Estimate*	*Std. Error*	*df*	*t*	*p*
Intercept	0.415	0.014	78.04	29.310	< .001
$\|\%\Delta HR\|$	6.05e-04	1.55e-03	653.2	0.391	.696
$\|\%\Delta SDNN\|$	5.84e-05	1.95e-04	774.7	0.300	.764
$\|\%\Delta BR\|$	6.00e-04	4.82e-04	192.9	1.244	.215

DEGREE OF PHYSIOLOGICAL RESPONSE VS. MUSICAL PREFERENCE

We create a linear mixed effects model to test if any of the physiological measures calculated are significant in modelling participants preference distances. Since we calculated the preference distance estimate on the level of genres and not songs (stimuli), we first average together the five stimuli per genre, per run, per participant (resulting in 800 datapoints, 20 participants x 8 runs x 5 genres, not accounting for NAN values). Absolute percent change in HR, SDNN, and BR are included as fixed effects in the model. Absolute percent change in RMSSD is not included because it correlates with absolute percent change in SDNN ($\rho = .87$). Correlations among the other three physiological variables are given in Table 1. Random intercepts for participant and run (nested within participant) are also included. The model is run in R using the *lmer* function from the *lmerTest* library. We use the maximum likelihood estimation method since our data are balanced and have nested random effects. Although we had hypothesized that degree of physiological response would correlate with preference, none of the physiological metrics in the model reach statistical significance (Table 2).

PHYSIOLOGICAL RESPONSE VS. MUSIC FACTORS

We use a repeated measures ANOVA to determine if characteristics of the musical stimuli—specifically how "mellow" or "intense" the stimulus was according to the MUSIC model—have a significant effect on physiological response. Stimulus genres are the within-subjects factor, which has five levels (one per genre) arranged in order from most to least "mellow." The order for "intense" is the same but reversed. No between-subjects factors are included. Results are summarized in Table 3. The repeated measures ANOVA test shows

TABLE 3. *Results From Repeated Measures ANOVA Within-subjects Effect of Stimulus Genre on Physiology*

Physiological measure	$df_{hypothesis}$	df_{err}	F	η_p^2	p	Power
%ΔHR[a]	2.491	27.401	1.334	0.108	.283	0.290
%ΔSDNN	4	44	1.229	0.100	.313	0.352
%ΔrMSSD	4	44	0.614	0.053	.654	0.186
%ΔBR	4	52	8.902	0.406	<.001	0.999

[a] Greenhouse-Geisser correction was used, assumption of sphericity was violated according to Mauchly's test, $\chi^2(9) = 20.23$, $p < .05$.

that stimulus genre did not have a significant effect on percent change in HR or HRV (SDNN or rMSSD). However, there was a significant effect on percent change in BR, $F(4, 52) = 8.902$, $p < .001$, $\eta_p^2 = 0.406$. The rock'n'roll genre had the highest average increase in breathing rate at 4.82% from baseline. Ambient was the only stimulus genre to cause an average decrease in breathing rate at -0.135% from baseline. Furthermore, tests of within-subjects contrasts reveal this effect to be linear with respect to the relative "mellowness" or "intensity" of the genres, $F(1, 13) = 20.65$, $p < 0.001$, $\eta_p^2 = 0.614$. Post hoc tests with the Bonferroni confidence interval adjustment show a significant difference between the most "mellow" (least "intense") genre, ambient, and every other genre ($p < .05$) (Figure 5). These results suggest that: 1) participants' breathing was affected to a greater degree than their heartrate in this study, and 2) considering higher-level features of the musical stimuli reveals interpretable breathing rate trends among stimulus genres.

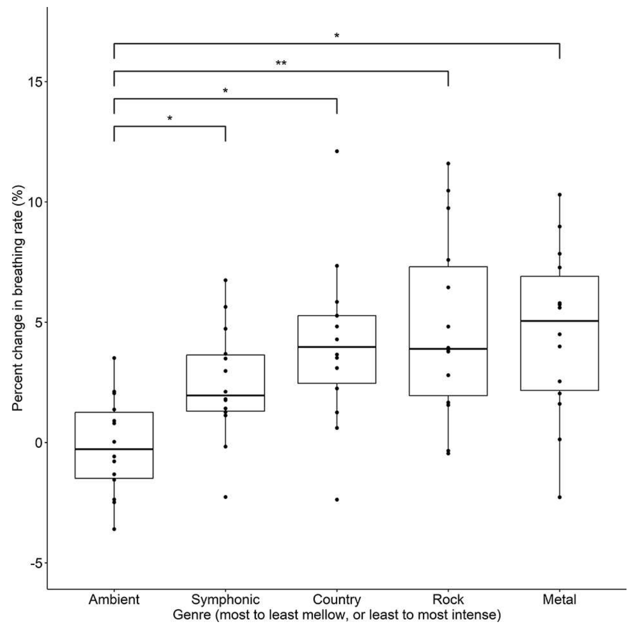

FIGURE 5. A repeated measures ANOVA test revealed a significant relationship between stimulus genre and breathing rate, analyzed as a percent change from baseline. Specifically, the relative "mellowness" of the genres (per the MUSIC model) had a positive linear relationship with breathing rate. Post hoc pairwise tests showed a significant difference between the most "mellow" genre, ambient, and every other genre (**p* < .05, ***p* < .01). Pairwise comparisons are Bonferroni-corrected.

Discussion

In this paper, we perform two post hoc physiological analyses on the "auditory perception" extension of the *Study Forrest* dataset. The first analysis addresses physiological correlates of individuals' estimated preference for the musical stimuli. Because participants did not directly rate their preference for stimuli they heard, we estimate this given the musical preference information they provided in the pre-study questionnaire. Additionally, since no physiological baseline data were provided, we calculate synthetic baselines for each participant in order to compare responses across individuals. Our novel application of the MUSIC five-factor model (Rentfrow et al., 2011) results in a quantitative "preference distance" estimate that maps participants' general musical preferences onto the five stimulus genres used in the experiment (Figure 3). Developing these methods was critical to adapt the existing dataset to our needs. However, given the provided data, we are unable to objectively evaluate the reliability of our "preference distance" estimate. Qualitative analyses of the metric suggest that it was more representative for participants that reported homogeneous favorite music styles, while it did not fully capture the range of those who reported varied genre preferences. Furthermore, the distribution of calculated preference distances across all participants (Figure 4) revealed an imbalance in preferred genres for this participant population. Comparing the estimated preference distributions to the pre-study questionnaire suggested that the way in which we interpolated the MUSIC factor correlations for the ambient genre influenced the preference distances for this genre. Future analyses should use caution and avoid such interpolations if possible. Studies with more participants would now need to be done to further develop the methods established here and determine the most accurate calculation for the preference distance. More ideas are given in the Future Work section.

Our hypothesis that degree of physiological response would correlate with preference for the experimental stimuli was unsupported: the analysis of the physiological data reveal no significant correlations with participants' estimated preferences. However, this is a result of limited reliability and may be significantly influenced by limitations of the secondary analysis. Musical preference is a complex, highly individualized phenomenon. One possible interpretation of the null result could be that the physiological metrics analyzed (HR, HRV, and BR) were not optimal selections for this particular question. Another possibility is that the design of the original study might not facilitate this preference analysis. Previous findings demonstrate that people experience higher-intensity emotional responses to familiar music, and that there is a correlation between familiarity and preference for music (Ali & Peynircioğlu, 2010). Given that all the stimuli in this experiment were researcher selected, and that participant familiarity was not measured, it is possible that the pieces of music were mostly unfamiliar to the participants and therefore did not induce a measurable response. We discuss this in terms of directions for future work in the following section. Additionally, since the primary focus of the original study was not musically induced emotion, the design did not allow for isolation of the possible emotion induction mechanisms. It is possible for multiple mechanisms to have been recruited simultaneously in a single participant's trial, as well as different mechanisms recruited across participants, leading to conflicting physiological results (Juslin et al., 2010). Furthermore, the relatively small sample size (N = 20 participants) of this dataset could be another limitation.

The second analysis addresses physiological correlates of high-level characteristics of the musical stimuli. Specifically, we compare the five stimulus genres according to their relative "mellowness" or "intensity" from the MUSIC model. We found significant differences between genres in precent change in BR, but not HR or HRV. This effect was linear with respect to the relative "mellowness" or "intensity" of the genres, with the most "mellow" (least "intense") genre (ambient) decreasing BR on average -0.135% from baseline, and the least "mellow" (most "intense") genre (metal) increasing BR on average 4.577% from baseline (Figure 5). An additional post hoc analysis supports that this is not explained by mimicking or entrainment to the stimulus tempo; a linear mixed model with fixed effect of tempo and random intercepts of participant and run (nested within participant) did not find tempo to be a significant term for percent change in BR, $t(2954, 3110) = 1.323$, $p = .186$, $AIC = 27746$ (t-test from the R *lmerTest* library uses Satterthwaite's method). However, we are unable to confirm that the change in BR is unrelated to the presence of vocals in the stimuli. Stimuli from two (ambient and symphonic) out of the five genres did not include vocals. These two genres also happen to be the most "mellow" (least "intense") in our dataset. Further research, likely using new stimuli controlled for vocals, would need to be conducted in order to investigate potential confounding effects.

As it stands, the current result supports part of our second hypothesis that physiological response would correlate with the "mellow" and "intense" MUSIC factors, and points to this being an unconscious effect of higher-level features in the musical stimuli. Breathing is a unique autonomic function, in that it is also able to be consciously controlled and is sensitive to external stimuli, such as music. Studies have previously shown that music can affect breathing both during focused (Etzel et al., 2006; Russo et al., 2013) and unfocused (Leslie et al., 2019) listening, substantiating our observation that breathing is not necessarily consciously entrained to the musical stimulus. While there is evidence that respiration can mimic low-level musical characteristics like tempo, many of the studies demonstrating this effect have used longer stimuli than the clips in the *Study Forrest* data (Bernardi et al., 2006; Etzel et al., 2006; Sakaguchi & Aiba, 2016). This difference in stimulus length could explain why we did not find a direct influence of tempo in our analysis. Few studies have investigated the effects of higher-level features (like style or instrumentation) while controlling for tempo. Our exploration of the *Study Forrest* dataset contributes an analysis of breathing rate over a range of musical genres, highlighting the roles such features can play in controlled breathing practices (Russo et al., 2017), as well as potential implications for other related autonomic activities such as heart rate and heart rate variability (Song & Lehrer, 2003).

Future Work

One potential limitation of the current preference distance estimate is that it reduces one's range of musical preferences to a single point by taking the centroid of all the reported preferences in the MUSIC model space. Thus, this would not capture the breadth of one's musical preferences, but rather interpolate a single most-representative point. This is illustrated by some of the examples shown in Figure 3 (e.g., Participant 2). The metric could be further developed by considering the convex hull formed by all a person's reported preferences in the MUSIC model space. The convex hull would create multidimensional surfaces connecting reported preferences, which might signify more nuanced interpolations and preserve more details compared to the single-point centroid representation. The preference distance between a person and some unknown stimulus would then be the minimum point-to-surface distance between the stimulus (point) and the person's preference (convex hull surface). This would theoretically maintain small distance calculations for stimuli that happen to be included in the individual's self-reported preferences (i.e., a distance of zero), addressing the issue we observed in the current implementation. Furthermore, while the current preference distance equation uses a five-dimensional Euclidean distance, different metrics (such as cosine similarity) should also be considered and may be more appropriate given the dimensionality of this problem. Additional analyses on data which include "ground truth" information would need to be done to determine if this would increase the accuracy of the preference estimate. Future work will also test this on a greater sample size and incorporate individualized, or participant-selected stimuli. This has been done in previous studies to induce stronger emotional and physiological responses in participants (such as Blood & Zatorre, 2001; Salimpoor et al., 2009). The greater preference contrast that could be provided by combining participant- and researcher-selected stimuli (or familiar and unfamiliar stimuli) may aid in revealing how preference distance can be modeled by physiological response.

Conclusion

In this paper we present a novel methodology for quantitatively estimating preference for unknown musical stimuli through the application of the MUSIC five-factor model. This "preference distance" metric was critical in facilitating our secondary analysis of the *Study Forrest* dataset, and with further development could increase feasibility of performing post hoc analyses on publicly available music perception data. While the null result from our preference and physiological response analysis suggests that such secondary analyses still have significant limitations when it comes to hypothesis-driven research, we demonstrate advantages of using such analyses to drive methodological developments. In support of our second hypothesis, our analysis revealed that breathing rate correlated with high-level musical features from the MUSIC five-factor model, rather than low-level features such as stimulus tempo. Given the substantial challenges posed by the secondary analysis, it is particularly remarkable for this result to be observed from a dataset whose primary studies were unrelated. Although no causation can be inferred from our analysis, this work adds to the large body of evidence demonstrating the influence of music on respiration by contributing a less-common analysis of changes on short timescales. While the methods we outline can certainly be improved upon, we provide these details and discussions as suggestions for future research utilizing existing, heterogeneous datasets.

Author Note

This work was done while the authors and the Brain Music Lab were part of the School of Music, Georgia Institute of Technology. The authors declare that there is no conflict of interest. They received no financial support for the research, authorship, and/or publication of this article.

Thank you to Dr. R. Michael Winters for his suggestions on this document.

Correspondence concerning this article should be addressed to Sophia Mehdizadeh, Roser ATLAS Center, 1125 18th St. 320, University of Colrado Boulder, Boulder CO 80309. E-mail: Sophia.Mehdizadeh@colorado.edu

References

Ali, S. O., & Peynircioğlu, Z. F. (2010). Intensity of emotions conveyed and elicited by familiar and unfamiliar music. *Music Perception*, *27*(3), 177–182. https://doi.org/10.1525/mp.2010.27.3.177

Baumgartner, H. (1992). Remembrance of things past: Music, autobiographical memory, and emotion. In J. F. Sherry, Jr. & B. Sternthal (Eds.), *NA - Advances in consumer research* (Vol. *19*, pp. 613–620). Association for Consumer Research.

Belfi, A. M., Kasdan, A., Rowland, J., Vessel, E. A., Starr, G. G., & Poeppel, D. (2018). Rapid timing of musical aesthetic judgments. *Journal of Experimental Psychology: General*, *147*(10), 1531–1543. https://doi.org/10.1037/xge0000474

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*(4), 445–452. https://doi.org/10.1136/hrt.2005.064600

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, *19*(8), 1113–1139. https://doi.org/10.1080/02699930500204250

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Bonneville-Roussy, A., & Eerola, T. (2018). Age trends in musical preferences in adulthood: 3. Perceived musical attributes as intrinsic determinants of preferences. *Musicae Scientiae*, *22*(3), 394–414. https://doi.org/10.1177/1029864917718606

Bonneville-Roussy, A., Stillwell, D., Kosinski, M., & Rust, J. (2017). Age trends in musical preferences in adulthood: 1. Conceptualization and empirical investigation. *Musicae Scientiae*, *21*(4), 369–389. https://doi.org/10.1177/1029864917691571

Casey, M. (2017). Music of the 7Ts: Predicting and decoding multivoxel fMRI responses with acoustic, schematic, and categorical music features. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01179

Casey, M., Thompson, J., Kang, O., Raizada, R., & Wheatley, T. (2012). Population codes representing musical timbre for high-level fMRI categorization of music genres. In G. Langs, I. Rish, M. Grosse-Wentrup, & B. Murphy (Eds.), *Machine learning and interpretation in neuroimaging* (Vol. *7263*, pp. 34–41). Springer. https://doi.org/10.1007/978-3-642-34713-9_5

Castaldo, R., Montesinos, L., Melillo, P., James, C., & Pecchia, L. (2019). Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life. *BMC Medical Informatics and Decision Making*, *19*(12). https://doi.org/10.1186/s12911-019-0742-y

Chamorro-Premuzic, T., Fagan, P., & Furnham, A. (2010). Personality and uses of music as predictors of preferences for music consensually classified as happy, sad, complex, and social. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 205–213. https://doi.org/10.1037/a0019210

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. https://doi.org/10.1037/a0024700

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, *930*(1), 28–42. https://doi.org/10.1111/j.1749-6632.2001.tb05723.x

Davis, M. (1984). The mammalian startle response. In R. C. Eaton (Ed.), *Neural mechanisms of startle behavior* (pp. 287–351). Springer. https://doi.org/10.1007/978-1-4899-2286-1_10

Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*, *8*(1), 14–25. https://doi.org/10.2174/157340312801215782

Elgendi, M., Fletcher, R., Norton, I., Brearley, M., Abbott, D., Lovell, N. H., & Schuurmans, D. (2015). On time domain analysis of photoplethysmogram signals for monitoring heat stress. *Sensors*, *15*(10), 24716–24734. https://doi.org/10.3390/s151024716

Elgendi, M., Jonkman, M., & De Boer, F. (2010). Applying the APG to measure heart rate variability. *The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, *3*, 514–517. https://doi.org/10.1109/ICCAE.2010.5452042

Elgendi, M., Norton, I., Brearley, M., Abbott, D., & Schuurmans, D. (2014). Detection of a and b waves in the acceleration photoplethysmogram. *BioMedical Engineering OnLine*, *13*(1). https://doi.org/10.1186/1475-925X-13-139

Etzel, J. A., Johnsen, E. L., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophysiology*, *61*(1), 57–69. https://doi.org/10.1016/j.ijpsycho.2005.10.025

Fricke, K. R., Greenberg, D. M., Rentfrow, P. J., & Herzberg, P. Y. (2018). Computer-based music feature analysis mirrors human perception and can be used to measure individual music preference. *Journal of Research in Personality*, *75*, 94–102. https://doi.org/10.1016/j.jrp.2018.06.004

Fricke, K. R., Greenberg, D. M., Rentfrow, P. J., & Herzberg, P. Y. (2021). Measuring musical preferences from listening behavior: Data from one million people and 200,000 songs. *Psychology of Music*, *49*(3), 371–381. https://doi.org/10.1177/0305735619868280

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), 1–22. https://doi.org/10.1371/journal.pone.0131151

Greenberg, D. M., Kosinski, M., Stillwell, D. J., Monteiro, B. L., Levitin, D. J., & Rentfrow, P. J. (2016). The song is you: Preferences for musical attribute dimensions reflect personality. *Social Psychological and Personality Science*, *7*(6), 597–605. https://doi.org/10.1177/1948550616641473

Greenberg, D. M., & Rentfrow, P. J. (2017). Music and big data: A new frontier. *Current Opinion in Behavioral Sciences*, *18*, 50–56. https://doi.org/10.1016/j.cobeha.2017.07.007

Hanke, M., Baumgartner, F. J., Ibe, P., Kaule, F. R., Pollmann, S., Speck, O., Zinke, W., & Stadler, J. (2014). A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Scientific Data*, 140003. https://doi.org/10.1038/sdata.2014.3

Hanke, M., Dinga, R., Häusler, C., Guntupalli, J. S., Casey, M., Kaule, F. R., & Stadler, J. (2015). High-resolution 7-Tesla fMRI data on the perception of musical genres – An extension to the studyforrest dataset. *F1000Research*, *4*, 174. https://doi.org/10.12688/f1000research.6679.1

Hargreaves, D. J., MacDonald, R., & Miell, D. (2008). Musical identities. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (1st ed., pp. 759–774). Oxford University Press.

Huron, D. (2013). On the virtuous and the vexatious in an age of big data. *Music Perception*, *31*(1), 4–9. https://doi.org/10.1525/mp.2013.31.1.4

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. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623. https://doi.org/10.1177/0305735613484548

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*(3), 217–238. https://doi.org/10.1080/0929821042000317813

Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L.-O. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford University Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Lamont, A., & Greasley, A. (2008). Musical preferences. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (1st ed., pp. 263–281). Oxford University Press.

Leslie, G., Ghandeharioun, A., Zhou, D. Y., & Picard, R. W. (2019). Engineering music to slow breathing and invite relaxed physiology. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–7. https://doi.org/10.1109/ACII.2019.8925531

Lex, E., Kowald, D., & Schedl, M. (2020). Modeling popularity and temporal drift of music genre preferences. *Transactions of the International Society for Music Information Retrieval, 3*, Article 1. https://doi.org/10.5334/tismir.39

Liikkanen, L. A., Jakubowski, K., & Toivanen, J. M. (2015). Catching earworms on Twitter: Using big data to study involuntary musical imagery. *Music Perception*, *33*(2), 199–216. https://doi.org/10.1525/mp.2015.33.2.199

Lingham, J., & Theorell, T. (2009). Self-selected "favourite" stimulative and sedative music listening – How does familiar and preferred music listening affect the body? *Nordic Journal of Music Therapy*, *18*(2), 150–166. https://doi.org/10.1080/08098130903062363

Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. *British Journal of Psychology*, *102*(1), 108–134. https://doi.org/10.1348/000712610X506831

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. https://doi.org/10.1177/0305735607086048

Nonin Medical (2005). *Nonin Operator's Manual: Models 8600FO and 8600FOM Pulse Oxymeters*. Nonin Medical, Inc. https://neurophysics.ucsd.edu/Manuals/Nonin/8600FO-English.pdf

North, A. C. (2010). Individual differences in musical taste. *The American Journal of Psychology*, *123*(2), 199–208. https://doi.org/10.5406/amerjpsyc.123.2.0199

North, A. C., & Hargreaves, D. J. (2007). Lifestyle correlates of musical preference: 1. Relationships, living arrangements, beliefs, and crime. *Psychology of Music*, *35*(1), 58–87. https://doi.org/10.1177/0305735607068888

Nyklíček, I., Thayer, J. F., & Van Doornen, L. J. P. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, *11*, 304–321.

Pettijohn, T. F., Eastman, J. T., & Richard, K. G. (2012). And the beat goes on: Popular billboard song beats per minute and key signatures vary with social and economic conditions. *Current Psychology*, *31*, 313–317. https://doi.org/10.1007/s12144-012-9149-y

Pinheiro, N., Couceiro, R., Henriques, J., Muehlsteff, J., Quintal, I., Goncalves, L., & Carvalho, P. (2016). Can PPG be used for HRV analysis? *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2945–2949. https://doi.org/10.1109/EMBC.2016.7591347

Quintana, D. S., & Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Frontiers in Psychology*, *5*, 805. https://doi.org/10.3389/fpsyg.2014.00805

Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, *100*(6), 1139–1157. https://doi.org/10.1037/a0022406

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*(6), 1236–1256. https://doi.org/10.1037/0022-3514.84.6.1236

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, *17*(3), 236–242. https://doi.org/10.1111/j.1467-9280.2006.01691.x

Russo, F. A., Vempala, N. N., & Sandstrom, G. M. (2013). Predicting musically induced emotions from physiological inputs: Linear and neural network models. *Frontiers in Psychology*, *4*, 468. https://doi.org/10.3389/fpsyg.2013.00468

Russo, M. A., Santarelli, D. M., & O'Rourke, D. (2017). The physiological effects of slow breathing in the healthy human. *Breathe*, *13*(4), 298–309. https://doi.org/10.1183/20734735.009817

Sakaguchi, Y., & Aiba, E. (2016). Relationship between musical characteristics and temporal breathing pattern in piano performance. *Frontiers in Human Neuroscience*, *10*, 381. https://doi.org/10.3389/fnhum.2016.00381

Salahuddin, L., Cho, J., Jeong, M. G., & Kim, D. (2007). Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4656–4659). https://doi.org/10.1109/IEMBS.2007.4353378

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. https://doi.org/10.1371/journal.pone.0007487

Schäfer, T., & Sedlmeier, P. (2009). From the functions of music to music preference. *Psychology of Music*, *37*(3), 279–300. https://doi.org/10.1177/0305735608097247

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*, 258. https://doi.org/10.3389/fpubh.2017.00258

Silver, D., Lee, M., & Childress, C. C. (2016). Genre complexes in popular music. *PLOS ONE*, *11*(5), 1–23. https://doi.org/10.1371/journal.pone.0155471

Siwiak, D., Berger, J., & Yang, Y. (2009). *Catch your breath—Musical biofeedback for breathing regulation*. Audio Engineering Society Convention 127. https://www.aes.org/e-lib/browse.cfm?elib=15065

Song, H.-S., & Lehrer, P. M. (2003). The effects of specific respiratory rates on heart rate and heart rate variability. *Applied Psychophysiology and Biofeedback*, *28*(1), 13–23. https://doi.org/10.1023/a:1022312815649

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380

van Gent, P., Farah, H., van Nes, N., & van Arem, B. (2018). Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the Taking the Fast Lane Project. *Journal of Open Research Software*, *7*(1), 32. https://doi.org/10.5334/jors.241

Wu, Y., Gu, R., Yang, Q., & Luo, Y. (2019). How do amusement, anger and fear influence heart rate and heart rate variability? *Frontiers in Neuroscience*, *13*, 1131. https://doi.org/10.3389/fnins.2019.01131