REVERBERATION TIME AND MUSICAL EMOTION IN RECORDED MUSIC LISTENING

HANNAH WILKIE & PETER HARRISON
University of Cambridge, Cambridge, United Kingdom

THE INFLUENCE OF ROOM ACOUSTIC PARAMETERS on musical emotion has to a degree been studied musicologically and empirically. However, there remain large gaps related to limitations in emotion measures and aspects of acoustic setting, with various iterations of digital acoustic reproduction represented in research. This psychological study explores the ways in which systematic alterations to reverberation time (RT) may influence the emotional experience of music listening over headphones. A quantitative approach was adopted, whereby musical stimuli with parametrically altered RTs were heard over user headphones. These were compared for domain-specific musical emotions on the Geneva Emotional Music Scale (GEMS). The main findings showed that the RTs and related acoustic features did not have a strong effect on "Unease" or "Vitality" components of the GEMS, but rather longer RTs had a significant positive effect on aspects of "Sublimity" (i.e., "Nostalgia," "Transcendence," "Wonder"). These results suggest that subjective percepts beyond pleasantness or emotional impact are affected by reverberation-based manipulations to room acoustic sound. The study outcomes have particular relevance to recorded music with artificial reverberation, and create scope for complex interactions between reverberation time and emotion more broadly.

Received: February 22, 2024, accepted July 20, 2024.

Key words: acoustics, emotion, musical acoustics, music technology, aesthetics

THE BROAD QUESTION OF WHETHER ROOM acoustic features influence subjective responses to musical sound has been broached in multiple branches of scholarship. In live listening, the concert hall acoustic has been speculated to influence musical value judgments by invoking "attentive" listening (Cressman, 2016; Johnson, 1995). It has also been argued that a "standardization" of acoustics for music—for example, a mid-frequency RT of c. 2s in concert halls (Beranek, 2011)—influences subjective responses to performance (Eidsheim, 2015; Thompson, 2002). Room acoustic features have exhibited prominent roles in medieval and renaissance music performance (Baumann & Haggh, 1990; Howard & Moretti, 2009), with acoustic modeling assisting research into historical performance spaces (Aletta & Kang, 2020; Boren, 2019, 2021). Features such as reverberation can also function as compositional devices, particularly in electroacoustic music (Barrett, 2017; Miller, 2012), while the manipulation of room acoustic effects is widespread in music production and sound engineering. In other words, acoustic parameters like reverberation play an active role in musical sound and subjective responses to performance.

Considering the role that room acoustic parameters may have in affecting subjective responses to music, it is logical to argue that musical emotion may be influenced by such parameters. Musical emotion is difficult to define, but it encompasses affective reactions whose sub-components include physiological arousal and subjective feeling (Juslin & Sloboda, 2010). In the domain of music psychology, emotions induced and perceived in music have been shown to be shaped by structural features such as mode and harmony (Gabrielsson & Lindström, 2010), performance-based features such as tempo and rhythmic articulation (Gomez & Danuser, 2007), and timbral or psychoacoustic cues (Eerola et al., 2012; Hailstone et al., 2009). Given these already multifaceted aetiologies, room acoustic features like reverberation are not an unusual addition. It has been shown that room acoustics can affect in-situ performance choices such as tempo (Fischinger et al., 2015; Schärer Kalkandjiev & Weinzierl, 2013), but acoustic settings may also influence more "fixed" effects, such as perceived timbral cues (Bech, 1996), perceived intimacy (Kaplanis et al., 2014), and subjective envelopment (Long, 2009).

Room acoustics and musical emotion are both vast fields, and to carry out a study combining the two, it is necessary to determine which room acoustic parameters will be studied, which measures of musical emotion will be used, and what the context of the listening

Music Perception, VOLUME 42, ISSUE 4, PP. 329–344, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312. DOI: https://doi.org/10.1525/MP.2025.2333445

environment will be. A limited number of empirical studies have already provided some support for the hypothesis that room acoustic parameters may influence musical emotion (summarized in Table 1), and these provided a basis for the present research.

As can be seen in Table 1, the previous studies have incorporated a range of room acoustic features, listening contexts, and emotion measures. Room acoustics have been shown to alter the subjective emotional "impact" of a musical performance in concert halls simulated by a loudspeaker array, which has been attributed to strong lateral reflections (Pätynen & Lokki, 2016) and dynamic responsiveness (Pätynen & Lokki, 2018). Lawless and Vigeant (2015) found altered reward responses to music in different simulated spaces, in conjunction with fMRI data. Larger RTs and anechoic conditions were generally disliked (see also Lawless, 2018). Extending beyond liking and impact, Västfjäll et al. (2002) and Tajadura-Jiménez, Larsson, et al. (2010) related virtual room size and reverberation amount to subjective valence and arousal—based on Russell's (1980) dimensional "circumplex" model of emotion—for a range of musical and nonmusical stimuli. Other studies have related valence-arousal ratings to source type and motion in acoustic settings simulated through loudspeakers (Hagman, 2010) and binaural headphones (Tajadura-Jiménez, Väljamäe, et al., 2010). In addition to "impact" and valence-arousal, effects have been found across some more fine-grained emotion categories (Algargoosh et al., 2022; Mo et al., 2015—see Table 1).

The studies in Table 1 have provided promising foundations for further research, with possible connections to be drawn between certain features of auditory processing and measured emotional responses. In Lawless (2018), strong reverberation degraded auditory cortex responses to sound sources. Emotional responses to acoustic space for music may interact with the subjective parsing of the auditory scene, in conjunction with attentional biases towards threat information in space (Ehret et al., 2021), resulting in a dislike for extreme reverberation and negatively valenced approaching sources (Hagman, 2010; Tajadura-Jiménez, Väljamäe, et al., 2010). Effects on emotion categories such as "Mysterious" (Mo et al., 2015) and "Inspired" (Algargoosh et al., 2022) also indicate a possible connection with "aesthetic" emotions. Aesthetic emotions occupy a contested status in psychological research and have been difficult to categorize (Menninghaus et al., 2019; Skov & Nadal, 2020). However, they have been attached to evaluative states such as "Wonder" or "Awe" (Juslin, 2013), aroused by perceived skill or beauty, and encouraging self-reinforcing exposure to a stimulus (Tschacher et al., 2012). An influence on such states is speculative; nevertheless, acoustic parameters like reverberation may engage a range of emotion induction processes, as outlined in the BRECVEMA model (Juslin, 2013; building on Juslin & Västfjäll, 2008)—between fundamental auditory processing and learned evaluations.

LIMITATIONS OF PREVIOUS WORK

The range of acoustic conditions, listening contexts, and emotion measures incorporated into previous studies means that there is no one method that serves as a model for the present study. Our methods are instead derived from an assessment of the gaps in previous work. In particular, we note a need to focus specifically on musical emotion, a need for a wider range of RTs, and an opportunity to explore alternative contexts of headphone listening.

Although the existing studies have many useful implications, particularly for acoustic modeling, there is often a lack of direct focus on musical emotion, which means that the emotion measures used have granted limited insights from this viewpoint. Subjective "impact" is a flexible concept that can be used as a pivot for exploring a wide range of acoustic parameters, but even in conjunction with physiological data it offers only a broad overview. For dimensional emotion measures (i.e., valence-arousal), results have been mixed. This was the case for pleasantness and "expressiveness" in Tajadura-Jiménez, Larsson, et al. (2010), relative to arousal. In Västfjäll et al. (2002), the largest RT induced the lowest arousal, whereas the largest RT in Tajadura-Jiménez, Larsson, et al. (2010) induced the highest arousal. Because the largest RT in the former was infinite, and 1.88s in the latter, it is difficult to draw definitive conclusions. Source motion is not the focus of the present study, but in Hagman (2010), negatively valenced sources generally induced high arousal regardless of direction, while in Tajadura-Jiménez, Väljamäe, et al. (2010), all approaching sources were associated with higher arousal.

For the few studies incorporating categorical emotion measures, it is difficult to combine the results and relate them to wider literature. In Algargoosh et al. (2022), emotion categories deemed suited to worship spaces were employed, while Mo et al. (2015) used categories derived from compositional expressive markings. These two choices of category do not overlap strongly, and alternative categories have been employed elsewhere in musical emotion research (e.g., Cowen et al., 2020; Zentner et al., 2008). Furthermore, Algargoosh et al. (2022) only compared two conditions: dry, or with an

TABLE 1. *Summary of Relevant Empirical Studies*

Study	Acoustic Parameters	Stimulus Type	Presentation	Measures	Relevant Findings
Algargoosh et al. (2022)	RT Anechoic / 1.7 – 2.7 s	Hymns/Quran recitations (c. 75 s)	Over-ear headphones; VR headset	Emotion categories (induced) Verbal self-report transcribed	With acoustic = stronger ratings for ‘Nostalgic’, ‘Spiritual’, ‘Inspired’, ‘Cheerful’, ‘Thrilled’; lower for ‘Sad’, ‘Tense’ Familiarity and audio-visual congruency determine intensity of emotional impact
Hagman (2010)	RT 0.2 – 0.9 s Source type/motion (CATT models)	Natural sounds	Loudspeaker array	Pleasantness-arousal ratings EMG, EDA	Approaching = higher arousal, less pleasant All negatively valenced sources = high arousal Larger room = higher arousal (no effect if source moving)
Lawless & Vigeant (2015)	RT 0.0 – 5.3 s (Modified ODEON model)	Orchestral (Bruckner 8, 11.2 s)	Binaural earbuds	Preference ratings fMRI	Altered reward responses Extreme RTs disliked (anechoic / 5.3 s)
Lawless (2018)	Multiple experiments, RT 0.0 – 7.2s. Also C80, EDT (Modified ODEON model)	Solo instrumental, Orchestral (11 – 17 s)	Binaural earbuds	Preference ratings fMRI Auditory cortex response	Altered reward responses Degraded cortex response for higher RTs Extreme RTs (> 2.8 s / anechoic) disliked
Mo et al. (2015)	RT Anechoic – 2.37 s Simple parametric reverb	Solo instrumental	Over-ear headphones	Emotion categories (perceived; paired comparison)	Higher RT = ‘Mysterious’, ‘Romantic’; Anechoic = ‘Comic’ Other emotions = moderate to mild effects
Pätynen & Lokki (2016)	RT 1.7 – 2.9 s (Concert hall IRs)	Orchestral (Beethoven 7, 28 s)	Loudspeaker array	Emotional ‘Impact’ (paired comparison) Skin Conductance Response (SCR)	Greater ‘impact’ = strong lateral reflections
Pätynen & Lokki (2018)	RT 1.7 – 2.9 s (Concert hall IRs)	Orchestral (Bruckner 8, 28 s)	Loudspeaker array	Emotional ‘Impact’ (paired comparison) Skin conductance response (SCR)	Greater ‘impact’ = strong dynamic responsiveness
Tajadura-Jiménez, Larsson, et al. (2010)	RT Outdoor/0.36/1.88 s (CATT models)	Animal, human, musical (5.5 – 6.5 s)	Loudspeaker array	Safety, valence-arousal ratings Electrodermal activity (EDA) Facial electromyography (EMG)	1.88 s most unpleasant, highest arousal, increased EDA, larger CS muscle activity Significant interaction with source valence
Tajadura-Jiménez, Väljamäe. et al. (2010)	One room Source motion (approaching/ receding)	Artificial tones, natural sounds	Binaural headphones	RTs for identifying photo valence Valence and arousal ratings EMG, EDA	Arousal dependent on source valence Approaching negatively valenced source = higher arousal, less pleasant, increased EDA, CS activity
Västfjäll et al. (2002)	RT Low (c. 1.6 s)/medium (c. 2.5 s) /infinite (CATT models)	Human voice, string quartet, clarinet	Over-ear headphones; images on screen	Pleasantness and arousal ratings ‘Expressiveness’ ratings	Infinite RT = lowest arousal Medium RT = highest arousal/expression Low RT = highest pleasantness

acoustic. The RT range in Mo et al. (2015) was also relatively narrow. An alternative emotion measure reproduced more widely in musical emotion research is needed.

With respect to the room acoustic parameters incorporated into studies, almost all of the experiments in Table 1 incorporate variations in RT; this is true even if RT is not the primary feature with which results are correlated (as with, e.g., Hagman, 2010; Pätynen & Lokki, 2016, 2018; Tajadura-Jiménez, Väljamäe, et al., 2010). This creates a useful model for the present study. However, the range of RTs employed has been relatively narrow—usually within a "concert hall" range of c. 1–3s (with the exception of Lawless, 2018; Lawless & Vigeant, 2015). This is a feature that can be expanded. RT transfers well across various listening technologies; even simple parametric reverb has perceivable effects in headphone listening (Mo et al., 2015), as well as the reverberation generated by more realistic room simulations (e.g., Lawless, 2018).

Many of the previous studies differ with respect to stimulus presentation. The studies incorporating a loudspeaker array generally aim for perceptual realism, synthesizing a concert hall space. However, headphones and earbuds generate a more complex percept. Binaural headphones continue to aim for an effect of immersion, but there are increasing overlaps with recorded music listening. All of the studies involved some form of technological simulation of an acoustic. Given that the aim has often been to simulate realistic space, albeit to varying degrees of accuracy, the present research decided to explore other directions by addressing an alternative, more subtle percept less tailored to realistic modeling. This is discussed further below.

THE PRESENT STUDY

In light of the limitations from previous studies, the present study employed a wide range of RTs as its main independent variable, a more domain-specific measure of musical emotion, and focussed on recorded music heard over personal headphone devices. Our study methods quantitatively measured self-reported induced emotion for musical stimuli whose reverberation times had been systematically altered and were heard over headphones. The aim was to identify whether emotions generally felt during music listening would be affected by manipulations to reverberation, keeping all other elements of the musical stimulus constant.

Our experiment used a simple acoustic model, based on a single impulse response function, which we parametrically manipulated to achieve different reverberation times. More complex methodologies would have achieved higher ecological validity with regards to realistic simulation, for example by using more complex acoustic models or employing in-situ recording. However, our simplified approach is advantageous for facilitating direct manipulation of stimulus parameters and for enabling easier interpretation of perceptual effects. Participants heard the stimuli at home on their own headphone devices. It was hypothesized that altered reverberation times and related acoustic features would have a systematic influence on certain musical emotions, with a null hypothesis of no effect.

The reasoning behind the use of individual participant devices and headphones warrants further attention. By choosing to use participant headphone devices, this study concentrated on everyday recorded music listening. If a fully immersive auralization experience were sought (following Kleiner et al., 1993), it would be desirable to use a three-dimensional loudspeaker array or high-quality binaural headphones. However, the uses of participant headphone devices and less naturalistic modeling in this study are not directed specifically at capturing a live experience, although the measured acoustic parameters are also relevant in these contexts. Music listening over user headphones is a prominent form of consumption, and everyday listening with personal technologies has formed the focus of a large number of musical emotion studies (e.g., Krause et al., 2015; Sloboda & O'Neill, 2001). There are limitations associated with a lack of control over user headphone devices, which will be discussed further. Nevertheless, the fact that trends in the results may still be observed, despite these variation in technological reproduction, is a useful outcome.

The incorporation of parametrically altered RTs within headphone listening creates a particular emphasis on the digital manipulations to reverb heard in music production. These manipulations generate subjective spatial percepts that form a valid and prevalent listening experience in their own way. The evolution and aesthetics of reverberation in recorded music have been explored relative to production and sound engineering (e.g., Brøvig & Danielsen, 2016; Doyle, 2005; Sterne, 2015). However, there is little psychological research on simple parametric reverb and its emotional effects (De Man et al., 2017; Mo et al., 2015), creating space for further elaboration. As with many acoustic experiments, this research exists along a spectrum of technological reproduction. Motivated by studies in acoustics and emotion across a wide range of listening contexts, this study narrowed down its focus to recorded music and the reverberation effects that can be found in headphone listening.

TABLE 2. *GMSI Scores*

	General Musical Sophistication	Emotions	Musical Training	Perceptual Abilities	Active Engagement	Start Age	Singing Abilities
Mean	3.23	5.49	2.36	4.80	3.71	10.02	3.39
SD	1.12	1.00	1.29	1.14	1.16	3.11	1.49

Before proceeding to the study methods, a more specific definition will be provided for the parameters incorporated in this research. As the main independent variable in this study, reverberation time (RT) is the time taken for sound to decay by 60 dB. RT alterations also bring changes in other parameters, which are noted in the results of the present study. These include objective clarity (C80; the logarithmic ratio of early to late sound energy, before and after 80 ms) and early decay time (EDT; the time taken for the early source energy to decay by 10 dB). Frequency attenuation was also explored. The RT range in the study was wide spanning from relatively dry conditions to longer RTs that might have parallels in, for example, a large worship space. For example, St Paul's Cathedral has exhibited mid-frequency FTs of around 11s (Lewers & Anderson, 1984), and Notre-Dame de Paris around 8s (Postma & Katz, 2016). It is beyond the scope of this study to examine whether the phenomenological experience of digital reverberation correlates with imagined subjective percepts of, for example, a cathedral. Nevertheless, the research establishes a set of relationships between digital RT manipulations and musical emotion, which is most relevant to sound engineering, although it has space for expansion to other forms of listening.

Method

PARTICIPANTS

One hundred and nine individuals participated in the experiment (males = 74, females = 27; age = 18–71, mean age = 37.6, *SD* age = 11.9). The experiment was distributed as an online survey on PsyNet (Harrison et al., 2020), and "Prolific," a crowdsourcing service for online psychology participants, was employed for recruitment. Those who participated received an average payment of £3.33 for a survey duration of approximately 20 minutes. Participants were not selected for musical abilities; note that there is only mixed evidence for musical experience impacting room acoustic sensitivity (Lawless, 2018; von Berg et al., 2021). Approval was obtained from the Cambridge University Faculty of Music Ethics Subcommittee prior to recruitment.

At the end of the experiment, participants provided data on self-reported musicality using the short-scale version of the Goldsmiths Musical Sophistication Index (Gold-MSI; Lin et al., 2021). This is a self-report inventory on differences in musical "sophistication," where participants report self-assessments on active musical engagement (e.g., resources/time/money spent on music), perceptual abilities (e.g., accuracy of listening skills), music training, singing abilities, and emotional engagement. GMSI scores are derived from a 7-point likert scale with responses range from 1 (*Completely Disagree*) to 7 (*Completely Agree*). Scores were slightly below midpoints (< 4) on most attributes, although emotional responses and perceptual abilities were above (see Table 2). Only one participant reported absolute pitch.

MATERIALS

Resource Availability

Stimuli, IRs, Pyroomacoustics code, code for the experiment implementation, code for data analysis, and Supplementary Materials are available on OSF at the following link: https://osf.io/976b3/?view_only=3d7425b602ad4917be42146755c189e3. Supplementary Materials are also available at online.ucpress.edu/mp.

Impulse Responses

Stimuli were created through the convolution of an anechoic audio file with an impulse response (IR) (a short signal describing a room acoustic), to simulate the audio file as it would sound in the acoustic described by the IR. IRs were created on Pyroomacoustics (Scheibler et al., 2018). Two initial IRs with long RTs were produced by modeling a 510 m^3 trapezoid-shaped "room," first using a glass material ("glass_3mm"), which attenuated low frequencies, then a brick material ("brick_wall_rough"), attenuating high frequencies. (For further details, see Pyroomacoustics code in Supplementary Materials). Decay manipulations were subsequently applied to each IR, following Cabrera et al. (2011): in particular, the original IR was multiplied by an exponential decay of varying exponent values. The first 80 ms were left unaltered, however, so as to preserve the direct sound and early reflections (see, e.g., Martellotta, 2010). The manipulation can be expressed mathematically as follows:

$$x_i{}' = \begin{cases} x_i, x_i \leq \alpha \\ \exp(-\lambda(x_i - \alpha)), \text{ otherwise} \end{cases}$$

where x_i is the ith element of the original IR, $x_i{}'$ is the ith element of the transformed IR, α is the duration of the initial unaltered period (80 ms), and λ is the decay constant.

The decay constant was altered to produce 8 RTs for each of the two frequency attenuation conditions (mid-frequency RTs were 0.5s; 1s; 1.5s; 2s; 3s; 5s; 6s; 8 s). Each of these had corresponding EDT and C80 values (see Supplementary Materials). The RTs were selected to represent a variety of acoustic conditions, ranging from “dry: settings typical of domestic spaces to longer RTs typical of cathedrals. As discussed, the systematic nature of these manipulations is not directly representative of a realistic spatial model; the resulting RTs are closer to artificial reverberation effects.

Convolutions

Convolutions of Pyroomacoustics IRs were created on ODEON v.17.00 (Odeon, 2021), using anechoic recordings of a viola and flute (available on the software). The flute was an excerpt from Debussy’s *Syrinx*, and the Viola was a variant on *Solveig’s Song* from Grieg’s *Peer Gynt*, available in ODEON’s library of sound files. The flute excerpt was cut slightly using REAPER (Cockos, 2004). Stimuli were selected for featuring some ambiguity regarding their emotional valence, although this does carry some subjectivity (limitations in stimuli will be discussed). After being convolved on ODEON, recordings were loaded onto REAPER. It proved necessary to set a threshold of -10 dB on the JS: 1175 Compressor (ratio 4)—this removed gain peaks that distorted the signal output over headphones, and the reverb tail was unaffected. We note that this kind of compression is common in recorded music. Stimuli were output in a 16-bit stereo format. The viola recording was 22–30s depending on RT, and the flute was 21–29s.

EMOTION MEASURES

The Geneva Emotional Music Scale (GEMS; Zentner et al., 2008) was used to measure self-reported felt emotion. This scale was chosen as a domain-specific measure of musically induced emotion. The focus of the GEMS is on felt rather than perceived musical emotion—significant differences can occur between the two (Juslin & Laukka, 2004). It was developed through a data-driven approach, employing emotion categories found through extended research to be relevant to music listening. These include possible “aesthetic” emotions such as “Wonder,” as well as “basic” emotions such as “Sadness.” While Russell’s circumplex model of emotion plots different emotions along the dimensions of valence and arousal, and categorical measures conceive emotion as discrete states, the GEMS groups emotions into different hierarchical levels. These culminate in “second-order” factors (“Sublimity,” “Vitality,” and “Unease”).

The GEMS-9 was used, which is an alternative to longer versions of the scale (the GEMS-45 and GEMS-25). Instead of requesting ratings of all emotion adjectives from the GEMS-45, nine primary factors are presented. Each is accompanied by related emotion terms (for example “Transcendence” oversees the additional explanatory adjectives “Fascinated, Overwhelmed, Feelings of transcendence and spirituality”). The GEMS-9 is psychometrically less robust than its longer versions, but it is more efficient, and its effectiveness has received support (Pearce & Halpern, 2015).

PROCEDURE

The experiment obtained paired comparisons of musical stimuli in different acoustic settings for emotions on the GEMS-9 and for preference. Paired comparisons enable efficient decision-making by participants, and their results have been shown to correlate well with psychophysiological responses to room acoustics (Pätynen & Lokki, 2016, 2018).

At the beginning of the experiment, participants completed a headphone test (Woods et al., 2017) before progressing to the paired comparisons. Each pair of recordings was drawn from the same instrument (i.e., Flute or Viola) and frequency condition (i.e., High Frequency Attenuation or Low Frequency Attenuation), and compared two out of eight possible RTs. Participants each reviewed 12 pairs, 3 for each combination of instrument and frequency condition. In each comparison, participants heard the complete recordings in succession and were presented with a random subset of four GEMS-9 emotions. Participants were instructed to select which recording, A or B, was better at making them feel each emotion, as well as selecting their preferred recording. At the end of the survey, participants filled out the short-scale version of the Gold-MSI (Lin et al., 2021) and had the option of providing written feedback.

Results

In total, there were 15,070 comparisons, with each RT undergoing an average of 1,884 comparisons, and

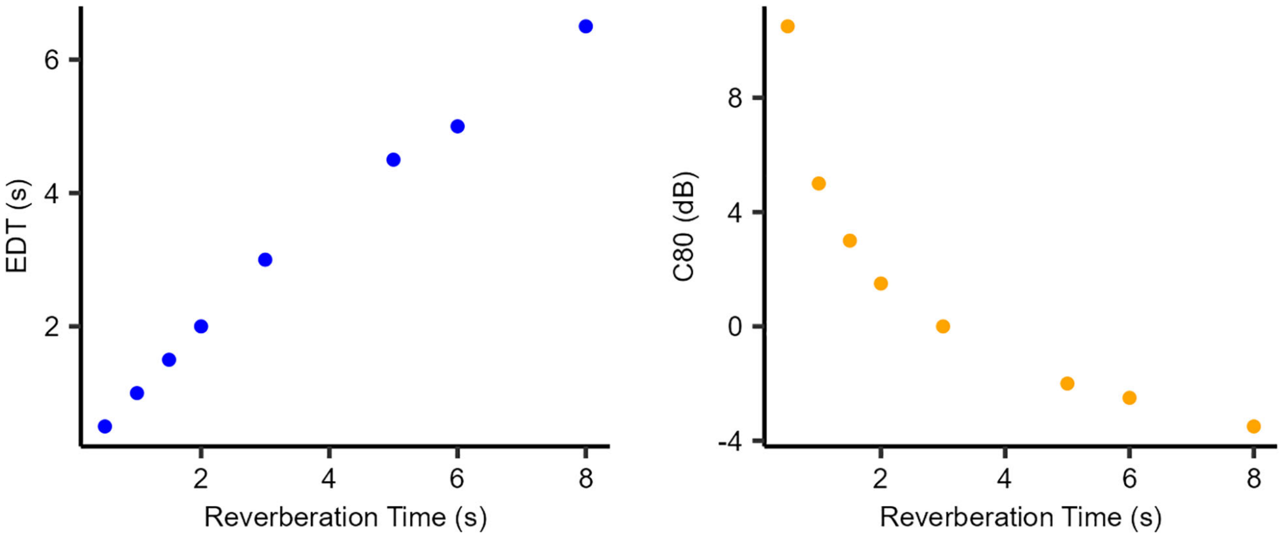

FIGURE 1. The relationship between RT and C80/EDT in the IRs for the experiment.

each emotion undergoing an average of 1,340 comparisons. Of the 109 participants who contributed results, 101 completed the entire experiment, while 8 others provided partial results. The total comparison count includes those from participants who did not complete the entire experiment but still reviewed some stimulus pairs.

Matrix tables of pairwise comparisons were analyzed using the Bradley-Terry-Luce (BTL) model (Bradley & Terry, 1952; Luce, 1959/2012), which takes as its input the results of a series of pairwise comparison trials, where the participant is asked which object (e.g., an audio stimulus) better satisfies a given criterion (e.g., felt "Tension"). The model analyzes these data, and returns a score for each object, summarizing how well the object satisfies that criterion. BTL models for paired comparison data have previously been incorporated in various auditory perception studies (e.g., Choisel & Wickelmaier, 2007; Mo et al., 2015; Pätynen & Lokki, 2016, 2018). The present analysis used the BTL model as implemented in the BradleyTerry2 package in R (Firth & Turner, 2012). For two stimuli (stimulus *i* and stimulus *j*) with scores α_i and α_j, the model states that the probability of choosing stimulus *i* over stimulus *j* is equal to:

$$\frac{e^{\alpha_i}}{e^{\alpha_i} + e^{\alpha_j}}$$

In general terms, the scores presented in the figures can be interpreted as values representing the ability of each acoustic condition to induce an emotion, when compared with another stimulus. BTL score is summarized in Figure 2 as a function of RT; it is labeled as "Emotion Score." Although "Preference" is not an emotion category and is not in the GEMS, it is included in the same graphs and tables for ease of presentation. The manipulations to the impulse response functions were primarily designed to influence RT, and hence the primary analyses use RT as the independent variable. However, it is worth noting that (under the given manipulations) RT is highly correlated with EDT, and highly negatively correlated with C80 (Figure 1; for precise values, see Supplementary Materials). The Supplementary Materials provide analogous plots and analyses using EDT and C80 as independent variables.

Figure 2 demonstrates that some emotion components do not relate clearly to RT, including "Tension" and "Sadness" (each in the "Unease" component of the GEMS), and "Joyful Activation" (in the "Vitality" component). "Power" may display a very weak negative correlation with RT. In other words, neither higher nor lower RTs are strongly associated with judgements that a stimulus is effective at inducing these emotion states.

However, for items in the "Sublimity" component of the GEMS, there are stronger results. While "Tenderness" and "Peacefulness" do not demonstrate clear correlations, the plots for "Nostalgia," "Wonder," and "Transcendence" indicate a positive correlation with RT. This seems particularly strong for "Transcendence."

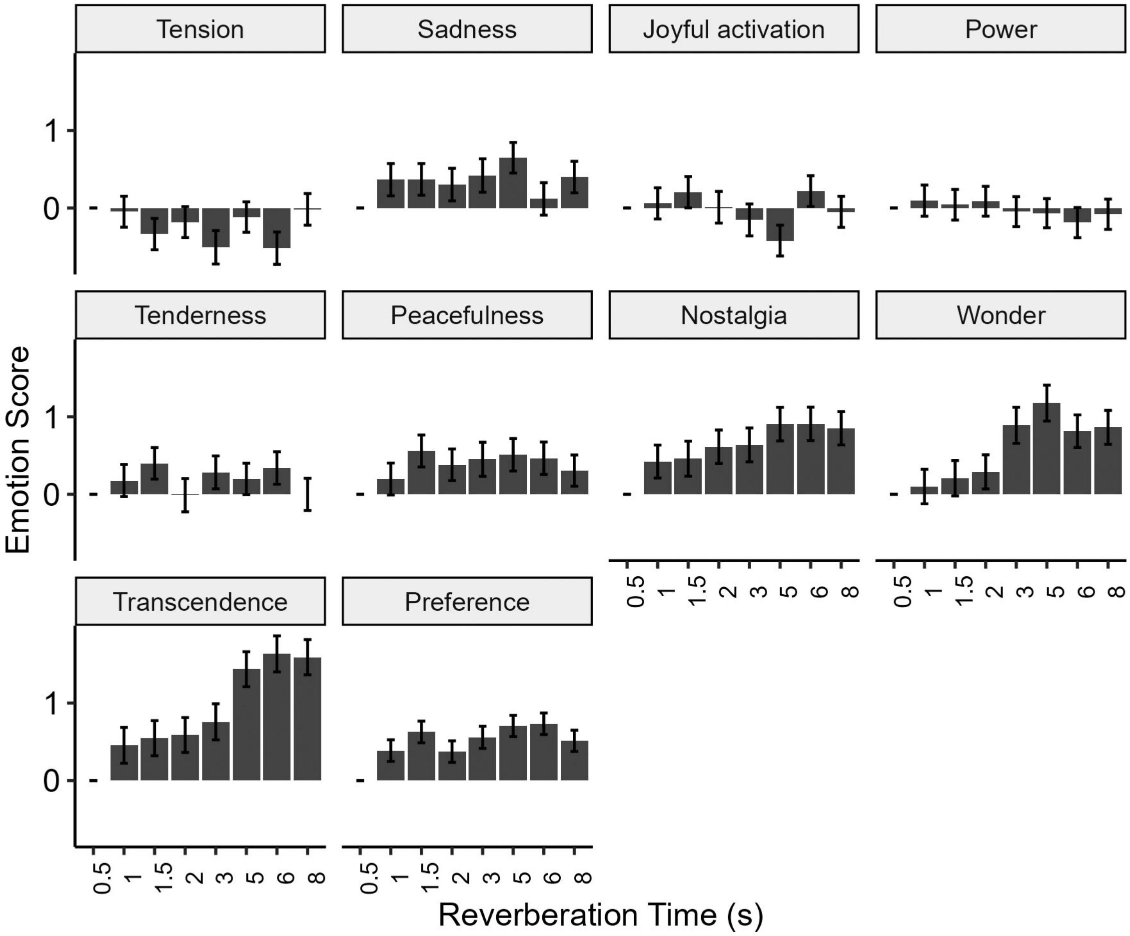

FIGURE 2. Emotion (BTL) Score vs. RT. *Note*: Error bars indicate standard errors as estimated in the BTL model. Higher scores for an RT indicate that it was more likely to be chosen in the recordings to induce the given emotion (or preference). The output for 0.5 s is set to 0 as an identifying condition. Negative values indicate that the RT will underperform the 0.5 s condition, and vice versa for positive values.

"Preference" scores may display a weak positive correlation. Between a 1–2s RT, results are more uniform, and it is in more extreme ranges that stronger patterns are observed. Similar results are observed for EDT, and the inverse is seen for C80 (see Supplementary Materials).

REGRESSION ANALYSIS

Linear regression models were constructed to analyze the effect of RT on different emotion components (+ preference). The analysis also tested whether these effects were moderated by instrument, frequency condition, and music training. Regression coefficients were calculated to identify significance levels. Since RT was highly correlated with C80 and EDT in the manipulations, it is possible to repeat the analyses using these measures as predictors instead and obtain essentially equivalent results (see Supplementary Materials).

Effects of RT on emotion judgments

Separate linear regression models were run for each emotion (+ preference), with Emotion Score as the response variable and RT as the (continuously treated) predictor variable. Regression coefficients are summarized in Table 3 and visualized in Figure 3. Linear models for "Nostalgia," "Wonder," and "Transcendence"—each in the "Sublimity" component of the GEMS-9—produce a significant positive coefficient between RT and Emotion Score. In other words, a stimulus with a higher RT was more likely to be selected for inducing "Nostalgia," "Wonder," and "Transcendence." The weakest relationships between RT and Emotion Score occur for "Tenderness," "Tension," and "Joyful Activation." While "Power" displays statistical significance, this is coupled

TABLE 3. *Separate Linear Models by Emotion, with RT as the Predictor Variable and Emotion (BTL) Score as the Response Variable*

Emotion (+ preference)	Regression Coefficients	Standardised Regression Coefficients	Standard Error	*t*-value	*p* value
'Tension'	−0.01	−0.11	0.03	−0.270	.790
'Sadness'	0.02	0.29	0.03	0.740	.480
'Joyful Activation'	−0.02	−0.21	0.03	−0.530	.610
'Power'	−0.03	−0.77	0.01	−3.000	.025*
'Tenderness'	0.00	−0.02	0.03	−0.044	.970
'Peacefulness'	0.03	0.36	0.03	0.950	.380
'Nostalgia'	0.10	0.84	0.03	3.800	.009**
'Wonder'	0.13	0.81	0.04	3.300	.016*
'Transcendence'	0.21	0.95	0.03	7.600	< .001***
Preference	0.05	0.59	0.03	1.800	.120

Note. $^*p < .05$ $^{**}p < .01$ $^{***}p < .001$

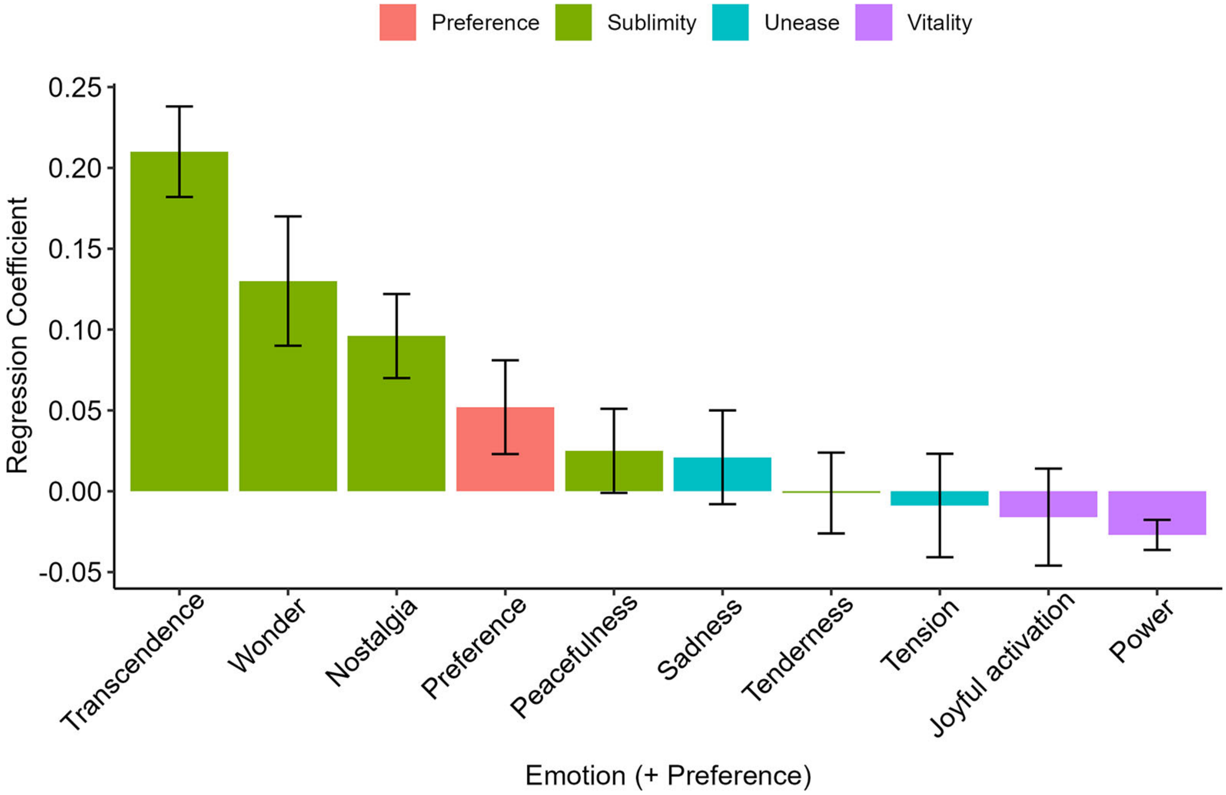

FIGURE 3. Regression coefficients for RT as a predictor of each emotion (and preference). *Note*: Error bars indicate standard errors as estimated in the regression model.

with a weak regression coefficient. "Preference" displays a non-significant positive relationship with RT. The strongest overall relationship is for "Transcendence."

Figure 4 provides a visualization of Table 3, with regression plots filtered by emotion and grouped by second-order factor in the GEMS (+ preference).

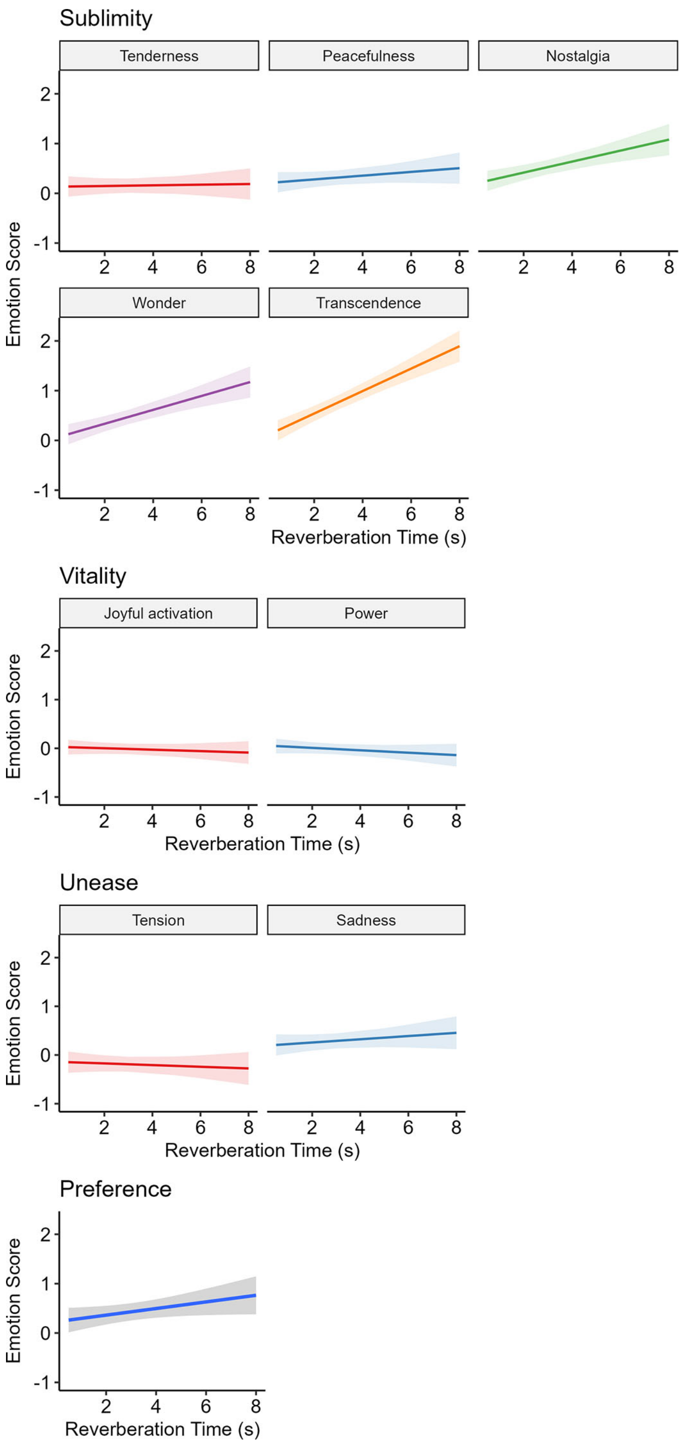

FIGURE 4. Linear regression plots for RT and Emotion Score for the GEMS-9. The shaded regions indicate the standard error of the regression line.

Moderating Effects of Instrument Type and Frequency Condition

An exploratory analysis was conducted to investigate whether these results might be moderated by other stimulus variables, in particular "instrument" (the instrumental excerpt played: Flute vs. Viola) and "frequency condition" (the frequency attenuating properties of the IR: emphasis on high frequencies vs.

emphasis on low frequencies). This was achieved by incorporating both instrument and frequency conditions as interaction variables in a linear model between RT and Emotion Score. The interaction term between RT and instrument/frequency was calculated, with the viola and low frequency conditions compared against a flute / high frequency intercept.

In general, interactions between RT and instrument or frequency condition did not display statistical significance ($p > .10$). However, for felt "Tension," there was a significant negative interaction between RT and the Viola stimulus when influencing Emotion Score ($p = .025$), suggesting that high RTs with the Viola stimulus tended to reduce felt "Tension" more than high RTs with the Flute stimulus. There was a trend towards a similar, inverse relationship for felt "Peacefulness" ($p = .09$), suggesting that high RTs with the Viola stimulus tended to increase felt "Peacefulness" more than with the Flute stimulus. For frequency conditions, there was a significant interaction between felt "Sadness" and an emphasis on low frequencies in the modeled IR ($p = .032$), implying that high RTs tended to induce "Sadness" in particular when the acoustics emphasised low frequencies. All these interaction effects should be treated as exploratory though given the larger number of effects tested.

Moderating Effects of Music Training

The potential moderating effect of music training was explored in a similar way. Here the original dataset was filtered into two groups, the first including participants who scored greater than or equal to the median level of music training, and the second including participants who scored below the median. This created a two-level "Music Training" variable for inclusion in the analysis. The interaction term between RT and Music Training" was calculated in linear models between RT and Emotion Score. This interaction did not display statistical significance ($p > .10$; see Supplementary Materials), implying that music training did not have a clear impact on evaluations.

Discussion

The purpose of this study was to explore the effect of systematic alterations to reverberation time on musical emotion, focusing on recorded music listening over headphones, with a wide range of RTs and a domain-specific musical emotion measure. The results build on previous research that has suggested that acoustic parameters alter emotional responses in music listening, showing in particular that RT manipulations can influence musical emotion for stimuli heard over personal headphone devices. For the specified stimuli, recordings with higher RTs and associated higher EDTs / lower C80 values, regardless of frequency attenuation, were better at inducing emotions of "Nostalgia," "Wonder," and "Transcendence." These are each in the "Sublimity" component of the GEMS. There are several key takeaways from these results.

First, the most important emotion terms for the acoustic manipulations are complex states that are difficult to map onto a two-dimensional valence-arousal emotion model. In the original model presented in Russell (1980), these three emotions are not mentioned. "Nostalgia" may have a mixed valence (Batcho, 2020), while "Wonder" and "Transcendence" are under-researched from a valence-arousal perspective. Algargoosh et al. (2022) found effects for similar emotions when a room acoustic was applied to a recording, compared against an anechoic condition including "Nostalgic" and "Spiritual." In Mo et al. (2015), "Mysterious" and "Romantic" were associated with higher reverberation and are again difficult to map onto valence-arousal dimensions. Compared to these two studies, this research demonstrated that even across a much wider range of RTs, using measures derived from musical emotion research, similar emotions maintain a significant interaction with RT.

Second, "Nostalgia", "Wonder", and "Transcendence" may have an important relationship with "aesthetic" appraisal. Juslin (2013) cites "Wonder" and "Nostalgia" when discussing "aesthetic" emotions. These states may relate to evaluative responses involving perceived skill, beauty, or awe. Keltner and Haidt (2003) discuss sensations of "Wonder" and "Transcendence" when describing "Awe," connected with perceptions of vastness and difficulties in comprehension. "Sublimity," with its heritage in Kant and Schopenhauer, has also been the topic of considerable psychological debate. It is for example conceived by Shusterman (2005) as an enhanced aesthetic response to a stimulus in dialogue with somatic responses. "Awe" and "Sublimity" have been speculatively associated with musical performances in "colossal" structures such as a medieval cathedral (Konečni, 2005). The results may support this, with cathedral spaces generally carrying higher RTs. Such acoustic properties have been seen as an asset in studies of large worship spaces (e.g., Penchteva, 2011). Sander et al. (2023), in a re-analysis of Hahn (2018), found an association between increased auditory immersion (e.g., musical stimuli heard over immersive 3D audio, compared against stereo and surround) and "Transcendence" on the GEMS. It is possible that there is an interplay between reverberation and concepts of immersion in the present results.

Third, the effect of RT on stimulus preference was surprisingly weak and contradicted previous research. In opposition to suggestions that high RTs are associated with decreased preference (Hagman, 2010; Lawless, 2018; Lawless & Vigeant, 2015; Västfjäll et al., 2002), preference was very weakly correlated with higher RTs and lower objective clarity. This may be specific to the two stimuli in question, though it is interesting to note that in live performance, these pieces would ordinarily be played in locations with lower reverberation times (e.g., concert halls) than those preferred here. See Limitations for further discussion. Although strong reverberation has been associated with degraded auditory cortex responses to sound sources (Lawless, 2018), possibly relating to a degraded sensitivity to threat information (Ehret et al., 2021), associations with dislike were limited in the present study. It may be that the headphone scenario undermines unease responses. Live music has been shown to engage stronger physical responses than recorded music (Swarbrick et al., 2019), which has been related to features such as interpersonal physiological synchrony in live events (Czepiel et al., 2021). A dissociation from these responses may be one consequence of everyday listening on user headphones.

Fourth, while effects for instrument type and musical features were generally weak, there were some key results related to felt "Tension" and "Sadness." The weaker effects for musical features may be due to certain correspondences between the two stimuli, which are discussed in the experiment limitations. However, felt "Tension" decreased with higher RTs for the viola stimulus, and "Sadness" was associated with high frequency attenuation; that is, an emphasis on lower frequencies in the IR. High RTs can have a "masking" effect, as the objective clarity of a stimulus decreases. The original viola stimulus had a rough tone quality. Speculatively, this may have increased felt "Tension," whereas for higher RTs, these performance details were masked. The association of "Sadness" with high frequency attenuation is paralleled in wider research, which has found that "Sadness," or low valence, is associated with an emphasis on lower frequency energy in the spectrum (e.g., Eerola et al., 2012; Scherer et al., 2017; Tan et al., 2020).

Finally, there was some suggestion that there was a "middleground" in the results where RT facilitated a focus on musical features, rather than overall acoustic sound. One participant provided the following written feedback:

> *"For me there was a middleground I reacted to more. Without reverb the pieces were flat and I didn't have a response. When there was a longer delay with high reflection on the reverb, there was an introduction of digital feedback that was distracting and irritating, and its unnatural sound pulled away from the piece being played. The middle ground reverb actually added to the music being played and made it more engaging and stimulating."*

Mid-range RTs (1–2 s) in Figure 2 generally produced more uniform effects. It may be that in this range, fixed musical features are the perceptual focus; conversely, at extreme levels, space may become more salient perceptually. The 1–2s RTs are closer to the standard "concert hall" RT range (Eidsheim, 2015; Thompson, 2002).

LIMITATIONS AND FUTURE WORK

There are some limitations to this research. In particular, in order to achieve high statistical power when testing a large number of emotion terms, the stimulus set was restricted to just two musical pieces. While there were key differences between the stimuli (different timbral qualities, harmonic properties, pitch range, event rate), there are also similarities (both Western classical, solo instrumental, relatively slow tempi). This means that we cannot be sure how the results would generalize to other musical pieces. With the experiment asking which recording from each pair was better at inducing an emotion, participants were still able to compare the emotional effects of acoustic alterations, even if the original piece was more sad than cheerful, for example. It would nevertheless be interesting to explore other genres in future research, given that genre is thought to be important for the perceptual effects of architectural acoustics (see, e.g., Forsyth, 1985). Much recorded music carries the additional complexity of multitracking, and it may be possible to analyze perceptions of reverb in a multitrack context (De Man et al., 2017).

Beyond genre, there are limitations to the GEMS-9 as an emotion measure. As noted, the GEMS-9 is the short version of the GEMS and can therefore be psychometrically less accurate (Zentner et al., 2008). Vuoskoski and Eerola (2011) found that a dimensional valence-arousal model outperformed the GEMS-9 when differentiating between musical stimuli, although the GEMS-9 was more nuanced. Self-report is more easily analyzed than behavioral data and physiological measurements (Mauss & Robinson, 2009), but it is limited by features such as demand characteristics—cues in participant instructions inducing hypothesis-consistent behavior—self-presentation biases, denoting an uneasiness to report undesirable emotion states, and limitations in emotional self-awareness (Zentner & Eerola, 2010, p.210). Efforts were made to mitigate these concerns by incorporating neutral language and randomizing the order of presentation of emotion terms.

In a related way, quantitative emotion ratings can only provide partial insights. Because the nature of "aesthetic" emotion states has been highly disputed (Menninghaus et al., 2019; Skov & Nadal, 2020), their experience in relation to acoustic space demands more detailed exploration. In future research, it would be useful to explore the feelings associated with acoustic space in a qualitative way across a wider range of contexts; for example, in critical discussion. While headphones enable the systematic influence of isolated RTs to be identified, the research does not examine the nuances of performance that would be expected in a live space. Such an examination would be better facilitated by a qualitative approach. A qualitative approach also enables closer analysis of individual differences in perception, whereas in this research, the listener sample was treated as an average of ordinary music listeners.

Finally, although the at-home context of the experiment reproduced normal headphone listening, it also generated more variability. In-person experiments have the advantage of enabling a controlled stimulus to be distributed to participants on identical devices. Sound reproduction technologies can have effects on frequency responses, source localization (Gutierrez-Parera & Lopez, 2016), and source internalization (Brimijoin et al., 2013). Although this research was not concerned with accurate positioning reproduction, such features may impact the overall perception of a signal. It may be that sensations of unease would have been more dependent on source positioning than RT (Hagman, 2010; Tajadura-Jiménez, Väljamäe, et al., 2010). Human auditory distance perception is under-researched (see Kolarik et al., 2016); future research may investigate thresholds for perception accuracy and discomfort.

Overall, the present study has contributed to an understanding of how alterations to reverberation can influence musical emotion, with an emphasis on recorded music listening, in tandem with a broad RT range and emotion measures directly targeted to music listening. The results demonstrate a clear relationship between higher RTs and emotion states associated with "Sublimity," particularly "Transcendence." Results also highlight a weakness in valence-arousal measures, and weak trends for preference. The identified effects bear associations with "aesthetic" emotions, which carry uncertainties in terms of their physiological correlates and relationships with auditory scene processing. Study methods limited the generalizability of these results to in-situ room acoustics and architectural modeling, instead focusing on parametric manipulations to reverberation. Future research might incorporate alternative listening scenarios, a different range of musical styles, and investigate, for example, free verbal discussion. Reverberation time can contribute to clarity and focus in recorded music listening, but it can also "overtake" the music to heighten concepts of "Sublimity."

Author Note

We have no conflicts of interest to disclose. We gratefully acknowledge the funding of the Cambridge Humanities Research Grant Scheme, and the Cambridge Trust and Clare Hall Boak Studentship. This research previously appeared in an unpublished Master's dissertation by the first author.

Correspondence concerning this article should be addressed to Hannah Wilkie (hw535@cantab.ac.uk) or Peter Harrison (pmch2@cam.ac.uk).

References

Aletta, F., & Kang, J. (2020). Historical acoustics: Relationships between people and sound over time. *Acoustics*, *2*(1), 128–130. https://doi.org/10.3390/acoustics2010009

Algargoosh, A., Soleimani, B., O'Modhrain, S., & Navvab, M. (2022). The impact of the acoustic environment on human emotion and experience: A case study of worship spaces. *Building Acoustics*, *29*(1), 85–106. https://doi.org/10.1177/1351010X211068850

Barrett, N. (2017). Trends in electroacoustic music. In N. Collins & J. D'Escrivan (Eds.), *The Cambridge companion to electronic music* (pp. 201–224). Cambridge University Press. https://doi.org/10.1017/9781316459874.013

Batcho, K. I. (2020). When nostalgia tilts to sad: Anticipatory and personal nostalgia. *Frontiers in Psychology*, *11*, 1186. https://doi.org/10.3389/fpsyg.2020.01186

Baumann, D., & Haggh, B. (1990). Musical acoustics in the middle ages. *Early Music*, *18*(2), 199–210. http://www.jstor.org/stable/3127808

Bech, S. R. (1996). Timbral aspects of reproduced sound in small rooms. II. *Journal of the Acoustical Society of America*, *99*(6), 3539–3549. https://doi.org/10.1121/1.413047

Beranek, L. (2011). Concert hall acoustics. *Architectural Science Review*, *54*(1), 5–14. https://doi.org/10.3763/asre.2010.0059

Boren, B. (2019). Computational acoustic musicology. *Digital Scholarship in the Humanities*, *34*(4), 707–715. https://doi.org/10.1093/llc/fqy073

Boren, B. (2021). Acoustic simulation of JS Bach's Thomaskirche in 1723 and 1539. *Acta Acustica*, *5*(14). https://doi.org/10.1051/aacus/2021006

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345. https://doi.org/10.2307/2334029

Brimijoin, W. O., Boyd, A. W., & Akeroyd, M. A. (2013). The contribution of head movement to the externalization and internalization of sounds. *PLOS One*, *8*(12), e83068. https://doi.org/10.1371/journal.pone.0083068

Brøvig, R., & Danielsen, A. (2016). *Digital signatures: The impact of digitization on popular music sound.* MIT Press.

Cabrera, D., Lee, D., Yadav, M., & Martens, W. L. (2011). *Decay envelope manipulation of room impulse responses: Techniques for auralization and sonification*. Proceedings of ACOUSTICS, Australia.

Choisel, S., & Wickelmaier, F. (2007). Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference. *Journal of the Acoustical Society of America*, *121*(1), 388–400.

Cockos (2004). *Reaper* (Version 6.80) [Computer Software]. https://www.reaper.fm/download.php

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, *117*(4), 1924–1934. https://doi.org/10.1073/pnas.1910704117

Cressman, D. (2016). *Building musical culture in nineteenth century Amsterdam: The concertgebouw.* Amsterdam University Press.

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M., & Merrill, J. (2021). Synchrony in the periphery: Inter-subject correlation of physiological responses during live music concerts. *Scientific Reports*, *11*(1), 22457. https://doi.org/10.1038/s41598-021-00492-3

De Man, B., McNally, K., & Reiss, J. D. (2017). Perceptual evaluation and analysis of reverberation in multitrack music production. *Journal of the Audio Engineering Society*, *65*(1/2), 108–116.

Doyle, P. (2005). *Echo and reverb: Fabricating space in popular music recording 1900–1960*. Wesleyan University Press.

Eerola, T., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Perception*, *30*(1), 49–70. https://doi.org/10.1525/mp.2012.30.1.49

Ehret, S., Schroeder, C., Bernet, J., Holzmüller, A., & Thomaschke, R. (2021) All or nothing: The interaction of musical and spatial atmosphere. *Psychology of Music*, *49*(3), 513–528. https://doi.org/10.1177/0305735619880288

Eidsheim, N. S. (2015). *Sensing sound: Singing and listening as vibrational practice*. Duke University Press.

Firth, D., & Turner, H. (2012). Bradley-Terry models in R: The BradleyTerry2 package. *Journal of Statistical Software*, *48*(9). http://www.jstatsoft.org/search

Fischinger, T., Frieler, K., & Louhivuori, J. (2015). Influence of virtual room acoustics on choir singing. *Psychomusicology: Music, Mind, and Brain*, *25*(3), 208–218. https://doi.org/10.1037/pmu0000117

Forsyth, M. (1985). *Buildings for music: The architect, the musician, the listener from the seventeenth century to the present day.* CUP Archive.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367–444). Oxford, Oxford Academic.

Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, *7*(2), 377–387. https://doi.org/10.1037/1528-3542.7.2.377

Gutierrez-Parera, P., & Lopez, J. J. (2016). Influence of the quality of consumer headphones in the perception of spatial audio. *Applied Sciences*, *6*(4), 117. https://doi.org/10.3390/app6040117

Hagman, F. (2010). *Emotional response to sound: Influence of spatial determinants* [Master's Thesis, Chalmers University of Technology].

Hahn, E. (2018). Musical emotions evoked by 3D audio. *Audio Engineering Society Conference: 2018 AES International Conference on Spatial Reproduction-Aesthetics and Science.* Audio Engineering Society. https://aes2.org/publications/elibrary-page/?id=19640

Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *Quarterly Journal of Experimental Psychology*, *62*(11), 2141–2155. https://doi.org/10.1080/17470210902765957

Harrison, P. M. C., Marjieh, R., Adolfi, F., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., et al. (2020). Gibbs sampling with people. *Advances in Neural Information Processing Systems*, *33*, 10659–10671.

Howard, D., & Moretti, L. (2009). *Sound and space in Renaissance Venice: Architecture, music, acoustics*. Yale University Press.

Johnson, J. H. (1995). *Listening in Paris: A cultural history.* University of California Press.

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., & 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., & Sloboda, J. A. (2010). Introduction: Aims, organization, and terminology. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 3–12). 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

Kaplanis, N., Bech, S., Jensen, S. H., & van Waterschoot, T. (2014). Perception of reverberation in small rooms: A literature study. In *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*. Audio Engineering Society. http://www.aes.org/e-lib/browse.cfm?elib=17348

Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, *17*(2), 297–314. https://doi.org/10.1080/02699930302297

Kleiner, M., Dalenbäck, B. I., & Svensson, P. (1993). Auralization-An overview. *Journal of the Audio Engineering Society*, *41*(11), 861–875.

Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, and Psychophysics*, *78*, 373–395. https://doi.org/10.3758/s13414-015-1015-1

Konečni, V. J. (2005). The aesthetic trinity: Awe, being moved, thrills. *Bulletin of Psychology and the Arts*, *5*(2), 27–44.

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. *Psychology of music*, *43*(2), 155–170. https://doi.org/10.1177/0305735613496860

Lawless, M. S. (2018). *Assessing the auditory and reward responses to room acoustics and music using functional magnetic resonance imaging* [Doctoral dissertation, Penn State University].

Lawless, M. S., & Vigeant, M. C. (2015). Investigating the emotional response to room acoustics: A functional magnetic resonance imaging study. *Journal of the Acoustical Society of America*, *138*(4), EL417–EL423. https://doi.org/10.1121/1.4933232

Lewers, T. H., & Anderson, J. S. (1984). Some acoustical properties of St Paul's Cathedral, London. *Journal of Sound and Vibration*, *92*(2), 285–297. https://doi.org/10.1016/0022-460X(84)90562-5.

Lin, H. L., Frieler, K., & Müllensiefen, D. (2021). *Development and validation of short forms of the Goldsmiths Musical Sophistication Index in three languages*. ICMPC/ESCOM.

Long, M. (2009). What is so special about shoebox halls? Envelopment, envelopment, envelopment. *Acoustics Today*, *5*(2), 21–25.

Luce, R. D. (2012). *Individual choice behavior: A theoretical analysis*. Courier Corporation. (Original work published 1959)

Martellotta, F. (2010). The just noticeable difference of center time and clarity index in large reverberant spaces. *Journal of the Acoustical Society of America*. *128*(2), 654–663. https://doi.org/10.1121/1.3455837

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*. *23*(2), 209–237. https://doi.org/10.1080/02699930802204677

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*. *126*(2), 171–195. https://doi.org/10.1037/rev0000135

Miller, P. (2012). An adventure into outer space: Stockhausen's Lichter-Wasser and the analysis of spatialized music. *Perspectives of New Music*, *50*(1), 342–392.

Mo, R., Wu, B., & Horner, A. (2015). The effects of reverberation on the emotional characteristics of musical instruments. *Journal of the Audio Engineering Society*. *63*(12), 966–979. http://www.aes.org/e-lib/browse.cfm?elib=18055

Odeon (2021). *ODEON* (Version 17) [Software]. https://odeon.dk/

Pätynen, J., & Lokki, T. (2016). Concert halls with strong and lateral sound increase the emotional impact of orchestra music. *Journal of the Acoustical Society of America*. *139*(3), 1214–1224. https://doi.org/10.1121/1.4944038

Pätynen, J., & Lokki, T. (2018). Dynamic responsiveness in concert halls as source of emotional impact. In *Auditorium Acoustics* (pp. 237–244). Institute of Acoustics.

Pearce, M. T., & Halpern, A. R. (2015). Age-related patterns in emotions evoked by music. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 248–253. https://doi.org/10.1037/a0039279

Pentcheva, B. V. (2011). Hagia Sophia and multisensory aesthetics. *Gesta*, *50*(2), 93–111. https://www.journals.uchicago.edu/doi/abs/10.2307/41550552

Postma, B. N., & Katz, B. F. (2016). Acoustics of Notre-Dame Cathedral de Paris. In *Proceedings of the International Congress on Acoustics (ICA)*, 5–9.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sander, K., Wycisk, Y., & Kopiez, R. (2023). Limited emotional value added by immersive 3D audio: A reanalysis. *Music and Science*, *6*, 20592043231194751. https://doi.org/10.1177/2059204323119

Schärer Kalkandjiev, Z., & Weinzierl, S. (2013). The influence of room acoustics on solo music performance: An empirical case study. *Acta Acustica United with Acustica*, *99*(3), 433–441. https://doi.org/10.3813/AAA.918624

Scheibler, R., Bezzam, E., & Dokmanić, I. (2018). Pyroomacoustics: A python package for audio room simulation and array processing algorithms. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 351–355). ICASSP. https://doi.org/10.1109/ICASSP.2018.8461310

Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S., & Eyben, F. (2017). The expression of emotion in the singing voice: Acoustic patterns in vocal performance. *Journal of the Acoustical Society of America*, *142*(4), 1805–1815. https://doi.org/10.1121/1.5002886

Shusterman, R. (2005). Somaesthetics and Burke's sublime. *The British Journal of Aesthetics*, *45*(4), 323–341.

Skov, M. & Nadal, M. (2020). There are no aesthetic emotions: Comment on Menninghaus et al. (2019). *Psychological Review*, *127*, 640–649. https://doi.org/10.1037/rev0000187

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in everyday listening to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 415–429). Oxford University Press.

Sterne, J. (2015). Space within space: Artificial reverb and the detachable echo. *Grey Room*, *60*, 110–131.

Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., & Trainor, L. J. (2019). How live music moves us: Head movement differences in audiences to live versus recorded music. *Frontiers in Psychology*, *9*, 2682. https://doi.org/10.3389/fpsyg.2018.02682

Tajadura-Jiménez, A., Larsson, P., Väljamäe, A., Västfjäll, D., & Kleiner, M. (2010). When room size matters: Acoustic influences on emotional responses to sounds. *Emotion*. *10*(3), 416–422. https://doi.org/10.1037/a0018423

Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion*, *10*(2), 216. https://doi.org/10.1037/a0018422

Tan, D., Diaz, F. M., & Miksza, P. (2020). Expressing emotion through vocal performance: Acoustic cues and the effects of a mindfulness induction. *Psychology of Music*, *48*(4), 495–512. https://doi.org/10.1177/0305735618809873

Thompson, E. (2002). *The soundscape of modernity: Architectural acoustics and the culture of listening in America, 1900–1935*. MIT Press.

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Trondle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96. https://doi.org/10.1037/a0023845

Västfjäll, D., Larsson, P., & Kleiner, M. (2002). Emotion and auditory virtual environments: Affect-based judgments of music reproduced with virtual reverberation times. *CyberPsychology and Behavior*, *5*(1), 19–32. https://doi.org/10.1089/109493102753685854

von Berg, M., Steffens, J., Weinzierl, S., & Müllensiefen, D. (2021) Assessing room acoustic listening expertise. *Journal of the Acoustical Society of America*, *150*(4), 2539. https://doi.org/10.1121/10.0006574. PMID: 34717486

Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173. https://doi.org/10.1177/1029864911403367

Woods, K. J., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Zentner, M., & Eerola, T. (2010). Self-report measures and models. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 187–221). Oxford Academic.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*, 494–521. https://doi.org/10.1037/1528-3542.8.4.494