DURMS
Durham Undergraduate
Research in Music & Science

Is it emotionally worth it to attend a Classical music concert? An empirical study on emotional responses to live versus recorded classical piano music experiences

Thomas Milnes
Durham University

ABSTRACT

Given the emotional power of music, explanations are needed for the decline in Western Classical music concert attendance within the domain of musically-induced emotion. This study considers the differences in audio-only and audiovisual stimuli as well as social setting, deploying a self-report measure to compare emotional responses to live versus recorded experiences of Franz Liszt's Consolation No. 3 in Db Major. Relevant participant demographics are also considered. Music listening habits, including whether or not participants listen to classical music, and musical expertise are examined with the objective of determining the extent to which these factors affect emotional responses to Western Classical music in both the live and recorded environments. The study found some significant differences with regards to high frequency music listeners; but overall, there was no significant difference in emotional responses across the live against recorded environments.

1. INTRODUCTION

Music is recognised as an emotionally powerful phenomenon in both live (Trost et al., 2024) and recorded formats (DeNora, 2000). This study's primary objective is to discern differences in emotional responses for live versus recorded music experiences of Western Classical Music; namely, Franz Liszt's Consolation No. 3 in D♭ Major. In this study, all genres of Romantic and Classical Western music from c. 1750 onward are included in the term 'classical.'

There has been a severe decline in Western Classical music concert attendance (Johnson et al., 2004), suggesting a need for targeted research, which justifies this study's Classical orientation. The objective is to reveal whether developments in recording and streaming technology, with its "numerous practical and financial advantages" (Wald-Fuhrmann et al., 2021, p. 2), might be related to the aforementioned decline. People may have grown tired of attending serious music events, such as Classical concerts, where there is a culture and ritual of how to react to 'high art,' (Wald-Fuhrmann et al., 2021) in favour of enjoying the same quality of aesthetic experience in the comfort of their homes. Data on the emotional responses in both contexts will offer insight into whether this is the case.

Two salient themes from current research offered a basis for hypotheses: audiovisual versus audio-only experiences; social versus isolated experiences. Focus on attention to the music, a potential variable, was controlled in this study.

1.1 Variables germane to live versus recorded music experiences.

1.1.1 Audio-only versus audiovisual experiences. Emotion in music performance is communicated through audio cues such as articulation, timbre, vibrato, dynamics and timing (Gabrielsson, 1999; Juslin & Timmers, 2010; Palmer, 1997; Vuoskoski et al., 2016); and also through visual cues: gestures, posture and facial expressions (Dahl & Friberg, 2007; Davidson, 1993, 1994; Thompson et al., 2008; Vuoskoski et al., 2016). Initially, then, given a combination of two contributing cues (audio and visual), a more intense emotional reaction would be expected. However, research in this area is contradictory (Czepiel et al., 2023).

Vuoskoski et al. (2016) took physiological and self-report measures from 19 participants in two different modal conditions: audiovisual (AV; bi-modal), audio-only (AO; uni-modal) and visual-only (VO; uni-modal). Both measurements concurred that the AV experience was no more emotionally provocative than the AO experience; but also that the AO experience was more provocative than the VO experience. These findings suggest that whether stimuli are unimodal or bimodal does not have a notable impact on emotional responses to Western Romantic music (Brahms' Intermezzo in B Minor). If AV against AO does not affect emotional responses as an independent variable, this predicts stable emotional responses across the live and recorded performances.

A contradiction with these 2016 findings, Pan et al. (2018) confirmed de Gelder & Vroomen's (2000) earlier study that music-induced emotion is consistent with the multisensory integration effect. McGurk & MacDonald (1976) concluded that changes or discoordination in audiovisual stimuli results in different perceptions of the same information, whereas multisensory integration (i.e., audiovisual integration) provides the most accurate perception. In Pan et al.'s (2018) and de Gelder & Vroomen's (2000) investigations, both concluded that audiovisual integration influences emotional reactions to music; whilst Pan et al. (2018) revealed that "emotion could be enhanced when the audio and visual emotions are consistent" (p. 18). Consistency, in this case, refers to the congruence of audio and visual emotional projection of a performer (e.g. facial expressions and articulation, as discussed).

Further, Czepiel et al. (2023) studied aesthetic and physiological effects of music under AV and AO conditions, attempting to replicate naturalistic settings. They took physiological measures to determine both dependent variables and are concerned with the enhancement of aesthetic experience, of which emotion is a significant contributor (Schindler et al., 2017). Czepiel et al.'s (2023) findings indicate that, although involuntary physical arousal in the somatic nervous system was higher in the AO modality, this reflected a more sensory, cognitive process in anticipation of what participants would hear next rather than an emotional aspect of aesthetic experience. Conversely, peripheral physiological responses with more degrees of control such as respiration and activity in the zygomaticus major – smile muscle – were higher in the AV modality, reflecting an enhanced aesthetic experience. An enhanced aesthetic experience is directly tangible with increased positive emotions and decreased negative emotions (Schubert et al., 2017), suggesting that the emotional reaction to an AV modality is indeed enhanced compared to an AO modality.

One reason for the inconsistencies between Vuoskoski et al.'s (2016) study and Czepiel et al.'s (2023) study is the difference in ecological concern. Vuoskoski et al.'s (2016) study was conducted fully in laboratory environments, with the AV stimulus being a video watched in a room; whilst Czepiel et al.'s (2023) investigation was undertaken in a concert hall. The experience of a concert inherits a plethora of cues that may affect emotional reactions, including the (potentially subconscious) vibrotactile perception of sounds, comfort of seating itself and lighting. The setting of a concert stretches in potential impact from material aspects such as these to sociocultural associations with concert halls (Wald-Fuhrmann et al., 2021). Furthermore, Sloboda et al. (2012) discuss how investigations into emotional reactions to music concerts *not* conducted in concert venues may inherit limited ecological validity. This might explain why Vuoskoski et al. (2016) did not find any significant differences between AO and AV cues.

Between Pan et al.'s (2018) and Vuoskoski's (2016) study, the discrepancies may be produced by the differences in repertoire exposure: whilst participants in Vuoskoski et al.'s (2016) study were exposed to only one piece, Pan et al.'s (2018) participants were exposed to a range of six pieces of Western Classical music. Pan et al.'s (2018) study is therefore more empirically grounded because emotional responses were tested with multiple stimuli within the domains of uni-modal and multi-modal cues.

In the domain of the AV versus AO variable, current research is inconclusive; but the majority suggests that the AV stimulus of the live performance should elicit a more intense emotional reaction than the AO recording.

1.1.2 Social versus isolated setting. Some studies have shown that listening in group contexts may cause distractions and reduce focus on attention (Egermann et al., 2011); leading to no increase in the intensity of emotional reactions but also no notable decrease.

In contrast, Liljeströhm et al. (2013) found that listening to music with a significant other does increase the intensity of emotional reactions. Pitts (2005) furthermore posits that there is a "close relationship between social and musical enjoyment that is at the heart of concert attendance" (p. 269), implying that the aesthetic experience of music (connected with emotional experience, as mentioned) observed in the study is closely associated with its social experience.

However, in the context of this study, social context was not deemed to be a significant factor from which to deduce hypotheses given (i) the finding that social experience does not affect emotional intensity (Egermann et al., 2011); (ii) the demographic of participants recruited not being couples or significant others so as to enhance the emotional experience (Liljeströhm et al., 2013), but a group of randomly selected individuals; and (iii) the nature of Pitts' (2005) findings being qualitative data, offering no empirical margin of emotional fluctuation across the two scenarios.

1.1.3 Focus on attention. Focus on attention is controlled in this study (see **Method**). Variation in emotional responses consequent to non-focussed attention or focussed attention on the music were not of concern for ascertaining the difference between emotional responses to live versus recorded music experiences.

1.2 Hypotheses.

1.2.1 Primary hypothesis. The primary hypothesis is formed resultant to the majority of findings suggesting that AV experiences elicit more intense emotional reactions than AO experiences; despite social setting being deemed not significant.

H1: The live music experience will elicit more intense emotional responses than the recorded music experience.

1.2.2 Secondary hypotheses. Three secondary hypotheses were tested in addition to the primary hypothesis in order to gauge whether individual participant characteristics affected any observed emotional fluctuation across the two scenarios.

(a) Musical expertise. Musical expertise is usually believed to refine processing of musical structure (Bigand et al., 2005). However, there is empirical evidence that emotional responses to music are not strongly correlated with musical training (Lynchner, 1998). Therefore, secondary hypothesis (a) is formulated:

H2(a): Musical expertise will not affect emotional responses to either live or recorded music experiences.

(b) Frequency of music listening. Sloboda (1991) found that in eighty-three British adults an increased frequency in music listening over a period of five years was linked with an increased number of reported physiological reactions, implying a positive relationship between frequency of music listening and emotional receptiveness. Secondary hypothesis (b) is generated:

H2(b): High frequency music listeners will experience more intense emotional reactions to the live and recorded experiences than low or moderate frequency music listeners.

(c) Western Classical music listeners. Schedl et al. (2018) found that Western Classical music listeners tend to disagree more on emotional reactions to Western Classical music than non-classical music listeners. However, evidence that more intense emotional reactions are elicited by either group is inconclusive. Exploratory secondary hypothesis (c) is generated:

H2(c): Western Classical music listeners will experience more intense emotional reactions to both live and recorded experiences than non-classical music listeners.

2. METHODS

2.1 Design. The investigation is an experiment with participants completing the same questionnaire (see **Materials**) after both live and recorded experiences of: the same work, performer and attempted musical expression (see **Materials**).

2.2 Participants. Participants were 15 Durham University students aged 19-21 years, 7 male; 8 female. They were recruited via word-of-mouth and through social media advertising. Given the already large quantity of within-participant variables investigated, it was decided that gender and age would be set aside from the investigated independent variables. Age and gender are branches from which future research can stem in this area.

An attrition rate was observed in that 3 participants did not partake in the live experience. However, data from all 15 participants was valuable for testing the secondary hypotheses for emotional responses to the recording. **Table 1** (overleaf) outlines participant demographics relevant to the secondary hypotheses (criteria for assignment forthcoming).

Table 1. *Participants per category.*

	Very high	High	Moderate	Low
Frequency of music listening	3	8	3	1
Musical expertise	N/A	3	4	8
Classical music listeners	7			
Non-classical music listeners	8			

2.3 Materials. Participants completed the Geneva Emotional Music Scale (GEMS; Zentner et al., 2008) twice. Participants rated 32 adjectival descriptors on a Likert scale of 1-5: 1 meaning they did not feel the emotion *at all*; 5 meaning they felt the emotion *very much*. See **Appendix A** for the full list of 32 descriptors, and **Data Analysis** for how they are processed.

GEMS was chosen because of its popularity (Juslin, 2016) and its recent use by experts in the field of music psychology (e.g. Baltes & Miu, 2014). It targets music-related dimensions of affect (Asmus, 1985); which is a crucial criterion for use in this study given that the emotional impact of music may be outside the parameters of measurement for everyday dimensions of affect (Collier, 2007).

In addition to the GEMS, the first questionnaire contained three more items. Modelled on Zhang & Schubert (2019), it determined level of musical expertise by asking for years of private lessons received; years of daily practice executed; and tertiary-level education enrolment (**Table 2**). This method of determining musical expertise was adopted so that classifications were not based on self-ratings of musicianship, which would have been subject to participant's definition of musicality. This method modelled on Zhang & Schubert (2019) allows for a consistent definition of musicianship to be applied.

Table 2. *Criteria for categories of musical expertise, modelled on Zhang & Schubert (2019, p. 464).*

	Level of musical expertise		
	High	Moderate	Low
Years of private lessons	>10	6-10	<6
Years of daily practice	>6	>6	<6
College-level music tuition	Not enrolled in tertiary-level music course	At least enrolled in 1 non-major music course	Enrolled in Bachelor of Music or above

The second item included in the first questionnaire aimed to deduce music listening habits based on Kähäri et al.'s (2011) criteria for categorisations of frequent and non-frequent listeners. Participants were asked for weekly listening habits; and if they listened to music daily, for hourly estimates of the time they spent listening to music. The criteria for assignment to each category are outlined in **Table 3** (overleaf). The present study further subdivided Kähäri et al.'s (2011) categorisations based on hourly listening habits to ascertain a more discrete continuum.

The third item included in the first questionnaire asked whether participants listened to Western Classical music.

A recording of the researcher playing Liszt's *Consolation No. 3 in Db Major* was made on 12/01/2024 using professional-grade equipment and qualified sound technicians at Collingwood College Music Studios, Durham University, United Kingdom. The link to the recording is made available as Appendix B. The live performance was made as similar as possible in all aspects of expression and articulation.

Table 3. *Criteria for categories of frequency of music listening.*

Category	Reported listening habit
Very high	4-8 hours, daily
High	2-4 hours, daily
Moderate	0-2 hours, daily
Low	A few times a week

2.4 Procedure. In the first part of the investigation, participants were asked to listen to the recording of the researcher playing Liszt's *Consolation No. 3 in Db Major*. The recording on questionnaire 1 interpolated the demographic questions and the GEMS questionnaire. Participants were asked to listen in a space where they felt comfortable and were alone; where they were able to completely focus on the music. They were asked not to analyse the music and specifically to rate emotions that they *felt* and not emotions that they *perceived* the music to evoke.

A one-week interval was left between when all the participants had completed the first questionnaire and when they attended the live experience. This was an attempt to ensure ecological validity for the responses to the live performance as self-report measures taken a greater time interval apart are more likely to yield accurate reports (Stone et al., 2007). A period of one week was considered an appropriate period to balance self-report validity and minimise the rate of participant attrition.

Participants attended the live performance in the Music Department Concert Hall, Durham University, UK. Participants were asked to remain silent during the performance and, as with the recorded version, focus attention completely on the music and the performer. Participants completed a final digitised version of the GEMS questionnaire immediately after the performance.

2.5 Data Analysis. Data was processed using Microsoft Excel version 16.82 for descriptive statistics; whilst RStudio version 2023.12.1+402 was used to ascertain inferential statistical values using ANOVAs and t-tests. The confidence interval is 95%.

Figure 1 (overleaf) is taken from Zentner et al. (2008, p.507) and shows how each value for the GEMS descriptors was assigned second and third order factors, which are the focus of the forthcoming results and discussion.

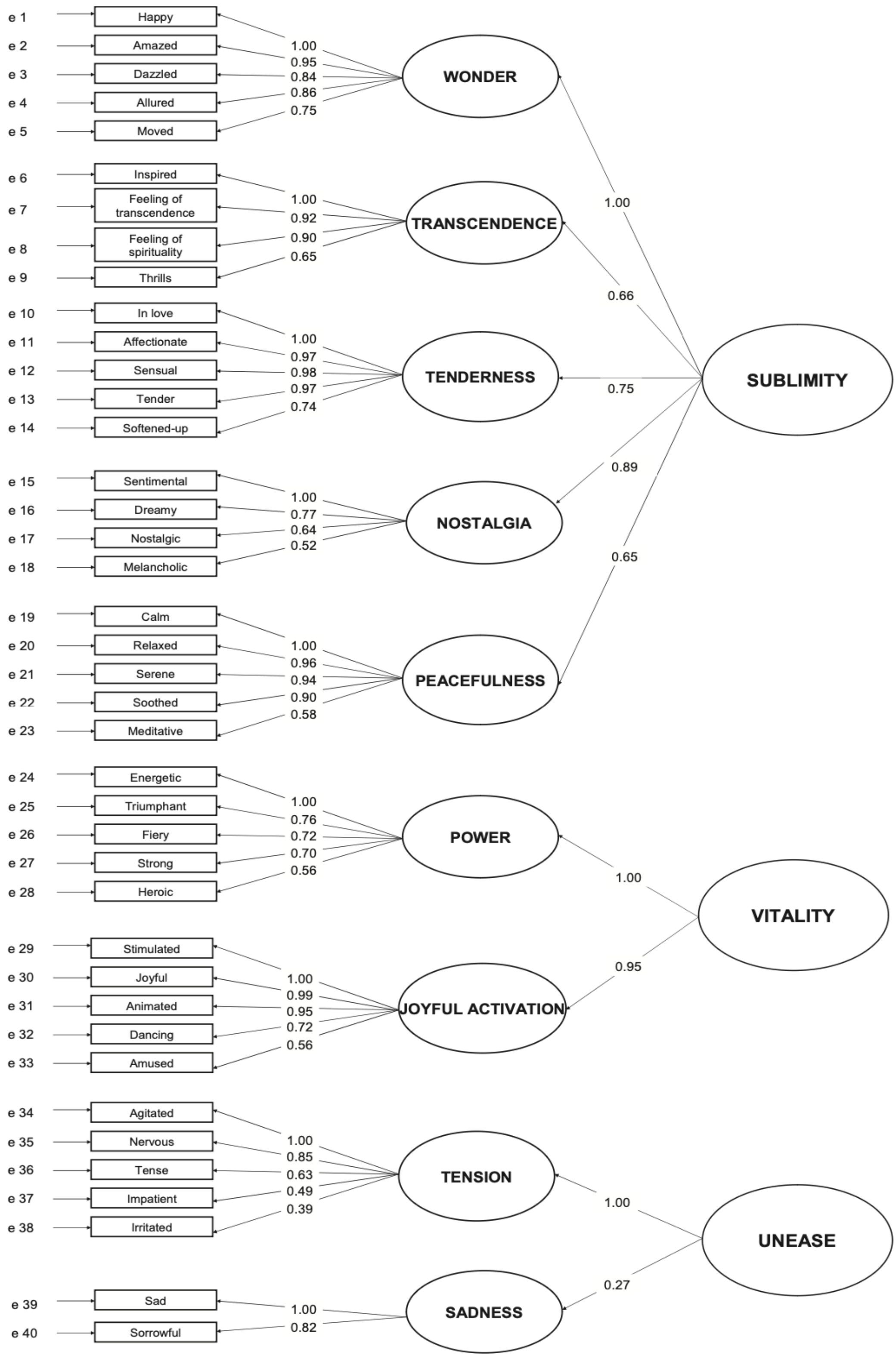

Figure 1. *GEMS Descriptors assigned first and second-order factors; taken from Zentner et al. (2008, p. 57).*

3. RESULTS

3.1 Insufficient values for statistical evaluation in some areas.

Running multiple one-way ANOVAs on categories musical expertise and frequency of music listening revealed it was not possible to evaluate the statistical significance of differences between participants within some subcategories. For example, there was only 1 participant who had a low level of musicianship and reported a value for *power*, from the GEMS model, in the live experience.

An ANOVA was performed to estimate the significance of variance for *power* against the three levels of musicianship in the live environment ($F(2, 2) = 66.38$, $p = .02$), but the resultant pairwise t-test comparisons cannot be run because the singular value (for low musicianship) is not a basis for a statistical comparison.

This was the case for several first-order factors from the GEMS model compared against musical expertise and frequency of music listening. Appendix D shows which factors this applies to and gives the ANOVA outcomes for these categories.

3.2 Secondary hypotheses.

3.2.1 Hypothesis 2(a): musical expertise will not affect emotional responses to either live or recorded music experiences. This hypothesis is supported by a null hypothesis in almost all areas: no statistically significant differences were recorded across musical expertise categories for the live or recorded environments (**Appendix D**); that is, bar one finding. ANOVA for *tenderness* across musical expertise for the recorded experience was significant, $F(2, 12) = 5.85$, $p = .02$. Pairwise t-test comparison outcomes are outlined in **Table 4**, revealing that the significant difference in tenderness was between low and moderate musical expertise for the recorded experience. For the live experience, there was no significant variation in tenderness across musical expertise categories where $F(2, 4) = 1.07$, $p = .42$. **Figure 2** outlines these differences which will be discussed further in the **Discussion** section.

Table 4. *Pairwise t-tests for tenderness against musical expertise, recorded environment.*

	Highest	
Low	*p* = .20	Low
Moderate	*p = .30*	***p* = .02***

*Statistical significance shown by asterisk and bold text, here and henceforth.

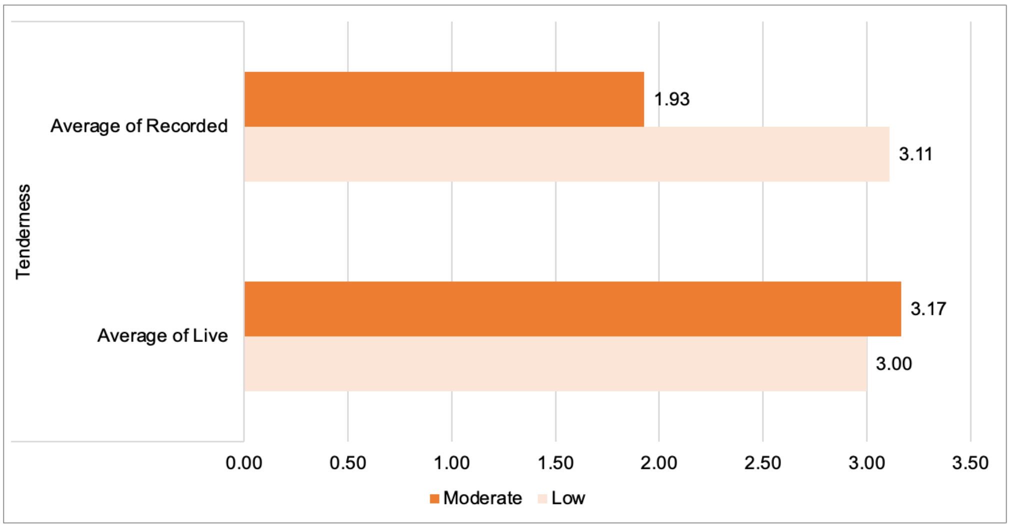

Figure 2. *Tenderness, live and recorded environments: moderate against low musical expertise.*

3.2.2 Hypothesis 2(b): High frequency music listeners will experience more intense emotional reactions to the live and recorded experiences than low or moderate frequency music listeners. Descriptive results show a trend in sublimity and unease which supports this hypothesis (**Figure 3**); whilst vitality shows an opposite trend. However, these trends are not statistically supported as running ANOVAs for second order factor variation against frequency of music listening shows non-significance in both the recorded and live environments, as outlined in **Table 5.** The same is true for first-order factors (**Appendix D**). Thus, a null hypothesis is supported over hypothesis 2(b): high (or very high) frequency music listening did not evidence any more intense emotional reactions than any other music listening habits.

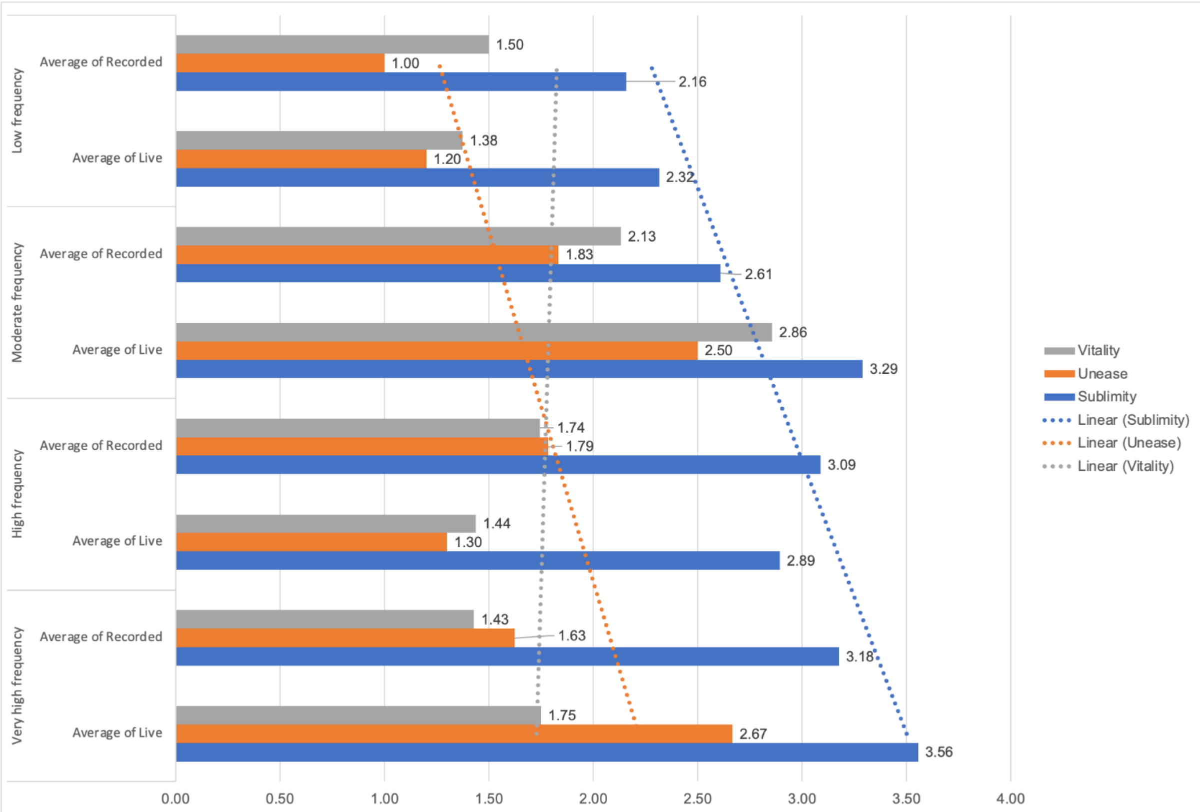

Figure 3. *Second order factors: descriptive trends for frequency of music listening, live and recorded environments.*

Table 5. *ANOVA outcomes against frequency of music listening, live and recorded environments.*

Second order factor	Environment	ANOVA Outcome against frequency of music listening
Sublimity	Recorded	F(3, 11) = 1.76, *p* = .21
	Live	F(3, 3) = 3.10, *p* = .21
Vitality	Recorded	F(3, 11) = .30, *p* = .82
	Live	F(3, 3) = 2.52, *p* = .26
Unease	Recorded	F(3, 10) = .40, *p* = .76
	Live	F(3, 3) = 3.10, *p* = .19

3.2.3 Exploratory hypothesis 2(c): Western Classical music listeners will experience more intense emotional reactions to both live and recorded experiences than non-classical music listeners. This hypothesis was rejected as classical music listening did not have any impact on emotional responses, which can be concluded using descriptive statistics only. **Figure 4** shows stability across the second order factors for classical and non-classical music listeners for both the recorded and live environments. This stability is confirmed in the first-order factors and t-tests (**Appendix D**). This exploratory hypothesis was unsupported.

3.3 Primary hypothesis: emotional response variation, live versus recorded environments.

3.3.1 Overall variation: all groups. Five participants did not attend the live experience. Data gathered from these five participants is also excluded from analysis of the recorded experience so as to maintain consistency in degrees of freedom and thus obtain more empirically valid outcomes.

Overall, no significant variation was evidenced for the live versus recorded experiences (**Figure 5**). Independent t-tests for variability within each first order and second order factor were also run, which confirm support for a null hypothesis (**Table 6**).

Table 6. *T-test outcomes, live versus recorded environments independent of secondary hypotheses variables.*

First & **Second** order factors	T-test outcome: live versus recorded environments independent of secondary hypotheses variables
Nostalgia	t(12.47) = .30, *p* = .77
Peacefulness	t(13.33) = 1.18, *p* = .26
Tenderness	t(14.22) = .41, *p* = .69
Transcendence	t(14.15) = -.53, *p* = .60
Wonder	t(9.78) = .52, *p* = .61
Sublimity	t(12.97) = .64, *p* = .53
Sadness	t(15.97) = .55, *p* = .59
Tension	t(8.59) = -.03, *p* = .33
Unease	t(12.92) = .25, *p* = .81
Joyful activation	t(13.23) = .62, *p* = .55
Power	t(6.39) = .55, *p* = .36
Vitality	t(11.26) = -.03, *p* = .98

Investigating factors related to the secondary hypotheses, however, did reveal variation across live versus recorded experiences for high frequency music listeners.

3.3.2 Frequency of music listening: salient variation across live versus recorded experiences. Although high frequency music listeners did not experience more intense emotional reactions within either environment than listeners of any other category, there was a statistically significant inhibition between live and recorded environments for high frequency music listeners in first order factors tension, t(5) = -2.52, *p* = .05, and power, t(6.36) = -3.38, p = .01. Second order factor vitality also displays a statistically significant decrease: t(7.82) = -2.32, p = .05 (**Figure 6**, overleaf). The second order factor corresponding to tension, unease, does not show significant variation across experiences, t(7.85) = -1.82, p = .11.

T-tests show that there are no other statistically significant differences between live and recorded experiences across all other participant subcategories (musical expertise, **Appendix E\1**; frequency of music listening, **Appendix E\2**; Y/N classical music listeners, **Appendix E\3**).

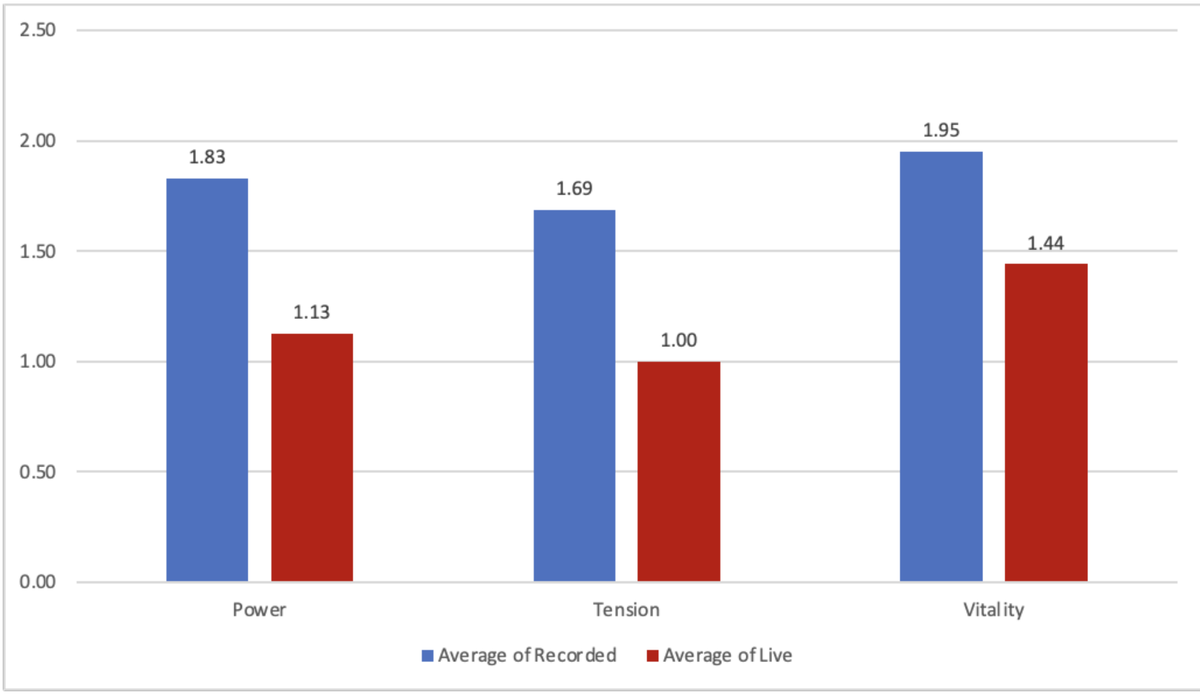

Figure 6. *High frequency music listeners: inhibition of power, tension and vitality, recorded versus live environment.*

4. DISCUSSION

4.1 Summary of results.

The primary hypothesis was rejected as there was no significant variation between live and recorded experiences of Liszt's Consolation No. 3 in Db Major. Secondary hypothesis (a) was supported by data in all areas except for variation in tenderness. Secondary hypothesis (b) and exploratory hypothesis (c) were rejected as (b) there was no variation resultant to frequency of music listening, with the exception that high frequency music listeners inhibited tension, power and vitality in the live experience compared with the recorded experience, nor (c) whether participants listened to classical music.

4.2 Discussion of findings in research context.

4.2.1 Overall conclusions. In line with Vuoskoski et al.'s (2016) findings, it appears multimodal (audiovisual) stimuli do not enhance the emotional experience of music. Moreover, these findings build on Vuoskoski et al.'s (2016) research by applying a comparison to a naturalistic setting (a concert venue) rather than a laboratory environment. The findings contradict the work of Pan et al. (2018) who found that the multi-sensory integration effect applied to music: in the present study, audiovisual stimuli were congruent for the live performance, whilst there was only an audio stimulus present for the recording, where visual information was absent. With regards to Czepiel et al.'s (2023) findings that the aesthetic experience was enhanced in the live performance, this study highlights that emotional responses are not necessarily as tangible with enhanced aesthetic experience within the domain of music performance reception. This underscores the need for further research into the variables affecting emotional responses to music performances and the resultant fluctuation in aesthetic appreciations.

In terms of social setting, attending a concert does not increase emotional intensity - in line with Egermann et al.'s (2011) finding that listening in a group setting may increase the overall enjoyment of the event, but does not intensify emotional experience. The findings contradict Pitts (2005) in that social experience is not a significant contributor to the emotional experience of a classical music concert. (The live experience was within a social setting and did not provoke more intense emotional reactions compared to the recording, an isolated setting.)

4.2.2 Significant findings in research context. The finding that, for moderate musical expertise, tenderness showed an increase from the recorded experience to statistically similar levels with low musical expertise suggests that the live environment is levelling and causes all levels of musicianship to experience similar emotional responses. This builds on Lynchner's (1998) study in adding a baseline comparison to a recorded experience, proving that musical expertise does not influence live emotional responses. This is furthermore supported by no statistically significant differences between all other factors in the live environment against musical expertise.

The inhibition of tension by high frequency music listeners in the live environment compared to the recorded environment might be explained by Linnemann et al. (2016) who suggest that stress is inhibited to a greater extent in a social environment. The fact that this is more likely to happen to high frequency music listeners is a new finding and indicates that people who listen to music with high frequency are more likely to be emotionally relieved by live performances. However, it is strange that very high frequency music listeners did not exhibit the same variation. It could be an issue of repeated overexposure to music inhibiting their emotional responses, although this does contradict current research surrounding the mere exposure effect (MEE) which induces an increased liking for a stimulus after repeated unenforced exposure (Barnstein & Craver-Lemley, 2022). Further research is needed on music listening habits and how these may affect recorded and live emotional responses to musical stimuli.

The inhibition of power and vitality overall by high frequency music listeners in the live compared to the recorded environment might be explained by the enhanced aesthetic experience (Czepiel et al., 2023) having a humbling effect on high frequency music listeners. The high frequency music listeners evaluated also held high musical expertise who, given the inclusion criteria for this category, would have an appreciation of the hours practice needed to perform the piece. This indicates that the experience of a live performance may earn a greater technical and aesthetic appreciation than a recorded experience, which is in line with current research on aesthetic experience being enhanced in a live environment (Czepiel et al., 2023).

4.2.3 Limitations. The number of participants recruited was undesirably low resultant to lack of incentive and the need to attend an in-person event. This complicated already heavy schedules for the participants. Moreover, the demographic of participants all being undergraduate students at Durham University does not necessarily reflect all age groups and social classes. Best attempts were made at recruiting participants but incentives (e.g., payment) were beyond the ethical scope of the study.

A second limitation was the non-consideration of trait characteristics of the participants. Future studies might deploy questionnaires such as the State-Trait Anxiety Inventory (STAI; Spielberger, 1983) or the Brief Symptom Inventory-18 (BSI-18; Derogatis, 2001) in order to ensure a consistent psychological state across both live and recorded environments. This would ensure that emotional reactions did not fluctuate as a consequence of significant events or trauma outside the scope of the study.

A third limitation is the use of self-report measures, and more specifically of GEMS which lacks relevant categories of *interest* and *surprise* and has further limitations that render it more valuable as "an example of how a scale might be developed than for the rating scale as such" (Juslin, 2016, p. 208). Self-report measures also do not detect underlying, subconscious mechanisms that may affect emotional experience (Cousineau & Shedler, 2006). Future research should consider incorporating physiological measures, as in Czepiel et al. (2023) and Vuoskoski et al. (2016), as well as self-reports in order to ascertain corroborative data.

4.2.4 Conclusion in line with study rationale. Given that the primary hypothesis was rejected, the decline in Western Classical music concert attendance is partly explained by a statistical similarity in emotional responses live versus recorded experiences. There is no emotional motivator to attend a live performance of Western Classical music when the same emotional responses can be elicited by listening at home.

REFERENCES

Asmus, E. (2005). The development of a multidimensional instrument for the measurement of affective responses to music. *Psychology of Music*, *13*(1), 19–30.

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

Barnstein, R. F., & Craver-Lemley, C. (2022). Mere exposure effect. In R. F. Pohl (Ed.), *Cognitive Illusions (3rd Ed.)*. Taylor & Francis.

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.

Collier, G. L. (2007). Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, *35*(1), 110–131.

Cousineau, T. M., & Shedler, J. (2006). Predicting physical health: Implicit mental health measures versus self-report scales. *The Journal of Nervous and Mental Disease*, *194*(6), 427–432.

Coutinho, E., & Scherer, K. R. (2016). The effect of context and audio-visual modality on emotions elicited by a musical performance. *Psychology of Music*, *45*(4), 550–569. https://doi.org/10.1177/0305735616670496

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2023). Aesthetic and physiological effects of naturalistic multimodal music listening. *Cognition*, *239*, 105537. https://doi.org/10.1016/j.cognition.2023.105537

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*(5), 433–454.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(2), 103–113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive invention to an audience? *Journal of Human Movement Science*, *26*(7), 279–301.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311.

DeNora, T. (2000). Music as a technology of self. In *Music in Everyday Life* (pp. 46–74). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9780511489433.004

Derogatis, L. R. (2001). *Brief symptom inventory-18: Administration, scoring and procedures manual*. Minneapolis, MN: NCS Assessments.

Egermann, H., Sutherland, M. E., Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2011). Does music listening in a social context alter experience? A physiological and psychological perspective on emotion. *Musicae Scientiae*, *15*(3), 307–323. https://doi.org/10.1177/1029864911399497

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 501–602). Academic Press.

Johnson, G., Pfrommer, P., Stewart, S., Glinkowski, P., Fenn, C., Skelton, A., & Joy, A. (2004). *New audiences for the arts*. Arts Council England. https://ualresearchonline.arts.ac.uk/id/eprint/1815/1/newaudiencessummary_phpVmLnJx.pdf

Juslin, P.N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559575. Cambridge Core. https://doi.org/10.1017/S0140525X08005293

Juslin, P. N. (2016). Emotional reactions to music. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology (2nd ed.)*. Oxford University Press.

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and emotion: Theory, research, Applications* (pp. 453–489). Oxford University Press.

Kähäri, K., Åslund, T., & Olsson, J. (2011). Preferred sound levels of portable music players and listening habits among adults: A field study. *Noise and Health*, *13*(50), 9. https://doi.org/10.4103/1463-1741.73994

Liljeström, S., Juslin, Patrik N, & Västfjäll, D. (2013). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychology of Music*, *41*(5), 579–599. https://doi.org/10.1177/0305735612440615

Linnemann, A., Strahler, J., & Nater, U. M. (2016). The stress-reducing effect of music listening varies depending on the social context. *Psychoneuroendocrinology*, *72*, 97–105.

Lynchner, J. (1998). An empirical study concerning terminology relating to aesthetic response to music. *Journal of Research in Music Education*, *46*(2), 303–320.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748.

O'Neill, K., & Egermann, H. (2022). Development of the social experience of a concert scales (SECS): The social experience of a live Western art music concert influences people's overall enjoyment of an event but not their emotional response to the music. *Music & Science*, *5*, 1–22. https://doi.org/10.1177/20592043221106478

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138.

Pan, F., Zhang, L., Ou, Y., & Zhang, X. (2019). The audio-visual integration effect on music emotion: Behavioral and physiological evidence. *PLOS ONE*, *14*(5), e0217040. https://doi.org/10.1371/journal.pone.0217040

Pitts, S. E. (2005). What makes an audience? Investigating the roles and experiences of listeners at a chamber music festival. *Music and Letters*, *86*(2), 257–269. https://doi.org/10.1093/ml/gci035

Schedl, M., Gómez, E., Trent, E. S., Tkalčič, M., EghbalZadeh, H., & Martorell, A. (2018). On the Interrelation Between Listener Characteristics and the Perception of Emotions in Classical Orchestra Music. *IEEE Transactions on Affective Computing*, *9*(4), 507–525. https://doi.org/10.1109/TAFFC.2017.2663421

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

Schubert, E., North, A. C., & Hargreaves, D. J. (2017). Aesthetic experience explained by the affect-space framework. *Empirical Musicology Review*, *11*(3-4), 330–345.

Sloboda, J. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*(2), 110–120.

Sloboda, J., Lamont, A., & Greasley, A. (2012). In S. Hallam, I. Cross, & M. Thaut (Eds.), *Choosing to hear music*. Oxford University Press.

Spielberger, C. D. (1983). *State-Trait Anxiety Inventory STAI (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Stone, A., Shiffman, S., Atienza, A., & Nebeling, L. (2007). *The Science of Real-Time Data Capture*. Oxford University Press.

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, *22*(8), 1457–1470.

Trost, W., Trevor, C., Fernandez, N., Steiner, F., & Frühholz, S. (2024). Live music stimulates the affective brain and emotionally entrains listeners in real time. *Proceedings of the National Academy of Sciences - PNAS*, *121*(10), e2316306121–e2316306121.

Vuoskoski, J. K., Gatti, E., Spence, C., & Clarke, E. F. (2016). Do visual cues intensify the emotional responses evoked by musical performance? A psychophysiological investigation. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 179–188. https://doi.org/10.1037/pmu0000142

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Zhang, J. D., & Schubert, E. (2019). A Single Item Measure for Identifying Musician and Nonmusician Categories Based on Measures of Musical Sophistication. *Music Perception*, *36*(5), 457–467. https://doi.org/10.1525/mp.2019.36.5.457

APPENDIX A

List of GEMS Items (Zentner et al., 2008)

Happy, filled with wonder, moved, fascinated, overwhelmed, thrills, feeling of transcendence, mellowed, tender, in love, affectionate, sentimental, melancholic, dreamy, calm, nostalgic, soothed, serene, meditative, triumphant, energetic, strong, joyful, fiery, animated, bouncy, tense, amused, agitated, sad, irritated, tearful

APPENDIX B

Recording of researcher playing Liszt's Consolation No. 3 in Db Major provided to participants for recorded evaluations.

Permanently linked here.

https://durhamuniversity-my.sharepoint.com/:u:/g/personal/fvws88_durham_ac_uk/EVMmhKIFZelLolJLKy0gR84BHd2wDZGyPNzr_ntxYvbjAA?e=htTNPi

APPENDIX C

ANOVA outcomes for categories with insufficient data for pairwise t-test comparisons

Category	Environment	First order factor	ANOVA Outcome
Frequency of music listening	Live	Peacefulness	$F(3, 3) = 9.27$, $p = .05$*
	Live	Tension	$F(2, 1) = 9.26^{e+31}$, $p = <.01$*
Musical expertise	Live	Power	$F(2, 2) = 66.38$, $p = .02$*
	Recorded	Tension	$F(2, 12) = 5.85$, $p = .02$*
	Live	Tension	$F(2, 1) = 9.26^{e+31}$, $p = <.01$*

* = statistically significant. NB significance is unconfirmed by pairwise t-test comparisons as discussed in main text.

APPENDIX D

ANOVA outcomes: secondary hypotheses variables against first and second-order factors for live and recorded environments.

First & Second Order Factors	Environment	ANOVA Outcomes		T-test Outcomes
		Frequency of Music Listening	Level of Musicianship	Classical Music Listener: Y/N
Joyful activation	Recorded	F(3, 10) = .15, *p* = .93	F(2, 11) = .18, *p* = .84	t(11.15) = .28, *p* = .79
	Live	F(3, 3) = 1.16, *p* = .45	F(2, 4) = 2.12, *p* = 0.24	t(2.06) = -0.64, *p* = 0.59
Nostalgia	Recorded	F(3, 11) = 1.51, *p* = .27	F(2, 12) = .58, *p* = .57	t(11.17) = .21, *p* = .83
	Live	F(3, 3) = 1.58, *p* = .36	F(2, 4) = .45, *p* = .67	t(4.45) = -.27, *p* = .80
Peacefulness	Recorded	F(3, 11) = 2.24, *p* = .14	F(2, 12) = .87, *p* = .45	t(10.40) = .07, *p* = .94
	Live	**F(3, 3) = 9.29, *p* = .05**[!]	F(2, 4) = 0.73, *p* = .54	t(3.48) = -.24, *p* = .83
Power	Recorded	F(3, 9) = 1.31, *p* = .33	F(2, 4) = 1.58, *p* = .25	t(7.39) = -.85, *p* = .42
	Live	F(3, 1) = 22.13, *p* = .16	**F(2, 2) = 66.38, *p* = .02**[!]	t(2) = -1.23, *p* = .34
Sadness	Recorded	F(3, 10) = .41, *p* = .75	F(2, 11) = .33, *p* = .73	t(10.38) = 1.49, *p* = .17
	Live	F(3, 3) = 1.10, *p* = .47	F(2, 4) = 0.63, *p* = .58	t(3.88) = -.08, *p* = .94
Tenderness	Recorded	F(3, 11) = .23, *p* =0.87	F(2, 12) = 5.85,*p* = .02*	t(12.98) = 1.14, *p* = .28
	Live	F(3, 3) = 1.05, *p* = .49	F(2, 4) = 1.07, *p* = .42	t(2.90) = .37, *p* = .74
Tension	Recorded	F(3, 11) = .23, *p* = .87	**F(2, 12) = 5.85, *p* = .02**[!?]	t(2.51) = .12, *p* = .91
	Live	**F(2, 1) = 9.25^{e+31}, *p* = <.01**[!]	**F(2, 1) = 9.25^{e+31}, *p* = <.01**[!]	Null
Transcendence	Recorded	F(3, 11) = 2.04, *p* = .17	F(2, 12) = .32, *p* = .73	t(11.86) = -.63, *p* = .54
	Live	F(3, 3) = 6.81, *p* = .07	F(2, 4) = .02, *p* = .98	t(4.03) = -.35, *p* = .75
Wonder	Recorded	F(3, 11) = 1.22, *p* = .35	F(2, 12) = .27, *p* = 0.77	t(12.75) = -.52, *p* = 0.61
	Live	F(3, 3) = .28, *p* = .84	F(2, 4) = .51, *p* = .64	t(2.77) = -.35, *p* = .75
Sublimity	Recorded	F(3, 11) = 1.76, *p* = .21	F(2, 12) = .72, *p* = .51	t(10.44) = .11, *p* = .91
	Live	F(3, 3) = 3.10, *p* = .21	F(2, 4) = .15, *p* = .86	t(4.95) = -.14, *p* = .89
Vitality	Recorded	F(3, 11) = .30, *p* = 82	F(2, 12) = .86, *p* = .50	t(12.95) = -.19, *p* = .86
	Live	F(3, 3) = 2.52, *p* = .26	F(2, 4) = 3.67, *p* = .13	t(2.40) = -.44, *p* = .70
Unease	Recorded	F(3, 10) = .40, *p* = .76	F(2, 11) = .57, *p* = .58	t(10.36) = 1.48, *p* = .17
	Live	F(3, 3) = 3.10, *p* = .19	F(2, 4) = 1.26, *p* = .38	t(4.24) = .50, *p* = .64

* = statistically significant; [!] = significance unconfirmed by pairwise t-tests (discussed in main text); [!?] pairwise t-tests show no significant variation.

APPENDIX E\1

T-test outcomes: live versus recorded environments for levels of musical expertise

First & Second Order Factors	Musical Expertise, t-tests live versus recorded		
	High	Moderate	Low
Joyful activation	t(2) = -.14, *p* = .90	t(1.23) = .83, *p* = .54	t(7.30) = .24, *p* = .81
Nostalgia	t(1.13) = .32, *p* = .80	t(3.62) = 2.43, *p* = .08	t(4.10) = -.64, *p* = .55
Peacefulness	t(2.99) = 1.4, *p* = .26	t(1.21) = .79, *p* = .55	t(3.87) = 0.78, *p* = .49
Power	t(2.89) = -1.81, *p* = .17	Null	t(5) = -1.70, *p* = .15
Sadness	t(2.99) = .20, *p* = .85	t(1.85) = 1.06, *p* = .41	t(5.48) = .04, *p* = .97
Tenderness	t(1.23) = -.32, *p* = .80	t(1.38) = 1.95, *p* = .24	t(5.10) = -0.30, *p* = .77
Tension	t(1) = -2.94, *p* = .21	Null	Null
Transcendence	t(2.98) = -.01, *p* = 1.00	t(3.22) = -.68, *p* = .54	t(3.31) = -.21, *p* = .84
Wonder	t(1.15) = -.04, *p* = .97	t(1.08) = .53, *p* = .68	t(8.03) = .98, *p* = .35
Sublimity	t(1.15) = -.04, *p* = =.97	t(1.36) = 1.34, *p* = .36	t(3.66) = .13, *p* = .91
Vitality	t(2.12) = -.62, *p* = .60	t(1.22) = .69, *p* = .60	t(7.31) = -.39, *p* = .71
Unease	t(2.63) = -.53, *p* = .64	t(2.97) = 1.06, *p* = .37	t(3.99) = .08, *p* = .94

APPENDIX E\2

T-test outcomes: live versus recorded environments for frequency of music listening. Only 1 participant reported low frequency music listening so t-test comparisons gave null outcomes.

First & Second Order Factors	Frequency of music listening, t-tests live versus recorded			
	Very high	High	Moderate	Low
Joyful activation	t(2) = 1, *p* = .42	t(7) = -.56, *p* = .59	t(1.12) = .73, *p* = .59	Null
Nostalgia	t(2.62) = .04, *p* = .97	t(1.28) = -.078, *p* = .95	t(2.24) = 1.70, *p* = .22	Null
Peacefulness	t(2) = 1.89, *p* = .20	t(7.97) = 1.81, *p* = .11	t(2.34) = 1.13, *p* = .37	Null
Power	Null	t(6.36) = -3.38, *p* = .01*	Null	Null
Sadness	t(1) = 1, *p* = .50	t(6.27) = -.68, *p* = .52	t(1) = 1, *p* = .50	Null
Tenderness	t(2.45) = 1.41, *p* = .27	t(2.06) = -.85, *p* = .49	t(2.05) = .94, *p* = .44	Null
Tension	Null	t(5) = -2.52, *p* = .05*	Null	Null
Transcendence	t(2.89) = 2.81, *p* = .07	t(6.96) = -2.04, *p* = .08	t(2.69) = .58, *p* = .60	Null
Wonder	t(1.23) = .63, *p* = .62	t(5.02) = -2.00, *p* = .10	t(1.53) = .66, *p* = .60	Null
Sublimity	t(2.14) = 2.92, *p* = .09	t(2.46) = -.60, *p* = .60	t(2.11) = 1.42, *p* = .29	Null
Vitality	t(2.75) = .29, *p* = .79	t(7.82) = -2.32, *p* = .05*	t(2.35) = 1.05, *p* = .39	Null
Unease	t(1.33) = 1.30, *p* = .37	t(7.85) = -1.83, *p* = .11	t(1.27) = 1.29, *p* = .38	Null

APPENDIX E\3

T-test outcomes: live versus recorded environments for Y/N classical music listeners.

First & Second Order Factors	Y/N Classical music listeners, t-tests live versus recorded	
	Yes: are classical music listeners	No: are not classical music listeners
Joyful activation	t(3.16) = .69, *p* = .54	t(3.16) = .69, *p* = .54
Nostalgia	t(2.83) = .48, *p* = .66	t(2.83) = .48, *p* = .66
Peacefulness	t(7.59) = 2.05, *p* = .08	t(7.59) = 2.05, *p* = .08
Power	t(2.40) = -.46, *p* = .68	t(2.40) = -.46, *p* = .68
Sadness	t(3.23) = 1.05, *p* = .37	t(3.23) = 1.05, *p* = .37
Tenderness	T(3.17) = .42, *p* = .70	t(3.17) = .42, *p* = .70
Tension	t(4.09) = -.61, *p* = .57	t(4.09) = -.61, *p* = .57
Transcendence	t(6.92) = -.72, *p* = .49	t(6.92) = -.72, *p* = .49
Wonder	t(2.40) = .33, *p* = .77	t(2.40) = .33, *p* = .77
Sublimity	t(3.01) = .67, *p* = .55	t(3.01) = .67, *p* = .55
Vitality	t(2.57) = .18, *p* = .87	t(2.57) = .18, *p* = .87
Unease	t(2.92) = .41, *p* = .71	t(2.92) = .41, *p* = .71