Audio Engineering Society

Conference Paper

Presented at the AES 2026 International Conference on
Audio for Virtual and Augmented Reality and Immersive Games
2026 June 30 – July 3, Paris, France

This conference paper was selected based on a submitted abstract and 750-word precis that have been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This conference paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

The Impact of Audio Spatialisation Reproduction on the Neurophysiological Responses of Music Listeners

Laney Haywood[1], Sam Hobern[1], Isao Nambu[2], Nicolas Epain[3], Helena Daffern[1], and Gavin Kearney[1]

[1]*AudioLab, School of Physics, Engineering and Technology, University of York, York, UK*
[2]*Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan*
[3]*L-Acoustics, 13 Rue Levacher Cintrat, 91460, Marcoussis, France*

Correspondence should be addressed to Laney Haywood (laney.haywood@york.ac.uk)

ABSTRACT

Research into listeners' emotional experience of different audio formats heavily relies on subjective, self-report measures. However, little is known about neural and physiological responses. As such, this feasibility study utilised electroencephalography (EEG), Heart Rate (HR) and Galvanic Skin Response (GSR), to explore the objective neurophysiological impacts of mono, stereo and spatial audio formats, across different music genres. In a within-subjects design, participants listened to 27 randomised stimuli, each comprising of a 30 second music excerpt across the three audio formats. Results were not significant but trends did arise in the data. While mono formats were shown to elevate cognitive load and arousal, spatial audio elicited a decrease in physiological arousal, promoting a more relaxed state. However, the effects overall were very genre-dependent. Differences in physiological response between static and dynamic spatial reproduction of different music genres are discussed. While limited by the lack of subjective validation and sample size, this study highlights interesting relationships between audio format and the physiological responses of music listeners.

1 Introduction

The study of music-induced emotion has become very popular in psychological and cognitive research. Many studies have used self-report measures to investigate the impact of specific musical features on the affective and attentional responses of music listeners. However, the use of neuroimaging and physiological measures are becoming more frequently employed to validate these psychological experiences. This pilot study expands beyond the traditionally compared mono and stereo formats to additionally explore how spatial audio formats compare and influence these neural and physiological responses.

2 Background

Electroencephalography (EEG) and physiological measures such as heart rate (HR) and galvanic skin response

(GSR) have been used to investigate listeners' emotional and attentional responses to music [1]. Specifically, activity in the central nervous system can be assessed using EEG, which can provide insight into cognitive load and emotional state through Power Spectral Density (PSD) analysis. For example, alpha power can reflect emotional regulation and relaxation [2], whereas a decrease in delta power has been linked to heightened alertness and dissipation of cognitive fatigue [3]. Conversely, HR and GSR can reflect autonomic nervous system activity. Increased HR and GSR have been associated with sympathetic activation during exciting or arousing experiences [4], while reduced HR has been linked to parasympathetic activation during more relaxed states [5].

Studies that have explored the mechanisms behind music-induced emotion and related neurophysiological responses typically focus on specific music features such as tempo, key, and mood as the determining factors [6]. However, the musical stimuli presented has commonly been in mono, or stereo formats, despite the growing prevalence of spatial audio in music-based applications. Spatial audio reproduction introduces additional perceptual cues related to envelopment, externalisation, and spatial presence, which may influence listeners' emotional engagement and embodied responses to music [7, 8].

Research into the relationship between audio formats and emotional responses of listeners have so far examined mono vs spatial sound (over headphones) [7, 9] and stereo vs spatial sound (over loudspeakers) [10, 11]. Limited research has examined all three together (mono, stereo, spatial) [12]. In addition, studies that primarily focus on musical stimuli often consist of a single audio track, usually the entire length of a piece [7, 9, 11, 12]. Musical stimuli selection is inconsistent in the literature. Some studies have included multiple western genres [11, 12, 13], whereas others focus on one (usually classical or orchestral) [7, 8]. Most studies rely on custom self-reports to measure emotional response, while the few using heart rate and EEG to validate these findings have yielded mixed results. Spatial audio can elevate heart rate over time [9] and modulate heart rate variability (HRV) [10]. However, this change is not always significant across different audio format conditions [14]. Some research has shown that compared to stereo, spatial audio formats significantly enhance frontal alpha wave power, associated with deep relaxation [10]. In addition, changing dynamics within the spatial audio itself (i.e. direction of sound sources), has been found to alter alpha, beta, and gamma oscillations depending on the perceptual and emotional context of the stimulus [13]. Very few studies have reported significant results for GSR measures. There is, however, a visible trend indicating that music induces a steady state of relaxation, regardless of the reproduction format [10].

Research into the emotional and physiological effects of music in spatial audio often uses multiple genres as stimuli [11, 12, 13] however, the influence of genre is under-reported. Due to the distinct acoustic profiles, different dynamics, and spatial contexts of different genres, they are likely to have specific interactions with spatial playback formats [12, 15]. Understanding these relationships is important both for music cognition research and in determining whether spatial audio consistently enhances physiological and neurological engagement, regardless of musical context [14].

Therefore, this study tests the feasibility of utilising genre as a moderating variable in addition to evaluating the impact of audio reproduction format (mono, stereo, spatial) on neural, (EEG) and physiological responses (HR and GSR) associated with listeners' music-induced emotion in an immersive loudspeaker environment.

3 Methods

3.1 Design

A within-subject, repeated measures design was used. The independent variable was the audio spatialisation (mono, stereo, spatial). The dependent variable was participants' neurophysiological responses to each format. The effect of music genre was a moderating variable.

3.2 Participants

A convenience sample of 12 male participants were recruited aged 21 - 24 years. Eleven were Japanese and one British. Eleven participants reported 6-8 hours of sleep the night before, while one reported 3-5 hours. Two drank a caffeinated beverage before participating.

3.3 Materials

The experiment took place at the Acoustic and Vibration Engineering Centre at Nagaoka University of Technology using the 41.2-channel 3D loudspeaker system. This comprises Genelec 8320 Smart Active Monitors and 7380 subwoofers. Audio stimuli were presented using mono, stereo or 7.1.4 channels. The experiment interface was developed in Python and audio playback used the Genelec Loudspeaker Manager (GLM).

Playback levels were set to -42 dBFS. Due to time constraints, sound pressure level measurements were not formally recorded; instead, playback loudness was perceptually calibrated by a trained audio engineer to maintain consistent listening levels across conditions. Playback levels for mono stimuli were reduced by 3 dB, as these were perceived to be significantly louder than stereo and spatial conditions.

The neurophysiological measuring devices were the Biosemi ActiveTwo 64-channel EEG System for neuroimaging (sampled at 2048 Hz), a Shimmer GSR+ device to measure GSR (Sampled at 128Hz, except 4 participants sampled at 51.2Hz in error) and a Polar H10 electrocardiogram (ECG) chest strap to measure heart rate (sampled at 128Hz).

3.4 Musical Stimuli

Three musical pieces representing distinct genres (Jazz, Electronic, and Pop-Rock) were selected [16]. While the complex instrumentation of jazz and electronic offer greater features for spatial audio manipulation, pop-rock was selected as a universally appealing genre. To minimise the impact of different musical features, the tracks were selected based on similarities in overall mood (happy/upbeat) and allegro tempo (120–168 BPM). They were independently mixed into three audio spatialisation formats: mono, stereo, and 7.1.4 surround sound. Based on the specific metrics used in this analysis, 30 second (s) excerpts were deemed effective to elicit measured neurophysiological responses [17, 18, 19], particularly because the cumulative exposure for each song was 90 s. Furthermore, as no visual stimulation was included, shorter durations were prioritized to maintain participants' engagement and prevent listener fatigue. Each piece was segmented into three 30 s excerpts representing key structural elements of the song; verse, bridge, chorus, with vocal stems removed. This resulted in nine excerpts and 27 stimuli across the three spatialisation conditions.

To minimise repetition effects, the 27 stimuli were split into three experimental blocks so an excerpt could not be presented consecutively. Stimuli within each block were randomised. All participants experienced every block and heard all 27 stimuli across the experiment.

3.5 Procedure

After completing a demographic questionnaire, participants were seated in the centre of the loudspeaker array equipped with the EEG cap, Shimmer GSR+ device, and Polar H10 ECG sensor. As a dark room wasn't available [14], and a blindfold was invasive and interfered with the EEG electrodes, participants were asked to keep their eyes closed for the duration of the experiment to minimise visual distraction. The experiment first began with a 60 s silent baseline. Participants were then presented with all 27 of the 30 s stimuli, with a 15 s silent rest period between each playback. The experiment duration did not exceed 30 minutes.

3.6 Ethical Approval

Ethical approval for this study was obtained from the President of Nagaoka University of Technology (National University Corporation) under Experiment 4 of the project; "Research on Adaptive BCI Using Auditory Stimuli." Further approval was given from the Physical Sciences Ethics Committee at the University of York, ensuring compliance with their 'Code of practice and principles for good ethical governance'. Participants provided informed consent and had the right to withdraw at any time.

4 Data Analysis

Analysis of the ECG, GSR and EEG data was performed on the University of York's Viking High Performance Compute facility. Physiological data was analysed using Python including the NeuroKit2 [20], MNE [21], and MNE-ICALabel [22] packages.

The difference between the baseline and the stimulus was calculated for each of the 27 stimuli, denoted as the Δ. Results were averaged across genre and participants, calculating a mean change for genre and condition.

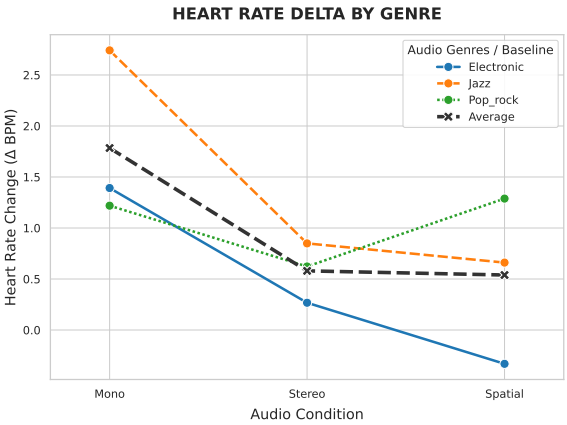


Fig. 1: Baseline-corrected heart rate (Δ BPM) by genre

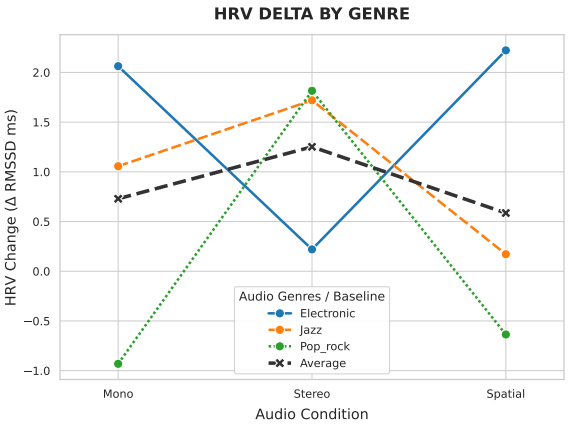


Fig. 2: Baseline-corrected heart rate variability (Δ RMSSD ms) by genre

4.1 ECG Analysis

The ECG data recorded by the Polar H10 devices was analysed using the NeuroKit2 Python library.

4.1.1 Heart Rate

HR in beats per minute (BPM) was calculated for each 30 s stimulus using the mean of the R-R interval in seconds where the R-R interval is the time difference between the peaks of the QRS complex [23]. The BPM was then expressed as the change from baseline and averaged across all participants and genre sections.

4.1.2 Heart Rate Variability

HRV was calculated to provide an indication of the response of the autonomic nervous system. HRV was expressed as the Root Mean Squared of Successive Differences (RMSSD) of the R-R interval in milliseconds.

4.2 GSR

The GSR data recorded by the Shimmer devices was also analysed using the NeuroKit2 library using the EDA tools. Due to a sensor malfunction, GSR data for one participant was not recorded correctly and thus excluded from the analysis.

The GSR data was separated into its Skin Conductance Level (SCL) Tonic and Skin Conductance Response (SCR) Phasic components. The mean tonic level in μS was extracted as well as the number of SCR peaks.

Due to the short stimuli, the Tonic was calculated for the entire 30 s duration and corrected against the 15 s baseline. Similarly, the SCR peaks are relative to the peaks which occurred during the baseline.

4.3 EEG

4.3.1 Pre-Processing

The EEG data was pre-processed using the MNE Python library. Bad channels identified during data capture were replaced through interpolation. Data was filtered between 1-100 Hz with an additional notch filter applied at 50 Hz to remove mains hum. Independent Component Analysis (ICA) was performed using the MNE-ICALabel library to remove undesirable components such as muscle artifacts and heart beats. Each epoch was baseline corrected using the 15 s rest prior to each stimuli onset.

4.3.2 Power Spectral Density

The analysis for the EEG data focused around the Power Spectral Density (PSD) averaged across all channels. The PSD was extracted in several distinct frequency bands using MNE: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-45 Hz). The mean of stimuli sections were averaged in order to provide a more stable indication of the effect of the genre.

5 Results

5.1 ECG Results

The Δ BPM shown in Figure 1 indicates a clear trend. Compared to the baseline, the mono condition exhibited the highest increase in BPM with an average increase of around 1.75 BPM. The change to the stereo

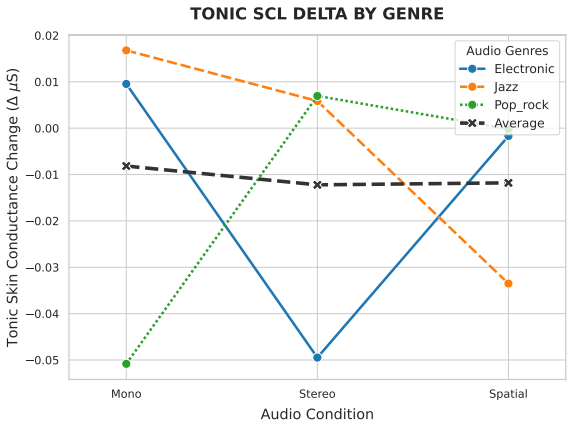


Fig. 3: Baseline-corrected SCL Tonic (Δ μS) by genre

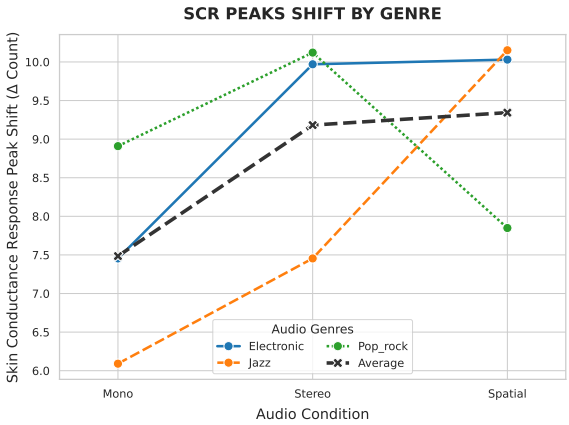


Fig. 4: Baseline-corrected SCR Peaks (Δ Count) by genre

condition results in a drop of around 0.6 BPM on average. The spatial condition showed the most genre dependent behaviour. Jazz closely following the global average whilst Pop-Rock resulted in an increase that peaked above the mono BPM change. Conversely, the Electronic genre resulted in a significant drop in BPM that reached below the baseline recording.

The Δ HRV in Figure 2 shows the mono and spatial conditions exhibit similar changes compared to baseline across all genres. However, genre does have an effect in the stereo condition. In both the Pop-Rock and Jazz genres, the stereo condition resulted in an increase of around 1.5 ms in HRV compared the mono or spatial conditions which resulted in either a minimal increase, or a notable decrease in HRV. The Electronic genre

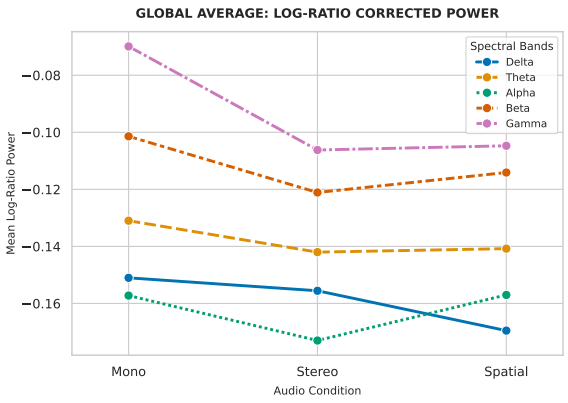


Fig. 5: Baseline-corrected log-ratio power (dB) for EEG frequency bands

showed the opposite effect where the stereo condition resulted in around a 0.3 ms increase from the baseline; both the mono and spatial conditions similarly increased at least 2 ms from the baseline.

5.2 GSR Results

Figure 3, shows the SCL Tonic is highly genre specific. The average across genres is similar for all three conditions. However, very different trends are observed for each genre. For the Electronic genre, both the mono and spatial conditions show minimal increases from the baseline. In contrast, the stereo condition saw a decrease in skin conductance which was five times greater in magnitude than the increase observed in the mono and spatial conditions.

Figure 4 indicates the change in the number of peaks between the baseline and stimulus. The average across all genres indicates a clear trend. From mono to stereo, an increase of approximately 2.5 peaks is observed. From stereo to spatial, the peak shift increases but only minimally. However, based upon the results, it appears that between stereo and spatial, the change in peaks is highly genre dependent with Pop-Rock seeing a large decline, Jazz increasing by a large margin, and Electronic seeing very little change.

5.3 EEG Results

From the log-ratio power between the baseline and stimulus period expressed in Figure 5, a slight trend is visible. Across all frequency bands, the mono condition

exhibits the highest power. There is a power reduction across all frequency bands between the mono and stereo conditions. The spatial condition exhibits similar power as the stereo condition across all bands with the only notable exceptions being the delta band which sees a greater reduction, and the alpha band which sees a greater increase compared to the stereo condition.

6 Discussion

Due to the small sample size, significance testing was not conducted. Consequently, the results are not conclusive but demonstrate that the method presented is feasible and the findings provide an insight into the trends associated with the neurophysiological impacts of spatial audio formats.

The spatial audio condition generally resulted in the lowest average HR (ΔBPM), just below stereo; compared to mono, which spiked arousal. This is contrary to previous findings [10, 9] and could be due to a methodological error. Spatial audio mixes were downmixed into their respective stereo and mono conditions. The perceived quality of the downmix depends heavily on how the instruments were spaced in the original mix [24]. An overloaded bus could have occurred in the mono condition increasing the perceived loudness of the mono stimuli, which has been associated with increased HR [25]. The decrease in arousal during the spatial condition is consistent with HRV and SCR peak data. While the stereo condition caused sudden jumps in ΔHRV, this remained steady for both mono and spatial audio. Similarly, SCR peaks were slightly more frequent during the stereo and spatial conditions compared to mono. The absence of a sharp upward trend from stereo to spatial suggests that arousal did not increase with spatial audio but instead, relaxed the sympathetic nervous system. Overall, these results suggest mono formats increase physiological arousal. This is likely due to the increased cognitive processing load required when multiple sound sources are produced from a single point source [12]. In contrast, spatial audio presents a more natural soundscape that requires less mental processing [26], allowing the sympathetic nervous system to relax.

Unexpectedly, the mono condition exhibited the highest overall log-ratio power across all measured spectral bands. This again suggests that the extreme spatial masking of temporally coincident sources in the mono mix may demand higher overall cognitive processing, or trigger a stronger orienting response, increasing brain activity. While other spectral bands were either similar or displayed a slight increase in power when comparing stereo to spatial, alpha and delta bands deviated from this trend. Alpha band power saw a large decrease from mono to stereo as observed across other frequency bands. However, from stereo to spatial the alpha band displayed a more prominent upward shift in power. This finding is consistent with [10] who found an increased shift in alpha power from stereo to spatial, suggesting a relaxation effect. In contrast, the lowest delta power was observed in the spatial condition; even though closed-eyes promote drowsiness and the delta activity typically associated with sleep [27]. Compared to the elevated delta power in mono, the decrease in the spatial condition counteracts this drift towards sleep, suggesting that the listener remains consciously alert and engaged with the environment. This is likely because spatial audio aims to mimic how we perceive sound in the real-world [3]. Much like the physiological data, the EEG results indicate that while the unnatural mono format demands high compensatory processing, the lifelike presentation of spatial audio facilitates a state of relaxed emotional regulation and conscious mental alertness.

The results suggest a strong interaction between genre and audio format. When comparing stereo to spatial, Jazz and Pop-Rock display opposite trends. For Jazz, spatial audio elicits a decrease in both average HR and tonic SCL. Conversely, Pop-Rock shows the inverse effect, indicating the largest increase in sympathetic arousal with the highest HR and lowest HRV in the spatial condition. However, this was not reflected in the GSR metrics. Although, sharp phasic SCR peaks were observed for Jazz in the spatial condition. These results might reflect that the compressed mix for Pop-Rock drives a more sustained cardiovascular arousal response [9], whilst Jazz's distinct instrumentation and wider dynamic range allow the spatial format to highlight localised acoustic elements [15, 12]. This may explain why Jazz produced more SCR peaks. Arousal increased at specific moments in the Jazz stimuli, rather than the continuous heightened arousal shown for Pop-Rock. Conversely, the physiological state for Electronic was highly responsive during the spatial condition. Average HR decreases as audio format increases. When paired with the dramatic increase in HRV for the spatial condition, this suggests parasympathetic activation which induces a relaxed state [5]. Interestingly, the

simultaneous increase for tonic SCL and phasic SCR peaks during the spatial condition suggests both sustained sympathetic arousal and heightened reactivity to specific moments within the stimuli [4]. The drastic decrease in SCL for the stereo condition, however, suggests that GSR only increased during peak moments in the stimuli, rather than a steady increase throughout. Interestingly, only Electronic had a moving sound source. Unlike the dramatic shifts in Phasic SCR peak counts between stereo and spatial for Pop-Rock and Jazz, counts remained consistently high for both formats in Electronic, where the sound source moved between the loudspeakers. Dynamic spatial audio can increase physiological arousal in alternative audio-visual media contexts [28]. This could explain the differences in physiological response between genres as Jazz and Pop-Rock were spatialized statically. Overall, the data suggests that the relationship between spatial audio and physiological arousal is not linear and depends on the dynamic range and spatial mix of the genre.

6.1 Limitations and Future Directions

The main limitation was the sample size and demographic, with a convenience sample of only 12 male participants. This limits the statistical power and generalisability of the results. Technical limitations regarding the stimuli design also impacted the study. The mono and stereo downmix of the musical stimuli were conducted under time constraints which raised concerns for the perceived loudness and standardised quality across all three conditions [24]. In future, sound pressure levels (SPL) should be measured during the trials to correct for these variations. The stimuli were kept short to minimise listener fatigue and were restricted to one track per musical genre. This brief exposure and lack of same-stimulus repetition may limit the depth of neurophysiological responses compared to studies that used longer stimuli [7, 9]. The limited stimuli prevents the validation of consistent results across different musical styles and in future the protocol should include multiple, longer tracks per genre [7, 9]. The laboratory was notably cold during testing which may have impacted physiological measures, particularly GSR [17]. However, given that GSR is more sensitive to warmer temperatures, the cold environment may have suppressed thermoregulatory activity, making measured GSR response more representative of specific effects of the audio stimuli.

Emotional engagement, attention and experience of the different audio formats can only be inferred from the neurophysiological measures collected. A language barrier between the researcher and participants and the large number of stimuli with short rest periods prevented the collection of self-report emotion data. In the future, time should be allocated to evaluate emotional response [12]. Incorporating the objective neurophysiological metrics with validated, subjective self-report measures similar to [9, 10] will provide a more robust and comprehensive understanding of how listeners' emotionally engage with different audio formats in music contexts.

To conclude, this study presents a method for assessing the neurophysiological responses to music of different genres in different audio formats. Whilst the sample size is too small for statistical power or generalisability, trends appear in the data that suggest while mono formats elevate cognitive load and arousal, spatial audio facilitates a more relaxed state by mimicking real-world acoustics. However, effects are genre-dependent. Future work incorporating perceived emotional response to this method will be valuable in understanding the impact of audio formats on music listeners.

7 Acknowledgements

This work was supported by CoSTAR Live Lab, funded by the Arts and Humanities Research Council, grant reference AH/Y001079/1 and by the Engineering and Physical Sciences Research Council, grant reference EP/W524657/1. This research was conducted as part of a student exchange between the University of York and Nagaoka University of Technology with funding from the JSPS Inter-University Exchange Project. Artist permission (Everything After Midnight, J P Lantieri and Ghostly Beard) was granted for audio materials and accessed via the Cambridge Music Technology "Mixing Secrets" library[16].

References

[1] Hodges, D. A., "Psychophysiological measures," in P. N. Juslin and J. A. Sloboda, editors, *Handbook of Music and Emotion: Theory, Research, Applications*, pp. 279–312, Oxford University Press, Oxford, UK, 2011.

[2] Schaefer, R. S., Vlek, R. J., and Desain, P., "Music Perception and Imagery in EEG: Alpha Band Effects of Task and Stimulus," 82(3), pp. 254–259, 2011, doi:10.1016/j.ijpsycho.2011.09.007.

[3] Harmony, T., "The functional significance of delta oscillations in cognitive processing," *Frontiers in Integrative Neuroscience*, 7, p. 83, 2013, doi:10.3389/fnint.2013.00083.

[4] Benedek, M. and Kaernbach, C., "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, 190(1), pp. 80–91, 2010, doi:10.1016/j.jneumeth.2010.04.028.

[5] Benarroch, E. E., "Parasympathetic System; Overview," in M. J. Aminoff and R. B. Daroff, editors, *Encyclopedia of the Neurological Sciences*, pp. 805–808, Academic Press, Oxford, 2 edition, 2014, doi:10.1016/B978-0-12-385157-4.00508-X.

[6] Schaefer, H.-E., "Music-Evoked Emotions—Current Studies," *Frontiers in Neuroscience*, 11, p. 600, 2017, doi:10.3389/fnins.2017.00600.

[7] Alam, T. A. and Dibben, N., "A Comparison of Presence and Emotion Between Immersive Virtual Reality and Desktop Displays for Musical Multimedia," in *Future Directions of Music Cognition*, pp. 97–101, Virtual Conference, 2021, doi:10.18061/FDMC.2021.0017.

[8] Hahn, E., "Musical emotions evoked by 3D audio," in *Audio Engineering Society Conference: 2018 AES International Conference on Spatial Reproduction-Aesthetics and Science*, Audio Engineering Society, 2018.

[9] Warp, R., Zhu, M., Kiprijanovska, I., Wiesler, J., Stafford, S., and Mavridou, I., "Moved By Sound: How head-tracked spatial audio affects autonomic emotional state and immersion-driven auditory orienting response in VR Environments," in *152nd Audio Engineering Society Convention*, The Hague, Netherlands, 2022.

[10] Hyodo, Y., Sugai, C., Suzuki, J., Takahashi, M., Koizumi, M., Tomura, A., Mitsufuji, Y., and Komoriya, Y., "Psychophysiological Effect of Immersive Spatial Audio Experience Enhanced Using Sound Field Synthesis," in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 1–7, Nara, Japan, 2021, doi:10.1109/ACII52823.2021.9597435.

[11] Lepa, S., Weinzierl, S., Maempel, H.-J., and Ungeheuer, E., "Emotional Impact of Different Forms of Spatialization in Everyday Mediatized Music Listening: Placebo or Technology Effects?" in *136th Audio Engineering Society Convention*, Berlin, Germany, 2014.

[12] Moiragias, G. and Mourjopoulos, J., "A listener preference model for spatial sound reproduction, incorporating affective response," *PLOS ONE*, 18(6), 2023, doi:10.1371/journal.pone.0285135.

[13] Filippi, E. D., Schmele, T., Nandi, A., Torres, A. G., and Pereda-Baños, A., "Emotional Impact of Source Localization in Music Using Machine Learning and EEG: A Proof-of-Concept Study," *TechRxiv*, 2022, doi:10.36227/techrxiv.21789866.v1.

[14] Ooishi, Y., Kobayashi, M., Kashino, M., and Ueno, K., "Presence of Three-Dimensional Sound Field Facilitates Listeners' Mood, Felt Emotion, and Respiration Rate When Listening to Music," 12, 2021, doi:10.3389/fpsyg.2021.650777.

[15] Conetta, R., Brookes, J., Rumsey, F., and Zieliński, S., "Spatial Audio Quality Perception (Part 1): Impact of Commonly Encountered Processes," *Journal of the Audio Engineering Society*, 62(12), pp. 831–846, 2014, doi:10.17743/jaes.2014.0048.

[16] Cambridge Music Technology, "Mixing Secrets For The Small Studio – Additional Resources," 2026, accessed: 2026-05-21.

[17] Boucsein, W., *Physiology of the Electrodermal System*, chapter 1.3, pp. 17–47, Springer Science+Business Media, 2012, doi:10.1007/978-1-4614-1126-0.

[18] Shaffer, F. and Ginsberg, J. P., "An overview of heart rate variability metrics and norms," *Frontiers in public health*, 5, p. 290215, 2017, doi:10.3389/fpubh.2017.00258.

[19] Gómez-Tapia, C., Bozic, B., and Longo, L., "On the minimal amount of EEG data required for learning distinctive human features for task-dependent biometric applications," *Frontiers in neuroinformatics*, 16, p. 844667, 2022, doi:10.3389/fninf.2022.844667.

[20] Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., and Chen, S. H. A., "NeuroKit2: A Python toolbox for neurophysiological signal processing," *Behavior Research Methods*, 53(4), pp. 1689–1696, 2021, doi:10.3758/s13428-020-01516-y.

[21] Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. S., "MEG and EEG Data Analysis with MNE-Python," *Frontiers in Neuroscience*, 7(267), pp. 1–13, 2013, doi:10.3389/fnins.2013.00267.

[22] Li, A., Feitelberg, J., Saini, A. P., Höchenberger, R., and Scheltienne, M., "MNE-ICALabel: Automatically annotating ICA components with ICLabel in Python," *Journal of Open Source Software*, 7(76), p. 4484, 2022, doi:10.21105/joss.04484.

[23] Bagliani, G., De Ponti, R., Gianni, C., and Padeletti, L., "The QRS Complex: Normal Activation of the Ventricles," 9(3), pp. 453–460, 2017, doi:10.1016/j.ccep.2017.05.005.

[24] Zieliński, S. K., Rumsey, F., and Bech, S., "Effects of downmix algorithms on quality of surround sound," *AES: Journal of the Audio Engineering Society*, 51(9), pp. 780–798, 2003.

[25] Shoushtarian, M., Weder, S., Innes-Brown, H., and McKay, C. M., "Assessing hearing by measuring heartbeat: The effect of sound level," *PLOS ONE*, 14(2), p. e0212940, 2019, doi:10.1371/journal.pone.0212940.

[26] Wilkie, H. and Harrison, P., "Reverberation time and musical emotion in recorded music listening," *Music Perception: An Interdisciplinary Journal*, 42(4), pp. 329–344, 2025.

[27] Long, S., Ding, R., Wang, J., Yu, Y., Lu, J., and Yao, D., "Sleep quality and electroencephalogram delta power," *Frontiers in Neuroscience*, 15, p. 803507, 2021, doi:10.3389/fnins.2021.803507.

[28] Larsen, J. J. and Pilgaard, M., *The Effect of Spatial Audio on Immersion, Presence, and Physiological Response in Games*, Master's thesis, Aalborg University, Department of Architecture and Media Technology, Aalborg, Denmark, 2015.