BMC Research Notes

DATA NOTE

Open Access

Data on neurophysiological and psychological responses to audio immersive experience in stereo and 3D audio formats

Norberto E. Naal-Ruiz[1,2*], Hyunkook Lee[2], Luz M. Alonso-Valerdi[1] and David I. Ibarra-Zarate[1]

Abstract

Objectives This dataset presents demographic, psychological, auditory and neurophysiological information of 31 volunteers, who participated in an experiment measuring the auditory immersive experience in two audio formats: stereophonic downmix and three-dimensional audio. This dataset could help understand the objectiveness (based on the nervous system response) behind the subjectiveness of immersion brought about by the audio format (based on the listener evaluation). The final objective of this dataset is to study the psychological and neurophysiological responses of immersive attributes in auditory events in future studies.

Data description Participants were randomly assigned to an audio format and attended two listening sessions, where soundscapes of a concert hall and urban park were randomly played. Electroencephalography, electrodermal activity, and blood volume pressure data were collected during auditory stimulation, and psychological responses to the auditory immersive experience were collected at the end of the session.

Clinical trial number Not applicable.

Keywords Immersion, Neuroacoustics, 3D Audio, Extended reality, Electrophysiology

Objective

The concept of immersion has been a compelling topic in the field of extended reality without reaching a consensus among researchers. Immersive audio has been used interchangeably with three-dimensional (3D) audio, assuming that a high-fidelity spatial audio system causes an immersive experience. However, this assumption lacks evidence [1].

Nervous system responses could help researchers objectively understand the attributes of the immersive experience [2]. In electroencephalography (EEG) data, research on the sense of presence has been associated with neural oscillations within clinical bands such as theta and beta [3, 4]. For blood volume pressure (BVP), heart rate variability, and its frequency composition have been related to involvement [5, 6]. For electrodermal activity (EDA), descriptive features in skin conductance response and levels can be extracted to monitor behavior in emotional activation [5, 7].

Therefore, understanding auditory immersive experiences through behavioral questionnaires and neurophysiological responses motivated the data collection, enabling the extraction of neural features and the study of their association with immersive attributes. The audio

*Correspondence:
Norberto E. Naal-Ruiz
a01281789@tec.mx
[1]Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias, Ave. Eugenio Garza Sada 2501, 64849 Monterrey, N.L, México
[2]Applied Psychoacoustics Lab (APL), University of Huddersfield, Huddersfield HD1 3DH, UK

recordings used are also included. Firstly, participants voluntarily registered for the experiment. Secondly, participants were selected based on exclusion criteria on auditory and neurological health. Thirdly, participants answered the Immersive Tendencies Questionnaire (ITQ) [8] to be assigned to stereophonic downmix (stereo) or 3D audio group. Lastly, participants attended two listening sessions, where the soundscape of a concert hall and an urban park were randomly played. During each listening session, EEG, BVP and EDA recordings of resting state and auditory stimulation were collected while participants were seated in the middle of a critical listening room with their eyes closed. After each soundscape, participants answered an immersive experience questionnaire.

Previous work using this dataset showed statistical significance between the mean heart rate variability and the immersive attribute of involvement [4], indicating that the volume of data may be suitable for further signal and statistical analyses.

Data description

Data corresponds to 31 subjects (stereo: 10 males, five females, mean ITQ=49.2. 3D: 10 males, six females, mean ITQ=48.3) aged 18–36 years, all university students. Participants were excluded if they: (1) were outside the age range to minimize the likelihood of presenting hearing loss [9], (2) reported auditory and neurological problems, and (3) reported never having been in a concert hall or an urban park. After registration, participants were randomly and blindly assigned to an audio format group (stereo or 3D) based on ITQ scores. Table 1 provides an overview of data files and data sets available in the repository.

DataSpecs excel file (Data file 1) contains six worksheets: (1) Demographic data: ID, audio format group, ITQ score, Sex, Age, Highest education level, Country of origin, and Handedness; (2) Biofiles: ID, audio format group, listening condition, biosignal, sample rate, number of events and total length in seconds; (3) ITQ: ID, audio format group, and responses to items of ITQ; (4) Responses: ID, listening condition, audio format group, responses to the Self-Assessment Manikin test [10], igroup Presence Questionnaire [11, 12], and system performance questionnaire designed by the authors based on the Wheel of Sound recommendation [13, 14]; (5) Items: questionnaires, item codes if available by the authors of the assessments, the question or statement of the item, type of scale, and minimum and maximum values; and (6) EEGchannels: electrode positions in Theta/Phi coordinates.

LoudspeakerInformation file (Data file 2) explains the alignment and positions of loudspeakers for stereo, PCMA-3D, and ESMA-3D audio playback. PCMA-3D and ESMA-3D are horizontally spaced and vertically coincident microphone techniques dedicated to music and sound field recordings, respectively [15].

RawData folder (Data set 1) has individual subfolders of participants with EEG, EDA, and BVP recordings in three conditions: (1) resting state (Bl), (2) concert hall (Music), and (3) urban park (Park) soundscapes. Neurophysiological data were collected with OpenViBE v3.3.1 and the following devices: (1) EEG: 24-channel electrode cap following the 10/20 international electrode system connected to a Smarting mbrain train amplifier at 500 Hz sample rate; (2) BVP and (3) EDA signals: Empatica E4 wristband connected via Lab Streaming Layer code [16] with 64 Hz and 4 Hz sample rates, respectively. Condition 8 code in GDF files indicates the start of audio playback and can be ignored for Bl files. Code 33031 refers to one second before the end of the experiment and is not indicated in Music and Park files since the recording was stopped a few seconds after the sound finished to avoid audio interruption.

LatencyAdjustment folder (Data set 2) mirrors Raw Data organization. Codes latencies were adjusted due to latencies caused by sound traveling through the air and digital delay. This process is explained in Data file 2.

AudioFiles folder (Data set 3) contains two subfolders: (1) Music: 2-minute-long WAV audio files of concert hall recordings [17]; and (2) Park: 2-minute-long WAV audio files of urban park recordings [18]. All files were

Table 1 Overview of data files and data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1	DataSpecs	Excel file (.xlsx)	Figshare (https://doi.org/10.6084/m9.figshare.25421464.v3) [19]
Data file 2	LoudspeakerInformation	PDF file (.pdf)	Figshare (https://doi.org/10.6084/m9.figshare.25421464.v3) [19]
Data set 1	RawData	General Data Format (.gdf)	Figshare (https://doi.org/10.6084/m9.figshare.25421464.v3) [19]
Data set 2	LatencyAdjustment	Metadata (.set) and signal data (.fdt)	Figshare (https://doi.org/10.6084/m9.figshare.25421464.v3) [19]
Data set 3	AudioFiles	Waveform Audio Format files (.wav)	Figshare (https://doi.org/10.6084/m9.figshare.25421464.v3) [19]

generated at 48k Hz sample rate and 24-bit depth. These files are organized for the three loudspeaker configurations detailed in Data file 2. Max 8 software was used for audio playback.

Limitations

- The data of three participants in specific soundscapes is missing because they did not attend the last session: Music – ID 20 and 40, and Park - ID 25.
- In some neurophysiological recordings, signal discontinuities and noisy segments may be present due to momentary device disconnection or subject movements, respectively, modifying the actual length of the recording as indicated in Data file 1.
- Stereophonic versions were created from the downmix process of PCMA-3D and ESMA-3D recordings. As a result, future findings may be applicable only to these specific playback conditions.

Abbreviations

3D	Three dimensional
Stereo	Stereophonic downmix
EEG	Electroencephalography
BVP	Blood volume pressure
EDA	Electrodermal activity
ITQ	Immersive tendencies questionnaire
BI	Resting state
Music	Concert hall
Park	Urban park

Acknowledgements
Thanks to members of the NeuroTechs Research group, Acoustic Innovation Research group, and Applied Psychoacoustics Laboratory for assisting in data management and audio calibration.

Author contributions
N.E.N.R collected, revised and organized the data presented in this work. He additionally produced the table, and wrote the original draft of the manuscript. H.L., L.M.A.V. and D.I.I.Z supervised, provided funds for the research, reviewed and edited the final version of the manuscript. All authors contributed to the conceptualization and methodology of data collection.

Funding
This work was supported by the Tecnologico de Monterrey, CONAHCYT and The University of Huddersfield.

Data availability
The data described in this Data note can be freely and openly accessed on Figshare under https://doi.org/10.6084/m9.figshare.25421464.v3. Please see Table 1 and reference [12] for details and links to the data.

Declarations

Ethics approval and consent to participate
The Ethics Committee of the School of Computing and Engineering at the University of Huddersfield granted approval for conducting the experiment and sharing the data (Reference: SCE_SREIC_2022/10/05_M_5.1). Informed consent was obtained from all volunteers before experimentation by signing a consent form. Personal information such as name, email and phone number of participants are not included in the data. All procedures were performed in accordance with the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 19 September 2024 / Accepted: 18 November 2024
Published online: 02 December 2024

References
1. Agrawal S, Bech S, de Moor K, Forchhammer S. Influence of changes in Audio Spatialization on Immersion in Audiovisual experiences. AES J Audio Eng Soc. 2022;70:810–23. https://doi.org/10.17743/jaes.2022.0034.
2. Zhang C. The why, what, and how of immersive experience. IEEE Access. 2020;8:90878–88. https://doi.org/10.1109/ACCESS.2020.2993646.
3. Zhang S, Feng X, Shen Y. Quantifying Auditory Presence using Electroencephalography. Appl Sci. 2021;11:10461. https://doi.org/10.3390/app112110461.
4. Naal-Ruiz NE, Lee H, Gonzalez A, Gonzalez R, Duville MM, Alonso-Valerdi LM et al. Towards the neurophysiological characterization of immersive attributes in auditory events. In: 2023 AES International Conference on Spatial and Immersive Audio. 2023. Huddersfield, UK. Audio Engineering Society. p. 9. https://www.aes.org/e-lib/browse.cfm?elib=22167
5. Kobayashi M, Ueno K, Ise S. The effects of Spatialized sounds on the sense of Presence in Auditory virtual environments: a psychological and physiological study. Presence Teleoperators Virtual Environ. 2015;24:163–74. https://doi.org/10.1162/PRES_a_00226.
6. Geronazzo M, Rosenkvist A, Eriksen DS, Markmann-Hansen CK, Køhlert J, Valimaa M et al. Creating an Audio Story with Interactive Binaural Rendering in Virtual Reality. Wirel Commun Mob Comput. 2019;2019. https://doi.org/10.1155/2019/1463204
7. Cuadrado F, Lopez-Cobo I, Mateos-Blanco T, Tajadura-Jiménez A. Arousing the sound: a field study on the emotional impact on children of arousing Sound Design and 3D Audio spatialization in an Audio Story. Front Psychol. 2020;11:1–19. https://doi.org/10.3389/fpsyg.2020.00737.
8. Witmer BG, Singer MJ. Measuring Presence in virtual environments: a Presence Questionnaire. Presence Teleoperators Virtual Environ. 1998;7:225–40. https://doi.org/10.1162/105474698565686.
9. Jin IK, Lee D, Jeong Y, Seo YJ, Kong TH, Suh MJ, et al. Trends in distributions of hearing threshold levels by ages: a comparison of the ISO 7029 and newly available country-specific data. J Audiol Otol. 2024;28:1–9. https://doi.org/10.7874/JAO.2023.00626.
10. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry. 1994;25:49–59. https://doi.org/10.1016/0005-7916(94)90063-9.
11. Oramus MT, Neubauer MP. Comparing perception of Presence between Binaural and Stereo Reproduction. 2022. https://aes2.org/publications/elibrary-page/?id=21872
12. Schubert T, Friedmann F, Regenbrecht H. The experience of Presence: factor Analytic insights. Presence Teleoperators Virtual Environ. 2001;10:266–81. https://doi.org/10.1162/105474601300343603.
13. Pedersen TH, Zacharov N. The development of a sound wheel for reproduced sound. 138th Audio eng soc Conv 2015. 2015;2:765–77. https://aes2.org/publications/elibrary-page/?id=17734
14. Pedersen TH. Perceptual characteristics of audio. 2015.
15. Lee H. Multichannel 3D microphone arrays: a review. AES J Audio Eng Soc. 2021;69:5–26. https://doi.org/10.17743/JAES.2020.0069.
16. Carral HL, Empatica, -E4-LSL. 2019. GitHub. https://github.com/HectorCarral/Empatica-E4-LSL
17. Lee H. 3D microphone array Recording comparison (3D-MARCo). Zenodo. 2019. https://doi.org/10.5281/zenodo.3474285.
18. Lee H. ESMA-3D immersive soundscape recordings. Zenodo. 2019. https://doi.org/10.5281/zenodo.3549050.

19. Naal-Ruiz NE, Lee H, Alonso-Valerdi LM, Ibarra-Zarate DI. Data on EEG, EDA, BVP, psychological responses and audio files used for the study of 3D Audio Immersive Experience. 2024. Figshare. https://doi.org/10.6084/m9.figshare.25421464.v3

Publisher's note