INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2025, VOL. 71, NO. 3, PP. 1-6
Manuscript received May 21, 2025; revised July 2025. doi: 10.24425/ijet.2025.153613

The angular position of a sound source affects the perception of scariness in a circular loudspeaker array

Antonina Stefanowska, and Sławomir K. Zieliński

***Abstract*—The quality of the user experience of virtual reality systems is contingent upon the emotional impact they provide. In this context, spatial audio technology emerges as a crucial area of focus. The existing literature provides evidence that sound source localization may influence the intensity of perceived emotions. However, the exact relationship between the localization of sound sources and the variation in listeners' emotional responses remains unclear. This study examines the impact of the angular position of selected audio sources on listeners' perception of a single emotion, namely scariness. The listening tests were conducted in an acoustically treated room equipped with a circular array of eight loudspeakers distributed in the horizontal plane. A total of 36 listeners participated in the experiment. The level of scariness perceived by the listeners was assessed subjectively using self-reports. According to the obtained results, the sound sources positioned outside of the listener's field of view were perceived as scarier. While the observed effect was statistically significant, its magnitude was relatively small. The potential future expansion of the study and its impact on the design of virtual reality systems is discussed further.**

***Keywords*—spatial sound; emotion; scariness; virtual reality systems**

INTRODUCTION

IN the entertainment industry, there is a particular focus on creating immersive experiences for listeners, particularly in the context of immersive multimedia systems such as virtual reality or streaming services. This has led to a rise in the use of audio technology and spatial sound reproduction devices in households. In light of this trend, sound design and playback device configuration that emphasize intensified experiences may prove to be a valuable area of study. Existing research indicates that there is a correlation between sound location and the emotions perceived and/or experienced by listeners. However, to the best of the authors' knowledge, there is a lack of literature providing directions on how to maximize the emotional impact of sound in multimedia systems.

Sight is regarded as the most developed sense among humans, constituting the primary source of the majority of sensory information received by the human brain [1]. It is therefore unsurprising that our primal survival instincts encourage us to maintain any potential threat in our field of view. The field of view for humans, including peripheral vision, has an approximate range of $\langle -60°, 60° \rangle$ around the fixation point [2]. Given the relationship between the visibility of potential threats and the feeling of fear, it can be hypothesized that listening to a scary sound coming from outside this angular range increases the perceived scariness of the sound.

The objective of this research is to expand the existing knowledge on this topic by conducting experiments performed on eight angular variants of sound source position and evaluating its impact on listeners' perception of scariness. The initial hypothesis is that the ratings of scariness for sound sources positioned outside the average human field of vision will be higher than the ratings of sound sources located within it. To date, the only studies that examined the influence of different angles of sound incidence on the listeners' emotions, including diagonal positions, were conducted via headphones utilizing a binaural format [3], [4]. However, it is known that sound reproduction systems using binaural techniques exhibit limited reliability in terms of sound externalization [5] and give rise to front-back localization errors [6]. In contrast, experiments in this study were performed in a laboratory setting utilizing an array of eight loudspeakers equidistantly arranged around the listener. The results of this study may prove to be useful to sound engineers, computer game developers, or designers of virtual reality systems who intend to create engaging experiences.

This paper is structured as follows: the subsequent section provides a summary of existing work within the topic of this research. The following section outlines the employed materials and the methodology of the testing procedure. The next chapter presents the analysis of the results obtained during the tests. The following section discusses the findings, with a detailed comparison to the existing studies on the subject. Finally, the last section concludes this paper and explores potential future directions.

I. RELATED WORK

There is an increasing body of evidence that suggests a correlation between spatial sound and the emotional responses of listeners. For instance, spatial sound, reproduced either using loudspeakers or headphones, evokes more intense emotions, with a greater positive valence and/or increased arousal, compared to mono [7]-[10]. Furthermore, spatial audio evokes stronger emotional responses in listeners than traditional two-channel stereo sound [11]-[13]. Nevertheless, some studies yield contradictory findings, indicating that spatial audio offers minimal or no discernible advantage over stereo in terms of evoking emotional responses [14]-[16]. Moreover, spatial sound reproduced over loudspeakers evokes stronger emotional

This work was supported by grants from Białystok University of Technology (WI/WI-IIT/3/2024, WZ/WI-IIT/5/2023) and funded with resources for research by the Ministry of Science and Higher Education in Poland.

A. Stefanowska and S. K. Zieliński are with Faculty of Computer Science, Białystok University of Technology, Poland (e-mail: antonina.stefanowska@sd.pb.edu.pl, s.zielinski@pb.edu.pl).

responses than the same audio content reproduced over headphones [17].

The existing literature provides evidence that the angular position of a sound source with respect to a listener may influence the intensity of perceived and/or experienced emotions. For instance, the angular location of sound sources outside the listeners' field of view resulted in increased arousal (intensity) [3], [4], [18] and decreased valence (positivity) ratings [3], [4]. Similarly, other study indicated that sound sources positioned on the sides of the listener evoke stronger emotional responses compared to sounds originating from the front [19], particularly in the context of driving a vehicle. The study conducted by Tajadura-Jiménez *et al*. [20] demonstrated that sounds located behind the listener evoke more intense negative emotions as well.

The study on the perception of scariness demonstrated that a sound source that is difficult to localize and positioned outside of the listener's view is perceived as scarier than otherwise [21]. Another study showed that listening to a sound source located outside the field of vision elicits specific brain electrophysiological patterns associated with emotional processing [22]. In contrast, the findings of the study that incorporated sound sources positioned on different sides of the listener at varying elevations [23] did not reveal any significant differences in the participants' ratings. However, notable differences emerged following the introduction of an accompanying subtle background foreboding track in the audio mix. This leads to the conclusion that the affective context of the sound potentially amplifies the effects of its position on the perception of emotions.

The available evidence suggests that the perceptual relationship between sound source position and listeners' emotions may be bidirectional. The findings of Pinheiro *et al.* [4] indicate that emotions experienced by listeners may influence their perception of the sound source location. A more comprehensive literature review on the relationship between spatial audio and listeners' affective responses can be found in [24].

II. Experimental Procedure

This section describes the conditions and configuration of the experimental setting, the utilized playback equipment, the audio recordings employed as stimuli, and the complete process of a listening test controlled by a custom graphical interface.

A. Room and apparatus

Subjective listening tests were conducted in the spatial sound laboratory of the Bialystok University of Technology with a noise rating level of NR22. The reverberation time in the room was approximately equal to 0.12 seconds. The audio stimuli were reproduced by an eight-channel active loudspeaker system comprising Genelec 8010AP-6 loudspeakers (frequency range: 74Hz–20kHz) arranged in a circular configuration with a radius of two meters around the listener. The angular distance between adjacent loudspeakers was 45°. A graphical illustration of the aforementioned loudspeaker array is presented in Fig. 1. The loudspeakers were positioned at a height of 1.2 meters, which is approximately the same as the height of a seated listener's head. Prior to the listening tests, the sound playback level was calibrated to 68.5dBA (± 0.5dB) utilizing pink noise. During the testing procedure, the light in the room was dimmed to encourage participants to focus on the sound. The ambient lighting was maintained at a consistent level throughout the duration of all listening tests. However, formal measurements were not conducted.

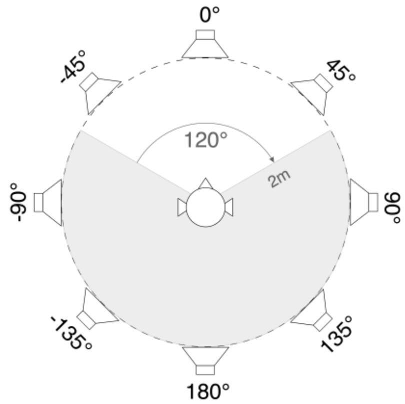

Fig. 1. The configuration of loudspeakers in relation to the listener. The white area in front of the listener represents the human peripheral field of view spanning 120° [2].

B. Stimuli

In this study, we utilized the IADS-E dataset, which comprises a collection of affective sound recordings of diverse origins, including animal, human, natural, urban, artificial, and musical sounds, among others [25]. Each sound is labeled with the mean and standard deviation of scores on six emotional dimensions: arousal, valence, dominance, fear, happiness, and sadness. Values range from 1 to 9. For the purposes of this study, three sounds from the IADS-E repository were selected based on their relatively high fear scores, namely glass shattering, door banging, and hog growling. They are listed in Table I, indicating their durations, identification numbers in the original IADS-E dataset, as well as their mean fear score values.

Table I
Recordings employed in the study.

Description	Duration after modifications	IADS-E identifier	Mean fear score in IADS-E dataset
Glass shattering	2.1 s	273	6.50 (± 2.28)
Door banging	2.7 s	746	6.36 (± 2.19)
Hog growling	3.4 s	553	5.96 (± 2.24)

The aforementioned recordings were trimmed to remove repetitions and their volume was normalized to the loudness level of -23.0 LUFS in accordance with EBU Recommendation R128 [26]. Following these modifications, the duration of a single recording ranged between 2 and 4 seconds (see Table I).

During the listening tests, sounds were played sequentially from the eight loudspeakers arranged around the listener, with each loudspeaker corresponding to a specific angle: -135°, -90°, -45°, 0°, 45°, 90°, 135°, 180° (as illustrated in Fig. 1). A total of 24 distinct sound samples were employed in the tests with three recordings and eight angular variants. The playback order was randomized for each listener, with samples grouped according to their recording type. Each sample was played twice, resulting in a total of 48 sound stimuli played for each participant.

C. Listening tests

During the listening test, participants were instructed to assess each sound stimulus based on their subjective perception of its scariness. A total of thirty-six students from the Bialystok University of Technology participated in the study. The participants were predominantly male ($n = 32$), with the majority falling within the 18-25 age range ($n = 34$). Two participants indicated an age range of 26-35.

The listening tests were conducted utilizing a custom desktop application developed in Max software. The developed application provided the participants with a graphical user interface in Polish, enabling them to enter the required data and control the playback of the sound stimuli. The interface was displayed on a monitor placed in front of the participant and was controlled using a computer mouse.

Prior to the main test phase, participants were requested via the application interface to complete a personal questionnaire. It encompassed questions about gender, age range, any preexisting hearing difficulties, and consent to participate in the study. Thereafter, the participants were presented with instructions outlining the procedure for the listening test. This included a request to refrain from moving their heads during the test. Next, the participants underwent a training session, during which three types of recordings were played in succession, each with one randomly selected angular variant.

The training part was subsequently followed by the main test with a randomized playlist comprising 48 sound samples. Both the training session and the main test involved the playback of the sound samples and their evaluation by the participant. However, the responses provided during the training phase were not recorded.

The listeners were required to evaluate each sound stimulus by finishing an incomplete statement displayed in the application interface: *The currently playing sound is…* (pol. *Odtwarzany dźwięk jest…).* The rating scale ranged from 1 to 7 and represented the following options: Extremely scary (pol. *Ekstremalnie straszny)*, Very scary (pol. *Bardzo straszny*), Pretty scary (pol. *Całkiem straszny*), Scary (pol. *Straszny*), Somewhat scary (pol. *Trochę straszny*), Not very scary (pol. *Niezbyt straszny*), Not scary at all (pol. *Wcale nie straszny*). The rating scale was inspired by the one employed in the study conducted by Hughes and Kearney [23]. It should be noted that the listener was not allowed to replay the sound stimulus nor to return to the previous ones. An exemplar screenshot of the application interface during the listening test session is provided in Fig. 2.

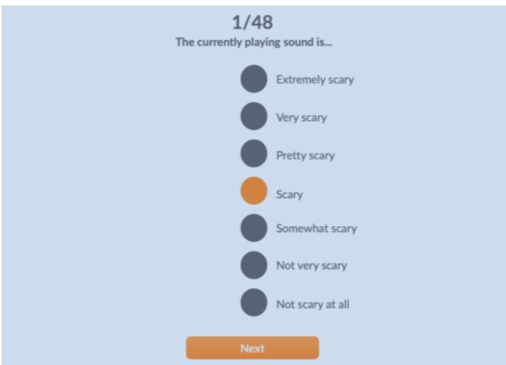

Fig. 2. Example view of the application interface during the listening test.

III. Results

This section presents a description of the listeners partaking in this study and characterizes them in terms of repeatability and discriminability. It also provides a rationale for the exclusion of some of the collected data. Moreover, it presents the main findings of the study.

A. Data screening

Out of 36 participants taking part in the study, three individuals reported a history of hearing difficulties. Nevertheless, these participants were not excluded from the study, as their level of repeatability and discriminability did not deviate notably from that of the remaining listeners.

During the course of the listening tests, each sound stimulus was played twice. A degree of repeatability for each participant was determined using the Root Mean Square Error (RMSE) value. It was calculated based on the differences between the scores assigned on the first playback of sound samples and those assigned upon their repeat. The resulting repeatability chart for the participants is presented in Fig. 3. It can be seen that listener number 13 exhibited the lowest RMSE value, being equal to 0. This outcome indicates that this particular listener was the sole individual who exhibited 'perfect' repeatability. However, it has been caused by the participant applying the same two rating scores throughout the entirety of the listening test. This observation was further supported by the low discriminability value for this participant discussed in the further part of this section. In contrast, participants number 15, 16, and 17 exhibited the highest RMSE values, indicating that their ratings of the repeated samples were the most inconsistent with the original scoring. Despite the relatively high RMSE values, the data obtained from these participants was not excluded from the results.

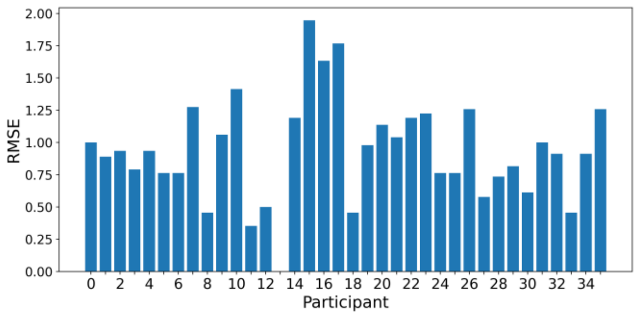

Fig. 3. Repeatability chart for the participants of the study. The lower the RMSE value, the higher the repeatability of results for a given participant.

The value of the F-statistic derived from a one-way ANOVA was employed as a measure of a degree of discriminability for the participants. The resulting discriminability chart is shown in Fig. 4. The calculated values indicate that some participants exhibited a particularly low degree of discriminability. Specifically, five participants of numbers 2, 11, 13, 25, and 28 showed an F-statistic value lower than 0.25. Further analysis revealed that these listeners rated all the stimuli using only one or two adjacent values from the rating scale. Consequently, the data collected from them were excluded from the final results.

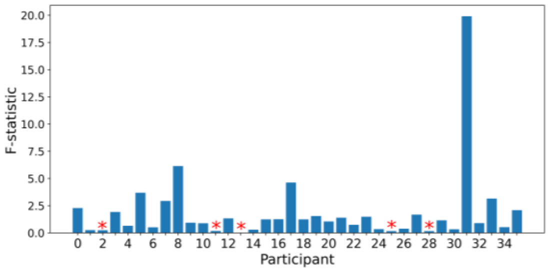

Fig. 4. Discriminability chart for the participants of the study based on *F*-statistic values from the one-way ANOVA. Data collected from participants marked with an asterisk were excluded from the results.

B. Analysis

The assumptions underlying a three-way ANOVA were tested for the collected data, which demonstrated compliance with the requisite conditions. Thereafter, the statistical tests were conducted. The effects of the angle, recording type, and participant variables on the score ratings were found to be statistically significant (p < 0.01, F = 8.82, 30.7, 91.27 respectively). The effect sizes expressed as partial eta-squared (η_p^2) for these three factors were equal to 0.077, 0.076, and 0.786, respectively. These values indicate that the individual differences between participants have a significantly greater effect on the rating of scariness compared to the effects of the angle and the recording content. The complete results of the three-way ANOVA are presented in Table II.

Subsequently, Tukey's HSD post-hoc pairwise tests were applied to the angle variants. The only statistically significant differences were identified between -90° and 0° ($p = 0.015$), as well as between -90° and 45° ($p = 0.012$). This finding was further confirmed by the mean score chart with 95% confidence intervals presented in Fig. 5. It can be seen in the figure that the confidence intervals for the aforementioned pairs of angles do not overlap.

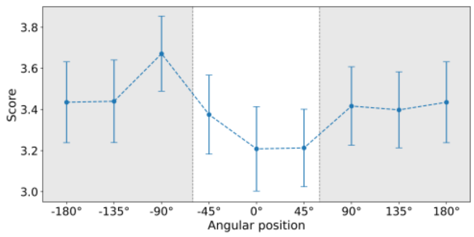

Fig. 5. Mean scariness scores with 95% confidence intervals for each studied angle. The white area indicates ⟨-60°, 60°⟩ range, which represents the approximate human peripheral field of view.

A similar series of pairwise tests was conducted on the type of recordings. The glass shatter sound received higher scariness ratings than both hog growling and door banging ($p < 0.01$), which may be attributed to the highest mean scariness score of the sound in the original dataset. The mean fear scores for recording types in the original repository and the mean scariness scores obtained in the current study were normalized to the same scale range and compared in Fig. 6. The relatively high values of fear scores reported in the original dataset compared to those obtained in our study appear to have been caused by the fact that originally the recordings were evaluated alongside numerous samples of varying emotional qualities, rather than exclusively those adjacent to fear. In contrast, in our study only three repeating recordings were employed and compared with each other, which would explain the comparatively lower scariness scores.

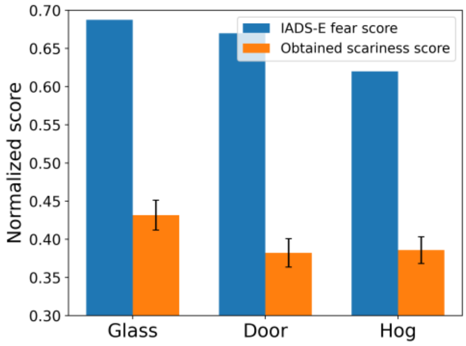

Fig. 6. Normalized mean fear scores from the IADS-E dataset (blue) and the scariness scores obtained in this study (orange). Error bars indicate 95% confidence intervals.

In consideration of the approximate human vision range of ⟨-60°, 60°⟩, the scores were further aggregated within the two groups: sound stimuli located within the field of view and sound samples localized outside of it. The first group comprised three angles: -45°, 0°, and 45°, while the second group encompassed angles of -135°, -90°, 90°, 135°, and 180°. Subsequently, the two groups of scores were compared using the t-test for independent samples, preceded by the Levene test. The homogeneity of data variance was assumed based on the result of the Levene test ($p = 0.84 > 0.05$). The results of the t-test indicate that the scariness ratings within the field of view were statistically significantly lower compared to those outside the field of view ($p = 0.0015$). A chart of the mean

TABLE II
THREE-WAY ANOVA RESULTS.

Factor	Degrees of freedom	Sums of squares	Mean squares	F	p	η_p^2
Angle	7	34.994	4.999	8.824	1.868×10^{-10}	0.077
Recording	2	34.784	17.392	30.697	1.546×10^{-13}	0.076
Participant	30	1551.267	51.709	91.273	2.332×10^{-226}	0.786
Angle × Recording	14	31.98	2.284	4.032	9.206×10^{-7}	0.071
Angle × Participant	210	331.819	1.58	2.789	3.295×10^{-24}	0.44
Recording × Participant	60	341.716	5.695	10.053	1.25×10^{-62}	0.448
Angle × Recording × Participant	420	353.52	0.842	1.486	1.552×10^{-6}	0.456
Residual	744	421.5	0.567			

scores with 95% confidence intervals for the two groups is shown in Fig. 7. The obtained results indicate the validity of the initial hypothesis.

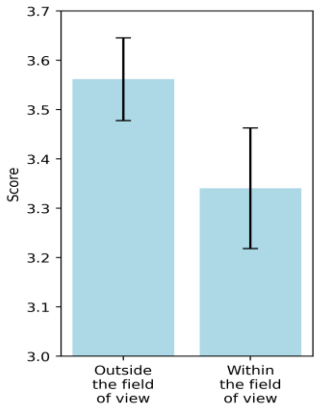

Fig. 7. Aggregated mean scariness scores with 95% confidence intervals within the human peripheral field of view and outside it.

The interaction between angles and participants was found to have a medium effect size ($\eta_p^2 = 0.44$), indicating the existence of different patterns by which participants perceived the scariness of sounds arriving from different directions. As an example, the interaction between angles and the five selected participants is illustrated in Fig. 8. Participant number 31 demonstrated the evaluation pattern of sound sources in accordance with our initial hypothesis. In contrast, participants 8 and 17 appeared to perceive sound sources located within the visual field as more scary than those outside it. Additionally, the participants also exhibited notable variability between the range of ratings. For instance, participants 8 and 33 seemed to rate all the sound samples relatively low using the limited range of the rating scale.

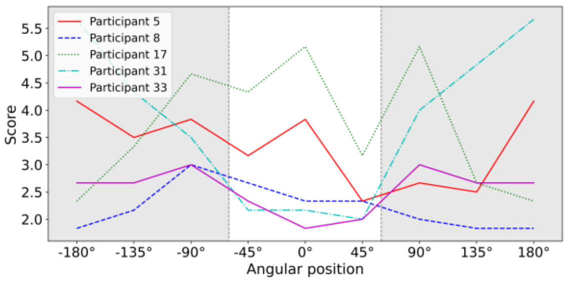

Fig. 8. Mean scariness scores from five selected participants with 95% confidence intervals grouped by studied angles. The white area indicates the approximate human peripheral field of view.

IV. Discussion

It is noteworthy that only the -90° (direct left) angular position from the outside the field of view group demonstrated significantly higher fear ratings compared to the two variants from the other group, which were 0° and 45°. Moreover, there appears to be an observable asymmetry between the left side versus the right side of the listener in terms of scariness perception. This phenomenon might be related to the observation reported in the literature that the left ear is more sensitive to emotional information than the right ear [27]-[30].

The results obtained in this study, along with their subsequent analysis, indicate that the group of sound sources positioned outside the bounds of vision was evaluated as scarier. These results are in line with the findings of Ekman and Kajastila [21] who demonstrated a correlation between a sound source positioned behind the back of a listener and the intensity of perceived scariness. This effect was further amplified when the sound sources were spread. The findings of Asutay and Västfjäll [18] as well as Drossos *et al.* [3], suggested that the sound arriving from behind the listener evokes more intense negative feelings than otherwise, which are also consistent with the results of this study.

In contrast, the analogous study conducted by Hughes and Kearney [23] did not identify any statistically significant difference in terms of perception of scariness between frontally positioned sound sources and sources placed at different azimuths and elevations in relation to the listener. However, the inclusion of a subtle background foreboding track in the audio resulted in a statistically significant difference in scores. Subsequently, audio played outside the front-center position in relation to the listener was evaluated as comparatively scarier.

A notable limitation of the study was the sequential playback of sound stimuli of the same type. According to the informal feedback obtained from some participants, this approach might have resulted in a diminished emotional impact. One potential solution to this issue in future experiments is the incorporation of a wider and more randomized playlist of recordings.

Conclusion

The findings of this study provide evidence that sound sources positioned outside of the listener's field of view are perceived as scarier than those positioned within the bounds of vision. However, the differences between the compared positions, while statistically significant, were relatively minor. Nevertheless, the results suggest that positioning physical or virtual sound sources behind or to the side of a listener could increase the perceived level of fear.

Future studies in this area could benefit from a more comprehensive examination of a wider range of emotions. The inclusion of physiological response data acquired from listeners through wearable sensors would also constitute a valuable contribution to the research, given the lack of similar studies incorporating emotional data other than self-reports.

Acknowledgments

We are grateful to the volunteering students of Białystok University of Technology for their participation in the study. The research was conducted with a positive opinion of the Ethics Committee of Białystok University of Technology dated 14.06.2024 – protocol no. 1/2024. It was supported by grants from Białystok University of Technology (WI/WI-IIT/3/2024, WZ/WI-IIT/5/2023) and funded by the Ministry of Science and Higher Education.

References

[1] F. Hutmacher, "Why Is There So Much More Research on Vision Than Any Other Sensory Modality?," Frontiers in Psychology, vol. 10, pp. 2246, 2019. https://doi.org/10.3389/fpsyg.2019.02246

[2] J. Sardegna, S. Shelly, A. R. Rutzen, and M. S. Scott, "The Encyclopedia of Blindness and Vision Impairment," Infobase Publishing, 2002.

[3] K. Drossos, A. Floros, A. Giannakoulopoulos, and N. Kanellopoulos, "Investigating the Impact of Sound Angular Position on the Listener Affective State," IEET Transactions on Affective Computing, vol. 6, no. 1, pp. 27–42, 2015. https://doi.org/10.1109/TAFFC.2015.2392768

[4] A. P. Pinheiro, D. Lima, P. B. Albuquerque, A. Anikin, and C. F. Lima, "Spatial location and emotion modulate voice perception," Cognition and Emotion, vol. 33, no. 8, pp. 1577–1586, 2019. https://doi.org/10.1080/02699931.2019.1586647

[5] V. Best, R. Baumgartner, M. Lavandier, P. Majdak, and N. Kopčo, "Sound Externalization: A Review of Recent Research," Trends in Hearing, 2020. https://doi.org/10.1177/2331216520948390

[6] U. Horbach, A. Karamustafaoglu, R. Pellergrini, P. Mackensen, and G. Theile, "Design and Applications of a Data-Based Auralization System for Surround Sound," in 106th Audio Engineering Society Convention, Munich, Germany, 1999.

[7] E. Hahn, "Musical Emotions Evoked by 3D Audio," in Audio Engineering Society International Conference on Spatial Reproduction – Aesthetics and Science, Tokyo, 2018. http://www.aes.org/e-lib/browse.cfm?elib=19640

[8] G. Moiragias and J. Mourjopoulos, "A listener preference model for spatial sound reproduction, incorporating affective response," PLoS One, vol. 18, no. 6, 2023. https://doi.org/10.1371/journal.pone.0285135

[9] R. Warp, M. Zhu, I. Kiprijanovska, J. Wiesler, S. Stafford, and I. Mavridou, "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, 2022.
http://www.aes.org/e-lib/browse.cfm?elib=21703

[10] Y. Wycisk, K. Sander, R. Kopiez, F. Platz, S. Preihs, and J. Peissig, "Development of the Immersive Music Experience Inventory (IMEI)," Frontiers in Psychology, vol. 13, pp. 951161, 2022. https://doi.org/10.3389/fpsyg.2022.951161

[11] F. Cuadrado, I. Lopez-Cobo, T. Mateos-Blanco, and A. Tajadura-Jiménez, "Arousing the Sound: A Field Study on the Emotional Impact on Children of Arousing Sound Design and 3D Audio Spatialization in an Audio Story," Frontiers in Psychology, vol. 11, pp. 737, 2020. https://doi.org/10.3389/fpsyg.2020.00737

[12] S. Lepa, S. Weinzierl, H.-J. Maempel, and E. Ungeheuer, "Emotional Impact of Different Forms of Spatialization in Everyday Mediatizes Music Listening: Placebo or Technology Effects?" in 136th Audio Engineering Society Convention, Berlin, 2014. http://www.aes.org/e-lib/browse.cfm?elib=17171

[13] M. Shin, S. W. Song, S. J. Kim, and F. Biocca, "The effects of 3D sound in a 360-degree live concert video on social presence, parasocial interaction, enjoyment, and intent of financial supportive action," International Journal of Human-Computer Studies, vol. 126, pp. 81–93, 2019. https://doi.org/10.1016/j.ijhcs.2019.02.001

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

[15] S. Oode and A. Ando, "Estimation of Kandoh Degree with Emphasis on Spatial Sound Impressions," in 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Kyoto, 2012.
https://doi.org/10.1109/SNPD.2012.98

[16] D. Västfjäll, "The Subjective Sense of Presence, Emotion Recognition, and Experienced Emotions in Auditory Virtual Environments," CyberPsychology & Behavior, vol. 6, no. 2, pp. 181–188, 2003. https://doi.org/10.1089/109493103321640374

[17] Y. Ooishi, M. Kobayashi, M. Kashino, and K. Ueno, "Presence of Three-Dimensional Sound Field Facilitates Listeners' Mood, Felt Emotion, and Respiration Rate When Listening to Music," Frontiers in Psychology, vol. 12, pp. 650777, 2021. https://doi.org/10.3389/fpsyg.2021.650777

[18] E. Asutay and D. Västfjäll, "Attentional and Emotional Prioritization of the Sounds Occurring Outside the Visual Field," Emotion, vol. 15, no. 3, pp. 281–286, 2015. https://doi.org/10.1037/emo0000045

[19] P. Larsson and D. Västfjäll, "Emotional and behavioural responses to auditory interfaces in commercial vehicles," International Journal of Vehicle Noise and Vibration, vol. 9, no. 1–2, pp. 75–95, 2013. https://doi.org/10.1504/IJVNV.2013.053818

[20] A. Tajadura-Jiménez, P. Larsson, A. Väljamäe, D. Västfjäll, and M. Kleiner, "When room size matters: acoustic influences on emotional responses to sounds," Emotion, vol. 10, no. 3, pp. 416-422, 2010. https://doi.org/10.1037/a0018423

[21] I. Ekman and R. Kajastila, "Localization Cues Affect Emotional Judgments – Results from a User Study on Scary Sound,", in 35th Audio Engineering Society International Conference: Audio for Games, London, United Kingdom, 2015. https://aes2.org/publications/elibrary-page/?id=15177

[22] E. De Filippi, T. Schmele, A. Nandi, A. G. Torres, and A. Pereda-Baños, "Emotional Impact of Source Localization in Music Using Machine Learning and EEG: a proof-of-concept study," TechRxiv, 2023. https://doi.org/10.36227/techrxiv.21789866.v2

[23] S. Hughes and G. Kearney, "Fear and localization: Emotional fine-tuning utilizing multiple source directions," in 56th Audio Engineering Society International Conference: Audio for Games, London, United Kingdom, 2015. https://aes2.org/publications/elibrary-page/?id=17596

[24] A. Stefanowska and S. K. Zieliński, "Spatial sound and emotions: A literature survey on the relationship between spatially rendered audio and listeners' affective responses," International Journal of Electronics and Telecommunications, vol. 70, no. 2, pp. 293–300, 2024. https://doi.org/10.24425/ijet.2024.149544

[25] W. Yang, K. Makita, T. Nakao, N. Kanayama, M. G. Machizawa, T. Sasaoka, A. Sugata, R. Kobayashi, R. Hiramoto, S. Yamawaki, M. Iwanaga, and M. Miyatani, "Affective auditory stimulus database: An expanded version of the International Affective Digitized Sounds (IADS-E)," in Behavior Research Methods, pp. 1–15, 2018.

[26] European Broadcasting Union, "R 128 – Loudness normalisation and permitted maximum level of audio signals," 2023, https://tech.ebu.ch/docs/r/r128.pdf

[27] A. D. Alzahrani and M. A. Almuhammadi, "Left ear advantages in detecting emotional tones using dichotic listening task in an Arabic sample," Laterality, vol. 18, no. 6, pp. 730–737, 2013. https://doi.org/10.1080/1357650X.2012.762373

[28] M. P. Bryden, R. G. Ley, and J. H. Sugarman, "A left-ear advantage for identifying the emotional quality of tonal sequence," Neuropsychologia, vol. 20, no. 1, pp. 83–87, 1982. https://doi.org/10.1016/0028-3932(82)90089-6

[29] R. G. Ley and M. P. Bryden, "A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content," Brain and Cognition, vol. 1, no. 1, pp. 3–9, 1982. https://doi.org/10.1016/0278-2626(82)90002-1

[30] L. Saxby and M. P. Bryden, "Left-ear superiority in children for processing auditory emotional material," Developmental Psychology, vol. 20, no. 1, pp. 72–80, 1984. https://doi.org/10.1037/0012-1649.20.1.72