forum acusticum 2026
8-12 September · Graz, Austria

EXPLORING THE INFLUENCE OF SPATIAL ATTRIBUTES AND VISUAL CUES ON PSYCHOACOUSTIC ANNOYANCE USING VIRTUAL REALITY

Samuel Dickinson, Hyunkook Lee

Applied Psychoacoustics Lab (APL), University of Huddersfield, HD1 3DH, UK

Abstract

The influence of spatial acoustic attributes on psychoacoustic annoyance remains understudied. This study investigated how source direction, acoustic condition, and visual cues affect perceived annoyance using a virtual reality (VR)-based experimental method. Four common domestic sounds were assessed in three simulated acoustic conditions: anechoic, small office, and large café. Higher-order ambisonic room impulse responses were captured and stimuli were presented binaurally with headtracking. 33 participants identified the most and least annoying source directions and rated annoyance levels. Results confirmed that lateral source positions ($\sim \pm 90°$) were perceived as most annoying, and frontal/rear positions were least annoying. Visual cues influenced the directional responses by providing visual anchors and reducing localisation uncertainty, but did not significantly influence annoyance ratings. Acoustic condition significantly affected annoyance levels for some sources, with the office environment eliciting greater annoyance than the café, potentially due to differences in acoustic characteristics and contextual associations. The findings suggest that, in addition to the sound source itself, annoyance perception depends on source direction, room acoustics, visual context, and psychological interpretation of the environment. The proposed VR-based method provides an ecologically valid approach for investigating spatial and multisensory factors of annoyance, with potential applications in soundscape evaluation and environmental design.

Keywords: *psychoacoustic annoyance, virtual reality, spatial attributes*

Corresponding author: *samuel.dickinson@hud.ac.uk.*

1. Introduction

Existing annoyance research focuses on spectral and temporal features of sounds, including psychoacoustic parameters such as loudness, sharpness, roughness, and tonality. This led to the development of theoretical psychoacoustic annoyance (PA) models, combining these parameters. Widmann's [1] model, often attributed to Zwicker and Fastl [2], first became widely adopted, forming the foundation of many subsequent models. These models are accurate annoyance predictors for mono audio signals, however binaural and ambisonics recording and playback methods are now common in soundscape quality studies [3], and recommended in ISO/TS 12913-2 [4]. Existing research on the influence of spatial acoustic attributes predominantly focuses on interaural cross-correlation (IACC) [5]–[8], highlighting the need for further investigation.

Previous studies [9], [10] identified a significant effect of sound source azimuth relative to the listener on perceived annoyance. The results showed that annoyance was greatest at lateral positions ($\pm 90°$) and lowest at the rear (180°), for static binaural audio stimuli. Dickinson and Lee [9] also identified a significant influence of acoustic condition, where annoyance was greater for a simulated reverberant (open-plan kitchen/living room) environment than for an anechoic environment. These findings demonstrated the potential for further investigation of the acoustic conditions effect, and presented opportunities to explore the impact of visual cues and natural head movement.

The present study employed a virtual reality (VR)-based experimental methodology to assess the effect of acoustic conditions and visual cues on perceived annoyance. A more natural method was used to obtain the most and least annoying source directions, in which participants freely rotated their head. This enabled the evaluation of variability in the source azimuths which elicited the highest and lowest annoyance levels, as a function of the acoustic and visual conditions. VR is becoming common in soundscape research, facilitating highly controlled and repeatable experiments, with accurate simulation of complex soundscapes, enabling more ecologically valid perceptual evaluation than traditional laboratory testing methods. The present study combined VR and higher-order ambisonics (HOA) audio stimuli, decoded for binaural playback with headtracking.

2. Methodology

2.1. Stimuli Creation

Four common domestic sound sources were assessed: a coffee grinder, mechanical keyboard, traffic noise, and simulated emergency alarm. 32-channel room impulse responses (RIRs) were captured using an Eigenmike EM32 spherical microphone array, in two indoor environments: a large room in a café and a small office room. The office environment was characterised by a large desk and a large window, whereas the café housed round tables and a large countertop. The dimensions and average reverb time of each room are shown in Table 1.

Table 1: Room dimensions (m) and average reverb times (T30) (s).

Room	Length	Width	Height	$T_{30,mid}$
Office	4.06	2.85	3.24	0.63
Café	7.50	6.00	3.43	0.98

Realistic positions were selected in each room at which each sound source would typically be situated. Each mono source signal was convolved with an encoded 3rd order HOA RIR, captured with the corresponding source positions. The EM32 was positioned near the centre of the café environment and close to the usual seated position at the desk in the office. Each source signal was also encoded to 3rd order HOA, creating an anechoic condition.

The listening test interface, with audiovisual stimuli presentation and response capture, was managed using Unreal Engine 5 and Wwise sound engine. Encoded audio stimuli were decoded for binaural playback. KU 100 dummy head HRTFs [11] and headphone EQ compensation were applied.

For the visual stimuli, 360° photos were captured at the receiver positions. Each 360° photo was mapped to the inside of a sphere in Unreal Engine, and the VR camera was placed at the centre of the sphere. Head tracking was enabled, but positional tracking was locked. 3D models representing each sound source were placed at the source positions. As the traffic noise stimulus was intended to simulate a road outside a closed window, no visual representation was included, and a window was selected as the source position.

Each visual stimulus was paired with its audio stimulus. Additional stimuli versions with no visual components were also created (blackout stimuli). The anechoic stimuli were also presented without visual components. A total of 20 audiovisual stimuli were created, each lasting 30 s.

The playback sound pressure levels (SPLs) of the audio stimuli were set to realistic levels in their usual context, using a calibrated MiniDSP EARS ear simulator. Direct sound levels were normalised across the three acoustic environments for a specified source position, maintaining distance-related differences in the RIRs within each room.

2.2. Subjects

33 participants took part in the experiment (25 male, 6 female, and 2 non-binary), aged 19 to 55 years (mean = 26.33), comprising both naïve and experienced critical listeners. All participants reported normal hearing. They were informed of the test procedure and use of personal information in compliance with EU General Data Protection Regulation (GDPR) 2018, and provided their consent to participate.

2.3. Listening Test Procedure

The listening experiment was conducted in the APL's ITU-R BS.1116-3 [12] compliant listening room at the University of Huddersfield. During the experiment, participants wore a Meta Quest 3 HMD and Sennheiser HD 650 headphones, and the audiovisual stimuli were presented in a random order. Each stimulus was presented twice in immediate succession. Participants were seated on an office swivel chair and instructed to perform rotational movements without adjusting their horizontal position.

For each stimulus, they were instructed to face the directions at which the sound caused the most and least annoyance independently. They also answered the question "How much does the sound bother, disturb, or annoy you?", rating their answer for each direction on a 0 (Not at all annoying) to 10 (Extremely annoying) scale, based on ISO/TS 15666 [13]. The experiment was split into 3 10-minute sections with two short breaks, to prevent fatigue, and participants could end the experiment at any time for any reason.

3. Results

Fig. 1 shows the circular kernel density estimates (KDE) of the most and least annoying azimuth responses. For all stimuli, the least annoying direction responses were generally split into two narrow clusters, centred at 0° and 180°. A third cluster was identified for most stimuli, with a lateral mean azimuth, although this comprised a small number of responses, or was distributed more evenly around the circle, indicating the absence of a dominant directional bias.

The most annoying source azimuths were grouped around ±90°. However, another group was present for most stimuli, centred around 0°. A potential influence of visual conditions on the most annoying source azimuth is also shown in Fig. 1. Some sources in the visual office and café conditions demonstrated a narrower and more elongated lateral (90° or 270°) lobe than their blackout counterpart. This was particularly clear for the coffee grinder, traffic noise, and mechanical keyboard sources in the office conditions;

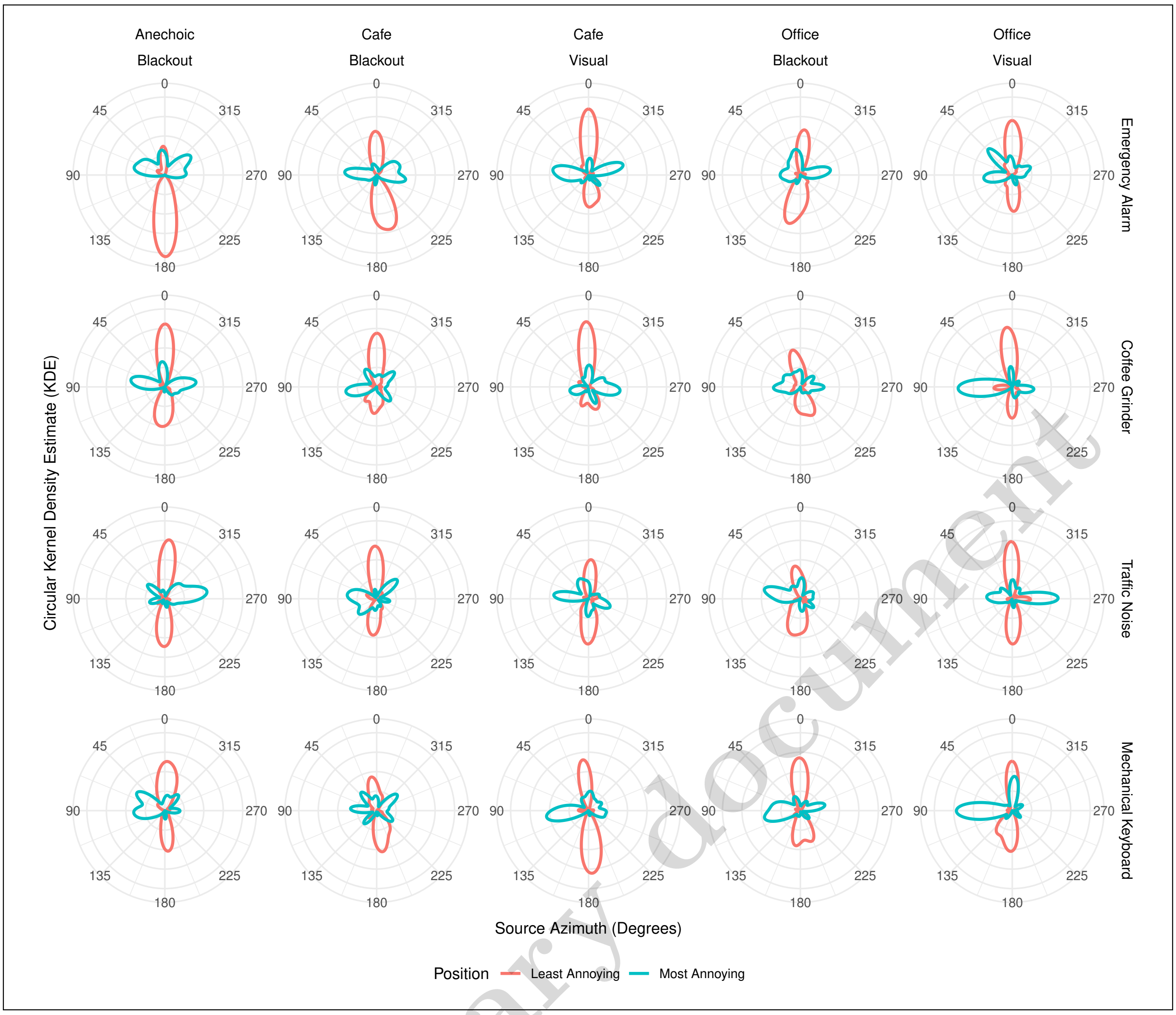

Figure 1: Circular kernel density estimates (KDE) of the most and least annoying azimuth responses for all stimuli.

the lateral café response lobes were less distinct. Visual conditions were also found to potentially affect the least annoying position responses, as a greater density was identified at 0° for the visual stimuli compared to the blackout stimuli, for all sources in at least one of the acoustic conditions. The anechoic stimuli responses more closely followed the trends seen for the visual conditions rather than the other blackout conditions, exhibiting high densities at distinct azimuths, especially for the least annoying position.

For all stimuli, mean annoyance levels were significantly greater at the most annoying source directions, than at the least annoying directions. The visual conditions did not appear to influence perceived annoyance levels, whereas a strong effect of acoustic conditions was identified. Greater annoyance was exhibited for the office condition than the café, for the traffic noise at both the most and least annoying directions in the visual and blackout conditions, as well as the coffee grinder in the visual condition and for the least annoying position in the blackout condition. Annoyance levels were also affected by the sound source, with the emergency alarm demonstrating the greatest overall annoyance and the traffic noise exhibiting the lowest annoyance levels.

4. Discussion

The azimuth responses were consistent with findings of [9]; most participants perceived the greatest annoyance for sources positioned at ±90°, with the least annoying position at 180° or 0°. However, in the present study, another smaller most annoying azimuth group was identified around 0°. This dissensus between participants for frontal sources may explain the annoyance increase compared to 180° identified by Dickinson and Lee [9]. The distribution of least annoying direction responses was mostly very narrow at both mean azimuths, suggesting that annoyance quickly increases as the interaural level difference (ILD) deviates from 0 dB.

The reverberant blackout stimuli exhibited wider variation in response clusters than the anechoic stimuli, particularly for the most annoying direction. This may be due to the reverberation spreading the sound, reducing localisation accuracy. This suggests that in larger, more reverberant rooms, the most annoying position is less defined and may become more distinct in smaller, more controlled acoustic environments. The more uniform response distribution for the most annoying position for some café stimuli also supports this.

The visual stimuli generally exhibited greater most annoying direction response density for the lateral azimuth clusters, than the blackout stimuli, suggesting that visible sound sources provide an anchor which aids in determining the most annoying direction. The dominant of the lateral most annoying azimuth response groups for each source in the office corresponded to the source location when facing the desk. This highlights the significance of contextual factors on annoyance, and suggests the interaction between visual cues and psychological context may strongly influence the source azimuth effect. Visual conditions also potentially influenced the least annoying azimuth responses. The visual stimuli elicited a greater response density at 0°, compared to the blackout stimuli, for all sources in at least one acoustic condition, implying that visible sources reduce annoyance.

The most annoying source azimuths elicited greater mean annoyance levels than the least annoying azimuths for all stimuli, which confirms that source azimuth significantly affects the perceived annoyance level for a range of common indoor sounds, across multiple audiovisual conditions.

Despite affecting the most and least annoying source directions, visual cues did not considerably influence annoyance levels. A potential effect of acoustic conditions was demonstrated, with a slight mean annoyance increase seen for all visual office stimuli, compared to the café condition, for both assessed directions. This was also observed in the blackout condition, though mostly for the least annoying direction. No considerable change was identified in responses between the anechoic and reverberant conditions. This opposes Dickinson and Lee's [9] findings, where annoyance increased significantly for reverberant stimuli. Frescura et al. [5] suggested that annoyance increases with IACC, which is supported by these findings. Although, the lack of response variation between anechoic and reverberant stimuli suggests reverb times and IACC do not fully explain this effect. Further analysis of acoustic characteristics will be carried out in future work.

The differences between the visual café and office responses may also relate to the visual context of the office environment. The mean annoyance levels for the coffee grinder and mechanical keyboard may also reflect this, remaining similar for both blackout reverberant conditions, but increasing for the visual office condition. These results suggest that both sound source and visual conditions influence the effect of acoustic conditions on annoyance.

The observed effect of sound source aligns with playback SPL and therefore the loudness of each source, which is widely shown to affect annoyance throughout existing literature.

The employed VR-based methodology provided a more ecologically valid testing environment than traditional laboratory methods. The use of 360° photos provided a good visual representation of the spaces, however this approach can result in limitations related to depth perception at close range. The assessment of most and least annoying directions with headtracking could be useful in a range of experimental applications, reducing the number of trials required for meaningful results. The system could be easily adjusted to include elements such as 6DOF or 3D visual environment simulations for future experiments.

5. Conclusions

The results verified the findings of a previous investigation [9], demonstrating a significant effect of sound source azimuth on perceived annoyance levels for a range of sounds. Laterally placed (±90°) sound sources caused the greatest mean annoyance, and the lowest mean annoyance levels were observed for rear and frontal sources. Although, some participants perceived the highest annoyance for frontal sources. This disagreement may explain a significant annoyance level difference between frontal and rear sources identified in [9].

Visual cues influenced the azimuths at which sounds caused the most and least annoyance, however they did not significantly affect mean annoyance levels. A potential influence of visual and contextual references on the most and least annoying source directions was identified.

Acoustic condition significantly affected mean annoyance levels, with differences identified between those for the office and café stimuli for some sources. Visual and contextual features of the environments may also influence the observed effects.

The novel experimental method performed well, providing meaningful results and additional insights that would be lacking for traditional methods. A realistic impression of the environments was achieved, and the method presents potential applications in future investigations.

Further assessment of the observed effects will be carried out, and the influence of interactions between these effects and contextual associations could also be investigated.

References

[1] U. Widmann, "Ein modell der psychoakustischen lästigkeit von schallen und seine anwendung in der praxis der lärmbeurteilung ("A model of the psychoacoustic annoyance of sounds and its application in noise assessment practice")," Ph.D. dissertation, Technische Universität München (Technical University of Munich), 1992.

[2] E. Zwicker and H. (Fastl, *Psychoacoustics : Facts and Models*. Springer, 1999, ISBN: 978-3-540-65063-8. (visited on 10/24/2023).

[3] Ö. Gök Tokgöz, M. S. Engel, C. Othmani, and M. E. Altinsoy, "Virtual Reality Application in Evaluating the Soundscape in Urban Environment: A Systematic Review," *Acoustics*, vol. 7, no. 4, p. 68, Oct. 2025, ISSN: 2624-599X. DOI: 10.3390/acoustics7040068. (visited on 03/30/2026).

[4] ISO, *Acoustics — Soundscape — Part 2: Data collection and reporting requirements*, Aug. 2018.

[5] A. Frescura, P. J. Lee, Y. Soeta, and A. Ariki, "Effects of spatial characteristics of footsteps sounds and non-acoustic factors on annoyance in lightweight timber buildings," *Building and Environment*, vol. 222, p. 109 405, Aug. 2022, ISSN: 03601323. DOI: 10.1016/j.buildenv.2022.109405. (visited on 01/24/2024).

[6] J. Y. Jeon, P. J. Lee, J. H. Kim, and S. Y. Yoo, "Subjective evaluation of heavy-weight floor impact sounds in relation to spatial characteristics," *The Journal of the Acoustical Society of America*, vol. 125, no. 5, pp. 2987–2994, May 2009, ISSN: 0001-4966, 1520-8524. DOI: 10.1121/1.3081390. (visited on 06/28/2024).

[7] T. Kitamura, R. Shimokura, S.-i. Sato, and Y. Ando, "Measurement of Temporal and Spatial Factors of a Flushing Toilet Noise in a Downstairs Bedroom," *Journal of Temporal Design in Architecture and the Environment*, vol. 2, no. 1, 2002.

[8] S.-i. Sato, T. Kitamura, and Y. Ando, "Annoyance of noise stimuli in relation to the spatial factors extracted from the interaural cross-correlation function," *Journal of Sound and Vibration*, vol. 277, no. 3, pp. 511–521, Oct. 2004, ISSN: 0022460X. DOI: 10.1016/j.jsv.2004.03.012. (visited on 02/12/2024).

[9] S. Dickinson and H. Lee, *Effect of Source Azimuth on the Perceived Annoyance of Domestic Sounds*, 2026. DOI: 10.2139/ssrn.6602723. (visited on 04/30/2026).

[10] H. ZHAO, K. CHEN, and H. LI, "Experimental study of the effect of horizontal orientation of sound source on annoyance," *Chinese Journal of Acoustics*, vol. 44, no. 1, pp. 66–83, Nov. 2025, ISSN: 0217-9776. DOI: 10.15949/j.cnki.0217-9776.2025.01.005.

[11] C. Armstrong, L. Thresh, D. Murphy, and G. Kearney, "A Perceptual Evaluation of Individual and Non-Individual HRTFs: A Case Study of the SADIE II Database," 2018.

[12] *RECOMMENDATION ITU-R BS.1116-3 – Methods for the subjective assessment of small impairments in audio systems*, 2015.

[13] ISO, *Acoustics — Assessment of noise annoyance by means of social and socio-acoustic surveys*, May 2021.