DOI: 10.61782/fa.2025.0373

EXPERIMENTAL DESIGN CONSIDERATIONS FOR PROXIMITY PERCEPTION OF BINAURAL ACTION SOUNDS IN ANECHOIC AND REVERBERANT ENVIRONMENTS

Mohammad MH. Rahmani[1*] **Marte Roel Lesur**[1,2] **Daniel de la Prida**[3]
Luis A. Apizcueta-Ruiz[1] **Joaquín R. Díaz Durán**[1] **Amar D'Adamo**[1]
Ana Tajadura-Jiménez[1,4]

[1] Universidad Carlos III de Madrid, Spain
[2] Association for Independent Research, Switzerland
[3] Grupo de Investigación en Acústica Arquitectónica, Universidad Politécnica de Madrid, Spain
[4] UCL Interaction Centre, University College London, United Kingdom

ABSTRACT

Manipulating spatial cues of self-produced sounds has been shown to change people's own body perception. For example, prior studies have demonstrated that tapping a surface at a distance and hearing the sound as if coming from twice that distance produces a sense of arm elongation. Such studies have relied on elaborate speaker setups or virtual cues that could be simplified using binaural recordings played back using headphones. To this end, multiple sounds of tapping on different surfaces were recorded both inside an anechoic chamber and a room with natural reverberation using a binaural microphone dummy head and torso. Tappings were produced at incremental lateral intervals of 15 cm beginning in front of the dummy head. To identify binaural sound characteristics that retain distance distinguishability, we conducted a study with 11 participants. Wearing headphones, they tapped on six impact-sensitive pads placed 15 cm apart and heard each of the recorded sounds. Two experimental blocks were presented respectively to evaluate the perception of each sound on a 7-point Likert scale questionnaire, and to assess the perceived distance of the sound on a visual analogue scale. A final debriefing questionnaire was presented. Our findings contribute to optimizing tools for modulating body perception.

Keywords: *auditory distance perception, anechoic, reverberant, binaural, bodily self-perception.*

1. INTRODUCTION

Auditory spatial perception plays an important role in bodily self-awareness. For example, the sounds we generate through our own actions can influence how we perceive the size of our body [1–6] (for a review, see [7]). Recent research has demonstrated that altering the apparent distance of self-produced sounds can recalibrate the mental representation of certain body parts [1, 2, 4]. In one study, tapping sounds were artificially presented as if originating from double the actual tapping distance, which led participants to perceive their arm as longer or tactile distances on the arm as smaller [1]. These findings suggest that auditory distance cues inform an internal body-representation. However, achieving such spatial manipulations in past experiments often required elaborate setups, such as arrays of loudspeakers or virtual audio rendering systems. For instance, distributed loudspeakers have been used to shift the apparent location of action sounds [2]. Virtual acoustics

***Corresponding author**: *mrahmani@inf.uc3m.es.*

techniques have also been used to simulate sound at different distances, but demand real-time processing power and careful calibration [1,4]. A simpler alternative is to use binaural recordings of action sounds played back over headphones. Binaural playback can recreate spatial cues (e.g., interaural differences and reverberation) authentically for the listener, potentially inducing similar perceptual effects without specialized hardware [8,9].

Given this context, the present study explores whether binaural recordings of action sounds can effectively recreate perceived distance in a simpler headphone-based setup. We focus on two key factors: the acoustic environment (anechoic chamber vs. reverberant room) and the sound source material that is being tapped (a cardboard box vs. a wooden box). The motivation is that reverberation and source spectrum might interact: a spectrally rich sound (e.g., a wooden box tap with sharp transients) might provide strong distance cues in anechoic conditions but become "blurred" by reverberation, whereas a more well-rounded sound (cardboard box) might have fewer high-frequency cues to lose. By comparing conditions, we examine how well participants can discern differences in distance when hearing sounds played back using their own tapping at uniformly increasing distances. If successful, this method would confirm that altering spatial cues of action-produced sounds using pre-recorded binaural samples can influence one's perception of distance and space, without the need for complex multi-speaker arrays or real-time audio systems, opening the possibilities for conducting field or at-home experiments more easily.

For this purpose we ran an experiment where participants judged the perceived qualities of different sound materials and environments, and estimated the apparent distance of the sound source while tapping at different locations on a surface.

2. METHODS

2.1 Participants

11 participants (four female, seven male, mean age 33 years, range: 29-44) took part in the study, two were left-handed and the rest were right-handed. All were university students or faculty members and reported having normal hearing and touch. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki and was approved by the Research Ethics Committee of the Universidad Carlos III de Madrid. Informed consent was provided. All participants took part in both experimental blocks.

2.2 Sounds and Apparatus

Binaural recordings were conducted using a binaural dummy Head and Torso Simulator (HATS, G.R.A.S. Sound and Vibration) placed within two distinct environments: an anechoic chamber and a typical room with natural reverberation. Sounds of tapping were generated on two different surfaces: a cardboard box (originally packaging for a pair of Sennheiser headphones) and a wooden box. The choice of tapping surface material is based on prior research employing the sound of tapping on a wooden surface in [2], while the sound of tapping on a cardboard box was employed in [1,4].

For each combination of material and environment, the experimenters recorded the sound of tapping at six incrementally increasing lateral distances from the HATS (at 0, 15, 30, 45, 60, and 75 cm, measured from the HATS' torso midline); and an additional six sounds at twice the original distance (0, 30, 60, 90, 120, and 150 cm), covering a total of nine distinct sounds (0, 30, and 60 cm are shared). Note that these distances were chosen based on prior studies [1,2,4]. This yielded a set of binaural audio clips corresponding to taps at specific distances under two acoustic environment conditions of anechoic chamber (anechoic), and a room with natural reverberation (reverberant); and two tapping surface materials of wooden box (wood), and cardboard box (cardboard).

To play back the sounds in sync with the participants' actions, we constructed a custom apparatus with six small force-sensitive pads (molded silicone, 15×15×2 cm), each embedded with a piezoelectric sensor and covered with mesh fabric to prevent sticking. The pads were placed on a table in front and extending to the right side of the participant, spaced 15 cm apart (center-to-center). This physical layout of the pads mimicked the geometry used during recording, i.e., the HATS had the boxes positioned along a similar front-right line at those same distances. The shock-absorbing qualities of silicone covered with mesh made sure that very little unwanted sound is produced from tapping on the pads. Participants were seated facing forward. A red LED was placed in front at eye-level, such that when they fixated on it, their head orientation aligned with the orientation the HATS had during recording. This ensured that the binaural playback over headphones would correspond directionally to the participant's perspective (i.e., the sound from the right-front side would perceptually align with the pad location). Participants wore circumaural headphones (Sennheiser HD 200 Pro) through which the binaural tap sounds were delivered.

Each silicone pad was associated with one of the prerecorded binaural tap sounds. Tapping a pad with a knuckle would trigger the corresponding sound to play through the headphones with minimal latency. The piezo sensors detected the tapping impact using an Arduino microcontroller, connected to a computer interfacing with TouchDesigner software, which triggers the playback of the audio sample assigned to that pad's distance, material, and environment condition.

2.3 Experimental Design and Procedure

The experiment comprised two blocks presented sequentially. Both blocks had four types of auditory stimuli differing in terms of material (cardboard and wood), and environment (anechoic and reverberant).

2.3.1 Block 1 (qualitative sound assessment)

Participants tapped sequentially from pads 1 to 6 (located laterally at 15 cm apart from each other) at their own pace. The pads triggered sounds corresponding exactly to their recorded spatial positions. The tapping was repeated for each of the four sound conditions, i.e., anechoic cardboard, reverberant cardboard, anechoic wooden, and reverberant wooden. The conditions were presented in pseudo-randomized order. Participants completed a 3-item questionnaire after each of the four sound conditions, evaluating subjective qualities of the sounds. Participants rated each statement on a 1–7 Likert scale to assess different aspects of the auditory experience. The items included in Tab. 1 respectively referred to the sense of agency (Q1), spatial congruence (Q2), and perceived naturalness (Q3).

2.3.2 Block 2 (perceived distance assessment)

Sounds were presented as if originating from twice their actual recorded distances (e.g., second pad played the 30 cm recording instead of 15 cm). Participants sequentially tapped pads from 1 to 6 for each of the conditions. After each sequence, they marked along a visual analogue scale the perceived location of the sound source. This process was repeated for all four material$\times$environment combinations. The implemented visual analog scale (Fig. 1) was an unmarked ruler with a humanoid mannequin that mirrors the participant's posture and hand positions at the 6 tapping distances, which is used as a visual scale reference to estimate distance from their own body (starting at their torso midline, increasing outward to their right, with the ruler being at least 4 stretched-hand mannequins long).

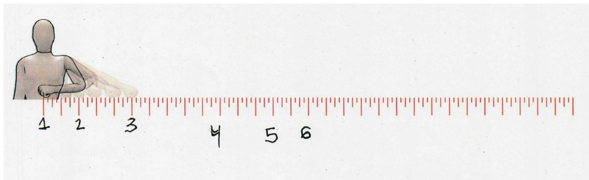

Figure 1. Visual analog scale for perceived distance.

3. DATA PROCESSING AND STATISTICAL ANALYSES

Data was processed using Python (numpy version 2.2.4) and statistical tests were performed using JASP (version 0.19.3). Due to the pilot nature of this experiment, the statistical analyses on a small sample have an exploratory purpose.

3.1 Block 1 - Qualitative Sound Assessment

Each of the three Questions (Q1, Q2, Q3) were evaluated individually. Medians and interquartile ranges (IQRs) for each question across conditions are shown in Tab. 1.

Separate 2×2 Analyses of Variance (ANOVAs) testing the effects of environment (anechoic, reverberant) and material (cardboard, wood) were conducted for each question. For Question 1, the analyses revealed a significant interaction between material and environment, $F(1,10) = 7.265$, $p = 0.022$. Post-hoc comparisons did not reveal any further significance.

For Question 3, a main effect of material was observed, $F(1,10) = 5.900$, $p = 0.036$; indicating that

Table 1. Survey Questions, Median summary of responses, and Interquartile Ranges (IQR).

	Q1: I felt that my hand was creating the sound I was hearing.		Q2: I felt the location of the sound was coming from my hand knocking on the surface.		Q3: I felt the sound I was hearing was natural.	
Condition	**Median**	**IQR**	**Median**	**IQR**	**Median**	**IQR**
anechoic–wood	2	1.5	4	2	2	1.5
anechoic–cardboard	2	2.5	3	2.5	4	3
reverberant–wood	3	3.5	3	2	2	1.5
reverberant–cardboard	2	3	4	2.5	3	3.5

participants perceived a significant difference in naturalness between the wood and cardboard sounds overall. Participants rated cardboard in anechoic environment as most natural (median = 4), and wood was rated least natural (median = 2) in both reverberant and anechoic environments.

ANOVA for Question 2 showed no significant main effects or an interaction. Although the median response from participants in Tab. 1 shows they generally rated the wood tap sounds as slightly more congruent with hand location (median = 4 in both environments), the effect was not statistically supported, which may reflect low variability in responses or insufficient sensitivity of the measure with the limited sample size ($N = 11$).

3.2 Block 2 - Perceived Distance Assessment

The perceived distances in Block 2 were measured from the participants' markings on the scale. The markings of each participant on the analog scale were normalized relative to the measurement of their furthest perceived distance, such that each participant had a maximum value of 1 for their farthest distance, and the rest are fractions of 1.

A repeated-measures (RM) ANOVA tested the main effects of environment (anechoic, reverberant), material (cardboard, wood), and distance (0, 30, 60, 90, 120, 150 cm), as well as their interactions on the average normalized distance. The first and last distances (0 and 150 cm) were not included because they are all 0 and 1 after normalization (though they are plotted in the figures for illustration). The RM ANOVA revealed a main effect of distance ($p < 0.001$, $F(3, 27) = 541.493$, $\eta^2 = 0.801$), indicating that participants clearly distinguished among the different distances. In contrast, the main effects of material and environment were not statistically significant. However, a statistically significant three-way interaction ($p = 0.020$, $F(3, 27) = 3.873$, $\eta^2 = 0.007$) emerged among material, environment, and distance, suggesting that perceptual distinguishability might depend on a combined influence of all three factors rather than any single factor alone.

In a post-hoc pairwise analysis of the material×environment interaction, a single comparison between reverberant wood and reverberant cardboard sounds exhibited a significant difference (Mean Difference = 0.084, $SE = 0.020$, $t(9) = 4.163$, $p = 0.015$).

All remaining comparisons (reverberant wooden vs. anechoic cardboard, reverberant cardboard vs. anechoic cardboard, and anechoic wooden vs. anechoic cardboard) were non-significant. These results indicate that the significant three-way interaction observed in the ANOVA is driven primarily by material differences in the reverberant environment, while differences in anechoic conditions or across environments appear negligible. It should be noted again that these results are exploratory given the small sample, and it is probable that more data would yield further significant findings.

To further explore the finer dynamics of overestimation and underestimation of perceived source location as a function of distance, a second-degree polynomial was fitted to the normalized perceived distances (0–150 cm), and the resulting slope at each point was compared to a 0.2 *baseline* slope derived from the actual distances (0, 30, 60, 90, 120, 150) in Tab. 2. A slope that remains close to 0.2 indicates that participants' perceived relative distances closely matches the actual relative distances. Fig. 2 shows that wood closely matched this 0.2 slope in the anechoic environment, suggesting high fidelity, whereas cardboard began below 0.2 (underestimation) and crossed above it around 40 cm (overestimation). Under reverberation, cardboard started below 0.2 before increasing, implying compression at near ranges and expansion at longer distances, while Wood showed the opposite trend (above 0.2 initially but decreasing with distance).

These findings indicate a strong divergence in perceptual distance estimation patterns between materials in reverberant environments. Wood shows high distinguishability at close range that diminishes with distance, while cardboard shows increasing perceptual spread, suggesting opposite bias mechanisms in auditory distance processing under reverberation.

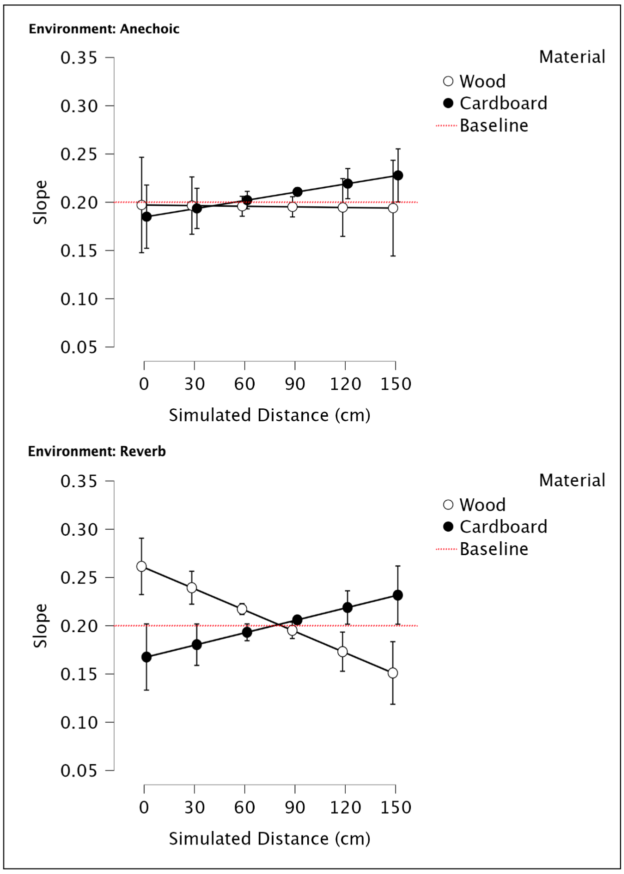

Figure 2. Mean slope of second-degree polynomial fits to normalized perceived distances (0–150 cm) for anechoic (*top*) and reverberant (*bottom*) environments.

4. DISCUSSION

Our exploratory pilot study aimed to assess how different materials and acoustic environments might influence the perceived distance of a sound source while tapping on a surface. Although this is a pilot study reliant on a small sample ($N = 11$, with one participant excluded from Block 2 analyses), our findings support the continuation of research using specific sound qualities, but our interpretations should be considered with caution.

Block 1 evaluated the subjective qualities of the sounds while interacting with distance-matched auditory feedback. The sense of agency, defined as the feeling of oneself producing the sensory consequences

Table 2. Standard deviation from baseline 0.2 slope in Fig. 2, represented by Root Mean Squared Error (RMSE) and Mean Absolute Deviation (MAD).

Condition	MAD	RMSE
anechoic–wood	0.00433	0.00447
anechoic–cardboard	0.01350	0.01597
reverberant–wood	0.03317	0.03826
reverberant–cardboard	0.01933	0.02198

(i.e., the heard sound), was significantly influenced by the interaction between material and environment.

Post-hoc comparisons revealed that wood in a reverberant environment was rated highest for agency (Q1), though results did not survive correction for multiple comparisons. Agency has a direct coupling with linking an action to a perceived consequence, which is highly influential on bodily self-perception [1, 10, 11]. Notably, wood-reverberant sounds also had the largest interquartile range in Tab. 1, suggesting individual variability, but also the potential to strongly elicit the impression of self-generated action.

The perceived naturalness of the sounds (Q3) seemed to be influenced by material alone, with cardboard in an anechoic environment, followed by reverberant, receiving the highest ratings. This could suggest that the inherent spectral characteristics of cardboard sounds closely match listeners' real-world expectations for natural sound. It is also possible that the act of tapping on soft silicone pads to trigger the sound influenced these expectations, leading participants to associate the tactile softness with more natural acoustic properties. Spatial congruence (Q2), or the perception that the sound originated from the participant's hand location, did not significantly differ across conditions.

In Block 2, we investigated participants' ability to perceive differences in sound distance under conditions where all sounds were presented from double the physical tapping distance. Fig. 2 and Tab. 2 results showed that the anechoic wooden box sound preserved distance separation most consistently. Similarly, the anechoic cardboard sound showed consistent perceived separation closer to the listener, but slightly increasing perceived separation at farther distances.

In contrast, reverberant wood sounds showed significant compression of perceived distance, with the smallest mean slope values at the farthest distances and a clear flattening pattern as distances increased. Conversely, the cardboard-reverberant sound behaves the opposite way, the gaps between sounds are perceived as smaller in the near space and increase as the sounds are presented from further away. The graphs of both wooden and cardboard reverberant sounds cross the 0.2 baseline at around 80 cm.

Our findings suggest that distance perception of sounds in anechoic environments is more consistent, an attribute that could be important when aiming to enhance the perceived elongation of limbs while tapping [1, 2, 4]. However, reverberant environments may be useful for settings in which greater ecological validity is desired, even if relative distance separation is compromised.

At large, our study suggests that binaural recordings can provide sufficient spatial fidelity for experimental manipulations of body perception using sound [2, 3, 12, 13].

5. ACKNOWLEDGMENTS

This project was funded by the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement number 101002711). We also thank the i_mBODY Lab for their support and collaboration.

6. REFERENCES

[1] A. Tajadura-Jiménez, M. Tsakiris, T. Marquardt, and N. Bianchi-Berthouze, "Action sounds update the mental representation of arm dimension: contributions of kinaesthesia and agency," *Frontiers in Psychology*, vol. 6, May 2015.

[2] A. Tajadura-Jiménez, A. Väljamäe, I. Toshima, T. Kimura, M. Tsakiris, and N. Kitagawa, "Action sounds recalibrate perceived tactile distance," *Current Biology*, vol. 22, pp. R516–R517, July 2012.

[3] A. Tajadura-Jiménez, M. Basia, O. Deroy, M. Fairhurst, N. Marquardt, and N. Bianchi-Berthouze, "As Light as your Footsteps: Altering Walking Sounds to Change Perceived Body Weight, Emotional State and Gait," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, (Seoul Republic of Korea), pp. 2943–2952, ACM, Apr. 2015.

[4] A. Tajadura-Jiménez, T. Marquardt, D. Swapp, N. Kitagawa, and N. Bianchi-Berthouze, "Action Sounds Modulate Arm Reaching Movements," *Frontiers in Psychology*, vol. 7, Sept. 2016.

[5] A. Tajadura-Jiménez, M. Vakali, M. T. Fairhurst, A. Mandrigin, N. Bianchi-Berthouze, and O. Deroy, "Contingent sounds change the mental representation of one's finger length," *Scientific Reports*, vol. 7, p. 5748, July 2017.

[6] A. Tajadura-Jiménez, O. Deroy, T. Marquardt, N. Bianchi-Berthouze, T. Asai, T. Kimura, and N. Kitagawa, "Audio-tactile cues from an object's fall change estimates of one's body height," *PLOS ONE*, vol. 13, p. e0199354, June 2018.

[7] T. R. Stanton and C. Spence, "The Influence of Auditory Cues on Bodily and Movement Perception," *Frontiers in Psychology*, vol. 10, p. 3001, Jan. 2020.

[8] A. J. Kolarik, B. C. J. Moore, P. Zahorik, S. Cirstea, and S. Pardhan, "Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss," *Attention, Perception, & Psychophysics*, vol. 78, pp. 373–395, Feb. 2016.

[9] B. G. Shinn-Cunningham, S. Santarelli, and N. Kopco, "Tori of confusion: Binaural localization cues for sources within reach of a listener," *The Journal of the Acoustical Society of America*, vol. 107, pp. 1627–1636, Mar. 2000.

[10] S. Gallagher, "Philosophical conceptions of the self: implications for cognitive science," *Trends in Cognitive Sciences*, vol. 4, pp. 14–21, Jan. 2000.

[11] Roel Lesur M., Gaebler M., Bertrand P., and Lenggenhager B., "The Plasticity of the Bodily Self: Head Movements in Bodily Illusions and Their Relation to Gallagher's Body Image and Body Schema," *Constructivist Foundations*, vol. 14, no. 1, pp. 94–105, 2018.

[12] M. Roel Lesur, E. Bolt, G. Saetta, and B. Lenggenhager, “The monologue of the double: Allocentric reduplication of the own voice alters bodily self-perception,” *Consciousness and Cognition*, vol. 95, p. 103223, Oct. 2021.

[13] A. Tajadura Jiménez, M. Fairhurst, and O. Deroy, “Sensing the body through sound,” in *The Routledge Handbook of Bodily Awareness*, pp. 230–246, London: Taylor & Francis, 2022.