

Finding My Way: A Comprehensive Evaluation of Auditory Cues for Outdoor Augmented Reality Navigation

Sina Hinzmann

sina.hinzmann@hshl.de

Hamm-Lippstadt University of Applied Sciences

Francesco Vona

Hamm-Lippstadt University of Applied Sciences

Tanja Kojic

Technische Universität Berlin

Juliane Henning

Hamm-Lippstadt University of Applied Sciences

Mohamed Amer

Hamm-Lippstadt University of Applied Sciences

Jan-Niklas Voigt-Antons

Hamm-Lippstadt University of Applied Sciences

Research Article

Keywords: Augmented Reality, Navigation, User Experience, Perceived Effectiveness, Auditory Cues

Posted Date: October 26th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11218193/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Finding My Way: A Comprehensive Evaluation of Auditory Cues for Outdoor Augmented Reality Navigation

Sina Hinzmann^{1*}, Francesco Vona¹, Tanja Kojic²,
Juliane Henning¹, Mohamed Amer¹, Jan-Niklas Voigt-Antons¹

^{1*}Immersive-Reality-Lab, Hamm-Lippstadt University of Applied
Sciences, Marker Allee 76-78, Hamm, 59063, Germany.

²Quality and Usability Lab, Technische Universität Berlin, Straße des
17. Juni 135, Berlin, 10623, Germany.

*Corresponding author(s). E-mail(s): sina.hinzmann@hshl.de;
Contributing authors: francesco.vona@hshl.de; tanja.kojic@tu-berlin.de;
juliane.henning@hshl.de;
mohamed-ahmed-mohamed-ali.amer@stud.hshl.de;
jan-niklas.voigt-antons@hshl.de;

Abstract

Auditory cues can support navigation in augmented reality (AR) without requiring continuous visual attention. However, different cue types may vary in their perceived usefulness, and individual preferences and user characteristics may further shape the user experience. This study evaluates five categories of auditory cues—Artificial Sounds, Nature Sounds, Spearcons, Musical Instruments, and Auditory Icons—in an outdoor AR navigation task using a Meta Quest 3. Fifteen participants evaluated each cue category using the User Experience Questionnaire (UEQ) and five task-specific questions. In addition to comparing cue categories, the extended analysis examines individual cue preferences, participant characteristics, and open-ended user feedback. Individually preferred cues were compared with non-preferred alternatives across the evaluation measures. Exploratory mixed-effects analyses investigated associations with age, gender, prior XR experience, and affinity for technology interaction. Qualitative responses were analyzed to identify recurring positive and negative experiences and suggestions for improvement. The results confirm differences between cue categories in UEQ Novelty and Stimulation. Preferred cues received significantly higher ratings for attractiveness, perspicuity, efficiency, novelty,

stimulation, and perceived effectiveness. Exploratory analyses further identified limited associations between participant characteristics and cue evaluations, including an age-by-cue interaction for perceived directional clarity and an association between technology-interaction affinity and perceived novelty. Qualitative findings revealed cue-specific patterns of recurring themes across positive and negative perceptions, as well as suggested improvements.

Keywords: Augmented Reality, Navigation, User Experience, Perceived Effectiveness, Auditory Cues

1 Introduction

Auditory cues offer a flexible means of conveying navigation information in augmented reality (AR), ranging from abstract sounds to more familiar or semantically meaningful cues. By presenting spatialized sounds within the surrounding environment, audio AR can communicate directions and destinations while allowing users to maintain visual attention on their surroundings [17, 20, 28]. This approach is relevant to mobile interaction, multitasking, and accessible navigation, where the suitability of visual guidance depends on users' abilities and situational demands [9, 24, 30].

The design of auditory guidance involves decisions about both how directional information is conveyed and which sounds represent navigation instructions. Spoken descriptions can support spatial learning and wayfinding [5], while spatialized sound has demonstrated advantages over verbal directional cues under concurrent cognitive load [33]. Non-speech cues, including artificial tones, natural sounds, musical signals, and auditory icons, provide additional possibilities through learned mappings or associations with familiar sound-producing events [1, 12, 13]. Spearcons, generated by compressing spoken phrases, offer a speech-derived alternative with demonstrated benefits for auditory menu navigation and cue learnability [3, 25, 26]. However, these findings do not establish which cue types provide the most favorable experience during outdoor navigation.

Beyond accurately conveying instructions, navigation sounds contribute to the perceived quality of the interaction. Previous work has examined intuitiveness, learnability, memorability, and preference in auditory interfaces [27], while outdoor audio AR studies have identified benefits for immersion and hedonic experience [7, 34]. Research on natural sounds also suggests that their affective properties warrant consideration, although findings from laboratory stress-recovery tasks cannot be directly generalized to brief navigation cues [23]. These perspectives motivate the evaluation of perceived clarity, usefulness, and pleasantness alongside navigation performance, since functional benefits demonstrated in one setting may not predict users' preferences in another.

Existing research has compared navigation beacon sounds, explored continuous musical guidance, and evaluated spatial auditory navigation in outdoor environments [29, 31, 32, 34]. These studies establish the potential of auditory guidance

but leave open the question of how users experience different non-speech and speech-derived cue sets when evaluated within a common outdoor AR navigation protocol. A comparison across several cue categories can therefore complement research focused on individual sound designs, navigation modalities, or spatial rendering techniques.

The present study examines five auditory cue categories—Artificial Sounds, Nature Sounds, Spearcons, Musical Instruments, and Auditory Icons—using an AR application developed for the Meta Quest 3 headset that provides audio-only navigation guidance. A pre-study identified user-preferred sounds for typical navigation scenarios, and the selected stimuli were subsequently evaluated in a within-subject outdoor navigation experiment. The evaluation focuses on perceived clarity, usefulness, pleasantness, overall user experience, and cue preference, complemented by navigation performance measures, qualitative feedback, and an analysis of participant characteristics. By comparing the five selected cue sets using a common platform and evaluation procedure, the study aims to inform sound selection for outdoor audio AR navigation.

2 Related Work

2.1 Auditory Cue Design and Learnability

Auditory interfaces can convey information when users cannot continuously attend to a visual display. Foundational research distinguishes between earcons, which communicate through structured abstract sound patterns, and auditory icons, which draw on associations between everyday sounds and the events they represent [1, 12]. An ecological account of auditory perception further emphasizes the perception of sound-producing events rather than acoustic properties alone [2]. These approaches provide complementary ways of designing navigation cues: users may learn an arbitrary sound–instruction association or interpret a cue through an existing association with its source. Consequently, recognizing a sound and understanding its intended navigational meaning are distinct requirements.

Empirical comparisons demonstrate the importance of these mappings. Research on hierarchical earcons investigated how they communicate a user’s position within a menu structure, establishing their potential for conveying structured information through learned acoustic patterns [4]. A longitudinal evaluation of mobile service notifications compared auditory icons and earcons across intuitiveness, learnability, memorability, and preference, favoring auditory icons on these measures [27]. However, the task involved identifying notifications rather than navigating physical space, so the findings do not establish a general advantage for auditory icons in pedestrian guidance.

Spearcons provide a speech-derived alternative. These cues were introduced as compressed spoken phrases, potentially shortened until they are no longer recognizable as speech. An initial menu study found that spearcons and speech supported faster and more accurate navigation than auditory icons and hierarchical earcons [3]. Subsequent experiments extended this evidence to learning and more complex auditory menus [26]. An additional comparison of cues representing environmental features found that spearcons were approximately as learnable as speech, whereas earcons required more learning [25]. These findings motivate including spearcons in navigation

studies, while leaving open whether their benefits for recognition and menu interaction extend to subjective experience during outdoor walking.

These distinctions also clarify the status of the five conditions examined here. Artificial Sounds, Nature Sounds, and Musical Instruments primarily describe sound sources or acoustic character, whereas auditory icons are defined by their representational relationship and spearcons by their speech-based construction. The categories therefore constitute operational stimulus groups rather than a universally exclusive taxonomy. A natural sound can function as an auditory icon if its association with the represented event is meaningful; similarly, an instrumental sound may form part of an earcon. This distinction is relevant when interpreting differences between the selected cue sets.

2.2 Spatial Audio for Navigation and Cognitive Support

Physical navigation introduces requirements beyond recognizing an instruction, including estimating direction, approaching waypoints, and monitoring the surrounding environment. AudioGPS explored non-speech spatial audio as a means of communicating direction and distance while reducing reliance on visual interaction [28]. Subsequent research investigated virtual navigation using different beacon sounds, waypoint capture radii, and levels of practice. The results showed that navigation depended on these design choices and their interactions, highlighting that cue effectiveness cannot be attributed to sound identity alone [29]. SWAN extended this line of work into a wearable system using non-speech sonification to communicate routes and environmental features [30].

Comparisons with language-based guidance provide a more specific account of attentional benefits. One study compared spatialized sound and verbal direction cues during virtual path following with and without a concurrent memory task [33]. Under the additional cognitive load, spatialized guidance supported better navigation and secondary-task performance; the navigation modes did not differ significantly in the no-load condition. Conversely, research on dynamically updated verbal descriptions demonstrated that they can support nonvisual learning and wayfinding in indoor layouts [5]. Together, these studies indicate that the relative benefits of spatial sound and language depend on the task and attentional demands, rather than implying that verbal guidance is intrinsically ineffective.

Research has also considered spatial learning beyond immediate route completion. An outdoor scavenger-hunt study compared auditory beacon navigation with turn-by-turn guidance [32]. Among participants unfamiliar with the environment, beacon navigation supported more accurate spatial representations on subsequent tests. However, participants retained access to vision, and the comparison concerned navigation strategies rather than categories of sound. This work therefore motivates distinguishing successful guidance, environmental learning, and user experience as related but separate outcomes.

2.3 Musical Guidance and Outdoor Audio Augmented Reality

Musical approaches show that navigational information can be embedded within an ongoing listening experience. ONTRACK continuously modified the spatial balance and volume of music to guide users toward destinations, with laboratory evaluation and field trials demonstrating the feasibility of the approach [31]. This differs from presenting short instrumental cues for individual instructions: continuous musical guidance combines the properties of the music with a persistent directional signal. Its results support investigating musical navigation, but cannot directly predict preferences for isolated instrument sounds.

Outdoor audio AR studies further demonstrate the importance of evaluating experience in context. A sound-garden study compared four auditory displays, varying earcon landmarks and spatialization [34]. Spatial audio combined with earcons supported exploration of simultaneous sound sources and increased immersion, illustrating that experiential benefits can accompany changes in exploratory behavior. An investigation of head, body, and device orientation for mobile audio AR found that tracking requirements depend on proximity to sound sources and the usage scenario [17]. These studies show that rendering and interaction design must be considered alongside the choice of cue.

Closer to the present setting, a comparison of spatial auditory and visual AR navigation outdoors reported higher hedonic quality for the auditory condition [7]. This provides a direct rationale for assessing experiential qualities, although it compares navigation modalities rather than multiple sound categories within an auditory interface. Further research examined environmental acoustic reproduction and playback devices in audio AR, addressing how virtual sounds are experienced in their listening context [21]. This focus on acoustic plausibility and presentation complements the question of which cue sets users find clear, pleasant, and useful for navigation.

Recent work also combines spatial cues with contextual spoken guidance. Navig-AI-tion pairs landmark-based instructions generated using a vision-language model with corrective spatial audio. A preliminary evaluation reported fewer route deviations with the combined system than with either model-generated instructions alone or an audio-only mapping baseline [35]. This evaluates a combination of guidance components, leaving the comparative experience of different cue categories as a separate design question.

3 Methods

3.1 Experimental Design

We employed a within-subjects experimental design where each participant completed five audio-only navigation tasks, each guided by a different sound cue category. The order of conditions was randomized using a Latin Square design to control for order effects. For each condition, participants were guided through a unique navigation route, ensuring that no path was repeated and that route familiarity did not bias the results. A subjective evaluation was obtained through post-task questionnaires, which

included the User Experience Questionnaire (UEQ) [15, 16] as well as custom questions developed for the present experiment. Participants responded to the statements: “I had a positive overall navigation experience with the audio cues (Q1),” “I was able to clearly perceive the direction of the audio cues (Q2),” “I understood the meaning of the audio cues (Q3),” “The audio cues helped me navigate effectively (Q4),” and “I found the audio cues pleasant to use (Q5).” If a task exceeded six minutes, the condition was marked as failed. In addition, three open-ended questions were included to elicit qualitative feedback: “What did you like about this type of audio cue?”, “What did you not like about this type of audio cue?”, and “How could this type of audio cue be improved for more efficient or pleasant navigation?”. After all conditions, participants were asked to indicate their preferred sound condition by selecting one of the five sound cues. The study received ethical approval from the local ethics committee.

3.2 Pre-Study: Identifying Preferred Navigation Sounds

Before the main study, we conducted a pre-study with 31 participants to identify the most intuitive and user-preferred sounds for each cue category. Participants were seated in a quiet lab environment and wore over-ear headphones connected to a desktop interface. They were presented with four prototypical navigation scenarios: i) “Turn left at the next intersection”, ii) “Obstacle ahead – change direction”, iii) “Continue straight for 100 meters”, iv) “You have reached your destination”.

For each scenario, participants listened to 15 audio cues (3 from each of the 5 categories: Artificial Sounds, Nature Sounds, Spearcons, Musical Instruments, and Auditory Icons), presented in randomized order. These categories were selected because they reflect the wide diversity in prior auditory interface research. After each group of three within a category, participants were asked to vote for the sound they believed was most effective and intuitive for the given instruction. The highest-rated sounds in each category formed the final stimulus set for the main study.

3.3 Participants

A total of 20 participants were initially recruited for the study. Five participants were excluded from the analyses for not completing at least two experimental conditions to allow for meaningful within-subject comparisons. The final sample thus consisted of 15 participants (3 female, 12 male), with ages ranging from 20 to 45 years ($M = 28.33$, $SD = 8.37$). Participants reported their prior experience with virtual, augmented, and mixed reality technologies: none ($n = 4$), little ($n = 6$), some ($n = 2$), and extensive ($n = 3$). The mean Affinity for Technology Interaction (ATI) score was $M = 4.27$ ($SD = 0.92$, $\alpha = 0.85$), indicating a moderate to moderately high level of technology affinity and a good internal consistency.

3.4 Equipment and Setup

The study took place in a controlled outdoor area on the university campus with low pedestrian and no vehicle traffic. The setting provided realistic conditions for auditory navigation, including typical ambient noise and natural acoustics. To focus solely on sound perception and avoid potential motion sickness, participants wore the

Table 1 Mean and standard deviation for each sound condition on the six UEQ subscales and five custom questionnaire items.

UEQ						
Condition	Attractiveness	Perspicuity	Efficiency	Novelty	Stimulation	Dependability
Artificial Sounds	1.26 (1.17)	1.05 (1.44)	1.20 (1.07)	0.85 (0.80)	1.17 (1.23)	0.76 (1.29)
Auditory Icons	0.72 (1.11)	0.89 (1.44)	0.83 (1.05)	0.53 (0.92)	0.58 (1.02)	0.58 (0.94)
Music Instruments	1.19 (1.13)	0.83 (1.39)	0.78 (1.21)	0.98 (0.65)	1.05 (1.03)	0.76 (1.22)
Nature Sounds	0.81 (1.27)	0.41 (1.67)	0.61 (1.38)	0.96 (1.10)	0.86 (1.00)	0.41 (1.14)
Spearcons	0.48 (1.28)	1.33 (1.13)	1.08 (1.25)	-0.36 (1.06)	0.26 (1.48)	0.98 (1.36)

Custom Questions					
Condition	Q1	Q2	Q3	Q4	Q5
Artificial Sounds	5.13 (1.85)	5.00 (1.81)	5.60 (1.45)	5.21 (1.63)	5.27 (1.22)
Auditory Icons	4.86 (1.75)	5.14 (1.96)	5.21 (1.76)	5.29 (1.33)	4.43 (1.60)
Music Instruments	5.20 (1.70)	5.20 (1.93)	4.93 (1.75)	5.53 (1.36)	5.27 (1.62)
Nature Sounds	4.47 (1.92)	4.60 (1.76)	4.33 (1.84)	5.00 (1.65)	5.20 (1.74)
Spearcons	4.87 (1.55)	5.00 (1.46)	6.07 (1.22)	5.40 (1.59)	4.33 (1.54)

Meta Quest 3 headset on their forehead, using it only for head tracking without visual display. Open-back high-fidelity headphones were used to enhance spatial accuracy and audio fidelity. A custom AR application, developed in Unity and deployed on the headset, guided the user along predefined paths using spatialized audio cues.

Each navigation route consisted of five sequential checkpoints, spaced approximately 11 meters apart, resulting in a total length of around 55 meters. All participants began at the same orientation, and no route was reused within a participant’s session to avoid familiarity effects.

3.5 Procedure

The full session lasted approximately 60 minutes and followed this structure: i) Onboarding, consent, and demographics questionnaires, ii) Navigation tasks and Post-task questionnaires (repeated for each condition), iii) Debriefing and final questionnaire.

For each navigation task, the participant follows five audio cues to five checkpoints, all from the same cue category. After completing the route, the participant removes the headset and completes a questionnaire assessing their experience. This process is repeated for all five cue categories.

4 Results

4.1 Quantitative Results

Due to the relatively small sample size, all analyses employed non-parametric statistics. Results are reported below for the comparison of sound conditions and the subsequent analysis of individual cue preferences.

4.1.1 Comparison of Sound Conditions

User experience ratings were compared across the five sound conditions—Artificial Sounds, Nature Sounds, Spearcons, Musical Instruments, and Auditory Icons—on

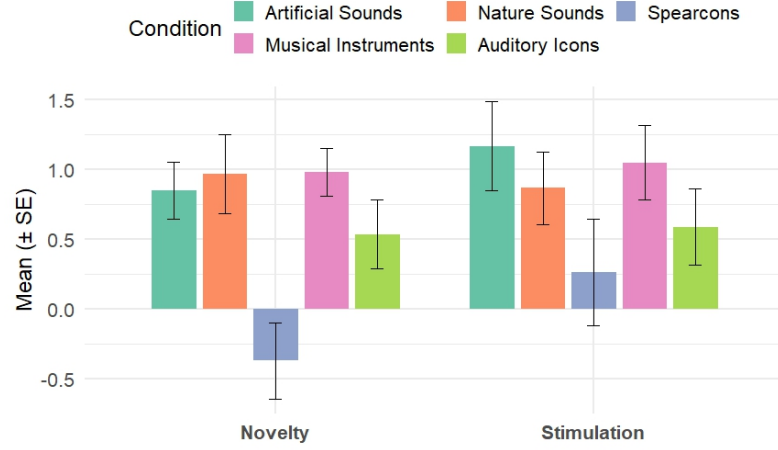


Fig. 1 Mean user ratings ($\pm$ SE) for the UEQ subscales *Novelty* and *Stimulation* across the five sound conditions. Artificial Sounds and Musical Instruments were rated significantly higher than Spearcons for *Novelty*, while Artificial Sounds also scored higher than Spearcons on *Stimulation*.

the six UEQ scales: *Attractiveness*, *Perspicuity*, *Efficiency*, *Novelty*, *Stimulation*, and *Dependability*. Friedman tests revealed significant effects for the scales *Novelty* ($\chi^2(4) = 14.20$, $p = .007$) and *Stimulation* ($\chi^2(4) = 9.52$, $p = .049$). Post hoc Wilcoxon signed-rank tests with Holm correction showed that both Artificial Sounds and Musical Instruments were rated significantly higher than Spearcons on *Novelty* ($p_{\text{adj}} = .023$ for both), while Artificial Sounds also scored significantly higher than Spearcons on *Stimulation* ($p_{\text{adj}} = .031$; see Fig. 1). Although the pairwise comparison between Nature Sounds and Spearcons on the *Novelty* scale was statistically significant before correction ($p = .006$), the result did not remain significant after Holm correction ($p_{\text{adj}} = .051$). No significant differences emerged for the other UEQ scales.

For the five custom questions (Q1–Q5), a Friedman test identified a significant effect for Q3 ($\chi^2(4) = 9.90$, $p = .042$), though subsequent pairwise Wilcoxon signed-rank tests with Holm correction revealed no significant differences between conditions.

4.1.2 Cue Preference Analysis

Regarding overall preference, participants' final votes showed a tie between Nature Sounds ($n = 5$) and Artificial Sounds ($n = 5$), with Musical Instruments the next most preferred ($n = 3$), followed by Auditory Icons ($n = 1$). Spearcons received no votes.

To examine the subjective characteristics associated with participants' final cue preferences, a paired within-participant delta analysis was conducted. Final preference data were available for $N = 14$ participants; for Q4, one additional response was missing, resulting in $N = 13$ for this comparison. For each participant, ratings of the selected preferred cue condition were compared with the mean ratings across all non-preferred alternatives.

Non-parametric paired Wilcoxon signed-rank tests were performed across five custom evaluation items and the six UEQ-scales. To control the family-wise error rate

across all $k = 11$ comparisons, p -values were adjusted using Holm's correction ($k = 11$). Descriptive statistics, Wilcoxon test statistics, Holm-adjusted p -values, and effect sizes are summarized in Table 2.

Among the custom items, the preferred cues received significantly higher ratings of perceived effectiveness (Q4; see Table 2), remaining significant after Holm adjustment. Perceived pleasantness (Q5) also showed higher ratings for preferred than for non-preferred cues; however, this difference did not remain statistically significant after Holm correction. Ratings of the other custom items did not differ significantly between preferred and non-preferred alternatives.

Across the UEQ dimensions, the preferred cues received significantly higher ratings of Attractiveness, Perspicuity, Efficiency, Novelty, and Stimulation (all $p_{\text{holm}} < .05$; see Table 2). Dependability also showed higher ratings for preferred cues, but this difference did not remain significant after Holm correction.

Table 2 Paired Within-Participant Delta Analysis Between Preferred and Non-Preferred Alternatives

Category	Dim.	Med. P.	Med. NP.	Med. Δ	V	p	p_{holm}	Effect r
Custom Items	Q1	5.50	4.75	+1.12	64.0	.207	.621	.370
	Q2	5.50	4.75	+1.00	73.0	.207	.621	.346
	Q3	6.00	5.12	+1.00	60.5	.309	.621	.295
	Q4	7.00	5.00	+1.25	88.5	.003	.026*	.835
	Q5	6.00	4.62	+1.38	79.0	.021	.083	.622
UEQ Scales	Attractiveness	1.92	0.52	+0.96	105.0	.001	.012*	.881
	Perspicuity	2.00	0.75	+1.31	96.0	.007	.048*	.730
	Efficiency	1.50	0.69	+0.78	103.5	.002	.015*	.856
	Novelty	1.25	0.34	+0.69	87.0	.004	.033*	.789
	Stimulation	1.50	0.53	+0.75	95.0	.008	.04995*	.714
	Dependability	1.75	0.50	+0.97	92.0	.014	.071	.664

Note. Med. P. = Median Preferred; Med. NP. = Median Non-Preferred; Δ = Median within-participant difference (Preferred minus Non-Preferred); V = Wilcoxon test statistic; p = uncorrected p -value; p_{holm} = Holm-adjusted p -value; r = Wilcoxon effect size ($Z/\sqrt{N}$) * $p_{\text{holm}} < .05$.

4.2 Analysis of Participant Characteristics

Friedman and Wilcoxon tests compare conditions within participants and cannot include between-subject factors. To examine whether participant characteristics were related to the evaluations, we therefore used mixed-effects models with sound condition and the position of the condition within the session as fixed effects and a random intercept per participant. UEQ scale scores were modeled with linear mixed models and the 7-point custom items (Q1 to Q5) with cumulative link mixed models for ordinal data. Age, gender, prior XR experience (none/little, $n = 10$, vs. some/extensive, $n = 5$), and ATI were tested in separate models because they were confounded; age

Table 3 Exploratory associations between participant characteristics and study outcomes ($N = 15$). (A) Rating outcomes showing effects significant after Holm correction. For Age and ATI, estimates are reported per SD; coefficients for Q1–Q5 are on the logit scale and UEQ coefficients are in scale points. (B) Performance and preference outcomes. p -values are Holm-adjusted within each analysis family. * $p_{\text{holm}} < .05$.

Outcome	Characteristic	Effect	Estimate [95% CI]	χ^2	p_{holm}
<i>(A) Rating outcomes</i>					
Q2 Direction perceived	Age $\times$ condition	Overall interaction		17.98	.015*
Nature Sounds	Age	Condition-specific slope	-1.60 [-2.59, -0.61]		.008*
Novelty	ATI	Main effect	-0.34 [-0.53, -0.14]	9.63	.047*
<i>(B) Performance and preference outcomes</i>					
Completion time	Age				.636
	Gender				.132
	XR experience				.636
	ATI				.295
Task failure	Age				.940
	Gender				1.000
	XR experience				.692
	ATI				.674
Final preference	Age				1.000
	Gender				1.000
	XR experience				1.000
	ATI				1.000

and ATI were z -standardized. For each characteristic and outcome, the main effect and the interaction with sound condition were tested with likelihood ratio statistics. As the χ^2 approximation of these tests is too liberal in small samples, p -values were obtained by permutation tests: the characteristic was permuted across participants, using all possible relabelings for gender (455) and XR experience (3,003) and 2,000 random permutations for age and ATI; p -values close to the significance threshold were recomputed with up to 100,000 permutations. Consistent with the analyses above, p -values were Holm-corrected across the eleven outcomes per characteristic and effect. Significant interactions were followed up by estimating the slope of the characteristic within each condition (Wald tests, Holm correction across the five conditions), and significant effects were checked with leave-one-participant-out analyses. Completion time of successful routes (log-transformed) was analyzed with the same model and permutation test, task failure (participants with at least one failed route) with Fisher’s exact and Mann-Whitney U tests, and the final preference with Fisher’s exact (Monte Carlo) and Kruskal-Wallis tests. As a sensitivity analysis, all tests were repeated with all 20 participants. Given the sample size, these analyses are exploratory.

Of the 88 tests of participant characteristics, two effects remained significant after correction (Table 3A, Fig. 2).

Age interacted with sound condition for perceived direction (Q2; $\chi^2(4) = 17.98$, $p = .001$, $p_{\text{holm}} = .015$; Fig. 2b). Older participants perceived the direction of Nature Sounds less clearly (slope = -1.60 logits per SD of age, 95% CI [-2.59, -0.61], $p_{\text{holm}} = .008$; Fig. 2c), whereas the slopes for the other conditions did not differ significantly

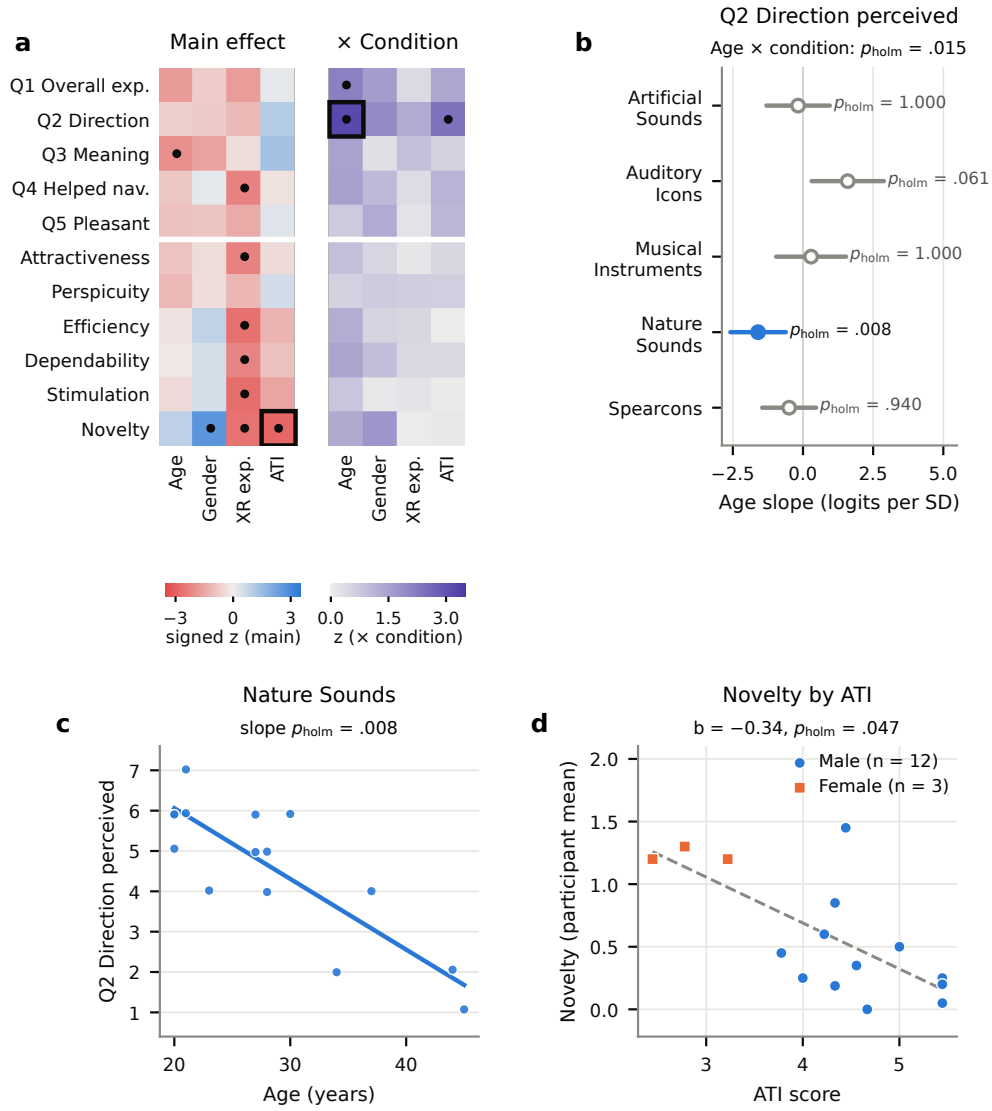


Fig. 2 Participant characteristics and ratings ($N = 15$). (a) Overview of all tests: permutation p -values expressed as z -values, signed by the direction of the main effect (red: lower ratings with higher age or ATI, or for female or experienced participants; blue: higher ratings) and unsigned for the characteristic × condition interaction. Outlined: $p_{\text{holm}} < .05$; dot: uncorrected $p < .05$. (b) Slope of age on Q2 Direction perceived within each condition (logits per SD, 95% CI; filled: $p_{\text{holm}} < .05$). (c) Q2 Direction perceived ratings of Nature Sounds by age (points: participants, slightly jittered; line: linear trend for illustration). (d) Novelty by ATI (participant means over the five conditions; dashed line: linear trend). Marker: gender.

from zero (Auditory Icons: 1.59, $p_{\text{holm}} = .061$; all others $p_{\text{holm}} \geq .94$). The Nature Sounds effect remained significant in all 15 leave-one-participant-out analyses. In the

full sample of 20 participants, the interaction pointed in the same direction but was not significant after correction ($p = .008$, $p_{\text{holm}} = .084$).

Higher ATI was associated with lower Novelty ratings across conditions ($b = -0.34$ per SD, 95% CI $[-0.53, -0.14]$, $\chi^2(1) = 9.63$, $p = .004$, $p_{\text{holm}} = .047$; Fig. 2d). This effect remained significant in all leave-one-participant-out analyses, but it lay close to the threshold, was not observed in the full sample ($p = .156$), and overlapped with gender and XR experience: female participants and participants with less XR experience also rated Novelty higher before correction (gender: $b = 0.80$, $p = .009$, $p_{\text{holm}} = .097$; XR experience: $b = -0.64$, $p = .010$, $p_{\text{holm}} = .090$). When the three characteristics were entered jointly, none retained a significant unique association (all $p \geq .07$).

Participants with some or extensive XR experience also rated the cues lower on Attractiveness, Efficiency, Dependability, and Stimulation ($b = -0.80$ to -1.07 scale points) and on perceived effectiveness (Q4) before correction ($p = .007$ to $.027$), but none of these effects remained significant after correction ($p_{\text{holm}} \geq .073$), and none depended on the sound condition. Completion time, task failure, and final preference were not significantly related to any participant characteristic after correction (Table 3B).

4.3 Qualitative Analysis of Open-Ended Responses

To complement the quantitative evaluation, the open-ended responses collected after each navigation condition were analyzed using thematic analysis. A common coding framework was applied across all five cue categories. The analysis focused on three aspects: (i) positive perceptions of the auditory cues, (ii) negative perceptions or difficulties encountered during navigation, and (iii) suggestions for improvement. Multiple themes could be identified within a single response, and themes were not considered mutually exclusive. Missing responses and equivalent entries were excluded from thematic coding. The analysis represents a single-coder thematic analysis and was intended to provide a structured descriptive summary of participants' subjective responses rather than a standalone qualitative study. Accordingly, no inter-rater reliability was calculated or claimed.

A theme was considered recurring if it was expressed by at least two participants within a given cue category. Each participant was counted only once per theme and cue category, even if the theme was mentioned multiple times.

The qualitative responses revealed several recurring patterns that differed across cue categories. For **Spearcons**, positive comments primarily concerned clarity, directness, and ease of use. At the same time, participants reported reduced intelligibility, fast pacing, low volume, and a lack of continuous feedback. Suggested improvements therefore focused on increasing volume and clarity and reducing the pace of the cues.

Nature Sounds were mainly associated with pleasantness and relaxation. However, some sounds were perceived as similar or difficult to distinguish, and participants also noted a lack of continuous feedback. Suggestions included providing an introductory explanation of the cues and using spatialized information to convey direction.

For **Musical Instruments**, participants positively described the cues in terms of pleasantness, clarity, and ease of use, whereas similarity or confusion between cues and

semantic ambiguity were recurring difficulties. No recurring improvement suggestion was identified for this category.

Auditory Icons were frequently described as familiar and recognizable, and participants also mentioned their clarity and continuous feedback positively. Negative comments concerned low volume or insufficient audibility and unclear meaning for some cues. Corresponding suggestions focused on increasing volume and clarity.

Artificial Sounds were commonly perceived as pleasant and easy to use or direct. However, participants reported ambiguity of individual cues and a lack of continuous feedback. Suggested improvements included providing more continuous or frequent feedback and improving the coherence of the sounds.

Across cue categories, positive comments therefore primarily concerned pleasantness, clarity, directness, familiarity, recognizability, and ease of use, whereas negative comments commonly related to ambiguity, difficulties distinguishing cues, insufficient audibility, inappropriate pacing, and interruptions or gaps in feedback. Participants' suggestions correspondingly focused on improving cue clarity and audibility, increasing differentiation between cues, adjusting the pacing of auditory information, and providing more continuous or frequent feedback.

Some participants also suggested using the same sound type for different directions while relying on spatialization to convey directional information. In the present study, the cues were already spatialized, but different sound identities were additionally assigned to different directions.

5 Discussion and Conclusion

This study compared five categories of auditory cues for outdoor AR navigation—Artificial Sounds, Nature Sounds, Spearcons, Musical Instruments, and Auditory Icons—and examined their effects on user experience, cue preferences, participant characteristics, and subjective experiences reported in open-ended responses.

5.1 Differences between auditory cue categories

The UEQ results showed significant differences between cue categories in Novelty and Stimulation. Spearcons were rated lower in Novelty and Stimulation than Artificial Sounds and lower in Novelty than Musical Instruments. One possible explanation is the familiarity of spearcons, which resemble speech-based navigation cues and may therefore be perceived as less novel than more unconventional sounds [3, 10]. This interpretation is consistent with qualitative feedback, in which Spearcons were described as clear and direct, suggesting that their speech-like character may have contributed to their familiarity rather than novelty. In contrast, Artificial Sounds and Musical Instruments were less strongly associated with conventional navigation cues and were perceived as more novel and engaging.

5.2 Individual preferences

Nature Sounds and Artificial Sounds were most frequently selected as preferred cue categories, followed by Musical Instruments and Auditory Icons, while no participant

selected Spearcons. The preference analysis further showed that individually preferred cues received significantly higher ratings for Attractiveness, Perspicuity, Efficiency, Novelty, Stimulation, and perceived effectiveness (Q4) than non-preferred alternatives. Perceived pleasantness (Q5) and Dependability did not remain significant after correction. The preference distribution suggests that Nature Sounds and Artificial Sounds were particularly frequently preferred, and these categories were also often described as pleasant in the qualitative responses. In contrast, the absence of preference votes for Spearcons may be related to the reported difficulties concerning pacing and intelligibility, although these qualitative observations do not establish the reasons for the observed preference pattern. Despite the qualitative relevance of pleasantness, the statistical results suggest that pleasantness alone does not account for the more positive evaluations of preferred cues, as Q5 did not remain significant after correction. Rather, individually preferred cues were associated with significantly more positive evaluations on most of the assessed user-experience dimensions, including Attractiveness, Perspicuity, Efficiency, Novelty, Stimulation, and perceived effectiveness. This suggests that preference was not related to a single characteristic, but to a broader pattern of user-experience evaluations. At the same time, the different preference patterns indicate that no single cue category was preferred by all participants, highlighting individual variation that may be obscured by aggregate ratings.

5.3 Qualitative experiences

The qualitative findings suggest that the suitability of an auditory navigation cue depends on several interacting properties rather than on pleasantness or familiarity alone. Across cue categories, participants especially valued clarity, directness, familiarity, and ease of use, while difficulties were commonly related to ambiguity, insufficient audibility, inappropriate pacing, and gaps in feedback. This suggests that auditory navigation cues need to balance experiential qualities with clear and sufficiently continuous communication of navigation information. The specific patterns varied between cue categories. Nature Sounds and Musical Instruments were often perceived as pleasant, but some cues were difficult to distinguish or confusing. Auditory Icons were often described as familiar and recognizable, although some were perceived as insufficiently audible or unclear in meaning. Artificial Sounds were frequently described as pleasant and easy to use, while ambiguity and insufficiently continuous feedback remained concerns. Spearcons were described as clear and direct by some participants, but were also associated with fast pacing, reduced intelligibility, low volume, and gaps in feedback. Some participants suggested using the same sound type for different directions while relying on spatialization to convey directional information. In the present study, cues were already spatialized, but different sound identities were additionally assigned to different directions. The suggestion therefore points to a possible simplification: rather than requiring users to learn and distinguish multiple sound identities in addition to interpreting their spatial position, directional information could potentially be conveyed primarily through spatialization. Whether this approach would improve the ease, pleasantness, and effectiveness of auditory navigation requires further experimental evaluation.

5.4 Participant characteristics

The analysis of participant characteristics indicates that the comparison of cue categories was largely independent of age, gender, prior XR experience, and technology affinity. The most consistent finding concerned Nature Sounds: older participants perceived their direction less clearly, whereas no comparable age effect appeared for the other cue categories. This is in line with the lowest ratings of Nature Sounds for perceived direction and understanding (Q2, Q3; Table 1) and suggests that the natural sounds selected here were harder to localize for some listeners. Several of these sounds, such as flowing water or thunder, are continuous or diffuse rather than short and clearly defined events, and age-related changes in hearing, which were not assessed in this study, may have contributed to this effect. Higher technology affinity was associated with a lower perceived novelty of the cues; however, this effect was at the significance threshold, was not confirmed in the full sample, and could not be separated from gender and XR experience. Participants with more XR experience tended to rate all cue categories lower, which may reflect a higher reference standard for immersive systems, but this tendency did not remain significant after correction.

5.5 Implications and limitations

From a design perspective, the quantitative findings suggest that auditory cue selection can shape different aspects of the navigation experience. The differences in Novelty and Stimulation indicate that the choice of cue type may be relevant when designing for the experiential qualities of an auditory navigation system. At the same time, the preference analysis showed that individually preferred cues were associated with more positive evaluations across most of the assessed user-experience dimensions, including Attractiveness, Perspicuity, Efficiency, Novelty, Stimulation, and perceived effectiveness. This suggests that cue design should consider a broader set of user-experience qualities rather than optimizing for a single experiential characteristic. The qualitative findings further point to practical properties that may contribute to these evaluations, including audibility, cue differentiation, appropriate pacing, and continuity of feedback. Pleasantness or familiarity alone may therefore not be sufficient to ensure clear and distinguishable navigation cues. The absence of clear differences in task completion time or perceived effectiveness at the cue-category level further suggests that experiential and functional aspects of auditory navigation may not vary in parallel. Cues that were experienced as more novel or stimulating were not necessarily associated with faster task completion or higher perceived effectiveness. This highlights the importance of evaluating auditory navigation using both experiential and functional measures. In particular, future studies should complement completion time and perceived effectiveness with objective measures such as navigation accuracy, directional errors, and route deviations. Several limitations should be considered. The final sample consisted of only 15 participants, limiting statistical power and generalizability. The participant-characteristic analysis was particularly constrained by the unbalanced gender distribution, narrow age range, and confounding between gender, ATI, and prior XR experience. Hearing ability was not assessed. The qualitative analysis was conducted by a single coder and was intended as a structured descriptive analysis

rather than an independently validated qualitative study. Finally, the outdoor setting was relatively controlled, and cue preferences or intelligibility may differ under higher background noise or in more complex navigation environments. Future studies should replicate these findings with larger and more diverse samples and systematically examine environmental noise, cue pacing, spatialization, and continuity of feedback.

5.6 Conclusion

The results show that auditory cue categories for outdoor AR navigation differ in selected user-experience dimensions, particularly Novelty and Stimulation, without clear differences in task completion time or perceived effectiveness. The preference analysis further shows that individually preferred cues were associated with more positive evaluations across several user-experience dimensions. The qualitative findings reveal additional cue-specific strengths and difficulties related to pleasantness, clarity, familiarity, ambiguity, audibility, pacing, and continuity of feedback. Exploratory analyses further suggest that some participant characteristics may be associated with cue evaluation, although these findings require confirmation in larger and less confounded samples. Overall, the findings highlight that auditory navigation cues can be evaluated at different levels: differences between cue categories, individual preferences, and users' subjective experiences. Considering these perspectives together provides a more comprehensive understanding of how auditory cues are experienced and evaluated.

Funding

Partial financial support was received from the project ARIadne (Grant ID: KK5644701RL4).

References

- [1] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.
- [2] W. W. Gaver, "What in the world do we hear? An ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.
- [3] B. N. Walker and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in *Proc. Int. Conf. Auditory Display (ICAD)*, 2006.
- [4] S. A. Brewster, "Using non-speech sounds to provide navigation cues," *ACM Trans. Comput.-Hum. Interact.*, vol. 4, no. 3, pp. 224–259, 1997.
- [5] N. A. Giudice, J. Z. Bakdash, and G. E. Legge, "Wayfinding with words: Spatial learning and navigation using dynamically updated verbal descriptions," *Psychological Research*, vol. 71, no. 3, pp. 347–358, 2007.

- [6] T. Lokki and M. Grohn, “Navigation with auditory cues in a virtual environment,” *IEEE Multimedia*, vol. 12, no. 2, pp. 80–86, 2005.
- [7] J.-N. Voigt-Antons, Z. Sun, M. Vergari, N. Ashrafi, F. Vona, and T. Kojic, “Impact of spatial auditory navigation on user experience during augmented outdoor navigation tasks,” in *Proc. DAGA 2023*, Hamburg, Germany, 2023.
- [8] B. Arons, “A review of the cocktail party effect,” *Journal of the American Voice I/O Society*, vol. 12, pp. 35–50, 1992.
- [9] S. Brewster, “Overcoming the Lack of Screen Space on Mobile Computers,” *Personal and Ubiquitous Computing*, vol. 6, no. 3, pp. 188–205, 2002.
- [10] S. Y. W. Li, T. L. Tang, A. Hickling, S. Yau, B. Brecknell, and P. M. Sanderson, “Spearcons for patient monitoring: Laboratory investigation comparing earcons and spearcons,” *Human Factors*, vol. 59, no. 5, pp. 765–781, 2017.
- [11] J. A. H. van Erp, H. van Veen, C. Jansen, and T. Dobbins, “Waypoint navigation with a vibrotactile waist belt,” *ACM Trans. Appl. Percept.*, vol. 2, no. 2, pp. 106–117, 2005.
- [12] W. W. Gaver, “Auditory icons: Using sound in computer interfaces,” *Human-Computer Interaction*, vol. 2, no. 2, pp. 167–177, 1986.
- [13] B. N. Walker and M. A. Nees, “Theory of Sonification,” in *The Sonification Handbook*, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Berlin: Logos Verlag, 2011, pp. 9–39.
- [14] D. McGookin and S. Brewster, “Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition,” *ACM Trans. Appl. Percept.*, vol. 1, no. 2, pp. 130–155, 2004.
- [15] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios,” in *Proc. Int. Conf. on Universal Access in Human-Computer Interaction (UAHCI)*, Heraklion, Greece, Jun. 2014, pp. 383–392, doi: 10.1007/978-3-319-07668-3_37.
- [16] B. Laugwitz, T. Held, and M. Schrepp, “Construction and Evaluation of a User Experience Questionnaire,” in *Proc. USAB 2008: HCI and Usability for Education and Work*, Graz, Austria, Nov. 2008, vol. 5298, pp. 63–76, doi: 10.1007/978-3-540-89350-9_6.
- [17] F. Heller, A. Krämer, and J. Borchers, “Simplifying orientation measurement for mobile audio augmented reality applications,” in *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI ’14)*, New York, NY, USA, 2014, pp. 615–624, doi: 10.1145/2556288.2557021.

- [18] J. R. Gordon, A. J. Fiannaca, M. Kneisel, E. Cutrell, A. Miller, and M. Gonzalez-Franco, "Hearing the Way Forward: Exploring Ambient Navigational Awareness with Reduced Cognitive Load through Spatial Audio-AR," in *Ext. Abstracts CHI Conf. Human Factors in Computing Systems (CHI EA '23)*, New York, NY, USA, 2023, Article 155, pp. 1–7, doi: 10.1145/3544549.3585800.
- [19] J. Li, "Beyond sight: Enhancing augmented reality interactivity with audio-based and non-visual interfaces," *Applied Sciences*, vol. 14, no. 11, Article 4881, 2024, doi: 10.3390/app14114881.
- [20] J. Yang, A. Barde, and M. Billinghamurst, "Audio augmented reality: A systematic review of technologies, applications, and future research directions," Empathic Computing Laboratory, Auckland Bioengineering Institute, The University of Auckland, New Zealand; Dept. of Computer Science, ETH Zurich, Switzerland, 2022.
- [21] J. Bhattacharyya, L. Picinali, A. Vinciarelli, and S. Brewster, "Investigating the influence of environmental acoustics and playback device for audio augmented reality applications," in *Proc. 6th Int. Conf. Audio for Games*, Tokyo, Japan, 2024.
- [22] E. H. Anne de Haas and L.-H. Lee, "Deceiving audio design in augmented environments: A systematic review of audio effects in augmented reality," in *Proc. 2022 IEEE Int. Symp. Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Singapore, 2022, pp. 36–43, doi: 10.1109/ISMAR-Adjunct57072.2022.00018.
- [23] J. J. Alvarsson, S. Wiens, and M. E. Nilsson, "Stress Recovery during Exposure to Nature Sound and Environmental Noise," *International Journal of Environmental Research and Public Health*, vol. 7, no. 3, pp. 1036–1046, 2010, doi: 10.3390/ijerph7031036.
- [24] Mattheiss, E., Regal, G., Vogelauer, C., Furtado, H. (2020). 3D Audio Navigation - Feasibility and Requirements for Older Adults. In: Miesenberger, K., Manduchi, R., Covarrubias Rodriguez, M., Peñáz, P. (eds) *Computers Helping People with Special Needs. ICCHP 2020. Lecture Notes in Computer Science()*, vol 12377. Springer, Cham. https://doi.org/10.1007/978-3-030-58805-2_38
- [25] T. Dingler, J. Lindsay, and B. N. Walker, "Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech," in *Proc. Int. Conf. Auditory Display (ICAD)*, Paris, France, 2008.
- [26] B. N. Walker, J. Lindsay, A. Nance, Y. Nakano, D. K. Palladino, T. Dingler, and M. Jeon, "Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus," *Human Factors*, vol. 55, no. 1, pp. 157–182, 2013, doi: 10.1177/0018720812450587.

- [27] S. Garzonis, S. Jones, T. Jay, and E. O'Neill, "Auditory icon and earcon mobile service notifications: Intuitiveness, learnability, memorability and preference," in *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI)*, 2009, pp. 1513–1522, doi: 10.1145/1518701.1518932.
- [28] S. Holland, D. R. Morse, and H. Gedenryd, "AudioGPS: Spatial audio navigation with a minimal attention interface," *Personal and Ubiquitous Computing*, vol. 6, no. 4, pp. 253–259, 2002, doi: 10.1007/s007790200025.
- [29] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," *Human Factors*, vol. 48, no. 2, pp. 265–278, 2006, doi: 10.1518/001872006777724507.
- [30] J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, and F. Dellaert, "SWAN: System for wearable audio navigation," in *Proc. 11th IEEE Int. Symp. Wearable Computers (ISWC)*, 2007, pp. 91–98, doi: 10.1109/ISWC.2007.4373786.
- [31] M. Jones, S. Jones, G. Bradley, N. Warren, D. Bainbridge, and G. Holmes, "ONTRACK: Dynamically adapting music playback to support navigation," *Personal and Ubiquitous Computing*, vol. 12, no. 7, pp. 513–525, 2008, doi: 10.1007/s00779-007-0155-2.
- [32] G. D. Clemenson, A. Maselli, A. J. Fiannaca, A. Miller, and M. Gonzalez-Franco, "Rethinking GPS navigation: Creating cognitive maps through auditory clues," *Scientific Reports*, vol. 11, Art. 7764, 2021, doi: 10.1038/s41598-021-87148-4.
- [33] R. L. Klatzky, J. R. Marston, N. A. Giudice, R. G. Golledge, and J. M. Loomis, "Cognitive load of navigating without vision when guided by virtual sound versus spatial language," *J. Exp. Psychol. Appl.*, vol. 12, no. 4, pp. 223–232, 2006, doi: 10.1037/1076-898X.12.4.223.
- [34] Y. Vazquez-Alvarez, I. Oakley, and S. A. Brewster, "Auditory display design for exploration in mobile audio-augmented reality," *Personal and Ubiquitous Computing*, vol. 16, no. 8, pp. 987–999, 2012, doi: 10.1007/s00779-011-0459-0.
- [35] M. N. Lystbæk, H. Adams, R. K. Ananda, E. J. Gonzalez, L. Ballan, Q. Wu, A. Colaço, P. Tan, and M. Gonzalez-Franco, "Navig-AI-tion: Navigation by contextual AI and spatial audio," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA)*, 2026, doi: 10.1145/3772363.3799295.