The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-9-2025
30th CIPA Symposium "Heritage Conservation from Bits:
From Digital Documentation to Data-driven Heritage Conservation", 25–29 August 2025, Seoul, Republic of Korea

Place-based Virtual Reconstruction of Heritage Soundscape: A Case Study of Magoksa Temple, UNESCO World Heritage Sites of Korea

Kyung Taek Oh[1], Rai Sato[1], Sungjoon Kim[1], Pooseung Koh[1], Sungyoung Kim[1, 2], Hyeseung Shim*[3]

[1] Graduate School of Culture Technology, KAIST, Daejeon, Republic of Korea - (kennyhola, s.rai, sungjoon.kim, seankoh754, sungyoung.kim)@kaist.ac.kr
[2] Department of Electrical and Computer Engineering Technology, RIT, Rochester, USA - sxkiee@rit.edu
[3] Department of Sociology, Hong Kong Shue Yan University, Hong Kong SAR, China - hsshim@hksyu.edu

Keywords: Virtual Reconstruction, Soundscape, Psychoacoustics, World Heritage, Sonic Heritage, Cultural Heritage Value.

Abstract

This study addresses the challenge of creating culturally resonant digital heritage by investigating user-driven virtual reconstructions of soundscapes. Focusing on Magoksa Temple, a Korean UNESCO World Heritage site, we employed a place-based interactive system where users could manipulate distinct sonic layers—ritual, architectural, natural, and visitor-related—to construct their preferred auditory environment. The research aimed to identify user priorities and uncover the factors that mediate a meaningful cultural experience. Findings reveal that users curate an idealized soundscape, prioritizing culturally significant sounds while filtering social background sound. Critically, the results show a distinct divergence between interaction and comprehension; while the system was highly engaging for all participants, achieving culturally meaningful experience was contingent upon their pre-existing cultural knowledge. This research provides empirical grounding for future heritage design, arguing that moving beyond sensory immersion to include integrated interpretive layers is essential for conveying intangible cultural value to diverse audiences.

1. Introduction

The digital representation of sound is increasingly pivotal for a holistic approach to heritage documentation, moving beyond visual reconstructions to encompass the acoustic dimensions integral to cultural understanding and appreciation (Champion and Rahaman, 2020, Skublewska-Paszkowska et al., 2022). While audio technologies offer potential for dynamic heritage representations (Pérez-Martínez et al., 2018, Privitera et al., 2024), conveying the distinct characteristics and cultural significance of these sonic environments remains challenging. A key methodological hurdle is balancing technical fidelity with the conveyance of cultural meaning. Conventional approaches that prioritize acoustic accuracy often yield what have been described as "acoustically accurate but culturally hollow" representations (Smith, 2006). This problem is exacerbated by soundscape reconstruction practices that often filter what is perceived as "background sound," inadvertently stripping away sonic layers vital for a site's authenticity and "aliveness" (Stockfelt et al., 2004, Brown and Muhar, 2004, Axelsson et al., 2010). At a heritage site like Magoksa Temple, a UNESCO World Heritage Buddhist monastery, sounds such as visitor murmurs, the rustle of wind through trees, or the creak of wooden architecture are integral to its evolving soundscape and the continuity of its intangible values. Their omission or downplaying in digital replicas risks creating partial experiences that diminish the site's experiential depth and cultural atmosphere.

This challenge is amplified by the evolving understanding of heritage itself, which has moved beyond a focus on static monuments to embrace the intangible, living practices that animate them (ICOMOS, 1994). This expanded perspective recognizes that the value of a physical site is deeply intertwined with its associated traditions, rituals, and social functions (UNESCO, 1972, UNESCO, 2003). In a living heritage site like a Buddhist temple, this intangible dimension is powerfully conveyed through its soundscape—a dynamic mix of ritual, social, architectural, and natural sounds. This dynamism often clashes with conventional digital reproductions reliant on static or idealized recordings, which can create a disconnect between the digital representation and the holistic experience of the site.

While enhancing cultural value transmission via digital media is a key goal for heritage practitioners, the common neglect of holistic soundscapes indicates a foundational knowledge gap. Prior investigations have often focused on the technical aspects of audio reproduction or used pre-filtered soundscapes, thereby overlooking how users engage with holistic sonic environments through interactive systems (Katz et al., 2020, Banfi and Bolognesi, 2021). Therefore, to develop robust methodologies for enriching cultural understanding through sound, it is first necessary to understand how general users interact with and prioritize heritage elements within a comprehensive, multi-layered soundscape when it is presented digitally and interactively. This research investigates user engagement within a context where individuals actively shape a soundscape composed of multiple, simultaneously present sound layers—including layers previously dismissed as 'background sound' but potentially crucial for a contextualized experience.

This study represents a preliminary investigation into user-driven soundscape construction within an interactive digital heritage environment. By examining how users configure and respond to different sounds, we aim to identify potential discrepancies between their instinctive mixing preferences and the prominence typically afforded to culturally significant elements in curated presentations. This is positioned as an empirical groundmaking effort: understanding these foundational user interactions is crucial before designing sophisticated strategies for guiding users toward specific cultural values through sound. It offers essential insight for future heritage sound design, informing culturally resonant digital reconstructions that represent the integrity of sonic elements. Therefore, this study focuses on three specific aspects: (1) What distinct categories of sonic elements (ritual, architectural, environmental, visitor-related), integral to the heritage character of Magoksa Temple, can be identified and derived for a digital reconstruction? (2) What relative sound levels and perceived spatial characteristics do users establish for these distinct sound layers, and what does this reveal about their priorities in constructing the soundscape? (3) Based on the outcome of second question, what design strategies can guide the creation of culturally-attuned digital heritage sound environments?

This paper is structured as follows: Section 2 reviews related work, followed by the methodology in Section 3. Section 4 presents the results, then Section 5 discusses the implications, and concludes in Section 6.

2. Related Works

2.1 Representing Dynamic Heritage Soundscapes

The evolving understanding of heritage as a set of dynamic, living practices, rather than a collection of static monuments, necessitates a corresponding evolution in digital representation methods. This broader perspective demands that heritage soundscapes be treated as evolving, performative environments, not merely as fixed artifacts to be captured (Titon et al., 2009, Seeger, 2004). Such soundscapes are connected to community, context, and the continuity of cultural traditions. While modern technology makes it possible to record a site's sounds with high fidelity, a key challenge lies in capturing and representing how the dynamic interplay between these sounds generates their cultural significance. Current digital documentation methodologies often fall short, frequently relying on discrete, decontextualized recordings or snapshots in time. This approach can "freeze" what is an inherently continuous acoustic environment, failing to make sense of the site's holistic values. The result is often a representation that, does not convey the crucial sense of presence, place, and cultural atmosphere that defines the living heritage (Wijesuriya, 2018). The core challenge, therefore, is not simply a technical one of recording, but a curatorial and interpretive one: how to represent the ongoing, place-based continuation of culture.

In response to this challenge, practices in heritage soundscape reconstruction often involve a difficult curatorial dilemma. Many approaches prioritize selected primary sounds (e.g., specific ritual sounds) while filtering out other sonic components deemed 'ambient noise' or 'distractions'. This reductive approach is frequently driven by a pursuit of acoustic clarity or by technical limitations (Privitera et al., 2024). However, such filtering can exclude sounds integral to the holistic character and perceived authenticity of a heritage site, including subtle environmental cues or the sounds of human presence that signify a living space (Mattioli and Díaz-Andreu, 2017, Guastavino, 2006)

This creates a critical tension between technical fidelity and holistic representation. Even as advanced audio capture techniques like Ambisonics document spatial information (Zotter and Frank, 2019), the subsequent translation of this data into user experiences often still undergoes these reductive curatorial choices. For example, the field of archaeoacoustics often strives for a veridical, evidence-based reconstruction of past sounds, which can be invaluable for research but may not align with the subjective experience of a contemporary visitor. Conversely, other approaches may prioritize an 'artistic' or 'evocative' soundscape, which can enhance atmosphere but may stray from documentary integrity (Yong Jeon et al., 2011, Dávid et al., 2024). Both expert-led approaches — the scientific and the artistic — make assumptions about what is important for the listener. Consequently, a gap persists in understanding how to effectively represent the entirety of a heritage soundscape in a way that preserves its complexity and the interrelationships between its diverse sonic layers.

2.2 User Engagement with Digitally Mediated Soundscapes

While research into user experience with digital heritage is growing, studies specifically focused on interactive engagement with sound remain less common. Existing audio-focused studies often investigate the impact of pre-composed soundscapes on presence, learning, or emotional response in virtual environments or through audio guides (Champion, 2021, Glaser et al., 2024, Maffei et al., 2015), or they evaluate technical aspects of sound reproduction quality (Vorlaender and Summers, 2008). Crucially, these approaches position the user as a passive recipient of a researcher-curated soundscape.

There is a notable lack of research exploring how users themselves engage with, prioritize, and balance the multiple layers of a holistic heritage soundscape when given direct, interactive control. Instead of simply measuring the effects of a predefined soundscape, understanding the choices users make during its active construction can reveal their implicit values and subjective sense of what constitutes cultural experience. This knowledge is fundamental for designing systems that can adapt to or guide user preferences effectively. This study addresses this gap by focusing on the user's active role in shaping a multilayered soundscape. By shifting from a passive listening model to an active construction task, this user-centered approach, employing the Method of Adjustment (Rumsey, 2002), provides an empirical basis for understanding how non-experts navigate the trade-offs between culturally significant sounds and ambient environmental elements.

3. Methods

This study employed a mixed-methods approach to investigate user-driven heritage soundscape construction, integrating on-site empirical documentation with a laboratory-based experiment and quantitative data analysis. The methodology was structured in four sequential phases: (1) Soundscape Documentation and Taxonomy, (2) System and Experimental Design, (3) Participant Recruitment and Procedure, and (4) Data Analysis.

3.1 Soundscape Documentation and Taxonomy

The foundational phase of this research was to deconstruct the complex sonic environment of the heritage site into meaningful, analyzable components.

3.1.1 Venue Selection: The study selected Magoksa Temple as an exemplar case. Founded in the 7th century and set within a mountain valley, the temple complex is bisected by the Taegeukcheon stream, creating a unique and sonically diverse environment. Its cultural significance was also recognized with its inclusion in the UNESCO World Heritage serial property, "Sansa, Buddhist Mountain Monasteries in Korea" (UNESCO World Heritage Centre, 2018), Figure 1. As a "living heritage site" (Smith, 2006), it is defined by the dynamic interplay between the continuous practice of Buddhist rituals and high levels of public visitation. This combination of distinct natural, architectural, ritual, and social sonic elements makes it an appropriate context for investigating the perceptual tensions between these sound layers, directly aligning with the research aims.

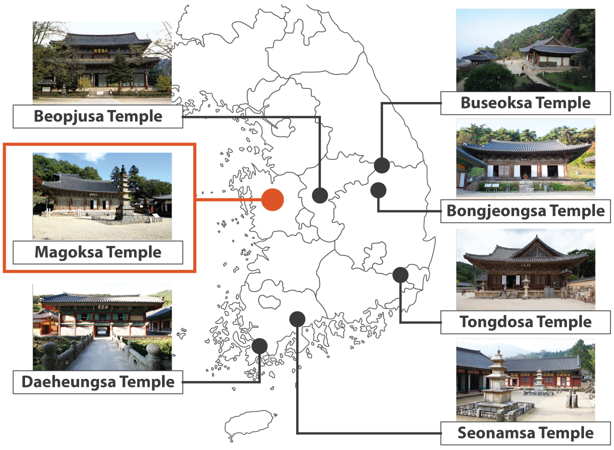

Figure 1. Seven components of Sansa, UNESCO World Heritage Serial Property of Korea. Image source: Korea Heritage Service.

3.1.2 Sound Taxonomy Development: A hybrid sound taxonomy was developed to ensure a comprehensive and culturally relevant categorization of the site's sonic elements. This approach was essential because a purely psychoacoustic framework would miss the cultural significance of the sounds, while a purely cultural one would lack the structural basis for sonic design. The taxonomy therefore combines insights from R. Murray Schafer's (1977) soundscape theory with UNESCO's cultural conventions. This allowed each sound to be classified by both its acoustic role in the soundscape (e.g., soundmark, sound signal) and its cultural function (e.g., ritual, social practice).

This dual-coding framework was refined through on-site observation and pilot recordings, resulting in four primary layers for the experiment. Each layer was designed to be both acoustically distinct and culturally meaningful:

- Architectural Soundmarks: This layer consists of sounds emanating from the tangible fabric of the temple, such as creaking wooden floors or the sliding of paper-and-wood doors. These recurring, site-specific sounds anchor listeners to the physical place and contribute to its perceived authenticity. In the experiment, they provided foundational elements contributing to the sense of place and realism.

- Ritual Sound Signals: This layer includes intentional sounds central to Buddhist practice and the temple's spiritual identity, such as chants, the large bronze bell (beomjong), and the wooden percussion instrument (moktak). As many of these sounds are recognized intangible cultural heritage (ICH) elements, they served as the key culturally significant content whose prominence in user-configured mixes was a primary focus of this investigation.

- Visitor Social-Practice: This layer comprises sounds generated by lay visitors, like footsteps on gravel and whispered conversations. These sounds reflect the temple's contemporary use as a living, accessible heritage site and contribute to its perceived vitality. Their inclusion allowed for an exploration of how users balance the sacred character of the site with its social, human aspect.

- Natural Keynotes: This layer contains the environmental sounds that form the temple's ambient backdrop, such as the adjacent stream and seasonal birdsong. These keynote sounds are fundamental to the temple's genius loci or "spirit of place." Experimentally, they provided the essential environmental context against which all other sounds were perceived.

A fifth category, "Exogenous Intrusions" (e.g., distant aircraft), was documented but excluded from the interactive system to maintain focus. To populate these layers, on-site First-Order Ambisonics (FOA) recordings were captured at 22 locations within Magoksa Temple (Figures 2-3). From these, a library of representative sound events was isolated for each taxonomic layer, Table 1.

Figure 2. Top: images of outdoor and indoor ambisonic microphone recording setup at Magoksa Temple. Bottom: a map of Magoksa Temple Complex with 22 recording locations (blue), and 3 reconstructed zones (red).

Table 1. Table of taxonomic layer and separated sound sources.

Layer Category	Description	Example Sound Sources
Architectural Soundmarks	Sounds emanating from the tangible fabric of the temple.	Creaking of wooden floorboards; sliding of paper-and-wood doors (changho); monks' footsteps on the main hall's wooden deck (maru); chime (punggyeong).
Ritual Sound Signals	Intentional sounds central to Buddhist monastic practice.	Morning/evening Yebul chants; striking of the large bronze bell (beomjong); rhythmic beats of the wooden percussion instrument (moktak); small cymbals.
Visitor Social-Practice	Sounds generated by lay visitors and tourists engaging with the site.	Footsteps of multiple people on gravel paths; distant, whispered conversations; isolated camera shutter clicks.
Natural Keynotes	Environmental sounds forming the site's characteristic ambient backdrop.	Flowing water from the adjacent stream; seasonal birdsong characteristic of the region; wind moving through pine trees.
Ambient Layer	A continuous, low-level background soundscape used only in the experiment.	A composite mix of diffuse room tone from inside temple halls and distant, unlocalizable environmental hum, designed to prevent unnatural digital silence.

3.2 System and Experimental Design

Building on the taxonomy, an interactive system was developed to empirically measure user preferences in a controlled environment.

3.2.1 Interactive Sound Environment: A within-subjects experimental design was chosen for its statistical observation and sensitivity in detecting differences across conditions while controlling for innate individual variability in listening habits and preferences. Using three distinct virtual zones: (1) Entrance Courtyard, (2) Pavilion by Stream, and (3) Main Worship Hall (Daegwangbojeon) was intended to investigate whether user preferences were static or context-dependent, shifting based on the implied location and its associated activities. This place-based approach allowed for a direct test of how spatial context, a key element of heritage sites, influences sonic perception. The presentation order of these zones was counterbalanced across participants to mitigate potential order effects.

A custom interactive environment was developed in Max 8 (see Fig. 4). The interface featured sliders allowing participants to adjust two parameters for each sound layer: (1) Intensity (Level): The audible loudness of each layer, adjustable in decibels (dB); and (2) Perceived Proximity (Space): The sense of a sound’s distance from the listener, manipulated by altering the balance between the direct sound (dry) and its simulated environmental reflections (wet) using a convolution reverb.

Moreover, the Ambient Layer was included in the experimental setup to prevent the perception of unnatural “digital silence” and to provide a foundational atmospheric presence upon which the other, more dynamic layers could be mixed.

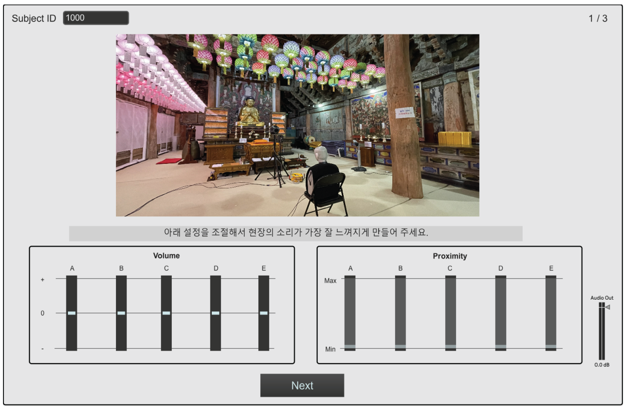

Figure 3. Snapshot of the user interface of the experiment.

3.2.2 Task and Measurement: The core task utilized the Method of Adjustment. Participants were instructed to freely manipulate the sliders for each sound layer until the overall soundscape felt most representative or most satisfactory to them for the depicted scene. This method, originating from psychoacoustics, was selected as it grants direct agency to the user, making it highly suitable for eliciting subjective preferences in complex, multi-stimulus environments (Rumsey, 2002)

3.3 Experimental Procedure

This phase details the recruitment of participants and the specific materials and procedures used to collect both behavioral and subjective data.

3.3.1 Participants: A total of 42 adults (24 female, 18 male; mean age = 34.4, SD = 4.5) were recruited for the study. To ensure the sample was suitable for investigating the role of prior knowledge, a purposive sampling strategy was employed. The goal was to include participants into two distinct strata based on their self-reported level of understanding of Korean Buddhist traditions: a ‘High-Context’ and a ‘Low-Context’ group. The ethical approval of this study was received from the authors’ institutional review board.

3.3.2 Questionnaire Design: To measure subjective experience, a post-task questionnaire was administered, centered on a modified version of the Museum Experience Scale (MES) (Othman, 2012). The MES was selected for its established validity in measuring key dimensions of a visitor’s engagement with heritage. Its four constructs—Engagement, Knowledge/Learning, Meaningful Experience, and Emotional Connection—were deemed highly relevant for assessing a digital heritage experience focused on intangible culture. These constructs allowed for an evaluation that went beyond usability to address the central problem of creating “culturally contextualized” versus “hollow” experiences. For this study, the scale items were adapted by rephrasing them for clarity in a sound-centric, non-physical context.

3.3.3 Procedure: Each participant attended a single session lasting approximately 40 minutes. Upon arrival, participants were briefed on the study’s general purpose and provided informed consent. Next, they received a tutorial on how to use the interactive sound adjustment interface, which was followed by a short practice trial with non-experimental sounds to ensure familiarity with the controls. For the main task, participants proceeded through three virtual temple zones in a counterbalanced order, adjusting the sound layers in each zone until they created a mix they found most satisfactory or representative. After completing all three zones, they filled out a demographic form, the modified MES, and several short open-ended questions.

3.4 Data Analysis

All quantitative data were analyzed using R statistical software. The analysis of soundscape settings began with descriptive statistics and a one-way Analysis of Variance (ANOVA) to assess overall layer preferences. The primary analysis used Linear Mixed-Effects (LME) models to assess the fixed effects of Zone and Sound Layer on user adjustments while controlling for individual differences by treating ‘Participant ID’ as a random effect (Baayen et al., 2008). The analysis of the questionnaire data involved assessing the internal consistency of the MES subscales with Cronbach’s Alpha, followed by one-sample t-tests to compare mean scores against the neutral midpoint. Finally, independent samples t-tests were used for an exploratory comparison between groups based on self-reported prior experience with Buddhist traditions.

4. Results

This section presents the quantitative results of the study. The analysis is structured to first report on general soundscape adjustment preferences across the entire participant sample, directly addressing which sounds were given prominence (4.1). It then examines how these adjustments were influenced by different contexts, both spatial and experiential (4.2 and 4.3). Finally, it concludes with a focused analysis of how prior knowledge with the subject matter influenced both behavior and subjective evaluation, providing foundational insights for heritage soundscape design (4.4).

4.1 Overall User Preferences in Soundscape Adjustment

To establish a baseline for user priorities, an initial analysis examined the overall prominence participants assigned to the five sound layers, aggregating data from all 42 participants across the three experimental zones. A one-way ANOVA confirmed that the choices for sound layer settings were statistically significant for both Intensity (F = 69.94, $p < .001$) and Proximity (F = 2.86, p = .023).

The mean settings for each layer, visualized in Figure 4, illustrate a clear and consistent hierarchy of preference. Participants afforded the highest mean intensity to the Ritual layer (M = 3.83, SD = 4.70), identifying it as the most important sound. In direct contrast, they actively suppressed the Visitor layer (M = -6.67, SD = 5.59), setting it to the lowest intensity. This finding provides a clear baseline for user priorities when context is not considered.

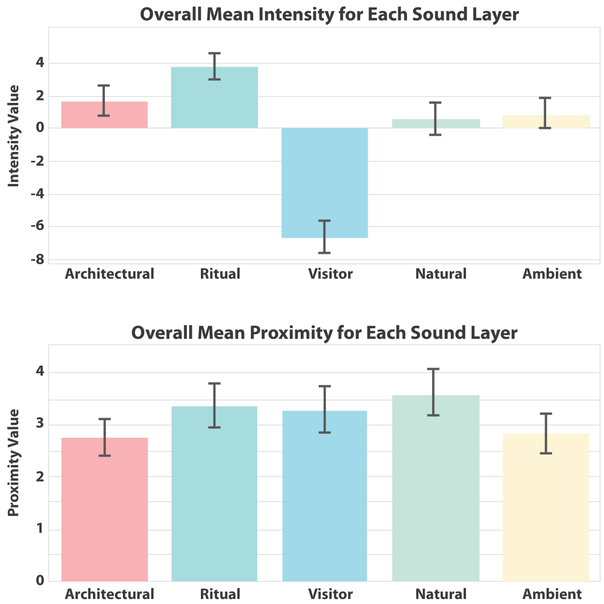

Figure 4. The mean Intensity (top) and Proximity (bottom) for each of the five sound layers. Error bars represent the 95% confidence interval.

4.2 The Influence of Virtual Zone on Contextual Adjustments

To investigate how mixing choices differ under various conditions, participants' adjustments were analyzed across the three zones. A LME model, which assesses the impact of Zone and Sound Layer on Intensity and Proximity, confirmed a significant interaction effect ($p < .001$). This statistical result means that participants did not just raise or lower all sounds together; they actively curated the soundscape differently based on the specific location.

The contextual nature of these user-curated soundscapes is evident in the descriptive statistics presented in Table 2 and scatter plot in Figure 5. The data reveals that participants were not simply adjusting volume but actively "placing" sounds within the virtual environment based on the implied location. For example: The 'Natural' layer was brought to the perceptual forefront in the streamside setting of Zone 2 ('Pavilion by Stream'). The 'Architectural' sounds were deemed most prominent in Zone 1 ('Entrance Courtyard'). Interestingly, the 'Ritual' layer's intensity was significantly lowered by participants when inside Zone 3 ('Main Worship Hall') compared to the other zones. The 'Visitor' layer was consistently relegated to a low-intensity, distant position across all three zones, confirming its status as undesirable "social background sound" that users actively filtered out regardless of the context.

Collectively, these spatial adjustments show that participants engaged in crafting an idealized, place-based soundscape that prioritized certain sounds while actively managing others to suit the character of each virtual location.

Table 2. Descriptive statistics of sound layer's Intensity and Proximity by zone.

		Intensity			Proximity		
Zone	Layer	Mean	SD	SEM	Mean	SD	SEM
1	Architectural	4.62	5.12	0.79	3.37	1.81	0.28
	Ritual	5.83	3.69	0.57	2.09	1.17	0.18
	Visitor	-6.55	5.64	0.87	3.66	2.72	0.42
	Natural	-0.45	6.16	0.95	3.8	2.53	0.39
	Ambient	1.62	5.83	0.9	3.48	2.53	0.39
2	Architectural	1.4	5.51	0.85	2.02	1.56	0.24
	Ritual	3.86	4.73	0.73	4.68	2.79	0.43
	Visitor	-7.64	6.22	0.96	3.36	2.92	0.45
	Natural	2.29	3.82	0.59	3.95	2.53	0.39
	Ambient	0.33	4.34	0.67	2.59	2.07	0.32
3	Architectural	-0.86	4.8	0.74	2.92	2.33	0.36
	Ritual	1.79	5.05	0.78	3.39	2.2	0.34
	Visitor	-5.81	5.12	0.79	2.9	2.53	0.39
	Natural	-0.07	5.51	0.85	3.05	2.14	0.33
	Ambient	0.9	5.31	0.82	2.51	1.75	0.27

Note: SD = standard deviation; SEM = standard error of the mean;

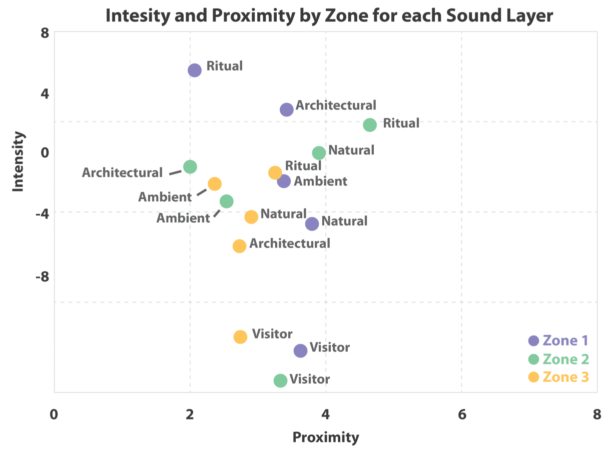

Figure 5. This visualization maps each sound layer's mean intensity on the y-axis against its mean proximity on the x-axis, with data points color-coded by the three distinct virtual zones.

4.3 Subjective Experience: Post-Experience Questionnaire Analysis

An analysis of the post-experience questionnaire was conducted to quantify the participants' subjective experience. First, the internal consistency of the four subscales from the modified MES was evaluated using Cronbach's Alpha. As shown in Table 3, all four subscales demonstrated acceptable to good reliability ($\alpha \geq .78$), confirming that they consistently measure their intended constructs (Nunnally, 1978).

Next, to determine if the experience was rated positively, a one-sample t-test was performed for each subscale, comparing the mean score against the neutral midpoint of 4.0. The results, detailed in Table 3, indicate a significantly positive user experience overall. The mean scores for Engagement (M = 6.29),

Table 3. Reliability and t-test results for questionnaire subscales.

Subscale	Mean	SD	t-statistic	p-value	Cronbach's α
Engagement	6.29	0.67	22.16	$<.001$	0.848
Knowledge/Learning	4.63	1.02	3.96	$<.001$	0.780
Meaningful Experience	4.36	1.25	1.88	$<.068$	0.792
Emotional Connection	4.98	1	6.35	$<.001$	0.836

Note: SD = standard deviation

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1105-2025

Knowledge/Learning ($M = 4.63$), and Emotional Connection ($M = 4.98$) were all statistically significantly higher than the neutral midpoint (all $p < .001$). However, the score for Meaningful Experience ($M = 4.21$, $p = .145$) did not significantly differ from the neutral point for the group as a whole.

4.4 The Role of Prior Knowledge on Behavior and Experience

A final analysis investigated the role of prior knowledge by comparing the High-Context (n = 15) and Low-Context (n = 27) groups. The results revealed a disconnect between interactive behavior and subjective experience. Independent samples t-tests on sound-mixing behaviors found no significant differences between the groups for the core heritage layers, indicating both groups curated the soundscape similarly.

In contrast, their subjective questionnaire scores were highly different. While both groups reported similarly high Engagement ($p = .350$), the High-Context group scored significantly higher on Knowledge/Learning ($t = 3.07$, $p = .004$), Meaningful Experience ($t = 6.83$, $p < .001$), and Emotional Connection ($t = 6.49$, $p < .001$) (Table 4). This suggests that while the system was equally engaging, prior knowledge was a critical factor in fostering a deeper cultural experience.

Table 4. T-test results for sound mixing behavior and MES questionnaire responses by prior knowledge

Variables	High-context Group (n=15) *Mean (SD)*	Low-Context Group (n=27) *Mean (SD)*	t-statistic	p-value
Sound Mixing				
Intensity - Architectural	2.47 (5.43)	1.29 (5.57)	0.99	0.328
Intensity - Ritual	4.98 (4.95)	3.19 (4.51)	1.9	0.065
Intensity - Visitor	-6.49 (5.83)	-6.77 (5.55)	0.37	0.714
Intensity - Natural	1.84 (5.55)	-0.16 (5.14)	1.58	0.123
Intensity - Ambient	**2.89 (4.57)**	**-0.05 (4.81)**	**2.67**	**.011***
Proximity - Architectural	2.99 (1.91)	2.65 (1.75)	0.69	0.497
Proximity - Ritual	3.34 (2.36)	3.41 (2.44)	-0.13	0.894
Proximity - Visitor	3.84 (2.79)	2.99 (2.61)	1.76	0.087
Proximity - Natural	3.28 (2.55)	3.79 (2.39)	-0.95	0.347
Proximity - Ambient	2.79 (2.19)	2.90 (2.20)	-0.39	0.699
MES Responses				
Engagement	6.42 (0.61)	6.21 (0.70)	0.95	0.35
Knowledge/Learning	**5.23 (0.78)**	**4.29 (0.97)**	**3.07**	**.004***
Meaningful Experience	**5.45 (0.63)**	**3.74 (1.19)**	**6.83**	**<.001***
Emotional Connection	**5.85 (0.52)**	**4.48 (0.90)**	**6.49**	**<.001***

Note: SD = standard deviation. Significant results p <.05 marked with *.

5. Discussion

The findings of this study reveal a clear hierarchy in users' sonic priorities and, more importantly, a critical gap between user engagement and meaningful cultural understanding within an interactive heritage soundscape. This discussion synthesizes these results to answer the three research questions, offering key interpretations of user behavior and providing foundational insights for the design of more culturally resonant digital heritage experiences.

5.1 The Deconstructed Soundscape: Identifying Core Sonic Categories

The first research question asked what distinct sonic categories could be derived for reconstruction. This study successfully identified and operationalized four culturally and sonically distinct layers—Architectural, Ritual, Visitor, and Natural—based on a hybrid conceptualization of psychoacoustic and cultural heritage theories. The fact that participants made distinct and consistent adjustments to each layer, as reported in the results, confirms that this method of deconstructing a complex heritage soundscape into important components is a viable and effective approach for both analysis and interactive design.

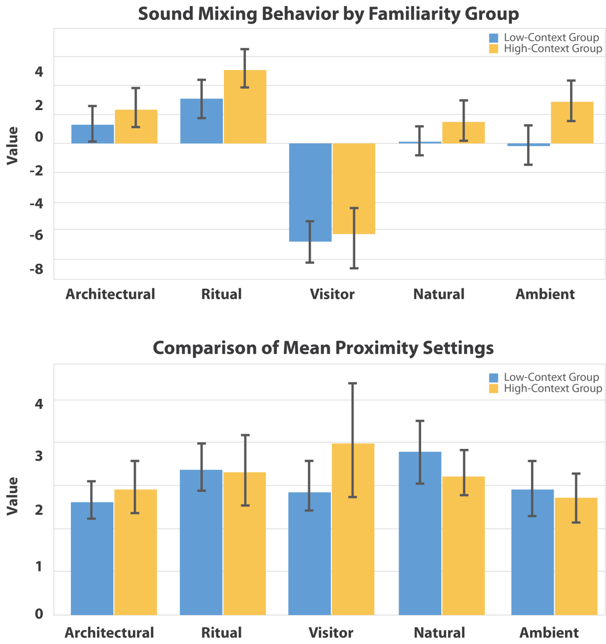

Figure 6. Sound Mixing Behavior by Knowledge Group Comparison of mean Intensity (top) and Proximity (bottom) settings shows that the groups behaved in a largely similar manner.

5.2 The User-Based Soundscape: Priorities and Context

The second research question addressed how users prioritize and balance these components when given control. The results indicate that users do not strive to replicate a literal, veridical acoustic snapshot of the site. Instead, they actively curate an idealized soundscape that aligns with their personal or perceived notion of the site's identity. This was evident in two key findings.

First, the consistent amplification of the Ritual layer demonstrates that users identify these sounds as the core carriers of the temple's cultural and spiritual identity. These sounds function as powerful "soundmarks" (Schafer, 1993), acoustic figures of high significance that define the character of a place. The user's choice to foreground these sounds suggests that, in a digital context, a genuine experience of the site is perceived to be most strongly tied to the intangible cultural practices for which it is known.

Second, and in stark contrast, was the consistent suppression of the Visitor layer. This finding presents a compelling challenge to academic notions that call for including all sounds of a place, or "social background sound," to represent its "aliveness" (Stockfelt et al., 2004, Axelsson et al., 2010). While these visitor sounds are an integral part of the contemporary soundscape, users in this study overwhelmingly perceived them as distractions rather than contributions. This points toward a user desire for a more personal, contemplative experience, where the digital environment serves as a space for focused engagement with the site's sacred character.

Furthermore, user preferences highly context-dependent. The finding that Ritual sounds were set to their lowest intensity in the main worship hall is particularly insightful. This does not suggest the sounds were unimportant in that context. Rather, it may reflect a sophisticated "perceptual realism," where users imagine the sound as it would be experienced from within a sacred, enclosed space—as an ambient, surrounding presence rather than a direct, foregrounded signal. This demonstrates a nuanced engagement where users are not simply adjusting volume but are actively shaping their interpretation.

5.3 Foundational Insights for Design: Bridging the Experiential Gap

The third research question sought foundational insights for designing culturally-attuned digital heritage environments. The study's core contribution stems from its empirical identification of disconnect between user engagement and cultural understanding, which has major implications for how interactive heritage experiences are designed and evaluated.

5.3.1 From Engagement to Understanding: The Critical Role of Prior Knowledge One of the key findings of this study is the divergent outcomes between the High-Context and Low-Context groups. Both groups reported equally high levels of engagement and exhibited nearly identical sound-mixing behaviors. However, this high level of engagement only translated into a meaningful and emotionally resonant experience for the High-Context group. This demonstrates a clear divide: an engaging user interface does not inherently foster a deep connection with the cultural content it presents.

This finding contributes to a growing conversation in digital heritage and Human-Computer Interaction that seeks to move beyond simplistic metrics of engagement (e.g., time on task) toward evaluating genuine "meaning-making" (McCarthy and Wright, 2004, Forlizzi and Battarbee, 2004, Schmitt and Labour, 2022). Our results provide empirical evidence that meaning is not automatically derived from sensory data or interactivity alone; it is co-created through an interplay between the digital artifact and the user's own pre-existing knowledge. The fact that user behavior did not correlate with the depth of their experience demonstrates that an intuitive interface is insufficient for a meaningful outcome. Therefore, to design a successful cultural experience, one must move beyond simple interaction to incorporate the cultural factors that allow users to form a genuine connection.

5.3.2 Recommendations for Designing for Cultural Understanding Our findings suggest a design recommendation for bridging the gap between user engagement and cultural understanding. This approach moves beyond simple interactivity to actively foster meaning-making, based on the following guidance:

Establish a Responsive and Coherent Sonic Foundation. A meaningful experience requires a robust technical and conceptual architecture. Our results demonstrate the value of a thematic, layer-based structure, which provides users with an intelligible and guideline for interaction. This structure must also be dynamic; the context-dependent adjustments made by our participants underscore the need for place-based, procedural, or context-aware audio systems where the soundscape evolves as a user navigates the virtual space. This creates an immersive cultural environment that encourages exploration.

Accommodate the User's Desire for Idealization. A technically robust foundation must also respect the user's experiential goals. The consistent suppression of visitor sounds indicates that users often seek a personal, contemplative connection with heritage content, free from social distractions. Designers can accommodate this preference by providing agency, perhaps through distinct listening modes such as a "Curator's Mix" (a complete, unfiltered soundscape) and a "Personal Mix" (which allows users to reduce social background sound). This acknowledges the user's potential role in curating their own experience.

Integrate a Subtle Interpretive Layer. This direction is the most critical for bridging the knowledge gap identified in our core finding. Instead of disruptive tutorials, designers should use "embedded narrative" techniques to offer optional, on-demand information (Jenkins, 2004). For a heritage soundscape, this could manifest as transient visual cues or optional audio annotations that explain the meaning of a bell or a chant when a user interacts with it. It is this carefully integrated semantic scaffolding that provides the necessary context for users to move from simply hearing the sounds to understanding their cultural significance.

6. Conclusion

This research investigated how users interact with the complex sonic layers of a living heritage site to inform the design of effective digital reconstructions. The study found that users actively curate an idealized experience, prioritizing the cultural sounds while suppressing social background sound. The core contribution of this work, however, is the empirical identification of a significant disconnect between user engagement and cultural understanding. While the interactive system was highly engaging for all, this engagement only translated into a cultural connection for those with prior knowledge, proving that interactive fidelity alone is insufficient to convey cultural value.

Building on these foundational insights, this research opens up clear and promising directions for the next phase of work. Having established a baseline for user-driven curation, future studies can now extend this inquiry to explore how interaction patterns might differ across more diverse user groups, such as international tourists or older visitors, to create universally accessible designs. More importantly, by empirically demonstrating the need for contextual information, this study provides the critical groundwork for a new, solution-oriented research agenda. The next logical step is to design and test the very interpretive layers proposed here, investigating which forms of subtle, embedded scaffolding are most effective at bridging the gap between engagement and understanding. This work moves the field from identifying a problem to actively designing its solution, heralding a future where digital heritage is not only immersive but also genuinely insightful for the widest possible audience.

Acknowledgments

We would like to express our sincere thanks to the Head Monk of Magoksa Temple, Venerable Wongyeong, for granting research access. We thank Dr. Hae Un Rii and Dr. Hyeseung Shim (*corresponding author) for their collaboration in connecting this research group with the religious communities, which made this research possible.

References

Axelsson, Ö., Nilsson, M. E., Berglund, B., 2010. A principal components model of soundscape perception. *The Journal of the Acoustical Society of America*, 128(5), 2836–2846.

Baayen, R. H., Davidson, D. J., Bates, D. M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4), 390–412.

Banfi, F., Bolognesi, C., 2021. Virtual reality for cultural heritage: New levels of computer-generated simulation of a unesco world heritage site. *From building information modelling to mixed reality*, Springer, 47–64.

Brown, A., Muhar, A., 2004. An approach to the acoustic design of outdoor space. *Journal of Environmental planning and Management*, 47(6), 827–842.

Champion, E., 2021. *Rethinking virtual places*. Indiana University Press.

Champion, E., Rahaman, H., 2020. Survey of 3D digital heritage repositories and platforms. *Virtual Archaeology Review*, 11(23), 1–15.

Dávid, L., Varga, I., Beták, N. et al., 2024. Soundscapes and sonicscapes in tourism: A decade of research insights from bibliometric analysis. *Journal of Infrastructure, Policy and Development*, 8(16), 8163.

Forlizzi, J., Battarbee, K., 2004. Understanding experience in interactive systems. *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, 261–268.

Glaser, M., Hug, L., Werner, S., Schwan, S., 2024. Spatial versus normal audio guides in exhibitions: Cognitive mechanisms and effects on learning. *Educational technology research and development*, 1–30.

Guastavino, C., 2006. The ideal urban soundscape: Investigating the sound quality of French cities. *Acta Acustica united with Acustica*, 92(6), 945–951.

ICOMOS, 1994. The nara document on authenticity. Adopted at the Nara Conference on Authenticity, Nara, Japan. Accessed: 2025-06-26.

Jenkins, H., 2004. Game design as narrative architecture. *Computer*, 44(3), 118–130.

Katz, B. F., Murphy, D., Farina, A., 2020. Exploring cultural heritage through acoustic digital reconstructions. *Physics today*, 73(12), 32–37.

Maffei, L., Brambilla, G., Di Gabriele, M. et al., 2015. Soundscape as part of the cultural and natural heritage. *Soundscape and the Built Environment*, CRC Press, Taylor and Francis Group, 215–242.

Mattioli, T., Díaz-Andreu, M., 2017. Hearing rock art landscapes: a survey of the acoustical perception in the Sierra de San Serván area in Extremadura (Spain). *Time and Mind*, 10(1), 81–96.

McCarthy, J., Wright, P., 2004. Technology as experience. *interactions*, 11(5), 42–43.

Nunnally, J. C., 1978. *Psychometric Theory: 2d Ed*. McGraw-Hill.

Othman, M. K., 2012. Measuring visitors' experiences with mobile guide technology in cultural spaces. PhD thesis, University of York Toronto, Canada.

Pérez-Martínez, G., Torija, A. J., Ruiz, D. P., 2018. Soundscape assessment of a monumental place: A methodology based on the perception of dominant sounds. *Landscape and Urban Planning*, 169, 12–21.

Privitera, A. G., Fontana, F., Geronazzo, M., 2024. The Role of Audio in Immersive Storytelling: a Systematic Review in Cultural Heritage. *Multimedia Tools and Applications*, 1–39.

Rumsey, F., 2002. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *Journal of the Audio Engineering Society*, 50(9), 651–666.

Schafer, R. M., 1993. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster.

Schmitt, D., Labour, M., 2022. Making sense of visitors' sense-making experiences: The REMIND method. *Museum Management and Curatorship*, 37(3), 218–234.

Seeger, A., 2004. *Why Suyá sing: A musical anthropology of an Amazonian people*. University of Illinois Press.

Skublewska-Paszkowska, M., Milosz, M., Powroznik, P., Lukasik, E., 2022. 3D technologies for intangible cultural heritage preservation—literature review for selected databases. *Heritage Science*, 10(1), 3.

Smith, L., 2006. *Uses of heritage*. Routledge.

Stockfelt, O., Schwarz, D., Kassabian, A., Siegel, L., 2004. Adequate modes of listening. *Popular music: Critical concepts in media and cultural studies*, 1, 375–91.

Titon, J. T., Cooley, T. J., Locke, D., McAllester, D. P., Rasmussen, A. K., Reck, D. B., Schechter, J. M., Stock, J. P., Suttun, R. A., 2009. *Worlds of music : an introduction to the music of the world's peoples*. 5th ed. edn, Schirmer Cengage Learning.

UNESCO, 1972. Convention concerning the protection of the world cultural and natural heritage. UNESCO General Conference.

UNESCO, 2003. Convention for the safeguarding of the intangible cultural heritage. UNESCO General Conference.

UNESCO World Heritage Centre, 2018. Sansa, buddhist mountain monasteries in korea. https://whc.unesco.org/en/list/1562/. Accessed: 2025-06-26.

Vorlaender, M., Summers, J., 2008. Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms, and Acoustic Virtual Reality. *The Journal of the Acoustical Society of America*, 123, 4028.

Wijesuriya, G., 2018. Living heritage. *Sharing conservation decisions: Current issues and future strategies*, 43–56.

Yong Jeon, J., Jik Lee, P., Young Hong, J., Cabrera, D., 2011. Non-auditory factors affecting urban soundscape evaluation. *The Journal of the Acoustical Society of America*, 130(6), 3761–3770.

Zotter, F., Frank, M., 2019. *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature.