architecture

MDPI

Article

Soundscape as Heritage Media Architecture: A Feng Shui-Informed Immersive VR Evaluation of a Pepper's Ghost Virtual Layer in Historic Kampung

Fransiskus Xaverius Teddy Badai Samodra [1,*], Sri Nastiti Nugrahani Ekasiwi [1], Audrey Tara Dianagri [1], Jeremy Lovendianto [1] and Juhee Nam [2]

[1] Department of Architecture, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; ekasiwi@its.ac.id (S.N.N.E.); audrey.dianagri@its.ac.id (A.T.D.); 6013231013@student.its.ac.id (J.L.)
[2] Department of Architecture, Pusan National University, Busan 46241, Republic of Korea; wngml7885@pusan.ac.kr
* Correspondence: samodra@its.ac.id

Abstract

Architectural heritage interventions often preserve visual form while leaving acoustic memory and everyday sound practices outside the design brief. This article positions soundscape as an architectural material for heritage media architecture and tests it through Sanggar Kungfu Kapasan, a historic Chinese-diaspora kampung in Surabaya, Indonesia. The research combined contextual and soundmark mapping, Pepper's Ghost prototyping, architectural translation of a semi-transparent virtual layer, and comparative evaluation using 2D visualization and immersive VR. After deduplication, 37 participants evaluated the 2D material, 31 evaluated VR, and 24 completed both modes for within-subject analysis. VR significantly improved perceived physical quality (Δ = +0.46, p = 0.0007, d_z = 0.80), communication quality (Δ = +0.42, p = 0.0218, d_z = 0.50), and overall evaluation (Δ = +0.20, p = 0.0285, d_z = 0.48). Sound-specific VR ratings were positive for place-meaning support (M = 5.87) and scene formation (M = 5.65), and the soundscape design index correlated with communication quality (r = 0.47, p = 0.007). The findings show that soundscape strengthens narrative legibility and architectural fit when deliberately coordinated with visual layering, scene sequencing, and cultural context. The study contributes a soundscape-informed, feng shui-readable workflow for pre-installation evaluation of multisensory heritage interventions. Because the respondent pool was predominantly composed of architecture students, the results are interpreted as a preliminary design-user evaluation rather than community validation.

Keywords: soundscape; heritage media architecture; immersive virtual reality; Pepper's ghost; digital placemaking; feng shui; psychoacoustics; adaptive reuse

Check for updates

Academic Editors: Miltiadis Katsaros, Konstantinos Serraos and Ioannis Timagenis

Received: 9 June 2026
Revised: 15 August 2026
Accepted: 21 August 2026
Published: 24 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Historic kampungs in Indonesian cities are living socio-spatial archives in which migration, ritual practice, informal economy, and collective memory remain intertwined with daily use. In Surabaya, Kampung Kapasan preserves a Chinese-diaspora urban fabric where Kung Fu training, lion-dance routines, temple-related activities, and street life continue to shape place attachment [1]. Such places require heritage interventions that go beyond mere conservation of visual objects. They must also support the experiential

conditions through which residents and visitors recognize the site, including the acoustic cues that make cultural practices audible.

Digital placemaking and media architecture offer a non-invasive way to reactivate such heritage environments without freezing them into museum-like artifacts. Recent studies define digital placemaking as the situated co-production of place meanings through digital media, spatial context, and human agency [2,3], while Indonesian heritage discourse frames conservation as a programmatic archive in which interpretation and everyday use remain part of urban continuity [4]. Yet media-architecture projects are still commonly judged by images, novelty, and spectacle, even though their effectiveness depends on multisensory interpretation, social readability, and contextual fit.

Soundscape theory provides the acoustic basis for this shift. ISO 12913 defines soundscape as the acoustic environment as perceived, experienced, and understood in context [5], and subsequent models show that appraisal is shaped by expectation, meaning, activity, and place identity as much as by sound-pressure signals [6–8]. For heritage environments, this is critical: sound can operate as intangible cultural heritage and as a medium through which historical buildings and cultural settings are encountered [9,10]. In media architecture, therefore, sound should be treated not as background ambiance but as a communicative layer that mediates between realism, symbolism, and spatial narrative [11].

Following Schafer, a soundmark is a community-recognized sound whose distinctiveness and social value make it important to preserve; however, situated sound studies caution that such meaning cannot be inferred from the signal alone [12,13]. Palmese et al. demonstrate how soundwalks, interview mapping, sound archives, and collective listening disclose memories and negotiated meanings that psychometric ratings cannot capture [13]. Thibaud's urban-ambiance research similarly treats atmosphere as a sensory, temporal, and social configuration, while Griffero conceptualizes atmosphere as an affective quality distributed across bodies and environments [14–16]. Ingold's critique of a detached soundscape perspective further emphasizes listening as embodied movement within a medium rather than observation of a bounded sonic object [17,18].

Accordingly, this study uses the terms experiential atmosphere (EA) and emotional resonance (ER) in a deliberately bounded psychometric sense. EA denotes perceived comfort, immersion, and coherence of audio-visual cues; ER denotes perceived cultural meaningfulness and affective connection to the represented heritage narrative. These constructs support comparative design review, but they do not constitute a full account of lived, relational, or community-interpreted sound experience. Such an account would require situated soundwalks, commented walks, interviews, and co-interpretation with residents and cultural practitioners [13–19].

To connect this acoustic perspective with architectural design, this article adapts the authors' architecture-entity framework and positions sound as an intangible but operative architectural layer (Figure 1). Architecture is understood as the integration of formal decisions—form, mass, proportion, material presence, and regulatory fit—and spatial decisions—room, context, system, atmosphere, sequence, and meaning. Sound crosses these domains because it modifies atmosphere, contextual fit, narrative legibility, and the perceived relation between body and place. Feng shui is used as a local-wisdom lens for discussing environmental harmony and relational balance, not as a deterministic rule set [20,21]. Its relevance derives from the site's Chinese–Indonesian cultural context and from feng shui's relational vocabulary for environmental fit; it is not used to claim causal, directional, or psychometric validity.

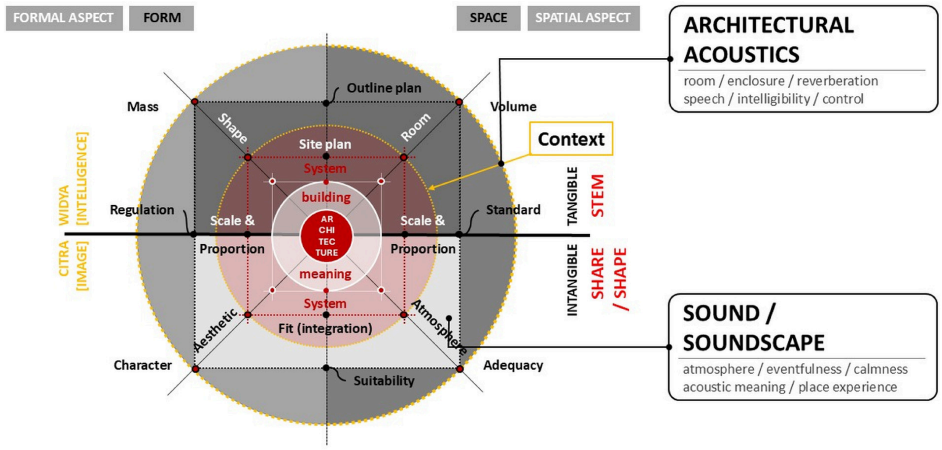


Figure 1. Original architecture-entity circle annotated to position sound/soundscape and architectural acoustics within the formal–spatial framework. Gray fields distinguish formal and spatial domains; the red core and nodes indicate integrated architectural relations; yellow outlines and arrows identify contextual and intangible soundscape links; black axes and callout lines organize the mapped attributes.

Comparable studies have used immersive environments to assess spatial judgment, urban scenarios, and sensory responses [22–30], while sonic placemaking research links identifiable auditory cues to attachment and public meaning [31,32]. Different from these strands, the present study connects a Pepper's Ghost optical prototype, candidate heritage soundmarks, paired 2D-VR evaluation, and an explicitly bounded feng shui interpretation within one pre-installation workflow. The novelty therefore lies in the integration and evaluative sequence, not in claiming that VR reproduces the full cultural life of the site.

This study investigates how a designed soundscape can be integrated into heritage media architecture for Sanggar Kungfu Kapasan through a Pepper's Ghost-based virtual layer and an immersive evaluation workflow. Building on immersive design-review research [22,23] and on the need for credible audio reproduction in VR soundscape studies [24–26], the paper addresses four questions: (1) does immersive VR improve perceived design quality compared with 2D visualization? (2) Which evaluation dimensions benefit most from immersion? (3) Does the designed soundscape support heritage meaning and scene formation? (4) How can feng shui help interpret the balance among form, atmosphere, resonance, and communication?

The contribution of this work is threefold. First, it expands architectural-acoustic discussion from comfort and noise control toward meaning-bearing soundscape design in a living heritage context. Second, it provides a reproducible design-research workflow that links site mapping, soundmark selection, Pepper's Ghost prototyping, 2D/VR representation, and paired statistical evaluation. Third, it translates feng shui from a symbolic reference into a bounded interpretive lens of harmony to examine whether an intervention balances xing (form clarity), qi (atmospheric continuity), shen (cultural resonance), and yi (narrative communication).

This positioning strengthens the manuscript beyond a descriptive design proposal. The study reports reliability, paired tests, effect sizes, sound-specific indices, and qualitative keyword evidence, allowing the media-architecture intervention to be evaluated as a multisensory architectural system rather than as a visual artifact with an added soundtrack. The resulting evidence is therefore framed as preliminary design-user appraisal rather than community-validated heritage evaluation.

2. Materials and Methods

2.1. Research Design and Theoretical Position

To answer these questions, the study used a design-research strategy in which inquiry, prototyping, and evaluation were treated as connected operations. Research-for-design established the spatial, cultural, and acoustic brief; research-through-design tested the feasibility of the virtual layer; and comparative evaluation examined how the representation mode changed users' judgments. Sanggar Kungfu Kapasan was selected because it combines active community practice, Chinese-diaspora heritage, narrow kampung spatial conditions, and a need for interpretation that does not physically overburden the existing setting (Figure 2).

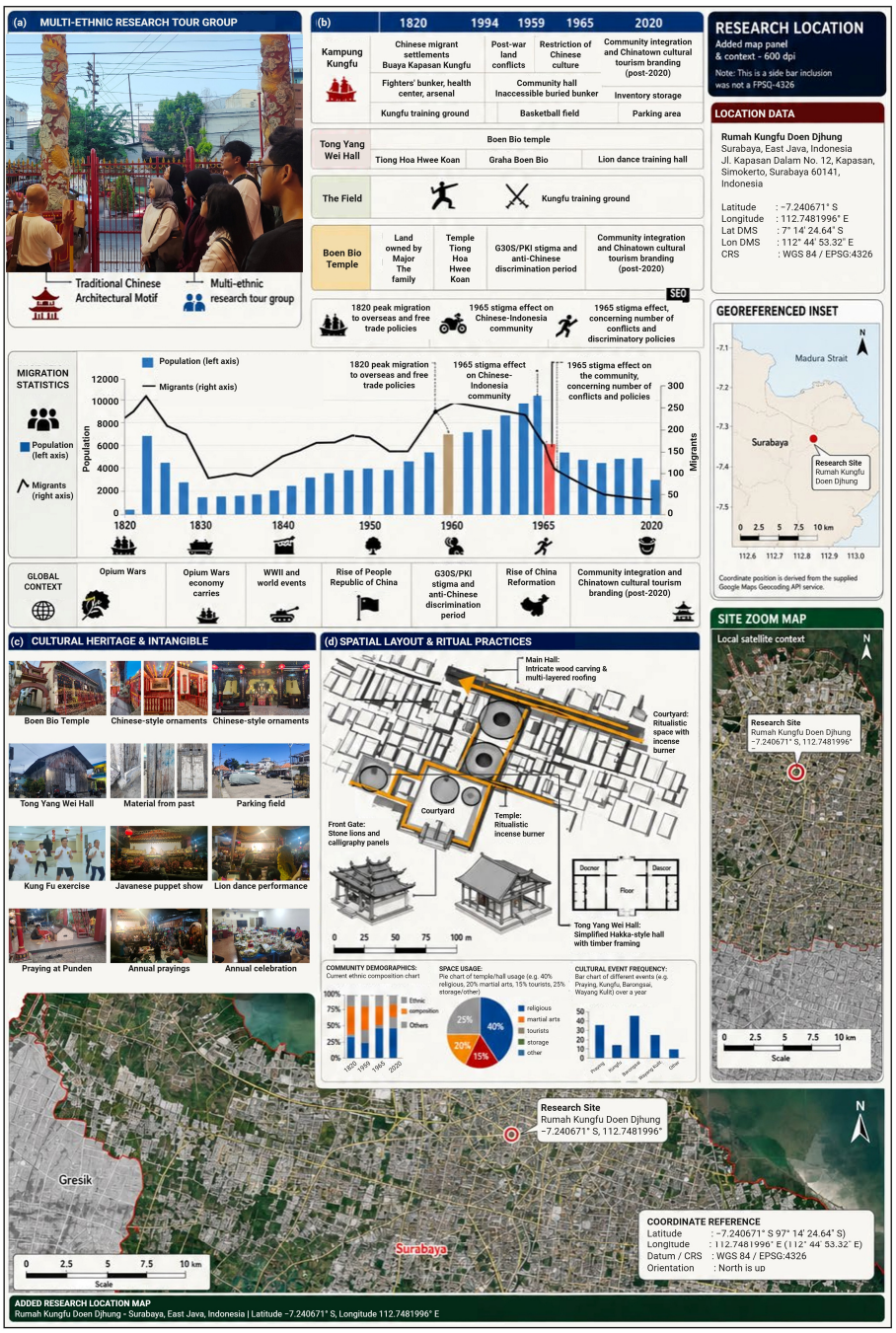


Figure 2. Geo-cultural context mapping of Sanggar Kungfu Kapasan within Kampung Kapasan, Surabaya. The analytical focus is the media-architectural threshold at Sanggar Kungfu Kapasan/Rumah Kungfu Doen Djihung; the surrounding timeline, practices, and maps provide contextual rather than separately tested variables.

Figure 2 locates Sanggar Kungfu Kapasan/Rumah Kungfu Doen Djihung at Jl. Kapasan Dalam No. 12, Kapasan, Simokerto, Surabaya 60141, East Java, Indonesia (WGS 84/EPSG:4326; lat. −7.240671° or 7°14′24.64″ S; long. 112.7481996° E or 112°44′53.32″ E), shown with north-up georeferenced and satellite maps. It summarizes the research tour context, Chinese architectural motifs, Kampung Kungfu–Tong Yang Wei Hall–The Field-Boen Bio Temple timeline (1820–2020), migration/global context, heritage practices, spatial–ritual layout, community demographics, space use, and event frequency as the geo-cultural basis for a soundscape media-architecture analysis.

Although Figure 2 summarizes the wider Kapasan heritage network, the unit of design intervention is narrower: the media-architectural threshold at Sanggar Kungfu Kapasan/Rumah Kungfu Doen Djihung and its perceptual relation to training, temple-associated practices, street movement, and the existing visual background. The study does not attempt a comprehensive digital reconstruction of the entire kampung.

The theoretical position is that sound is a design operation embedded in architectural decision-making. Within the framework of Figure 3, sound primarily affects atmosphere, context, meaning, fit, and system. Yet, these perceptual effects feed back into formal choices such as interface thickness, viewing threshold, media opacity, timing, and the alignment between visible and audible cues. This position is compatible with sonic placemaking approaches that connect identifiable soundmarks to place attachment, public meaning, and urban vibrancy [31,32].

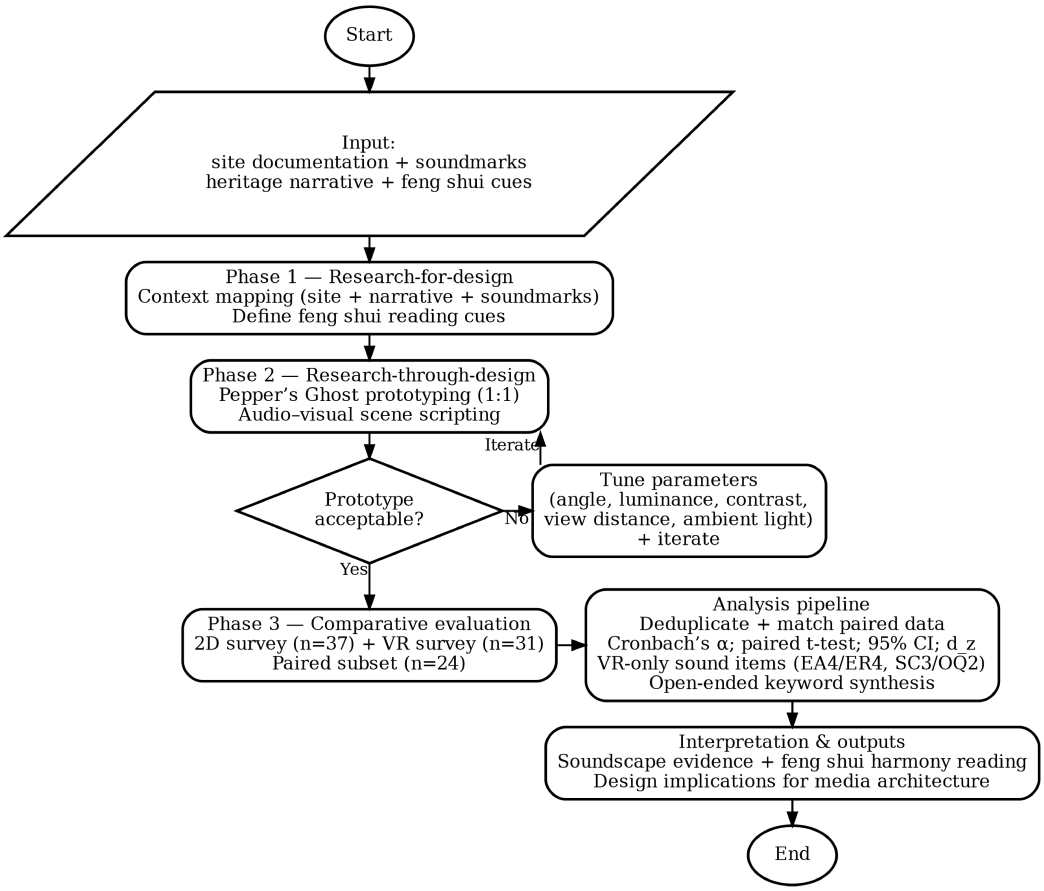


Figure 3. Algorithmic flowchart of the research workflow and analytic pipeline (soundscape + feng shui lens; 2D visualization vs. immersive VR evaluation).

2.2. Research-for-Design: Site Context Mapping and Heritage Soundmarks

Field observation and document review were organized into four mapping layers: (1) spatial boundaries, access points, and visual landmarks; (2) activity rhythms associ-

ated with training, ritual, commerce, and movement; (3) heritage narratives connected to Sanggar Kungfu Kapasan and its Chinese-diaspora setting; and (4) soundmarks, including martial-arts impacts, lion-dance or procession cues, temple-related sounds, market ambiance, and traffic. These layers generated a design brief in which the virtual layer had to communicate heritage meaning while remaining compatible with everyday circulation, background visibility, and the sonic life of the kampung. This mapping was conducted to generate a design brief; it was not an ethnographic sound-history study or a community soundwalk. The listed sounds are therefore treated as candidate soundmarks for prototyping, pending community confirmation.

2.3. Research-Through-Design: Pepper's Ghost Prototyping and Architectural Translation

The research-through-design phase converted the concept into testable interface parameters. Two Pepper's Ghost reflector configurations were compared in terms of distortion, viewing cone, footprint, legibility, and ease of architectural integration. The prototype was then refined by adjusting reflector angle, projection luminance, background contrast, viewing distance, and ambient-light control (Figure 4). These variables were treated as architectural criteria because they determine whether the virtual image reads as an inhabitable layer, a threshold, or merely a detached projection artifact.

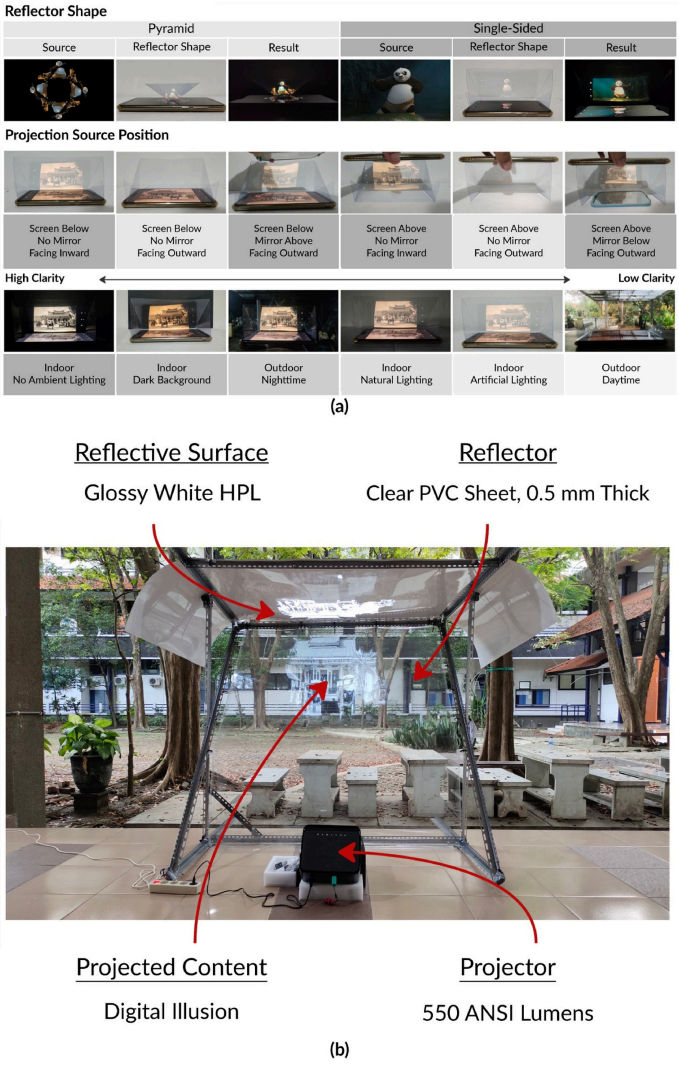


Figure 4. (**a**) Bench-scale Pepper's Ghost feasibility experiment; (**b**) Reflector geometry, projection-source position, ambient-light/background clarity tests, and the architectural prototype identifying the reflective surface, clear PVC reflector, projected content, and projector.

The experiment was a bench-scale feasibility test rather than a participant outcome test. It compared reflector form and projection-source position under contrasting background and ambient-light conditions, establishing the optical envelope later translated into the architectural proposal. Its importance lies in identifying the viewing, contrast, and installation constraints that the immersive evaluation subsequently represented.

The architectural translation paired these optical criteria with soundscape composition. Scene-specific soundmarks and ambiance were assigned to narrative segments to reinforce place meaning, foreground key actions, and control masking between the virtual layer and the existing kampung environment. In this sense, the design workflow connects perceptual soundscape thinking with architectural-acoustic evidence that material configuration, masking, and physical audio conditions must be considered together when evaluating built-space experience [33,34].

Figure 5 synthesizes the prototype observations into architectural translation criteria. Internal factors—human scale, field of view, eye-level height, and distance to users—determine whether the reflection can be perceived as part of the surrounding envelope. External factors—background brightness, contextual noise, content contrast, and installation integration—determine whether the same layer remains legible in the kampung. The figure, therefore, establishes the virtual surface as a threshold element: visible enough to carry heritage content, yet transparent enough to keep the existing site, movement, and ambient activities perceptually present.

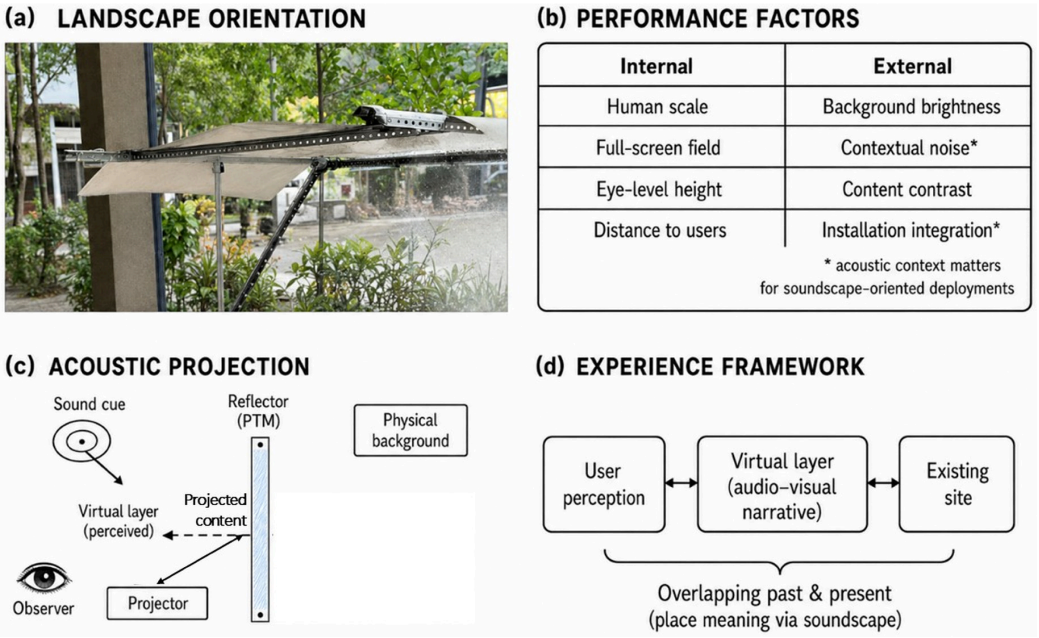


Figure 5. Architectural translation of the outdoor Pepper's Ghost prototype: (**a**) landscape orientation; (**b**) internal and external performance factors; (**c**) acoustic-projection relation; and (**d**) experience framework linking user perception, the audio-visual virtual layer, and the existing site.

In Figure 5, panel (a) records the landscape-oriented installation, panel (b) separates internal and external performance factors, panel (c) shows the relation among observer, projector, reflector, sound cue, and perceived virtual layer, and panel (d) states the experiential mechanism. In panel (d), user perception mediates a two-way relation between the audio-visual virtual layer and the existing site; their overlap is intended to connect past and present without erasing the physical background. This mechanism is the conceptual bridge between the prototype and the place-meaning measures used in the evaluation.

The same synthesis informed the soundscape strategy. Because the virtual image is semi-transparent and sensitive to ambient light, soundmarks act as stabilizing cues that

help users bind the image to the physical place. Martial-arts impacts, ritual or procession sounds, and street ambiance were therefore assigned as scene-specific anchors rather than continuous background layers. This reduces visual overload and allows sound to clarify spatial intention, narrative timing, and user orientation.

The prototype established three translation rules for the architectural proposal. First, visual content should remain within a stable viewing cone and align with architectural surfaces so that the reflection reads as an inhabitable layer. Second, audio cues should be deliberately coordinated with the intended visual scene because unrelated or mistimed sound can weaken illusion and narrative credibility. In the reported evaluation, this coordination was implemented manually at the scene level: the assigned backing track was launched separately for each visual-only 360° panorama rather than embedded or head-tracked. Third, the physical installation should preserve everyday circulation and background visibility because the kampung context is part of the intervention, not a neutral backdrop.

For reproducibility, the evaluated audio was not a bespoke multitrack or spatial-audio composition. Each visual-only 360° panorama was paired with a scene-matched YouTube backing track selected from the project cue list. Across the panorama set, the source materials represented four cue families identified during research-for-design: martial-arts impacts/training, lion-dance or processional cues, temple-related sounds, and street/market ambiance. The panorama appeared first in the Meta Quest 3S HMD at visual onset (t_0); after an unmeasured manual delay (Δ_m), the corresponding track was started on a separate device and played through a fixed room loudspeaker until the end of the scene (t_e). Audio was therefore independent of head rotation, not embedded in the panorama or HMD, and not binaural, ambisonic, or head-tracked. Supplementary Material S1 provides representative session captures, the source/cue categories, and the relative timeline. Exact start/stop times, playback level, loudspeaker coordinates, and online-source timecodes were not logged and are not reconstructed.

2.4. Evaluation Media and Participants

Two evaluation media were prepared from the same design proposal. The 2D condition used renderings and storyboard sequences to represent form, narrative order, and interface placement (Figure 6). The immersive condition used a Meta Quest 3S head-mounted display (HMD) to present animation videos and visual-only 360° panoramas for first-person inspection of scale, transparency, and scene sequence. For each 360° panorama, the assigned backing track was manually started and played separately through a loudspeaker while the participant viewed the panorama; the audio was co-presented with, rather than embedded in, the panorama. The 2D questionnaire yielded 38 submissions; after deduplication of repeated names by retaining the latest timestamp, 37 cases remained valid. The VR questionnaire yielded 31 valid cases. Twenty-four respondents completed both conditions and were retained for paired within-subject analysis (Table 1). Figure 6 presents the shared design content rather than a second experiment: panel (a) shows the daytime exterior and virtual-layer alternatives; panel (b) the nighttime site condition; panel (c) the eye-level heritage projection; panel (d) the viewing-field and coverage logic; panel (e) the heritage-content composition; and panel (f) the interior/sectional relationship among projector, reflector, observer, and projected layer.

The paired subset was analytically prioritized because within-subject comparison reduces interpersonal variability when testing representation-mode effects. Respondents were predominantly young adults, and previous HMD-VR experience was more frequent in the VR dataset than in the 2D dataset (Figure 7). This profile is consistent with recent VR design-review research, which notes that immersive evaluation often attracts participants

with at least some familiarity with digital interfaces [23]. The imbalance is acknowledged as a sampling limitation, while the paired analysis helps control part of its influence. CAVE refers to a cave automatic virtual environment, a room-scale immersive projection system; it is reported only as prior-experience background. Because most respondents were architecture students or design-literate young adults, the sample is interpreted as a preliminary design-user/expert-user group, not as a proxy for residents, kungfu practitioners, temple users, or other community stakeholders.

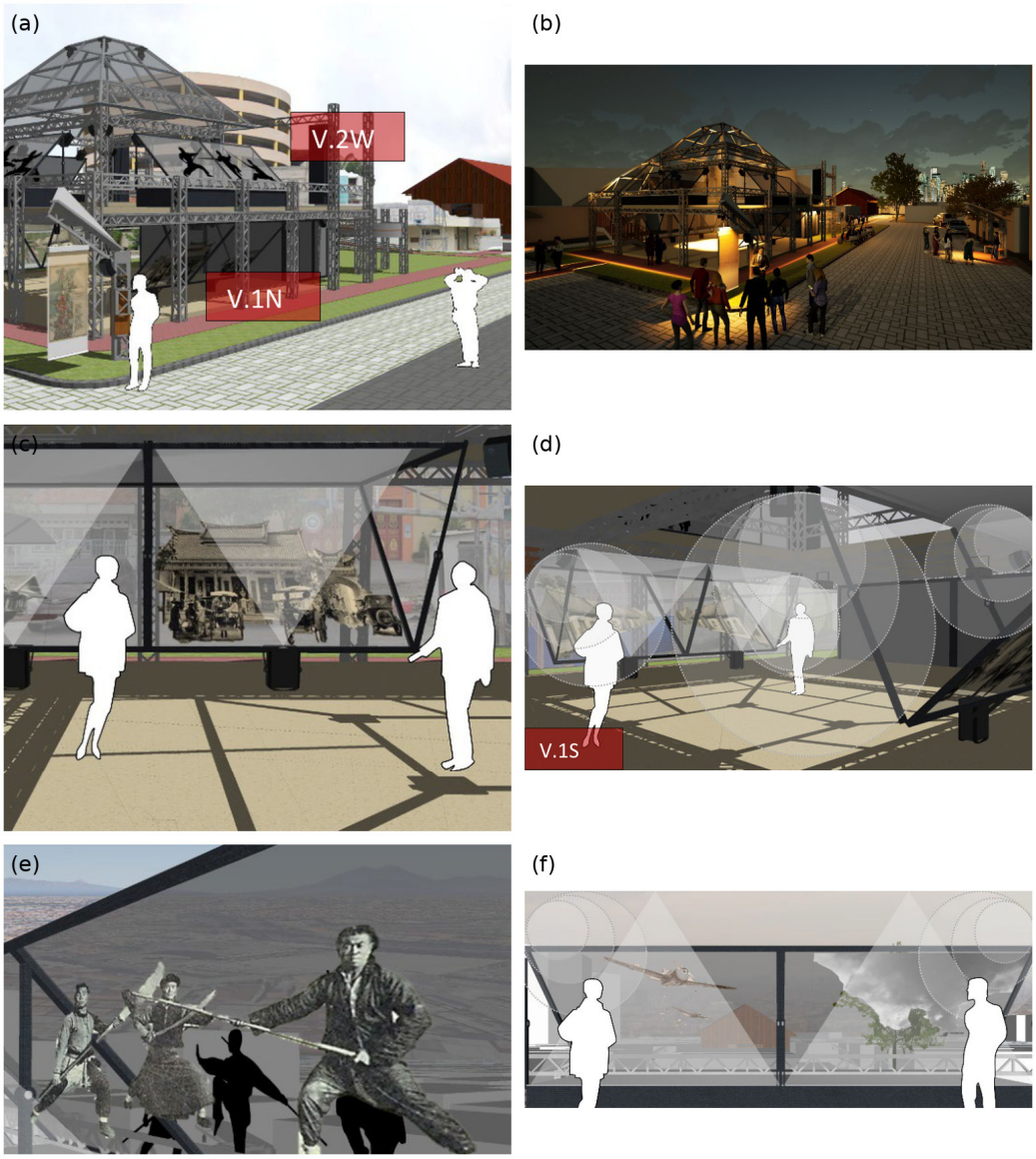


Figure 6. Shared design proposal used for 2D evaluation: (**a**) daytime exterior and virtual-layer alternatives; (**b**) nighttime site condition; (**c**) eye-level heritage projection; (**d**) field-of-view and coverage logic; (**e**) heritage-content composition; and (**f**) interior/sectional projection relationship. White silhouettes indicate observer scale; translucent cones and circles show schematic viewing fields and projection/coverage extents; red tags identify the tested visual alternatives. These graphic elements are explanatory, not additional outcome variables.

Table 1. Participant profile and dataset size (after deduplication).

Variable	2D (n = 37)	VR (n = 31)	Paired (n = 24)
Valid respondents (n)	37	31	24
Female	27 (73.0%)	22 (71.0%)	19 (79.2%)
Age, M (SD), years	22.89 (2.31)	23.00 (2.37)	22.96 (2.27)
Prior HMD-VR experience	Never: 20 (54.1%); ever: 17 (45.9%)	Never: 8 (25.8%); ever: 23 (74.2%)	Never: 13 (54.2%); ever: 11 (45.8%)
Prior CAVE experience	Never: 29 (78.4%); ever: 8 (21.6%)	Never: 21 (67.7%); ever: 10 (32.3%)	Never: 18 (75.0%); ever: 6 (25.0%)

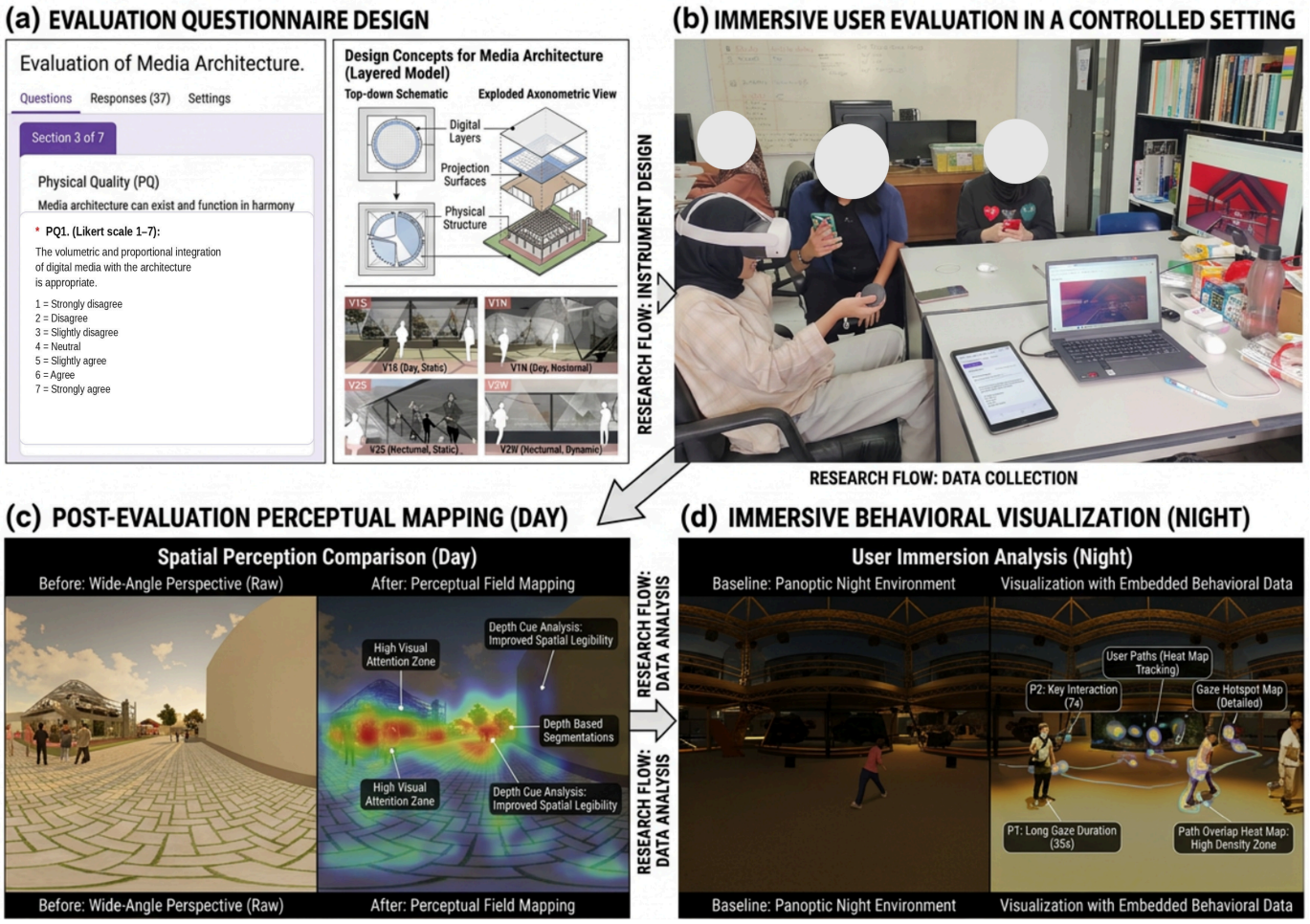


Figure 7. Evaluation workflow: (**a**) questionnaire design and shared media-architecture concepts; (**b**) immersive user evaluation in a controlled setting, with participant identities concealed in the published image; (**c**) post-evaluation perceptual mapping of the daytime scene; and (**d**) nighttime immersive behavioral visualization with gaze and movement annotations. In panels (**c**,**d**), heat-map colors and path/circle overlays are schematic annotations of relative attention and movement features rather than calibrated eye-tracking measurements.

2.5. Instruments and Variables

Both questionnaires operationalized four shared design-quality constructs on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree): physical quality (PQ1–PQ3), experiential atmosphere (EA1–EA3), emotional resonance (ER1–ER3), and communication quality (CQ1–CQ3) (Table 2). Item wording was kept parallel across 2D and VR where possible so that representation-mode effects could be interpreted at the construct level rather than as changes in survey language.

Table 2. Questionnaire constructs and sound-focused additions (all items rated on a 1–7 Likert scale).

Construct	Item Codes	Operational Definition (English)	Mode
Physical Quality (PQ)	PQ1–PQ3	Legibility, clarity, and physical plausibility of the media-architecture layer and its integration with the existing setting.	2D & VR
Experiential Atmosphere (EA)	EA1–EA3	Perceived atmosphere and spatial experience (e.g., comfort, immersion, and coherence of cues).	2D & VR
Emotional Resonance (ER)	ER1–ER3	Affective and cultural resonance (e.g., perceived meaningfulness and connection to heritage narratives).	2D & VR
Communication Quality (CQ)	CQ1–CQ3	Clarity of design intent and narrative communication (how well the proposal communicates place meaning).	2D & VR
Soundscape Design (Explicit)	EA4, ER4	EA4: soundscape supports place meaning; ER4: soundscape helps form scenes in the media-architecture narrative.	VR only
VR-Session Audio Evidence (System-Level)	SC3, OQ2	SC3: ability to hear the accompanying soundscape during VR; OQ2: perceived output quality of the separately reproduced backing audio.	VR only

The construct labels EA and ER derive from the original survey wording, but are interpreted here as experiential atmosphere and emotional resonance/responsiveness, as the items address mood, scene formation, cultural meaningfulness, and placemaking. Feng shui is used to interpret harmony among these constructs—especially the balance among form clarity, atmospheric continuity, resonance, and communication—in line with empirical and Indonesian discussions of feng shui as an environmental-fit and well-being lens [20,21,35–37].

2.6. Statistical Analysis

To improve statistical transparency, construct means, reliability, paired inference, effect size, correlation, and study-specific indices were computed using the standard forms summarized in Equations (1)–(6). For each respondent i and each construct g, the construct mean was calculated from item scores, and the overall score was the mean of the four shared construct scores. Internal consistency was estimated with Cronbach's alpha. Representation-mode effects were tested with two-tailed paired-samples t-tests at $p < 0.05$, supported by 95% confidence intervals and the within-subject effect size d_z. A soundscape design index (SDI) was defined from EA4 and ER4, and open-ended responses were coded through keyword counts to compare the salience of sound-related comments across modes.

$$C_{ig} = \frac{1}{k_g}\sum_{j=1}^{k_g} x_{ijg}, \quad O_i = \frac{\mathrm{PQ}_i + \mathrm{EA}_i + \mathrm{ER}_i + \mathrm{CQ}_i}{4} \tag{1}$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{2}$$

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{j=1}^{k} s_j^2}{s_T^2}\right) \tag{3}$$

$$d_i = x_{\mathrm{VR},i} - x_{\mathrm{2D},i}, \quad t = \frac{\overline{d}}{s_d/\sqrt{n}}, \quad \mathrm{CI}_{95\%} = \overline{d} \pm t_{0.975,n-1}\frac{s_d}{\sqrt{n}} \tag{4}$$

$$d_z = \frac{\overline{d}}{s_d} = \frac{t}{\sqrt{n}}, \quad r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{5}$$

$$\mathrm{SDI}_i = \frac{\mathrm{EA4}_i + \mathrm{ER4}_i}{2}, \quad G_i = \mathrm{PQ}_i - \mathrm{ER}_i, \quad D_m = \mathrm{SD}\{\overline{C}_{\mathrm{PQ}}, \overline{C}_{\mathrm{EA}}, \overline{C}_{\mathrm{ER}}, \overline{C}_{\mathrm{CQ}}\}, \quad K_w = \sum_i \mathbf{1}(w \in c_i) \tag{6}$$

Notation: x_{ijg} is respondent i's score on item j in construct g; C_{ig} is the respondent-level construct mean; O_i is the overall score; M and $\overline{x}$ denote sample means; s and SD denote sample standard deviations; α is Cronbach's alpha; d_i, $\overline{d}$, and s_d denote the paired difference, mean paired difference, and standard deviation of paired differences; d_z is the within-subject effect size; r is the Pearson correlation coefficient; SDI_i is the soundscape design index; G_i is the form–shen gap; D_m is the profile-dispersion index; and K_w is the keyword count for keyword w.

2.7. Validity Controls and Ethical Considerations

Three safeguards were used to reduce overinterpretation. First, the paired subset was used for the main representation-mode comparison so that each participant served as their own control. Second, reliability values were reported for each multi-item construct, and constructs with modest alpha values were interpreted together with item-level patterns. Third, sound-specific claims were separated into design-level evidence (EA4 and ER4) and system-level evidence (SC3 and OQ2), preventing positive soundscape interpretation from being confused with the perceived output quality of the combined HMD-visual and separate-loudspeaker presentation.

The open-ended keyword counts were used only as supplementary evidence of salience. They do not replace qualitative accounts of memory, affect, cultural interpretation, or interpersonal negotiation; no claims of phenomenological or ethnographic saturation are made.

Participation was voluntary, and informed consent was obtained from all participants. Names were used only for deduplication and paired-case matching; the reported analyses contain no personal identifiers, and participant identities are concealed in Figure 7. The evaluation consisted of perception-based questionnaires and non-invasive VR viewing, without physiological measurements, deception, or the collection of sensitive personal data.

3. Results

3.1. Descriptive Results and Reliability

Table 3 summarizes the descriptive statistics and reliability estimates for the four shared constructs in the 2D and VR datasets. All construct means were above the neutral midpoint in both modes, indicating favorable baseline evaluations. Internal consistency was acceptable to high for most constructs and for the overall composite, whereas physical quality showed modest reliability in both modes. Consequently, the physical-quality findings are interpreted in terms of the item pattern and paired differences rather than as a stand-alone latent scale.

3.2. Representation-Mode Comparison

In the paired sample (n = 24), immersive VR produced statistically significant gains in physical quality, communication quality, and overall evaluation relative to 2D visualization (Table 4; Figures 8 and 9). The strongest gain was physical quality (Δ = +0.46, p = 0.0007, d_z = 0.80), followed by communication quality (Δ = +0.42, p = 0.0218, d_z = 0.50) and the overall composite (Δ = +0.20, p = 0.0285, d_z = 0.48). By contrast, experiential atmosphere increased only slightly (Δ = +0.10, p = 0.4748), and emotional resonance decreased non-significantly (Δ = −0.18, p = 0.2794).

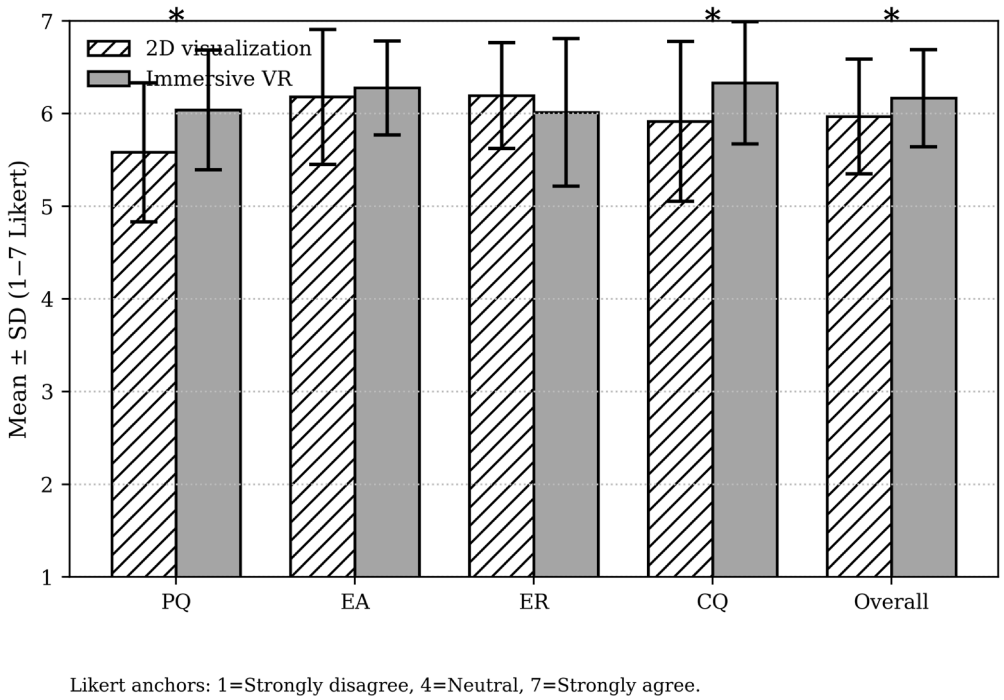
Table 3. Descriptive statistics and reliability (Cronbach's α) for shared constructs (2D: $n = 37$; VR: $n = 31$).

Dimension/Item	2D *M* (SD)	VR *M* (SD)	Cronbach's α (2D/VR)
Physical Quality (PQ1–PQ3)	5.54 (0.85)	5.95 (0.69)	0.51/0.58
Experiential Atmosphere (EA1–EA3)	6.03 (0.86)	6.19 (0.50)	0.75/0.66
Emotional Resonance (ER1–ER3)	6.18 (0.63)	5.92 (0.86)	0.74/0.72
Communication Quality (CQ1–CQ3)	5.86 (0.91)	6.27 (0.63)	0.83/0.71
Overall Composite (12 Common Items)	5.90 (0.67)	6.08 (0.55)	0.88/0.86
Soundscape Supports Place Meaning (EA4)	—	5.87 (1.20)	—
Soundscape Helps Form Scenes (ER4)	—	5.65 (1.25)	—
Audibility of Accompanying Soundscape During VR (SC3)	—	5.13 (1.45)	—
Separate Audio Playback Quality During VR (OQ2)	—	4.90 (1.27)	—

Table 4. Paired comparison between 2D visualization and immersive VR (paired sample: $n = 24$; two-tailed paired *t*-test; 1–7 Likert scale).

Dimension	2D *M* (SD)	VR *M* (SD)	Δ (VR − 2D)	*t* (df)	*p*	d_z
PQ	5.58 (0.75)	6.04 (0.65)	+0.46	3.93 (23)	0.0007 *	0.80
EA	6.18 (0.73)	6.28 (0.51)	+0.10	0.73 (23)	0.4748	0.15
ER	6.19 (0.57)	6.01 (0.80)	−0.18	−1.11 (23)	0.2794	−0.23
CQ	5.92 (0.86)	6.33 (0.66)	+0.42	2.46 (23)	0.0218 *	0.50
Overall	5.97 (0.62)	6.17 (0.53)	+0.20	2.34 (23)	0.0285 *	0.48

Note: * indicates statistical significance at $p < 0.05$ (two-tailed paired *t*-test).



Likert anchors: 1=Strongly disagree, 4=Neutral, 7=Strongly agree.

Figure 8. Mean ratings by representation mode (paired sample: $n = 24$; $M \pm$ SD). * $p < 0.05$.

Figures 8 and 9 show that the effect of immersion was selective rather than uniform. VR clarified the physical plausibility of the virtual layer and strengthened narrative communication, but it did not automatically intensify affective response. This pattern is important because it prevents overclaiming: the evidence supports VR as a tool for improving architectural legibility and communication, while emotional resonance remains a design problem requiring further audio-visual tuning. Two psychoacoustic mechanisms may contribute to

this pattern [38]. Static 2D images can invite participants to imaginatively complete missing acoustic detail, whereas VR exposes mismatches among source action, timing, masking, spectral balance, and apparent source location more directly. In addition, the moderate system-level audio ratings may partly reflect the dual-device presentation: participants viewed visual-only panoramas in the HMD. At the same time, the assigned backing track was reproduced from a separate loudspeaker. Because the audio was neither embedded nor head-tracked and was launched manually rather than by a common timecode, this arrangement may have constrained audio-visual coherence and ER. More fundamentally, the non-significant ER decrease may signal that short, individual VR exposure cannot reproduce the socially situated memories and shared practices through which cultural–emotional depth is formed; this possibility is treated as a methodological constraint rather than only a tuning problem.

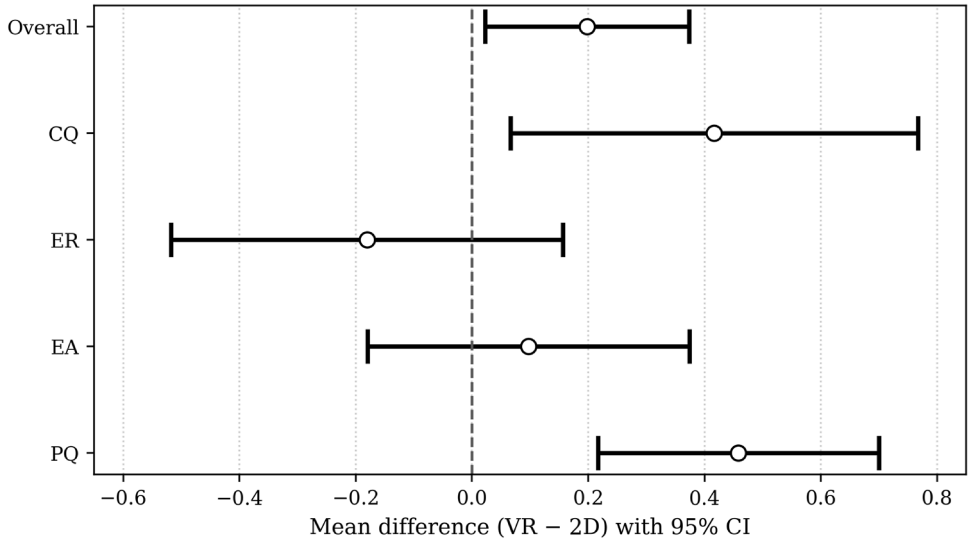


Figure 9. Paired mean differences (VR − 2D) with 95% CIs across evaluation dimensions (n = 24).

3.3. Sound-Specific Findings in the Immersive Condition

Figure 10 isolates the sound-specific results from the VR dataset. Respondents agreed that the soundscape supported place meaning (EA4: M = 5.87) and helped form narrative scenes (ER4: M = 5.65). At the system level, audibility was positive but lower (SC3: M = 5.13), and audio-output quality was acceptable but moderate (OQ2: M = 4.90). The distinction between design-level and system-level items indicates that the soundscape concept was stronger than the separately reproduced loudspeaker playback used to accompany the visual-only panoramas.

Open-ended comments reinforce this interpretation. Figure 11 shows that sound-related keywords were more salient in the immersive condition than in the 2D comments, even within the same paired subset. VR, therefore, changed not only what participants saw but also what they attended to and verbalized. In architectural terms, the soundscape entered the proposal's interpretation as a placemaking and narrative device rather than an optional atmospheric layer.

3.4. Exploratory Feng Shui-Soundscape Reading of the Results (Non-Inferential)

Rather than using feng shui as a stand-alone proof, Table 5 remaps the evaluation constructs into a harmony reading of form (xing), qi, shen, yi, and balance. Figure 12 shows the paired quality profile, and Figure 13 plots physical quality against emotional resonance as a form–shen balance plane. The paired means indicate that the negative form–shen gap in 2D (PQ < ER) approached zero in VR, meaning that immersion improved the perceived

fit between the virtual layer's physical clarity and its cultural reading. This remapping was specified after the questionnaire constructs had been analyzed; it was not a preregistered scale, a causal model, or an inferential test of feng shui.

Profile dispersion (D_m) across PQ, EA, ER, and CQ decreased from 0.25 in 2D to 0.14 in VR, indicating a more internally balanced appraisal profile. The soundscape design index was positive (M = 5.76, SD = 1.15) and correlated with communication quality (r = 0.47, p = 0.007). These results support a specific interpretation: the soundscape strengthened yi, or narrative legibility, more clearly than it strengthened shen, or emotional resonance. Accordingly, the following reading is descriptive and hypothesis-generating.

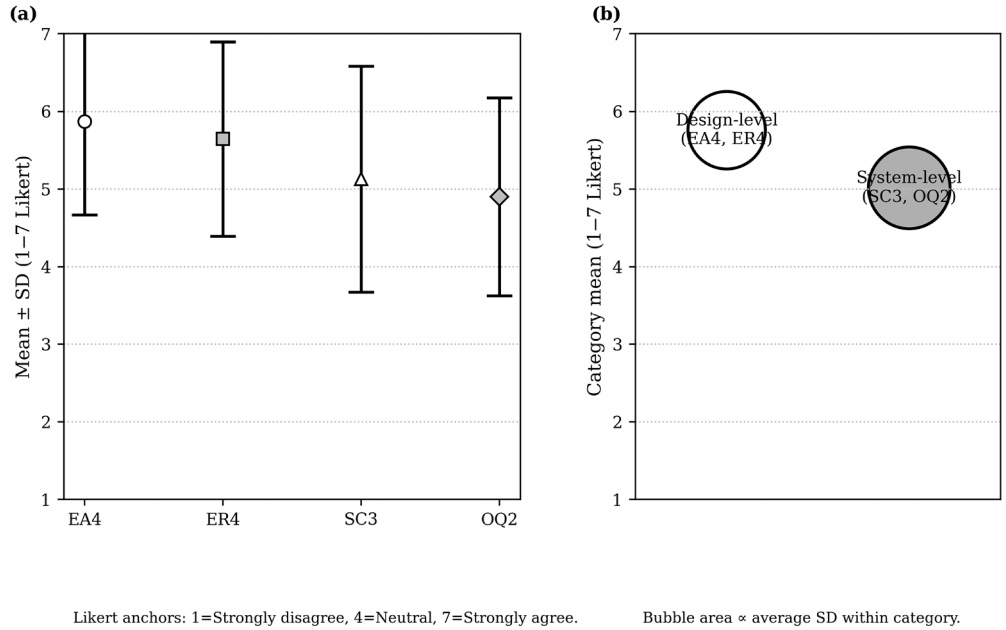


Figure 10. Sound-focused ratings in immersive VR: (**a**) item ratings (n = 31, $M \pm$ SD) and (**b**) category bubbles separating design-level (EA4/ER4) from system-level (SC3/OQ2) evidence; bubble area is proportional to the category-average SD. Open circle, square, triangle, and diamond markers identify EA4, ER4, SC3, and OQ2, respectively; white and gray fills distinguish design-level and system-level evidence. Labels do not obscure the plotted values.

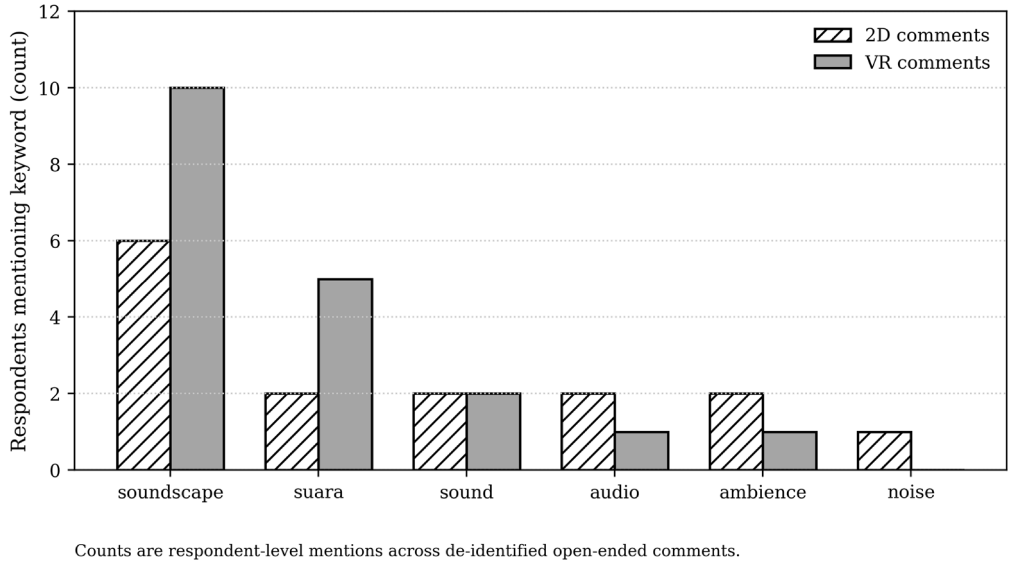


Figure 11. Sound-related keyword counts in open-ended comments for paired participants (n = 24).

Table 5. Exploratory feng shui-soundscape remapping of evaluation results (paired sample: $n = 24$; VR-only items: $n = 31$). The mapping is post hoc and non-inferential.

Feng Shui Lens	Proxy in This Study	Empirical Evidence	Interpretation/Contribution
Form (Xing)	PQ (paired, $n = 24$)	PQ: 5.58 (0.75) → 6.04 (0.65); $\Delta = +0.46$, $p = 0.0007$	VR increases form clarity, supporting coherent perception of the virtual layer.
Qi/Atmosphere	EA (paired, $n = 24$)	EA: 6.18 (0.73) → 6.28 (0.51); $\Delta = +0.10$, $p = 0.47$	The atmosphere is high in both modes; VR provides marginal uplift.
Shen/Resonance	ER (paired, $n = 24$)	ER: 6.19 (0.57) → 6.01 (0.80); $\Delta = -0.18$, $p = 0.28$	Affective resonance is not improved; it points to the need for audio-visual tuning.
Yi/Communication	CQ (paired, $n = 24$)	CQ: 5.92 (0.86) → 6.33 (0.66); $\Delta = +0.42$, $p = 0.022$	Narrative clarity increases; sound serves as a carrier of meaning for place identity.
Harmony Balance	Overall + profile balance	Overall: 5.97 (0.62) → 6.17 (0.53); $\Delta = +0.20$, $p = 0.029$; D_m: 0.25 → 0.14	In addition to mean gains, the perceived profile becomes more balanced in VR.
Yin–Yang Sonic Balance	Soundscape index + system quality (VR-only)	SDI (EA4/ER4): 5.76 (1.15), $n = 31$; correlated with CQ ($r = 0.47$, $p = 0.007$); SC3: 5.13 (1.45); OQ2: 4.90 (1.27)	Sonic fitting supports communication; different from purely visual cues, audibility and audio quality constrain full resonance.

Read together with Table 5, Figure 12 shows that immersive VR compressed the distance among xing, qi, shen, and yi into a more coherent appraisal profile. In the 2D condition, emotional resonance remained high. At the same time, physical quality and communication quality lagged, suggesting that respondents could sympathize with the heritage idea before the virtual layer became fully legible as architecture. In VR, first-person scale, scene sequence, and sonic cueing helped close that gap. Here, 'more coherent' refers only to lower descriptive dispersion among the four questionnaire means; it is not a validated feng shui threshold.

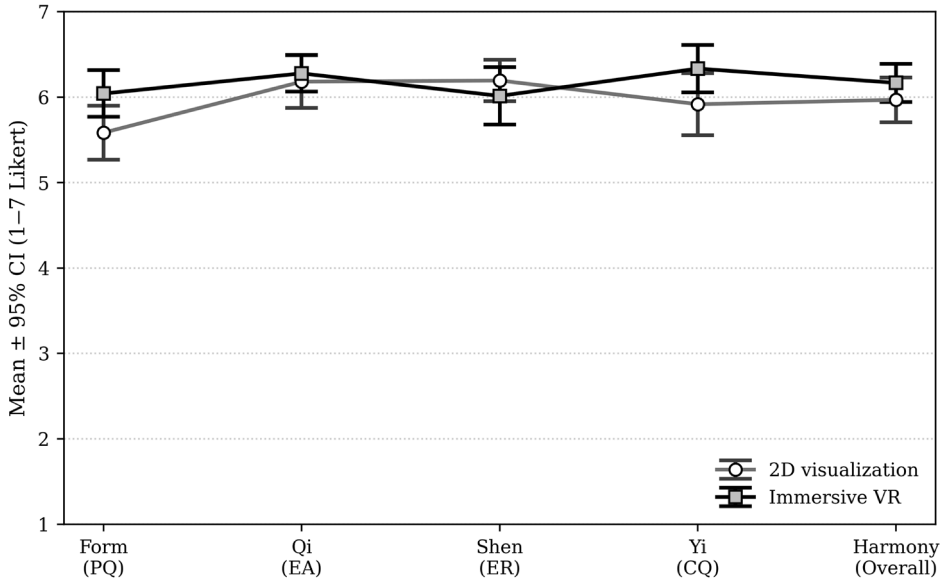


Figure 12. Feng shui-coded quality profile by representation mode (paired sample: $n = 24$; $M \pm 95\%$ CI; anchors: 1 = strongly disagree, 4 = neutral, 7 = strongly agree).

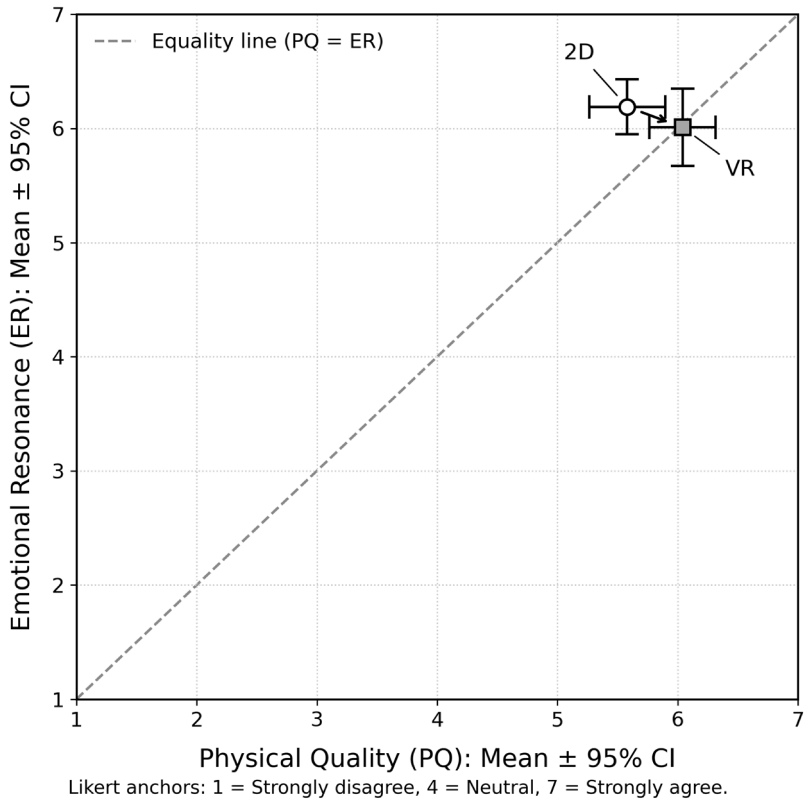


Figure 13. Exploratory form–shen plane (paired sample: n = 24): mean physical quality (PQ) versus emotional resonance (ER) with 95% CIs. The dashed diagonal is a descriptive equality reference (PQ = ER), not a validated harmony threshold. The open circle denotes 2D, the gray square denotes VR, horizontal and vertical bars show 95% CIs, the solid arrow shows the descriptive shift from 2D to VR, and the dashed diagonal denotes PQ = ER.

In practical terms, the profile points to the next design task: deepen shen without weakening the gains in xing and yi. Later iterations should refine the timing of soundmarks, the density of ambient layers, the strength of visual overlays, and the transition between calm and lively segments. For heritage media architecture, selective calibration is more productive than simply increasing volume, contrast, or narrative complexity.

Figure 13 clarifies why the paired means matter architecturally. The 2D point sits above the balance line because respondents could appreciate the intended cultural resonance even when the interface still lacked physical clarity. The VR point moves closer to the diagonal by elevating physical quality while maintaining emotional resonance within the same positive range, indicating a more proportionate relationship between visible structure and affective reading. In feng shui terms, the immersive condition reduces disequilibrium between xing and shen; in design terms, it shows that the proposal becomes easier to trust as an architectural insertion when form, audio cueing, and narrative sequence are perceived together. The diagonal in Figure 13 is therefore a visual equality reference, not a statistically validated criterion of harmony.

4. Discussion

4.1. Why Sound Must Be Treated as Architecture, Not Soundtrack

The main conceptual implication is that sound becomes architectural only when it participates in spatial organization, interface timing, and cultural interpretation. In the Kapasan proposal, sound did not operate as a decorative soundtrack added after visual design; it guided scene recognition, stabilized the semi-transparent image, and helped respondents understand the virtual layer as a contextual threshold. This mechanism

explains why the clearest empirical gain appeared in communication quality rather than in a generic atmosphere score.

4.2. Relation to Immersive Design Review and VR Soundscape Research

The results also refine the argument for immersive design review. Previous benchmarking and end-user studies indicate that immersive environments can approach or outperform 2D and physical prototypes for spatial judgment and user-built environment exploration [22,23]. The present study extends that evidence to heritage media architecture by showing that immersion can clarify not only scale and physical plausibility but also the communicative status of a digital layer attached to an existing setting.

The non-significant findings for experiential atmosphere and emotional resonance are equally important. They suggest that VR representation alone is not sufficient to guarantee multisensory coherence. The separate-loudspeaker setup provided sufficient audibility for narrative communication, but the moderate system-level ratings indicate that manual dual-device playback, the absence of documented level calibration, and the lack of head-tracked spatialization may have limited multisensory coherence. Calibrated binaural or ambisonics reproduction, or a controlled multichannel loudspeaker condition, should therefore be compared before affective resonance is tested more decisively. This interpretation aligns with research on VR acoustic reproduction quality, VR soundscape applications, and ecological validity in virtual soundscape experiments.

This interpretation also limits the claim that ER can be solved by technical refinement alone. Calibrated binaural or ambisonic reproduction may improve acoustic external validity. However, cultural–emotional resonance may still require collective listening, prior familiarity, memory work, and social interaction that an individual short-duration VR session cannot reproduce.

4.3. Relation to Architectural Acoustics and Heritage Adaptation

The paper complements rather than replaces building-physics approaches to architectural acoustics. Material performance, noise control, and masking remain essential for determining the physical conditions under which a media layer can function [33,34]. However, heritage media architecture also requires a perceptual criterion: whether the audible layer fits the place, supports the intended narrative, and remains compatible with everyday sound activity.

This issue is especially relevant in preserved or historically sensitive settings, where sound is part of the heritage record and where intervention should not reduce the site to a mute visual shell [9,10]. Treating soundmarks as design material shifts the brief from isolated acoustic comfort toward a richer architectural problem involving memory, rhythm, use, and social recognition.

Methodologically, the present study is closer to quantitative environmental psychology and immersive design review than to sensory ethnography. It complements rather than substitutes for situated approaches: soundwalks can reveal embodied routes and attention; commented walks and interviews can elicit memory and emotion; and community sound archives or co-listening workshops can test whether proposed soundmarks are culturally recognized [13–19]. This distinction clarifies both the contribution and the boundary of the reported evidence.

4.4. Feng Shui as a Harmony Lens and Design Implications

Feng shui strengthens the interpretation as a harmony lens rather than an independent proof system. Prior empirical and quantitative studies connect feng shui with environmental fit, emotional response, and well-being [20,21,35]. At the same time, Indonesian applications have used it to discuss domestic comfort and indoor environmental judgment

in Surabaya [36,37]. In this study, the feng shui reading clarifies that the soundscape most strongly supported yi, or narrative legibility, while helping rebalance xing and shen without claiming that emotional resonance was statistically improved. Accordingly, xing, qi, shen, and yi are interpretive labels applied to observed questionnaire patterns, not independently validated constructs.

Design implications follow from this differentiated result. First, sound should be composed scene by scene and anchored to visible actions or places. Second, prototypes should test audio-visual congruence rather than visual fidelity alone. Third, immersive walkthroughs should be used with community stakeholders in subsequent phases to examine whether narrative cues remain intelligible, comfortable, and culturally precise before installation [39]. Table 6 translates these principles into a practical feng shui-soundscape checklist.

Table 6. Exploratory feng shui-soundscape design checklist for media-architecture virtual layers (synthesized from this study; proposed for subsequent community verification).

Feng Shui Principle	Media-Architecture Strategy (Visual/Spatial)	Soundscape Strategy	VR Evaluation Check
Form clarity (Xing)	Align projections/virtual objects to existing surfaces; maintain scale cues and legible edges.	Anchor sound sources to visible actions/locations (spatial congruence).	PQ items: check audio-visual alignment issues during walkthrough.
Qi continuity (flow)	Design smooth scene transitions along kampung paths and thresholds; avoid abrupt visual jumps.	Layer ambiance that evolves with movement; use masking to manage intrusive noise.	EA items; presence/immersion; note transition points that break atmosphere.
Shen cultural resonance	Use locally meaningful symbols and temporal cues (ritual vs. everyday scenes); avoid generic assets.	Foreground culturally specific soundmarks (e.g., temple/martial-arts activity) while keeping comfort.	ER items; qualitative feedback on cultural fit and emotional engagement.
Yi narrative legibility	Synchronize projection beats with heritage story nodes; highlight cause–and–effect across spaces.	Use sonic cues as narrative signposts (call-and-response, emphasis at story beats).	CQ items: verify that adding sound increases interpretation clarity.
Yin–yang balance	Compose alternation between calm (yin) and lively (yang) segments without overstimulation.	Balance pleasantness and eventfulness; avoid persistent high eventfulness or long monotony.	SDI (EA4/ER4); OQ2/SC3; cybersickness/comfort checks.

Table 6 should therefore be read as a prototype-governance matrix rather than as a generic checklist. Each row links a feng shui principle to a concrete media-architecture action, a corresponding soundscape strategy, and a VR verification step. This structure prevents symbolic intentions from remaining abstract: clarity, flow, resonance, legibility, and balance are each translated into a visible-spatial decision, an audible decision, and an evaluative criterion.

4.5. Comparative Contribution Within Recent Architecture Discourse

The principal comparative literature is now introduced before the research questions; this section therefore focuses on how the present findings extend the literature rather than serving as a stand-alone literature review.

Recent papers in *Architecture* confirm that immersive media is becoming a serious evaluative instrument rather than a novelty visualization tool. Saleh et al. show that adoption among emerging architectural professionals depends on both functional expectations and experiential presence [27]. This supports the present finding that VR improved

physical quality and overall evaluation because participants could inspect the proposal as an experienced spatial event rather than a static image.

Seghetto et al. similarly demonstrate that VR can support understanding of tactical urbanism by clarifying spatial perception and social use [28]. The Kapasan case extends this line of work from altered public-space scenarios to a heritage media-architecture layer, where the evaluative problem is whether a digital insertion remains legible, culturally compatible, and acoustically meaningful within an existing urban village.

Abdallah's work on anti-misophonia and anti-ADHD spaces further reinforces the claim that acoustic triggers and sensory stressors should be treated as primary spatial criteria [29]. Although the program differs, the methodological implication is parallel: acoustic conditions are not secondary corrections but variables that shape material choice, spatial behavior, and user experience. The Kapasan study translates this insight from stress reduction to meaning-bearing auditory design.

Anwar et al. demonstrate, through EEG-based research, that spatial configuration and environmental stimuli can measurably influence restoration and attention in educational settings [30]. The present study uses a different method—paired perception questionnaires—but it supports the same multisensory orientation: architectural judgment emerges from the calibration of geometry, image, sound, and bodily orientation rather than from visual form alone.

Taken together, these *Architecture* studies position the present article within an emerging discourse on immersive media, sensory design, and user-centered evaluation [27–30]. The additional contribution here is a case-based workflow in which soundscape is mobilized across the full design chain: site reading, soundmark selection, optical prototyping, immersive testing, and architectural translation.

Therefore, immersive media in this article is not only a tool for readiness assessment or urban scenario comprehension [27,28]. It is a pre-installation test bed for determining whether a combined visual–auditory layer can be read as culturally fitting architecture before physical construction. The significant gain in communication quality indicates that VR can answer this question more precisely than 2D visualization alone.

4.6. Exploratory Period 9 Heuristic (Optional and Non-Empirical)

The Period 9 material is retained only as an optional, hypothesis-generating design heuristic. It was not included in the questionnaire, no participant response was linked to direction, and no inferential test evaluated the proposed sectors; consequently, it must not be read as an analytical finding.

Recent work by Samodra and colleagues strengthens the continuity of the argument. Pepper's Ghost-based virtual layering has been shown to reinforce place meaning in historical urban-village settings [40]. At the same time, studies on mosque acoustics, sensory sound accentuation, and dynamic soundscape reconstruction demonstrate that acoustic shaping can structure heritage legibility, adaptive occupancy, and temporally responsive spatial experience [41–43]. These studies frame sound as a controllable architectural medium that changes how space is read, remembered, and used.

The Period 9 reading is included as a supplementary design-calibration layer, not as prescriptive geomancy. The 2025 Fire Snake chart in Rahardjo's Period 9 material locates star 2 at the center, star 9 in the east, star 1 in the southeast, star 6 in the south, star 8 in the southwest, star 4 in the west, star 3 in the northwest, star 7 in the north, and star 5 in the northeast [44]. Read cautiously: this distribution suggests that the east–southeast arc can carry clearer narrative and social cues, while the northeast–center axis should be treated with sonic and visual restraint.

For the Kapasan proposal, the implication is selective calibration rather than literal directional prescription. Sonic emphasis, interactive triggers, and scene transitions can be intensified in communicative sectors, while dense overlays, persistent low-frequency buildup, or cognitively heavy content can be moderated in restraint sectors. This reading is consistent with the empirical pattern: communication quality improved more strongly than emotional resonance, indicating that harmony depends on calibrated correspondence among directionality, narrative sequence, and architectural fit rather than on sensory intensity alone. These allocations are proposed scenarios for a future controlled comparison, not recommendations derived from the present data.

Figure 14 translates this logic into three cue zones: an east–southeast communicative arc, a north–northeast restraint arc, and a stabilizing center. The first zone is suited to clearer story beats and lighter sonic emphasis; the second calls for lower visual density and softer ambiance; and the center holds the sequence together rather than competing for attention. The diagram is therefore a sector-based architectural reading for cue intensity, visual emphasis, and ambient sound density. The labels describe untested cue-allocation hypotheses.

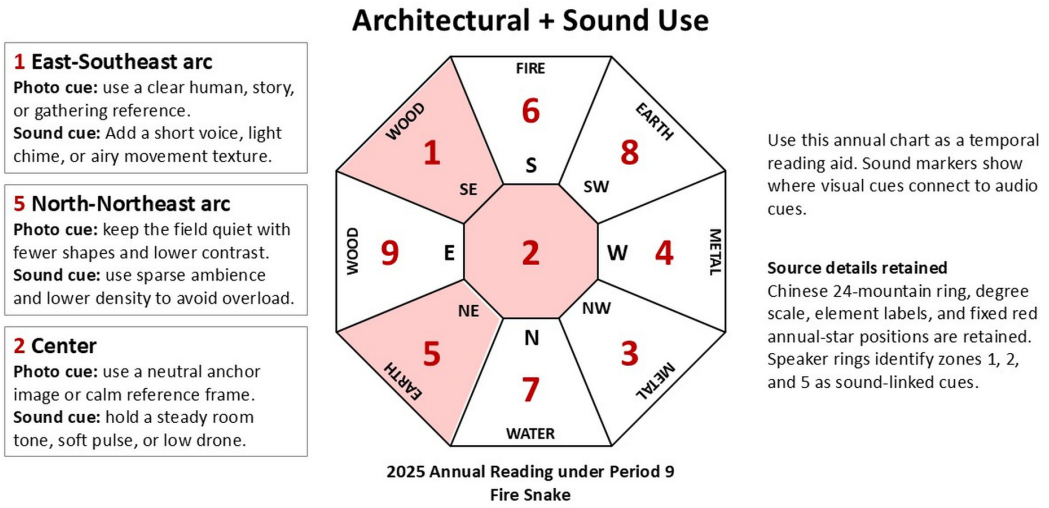


Figure 14. Optional and untested Period 9 cue-allocation heuristic, adapted from the 2025 directional chart and redrawn as an east–southeast communicative arc, a north–northeast restraint arc, and a stabilizing center. The diagram formulates a future design hypothesis; it does not represent an empirical result of this study.

Figure 14 converts the annual 2025 Period 9 chart into a simplified three-zone diagram for architectural and sound cue calibration. The redraw is used as an interpretive design aid for cue placement, sequencing, and balance, rather than as a literal prescription.

Read in conjunction with Table 6, Figure 14 shows how the annual chart can serve as a selective calibration layer. It helps the designer decide where to intensify sonic signposts, where to simplify the image field, and where to maintain a neutral core so that the virtual layer does not become overdetermined. It also clarifies a key design lesson from the results: adding more cultural content does not automatically produce stronger resonance. Communication improved when sound and image were coordinated into an intelligible sequence. The figure should therefore be used to formulate a later experiment, not to explain the observed VR effects.

4.7. Limitations and Future Work

This study has six limitations. First, the paired sample (n = 24) offers useful within-subject control but remains modest for subgroup analysis. Second, most respondents were architecture students rather than residents, martial-arts practitioners, temple users,

or heritage stakeholders; therefore, the findings represent early-stage expert-user perception rather than community validation. Third, the 360° panoramas were visual-only and were accompanied by scene-assigned backing tracks launched manually and reproduced through a separate loudspeaker while participants viewed the panorama in the HMD. This dual-device condition was not frame-accurate, embedded, head-tracked, binaural, or ambisonic, and no calibrated playback level or loudspeaker geometry was documented. Fourth, the evaluation was pre-installation and short-term, so it cannot yet establish long-term acceptance, maintenance effects, or everyday neighborhood impact. Fifth, although representative visual captures and the relative audio timeline are embedded in Supplementary Material S1, absolute onset and offset timecodes, track-level edits, gain settings, and measured loudspeaker positions were not logged; replication is therefore procedural rather than signal-exact. Sixth, keyword counts cannot disclose the lived memories, emotions, or contested cultural meanings that situated qualitative methods would reveal.

Future work should therefore extend the method in three directions: participatory evaluation with residents and cultural actors; comparison of calibrated binaural/ambisonic or controlled multichannel playback with the present separate-loudspeaker condition; and post-installation assessment after a physical or hybrid prototype is built. The Period 9 reading should also be tested as an independent cue-allocation variable by comparing purely visual, sound-led, and feng shui-tuned versions across different heritage morphologies and spatial-audio configurations. Future community phases should combine co-designed soundwalks, commented listening, interviews, and an openly documented audio archive, subject to participant privacy and source licensing. Future VR iterations should also use scene-embedded, time-coded playback with documented level, loudspeaker or headphone configuration, and stable source versions. The Period 9 heuristic should be tested against a non-directional baseline before any design guidance is inferred.

5. Conclusions

This preliminary design-user evaluation indicates that the soundscape can be designed as an architectural material for heritage media architecture, rather than merely added as decorative audio. Through Sanggar Kungfu Kapasan, the paper links soundmark mapping, Pepper's Ghost prototyping, immersive VR evaluation, and feng shui-based interpretation into a single design-research workflow for a historic kampung.

Empirically, immersive VR outperformed 2D visualization in perceived physical quality, communication quality, and the overall evaluation. The strongest effects were not found in general atmosphere or emotional resonance, but in the dimensions that make a virtual layer legible as architecture: form clarity, scale plausibility, and narrative communication. Positive sound-specific ratings and the correlation between SDI and communication quality indicate that sound chiefly strengthened place-meaning communication. EA did not improve significantly, and ER decreased non-significantly; thus, the evidence does not demonstrate enhanced cultural–emotional depth.

The main contribution to architectural acoustics is therefore methodological and conceptual. Methodologically, the study demonstrates a pre-installation evaluation workflow that combines multisensory design prototyping, paired statistical testing, and qualitative evidence. Conceptually, it reframes soundscape as a layer that can coordinate interface design, scene sequencing, heritage interpretation, and subsequent community-stakeholder evaluation.

For Kapasan and similar heritage kampungs, the design implication is clear: media architecture should be judged as a calibrated multisensory insertion whose success depends on proportional balance among visual legibility, sonic cueing, cultural specificity, and contextual fit. Immersive VR is especially valuable when a proposal contains narrative

and atmospheric qualities that cannot be reliably inferred from 2D images alone. The feng shui and Period 9 components remain optional interpretive hypotheses until tested independently and with community participation.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/architecture6030146/s1: Figure S1: Representative visual captures aligned with the operator-controlled audio-presentation timeline; Table S1: Direct documentation of the evaluated audio composition, source materials, relative timing, spatialization, and replication boundary.

Author Contributions: Conceptualization, F.X.T.B.S.; methodology, F.X.T.B.S., S.N.N.E., A.T.D., J.L. and J.N.; formal analysis, F.X.T.B.S.; investigation, F.X.T.B.S. and J.L.; resources, F.X.T.B.S.; data curation, F.X.T.B.S.; writing—original draft, F.X.T.B.S.; writing—review and editing, S.N.N.E., A.T.D., J.L. and J.N.; visualization, F.X.T.B.S. and J.L.; supervision, S.N.N.E., A.T.D. and J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of "Penelitian Keilmuan, Institut Teknologi Sepuluh Nopember No. 1696/PKS/ITS/2025". The authors gratefully acknowledge this financial and technical support.

Institutional Review Board Statement: This study was approved by the Research Ethics Committee of Institut Teknologi Sepuluh Nopember (ITS). The study involved voluntary, non-invasive perception-based questionnaires and VR viewing. No biomedical intervention, deception, physiological monitoring, or collection of sensitive personal data was involved. Informed consent and participant privacy were maintained throughout.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The datasets and analytical worksheet supporting the presented results are available from the corresponding author upon reasonable request. Respondent privacy limits public dissemination. AI-assisted tools were used only to support figure visualization and language refinement; the authors conducted the original data collection, analysis, interpretation, and manuscript review. Supplementary Material S1 is self-contained and embeds representative visual captures from the project materials used during the VR session together with the relative audio-presentation timeline. The authors retain the full animation, panorama, and scene-to-audio mapping files and may be consulted through the corresponding author upon reasonable request.

Acknowledgments: We thank all volunteers who completed the questionnaires and immersive VR evaluation, as well as the local information providers who supported contextual documentation of the study site.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

2D	Two-dimensional visualization
ADHD	Attention-deficit/hyperactivity disorder
CAVE	Cave automatic virtual environment
CC BY	Creative Commons Attribution license
CI	Confidence interval
CQ	Communication quality
df	Degrees of freedom
EA	Experiential atmosphere
EA4	Soundscape-supports-place-meaning item
EEG	Electroencephalography
ER	Emotional resonance/responsiveness

ER4	Soundscape-helps-form-scenes item
HMD	Head-mounted display
ITS	Institut Teknologi Sepuluh Nopember
M	Mean
OQ	Output quality
OQ2	Separate audio-playback-quality item during VR
PQ	Physical quality
PQ1–PQ3	Physical-quality item set
SC	System characteristic
SC3	Ability-to-hear-accompanying-soundscape item during VR
SD	Standard deviation
SDI	Soundscape-design index
VR	Virtual reality
α	Cronbach's alpha
Δ	Paired mean difference (VR − 2D)
d_z	Within-subject effect size
D_m	Profile-dispersion index
G_i	Form–Shen gap
K_w	Keyword count for keyword w
n	Sample size
p	p-value
r	Pearson correlation coefficient
t	t statistic

References

1. Tutuko, P. Community attachment pada transformasi desain bangunan permukiman di sekitar kawasan pecinan. *Local Wisdom J. Ilm. Kaji. Kearifan Lokal* **2010**, *2*, 10–19.
2. Chen, K.; Guaralda, M.; Kerr, J.; Turkay, S. Digital intervention in the city: A conceptual framework for digital placemaking. *Urban Des. Int.* **2024**, *29*, 26–38. [CrossRef]
3. Canelas, P.; Hoehnk, J. Practices of digital placemaking: Accelerating place attachment. *Urban Res. Pract.* **2025**, *18*, 378–396. [CrossRef]
4. Redyantanu, B.P.; Otto, F. Architecture as Programmatic Archive: Sustainable Conservation on Tunjungan Street History. *DIMENSI (J. Archit. Built Environ.)* **2025**, *52*, 83–94. [CrossRef]
5. *ISO 12913-1:2014*; Acoustics—Soundscape—Part 1: Definition and Conceptual Framework 2014. International Organization for Standardization: Geneva, Switzerland, 2014.
6. Axelsson, Ö.; Nilsson, M.E.; Berglund, B. A principal components model of soundscape perception. *J. Acoust. Soc. Am.* **2010**, *128*, 2836–2846. [CrossRef] [PubMed]
7. Bruce, N.S.; Davies, W.J. The effects of expectation on the perception of soundscapes. *Appl. Acoust.* **2014**, *85*, 1–11. [CrossRef]
8. Zhang, R.; Ma, H.; Wang, C.; Zhang, Y.; Kang, J. Soundscape and its context: A framework based on a systematic review. *J. Acoust. Soc. Am.* **2025**, *157*, 4417–4436. [CrossRef] [PubMed]
9. Álvarez-Morales, L.; Díaz-Andreu, M. Acoustics, soundscapes and sounds as intangible heritage. *Acoustics* **2024**, *6*, 408–412. [CrossRef]
10. Zhang, D.; Shan, Y.; Chen, X.; Wang, Z. Soundscape in religious historical buildings: A review. *Herit. Sci.* **2024**, *12*, 45. [CrossRef]
11. Pietroni, E. Mapping the soundscape in communicative forms for cultural heritage: Between realism and symbolism. *Heritage* **2021**, *4*, 4495–4523. [CrossRef]
12. Schafer, R.M. *The Soundscape: Our Sonic Environment and the Tuning of the World*; Destiny Books: Rochester, VT, USA, 1994.
13. Palmese, C.; Carles Arribas, J.L.; Rodríguez Antolín, A. The Soundscape and Listening as an Approach to Sensuous Urbanism: The Case of Puerta del Sol (Madrid). *Urban Plan.* **2024**, *9*, 7206. [CrossRef]
14. Thibaud, J.-P. The Sensory Fabric of Urban Ambiances. *Senses Soc.* **2011**, *6*, 203–215. [CrossRef]
15. Thibaud, J.-P. A Sonic Paradigm of Urban Ambiances. *J. Sonic Stud.* **2011**, *1*, 1–14. [CrossRef]
16. Griffero, T. The Atmospheric "Skin" of the City. *Ambiances* **2013**, 1–16. [CrossRef]
17. Ingold, T. *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*; Routledge: London, UK, 2000.
18. Ingold, T. Against Soundscape. In *Autumn Leaves: Sound and the Environment in Artistic Practice*; Carlyle, A., Ed.; Double Entendre: Paris, France, 2007; pp. 10–13.

19. Augoyard, J.-F.; Torgue, H. (Eds.) *Sonic Experience: A Guide to Everyday Sounds*; McGill-Queen's University Press: Montreal, QC, Canada, 2006.
20. Mak, M.Y.; Ng, S.T. The art and science of Feng Shui—A study on architects' perception. *Build. Environ.* **2005**, *40*, 427–434. [CrossRef]
21. Han, Y.; Lin, S. Empirical and quantitative studies of Feng Shui: A systematic review. *Heliyon* **2023**, *9*, e19532. [CrossRef] [PubMed]
22. Heydarian, A. Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations. *Autom. Constr.* **2015**, *54*, 116–126. [CrossRef]
23. Ma, J. A comparative analysis of the effectiveness of immersive virtual reality as a method for end-user design reviews compared to 2D and physical prototypes. *Build. Environ.* **2025**, *267*, 112237. [CrossRef]
24. Hong, J.Y.; Lam, B.; Ong, Z.-T.; Ooi, K.; Gan, W.-S.; Kang, J.; Feng, J.; Tan, S.-T. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Build. Environ.* **2019**, *149*, 1–14. [CrossRef]
25. Tokgöz, Ö.G.; Engel, M.S.; Othmani, C.; Altinsoy, M.E. Virtual reality application in evaluating the soundscape in urban environment: A systematic review. *Acoustics* **2025**, *7*, 68. [CrossRef]
26. Lu, Y.; Lau, S.-K. Examining ecological validity of virtual reality experiments in soundscape research: From soundwalk to virtual reality experiments. *Comput. Hum. Behav.* **2025**, *162*, 108462. [CrossRef]
27. Saleh, M.S.; Alalouch, C.; Al-Saadi, S. From Skepticism to Adoption: Assessing Virtual Reality Readiness Among Emerging Architectural Professionals in a Developing Economy. *Architecture* **2025**, *5*, 86. [CrossRef]
28. Seghetto, I.; Lopes, R.; Lima, F.; Boffi, M.; Rainisio, N.; Stancato, G.; Piga, B.E.A. Virtual Reality as a Tool for Enhancing Understanding of Tactical Urbanism. *Architecture* **2025**, *5*, 26. [CrossRef]
29. Abdallah, Y.K. Designing for Special Neurological Conditions: Architecture Design Criteria for Anti-Misophonia and Anti-ADHD Spaces for Enhanced User Experience. *Architecture* **2025**, *5*, 85. [CrossRef]
30. Anwar, R.; Elkhateeb, S.; Afifi, S.; Bayoumi, K. Brainwave Dynamics: Neurophysiological Responses to Enclosed Courtyards for Mental Wellbeing in Educational Contexts. *Architecture* **2025**, *5*, 76. [CrossRef]
31. Lacey, J. Sonic placemaking: Three approaches and ten attributes for the creation of enduring urban sound art installations. *Organised Sound* **2016**, *21*, 147–159. [CrossRef]
32. Aletta, F.; Kang, J. Towards an urban vibrancy model: A soundscape approach. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1712. [CrossRef] [PubMed]
33. Samodra, F.X.T.B.; Irvansyah, I. Architectural acoustics experiment for tropical and eco-house prototype design. *AIP Conf. Proc.* **2020**, *2255*, 070012. [CrossRef]
34. Soegiono, P.D.; Elsiana, F.; Arifin, L.S. The Sound Masking Assessment of Recycled Partition. *DIMENSI (J. Archit. Built Environ.)* **2024**, *51*, 61–72. [CrossRef]
35. Jin, J.; Juan, Y. Is feng shui a science? An empirical study on the built environment and occupants' emotional responses. *Build. Environ.* **2021**, *204*, 107992. [CrossRef]
36. Linda, O.; Tanuwidjaja, G. Feng Shui in modern house design searching for the rationale and possible impacts assessment. *DIMENSI (J. Archit. Built Environ.)* **2014**, *41*, 43–50. [CrossRef]
37. Hogianto, J.; Mediastika, C.E. Relevancy of indoor temperature and Feng Shui Luo Pan compass: The case of domestic buildings in Surabaya, Indonesia. *DIMENSI (J. Archit. Built Environ.)* **2019**, *46*, 35–42. [CrossRef]
38. Bregman, A.S. *Auditory Scene Analysis: The Perceptual Organization of Sound*; MIT Press: Cambridge, MA, USA, 1990.
39. Lo, M.; Wünsche, B.C.; Manfredini, M.; Hooper, S. (Semi-)immersive digital placemaking: A systematic literature review on augmented reality, virtual reality, and mixed reality applications in placemaking. *Converg. Int. J. Res. Into New Media Technol.* **2025**, *31*, 2025–2045. [CrossRef]
40. Lovendianto, J.; Samodra, F.X.T.B.; Ekasiwi, S.N.N. Virtual Layer: Pepper Ghost Illusion and Digital Projection to Reinforce Place Meaning. *Sinektika J. Arsit.* **2026**, *23*, 96–107. [CrossRef]
41. Setiyowati, E.; Ekasiwi, S.N.N.; Samodra, F.X.T.B. Acoustic Performance Comparison Between the Historic Traditional and Contemporary Forms of Tropical Mosques. *E3S Web Conf.* **2025**, *674*, 05008. [CrossRef]
42. Samodra, F.X.T.B.; Tyas, K.N.; Nam, J. Bus Sound Accentuation through Sensory Technology in Adaptive Transit Architecture. *E3S Web Conf.* **2025**, *674*, 05007. [CrossRef]
43. Samodra, F.X.T.B.; Tyas, K.N. Temporary Transit Architecture: Reconstruction of Dynamic Soundscape through Moment-Based Adaptive Spaces. *E3S Web Conf.* **2026**, *685*, 02004. [CrossRef]
44. Rahardjo, M.P. *Feng Shui Periode 9: Mengubah Energi, Meningkatkan Peruntungan*; Bhuana Ilmu Populer: Jakarta, Indonesia, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.