sustainability

MDPI

Article

Museum-Grounded Generative-AI Guide Affordances and Heritage Stewardship Intention: Evidence from Jinghui

Qi Wu [1], Qinchuan Zhan [1], Yang Zhao [2,*] and Mingxi Shi [3]

[1] School of Design and Art, Shaanxi University of Science & Technology, Xi'an 710021, China; zhanqc@sust.edu.cn (Q.Z.)
[2] School of Art and Design, Zhengzhou University of Light Industry, Zhengzhou 450003, China
[3] College of Materials Science and Art Design, Inner Mongolia Agricultural University, Hohhot 010018, China
* Correspondence: 2025076@zzuli.edu.cn; Tel.: +86-15829471020

Abstract

Museum-grounded generative-AI guides can extend heritage interpretation from information delivery to a participatory encounter that connects what visitors understand, experience, and intend to protect. Existing research, however, rarely distinguishes whether system affordances are associated with desirable outcomes, predictively important, or necessary for those outcomes to reach high levels. Drawing on the stimulus–organism–response framework, this study develops an integrated model of how dialogic, epistemic, social, cultural, adaptive, and participatory affordances relate to perceived heritage-value enhancement, flow, stewardship self-efficacy, and pro-heritage stewardship intention. Post-interaction evidence from adults reporting prior Jinghui use was examined through partial least squares structural equation modeling, artificial neural networks, a matched Basis-KAN predictive benchmark, and necessary-condition analysis. The findings identify outcome-specific profiles: responsive dialog is prominent for perceived value enhancement; visitor agency has the highest mean ANN importance for flow; epistemic transparency and cultural authenticity are closely ranked for self-efficacy; and perceived value enhancement and flow retain necessity evidence for stewardship intention. The direct value-enhancement–self-efficacy path is unsupported. Close coefficient differences do not establish superior design features, and ANN ranks vary across fitted runs. Comparison of nonlinear importance with held-out prediction further shows that a flexible representation does not automatically yield stronger predictive performance. By linking accountable system design to visitors' meaning-making and stewardship motivation, this study provides an evidence-informed foundation for designing museum generative AI as public interpretive infrastructure that connects cultural accountability with informed participation in heritage stewardship.

Keywords: museum-grounded generative AI; cultural heritage; stewardship intention; PLS-SEM; artificial neural network; Basis-KAN; necessary condition analysis

Check for updates

Academic Editor: Mark A. Bonn

Received: 20 September 2026
Revised: 22 October 2026
Accepted: 23 October 2026
Published: 29 October 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

Museums are increasingly using generative artificial intelligence in public-facing interpretation. Compared with fixed-content guides, a museum-grounded generative-AI virtual guide can support question-driven dialog, adapt explanations to visitors' interests and prior knowledge, and present heritage content through an embodied interface. These capabilities may widen access to museum knowledge and enrich visitor experiences, but they also increase the museum's responsibility for the accuracy, provenance, and cultural framing of AI-mediated

interpretations. UNESCO's Recommendation on the Ethics of Artificial Intelligence emphasizes transparency, human oversight, cultural diversity, and the prevention of harm in the development and use of AI systems [1]. Research on AI ethics in cultural heritage similarly highlights the risks of inaccurate interpretation, cultural misrepresentation, marginalization of community perspectives, and unclear responsibility for AI-generated content [2].

Research on AI and cultural heritage extends well beyond public-facing guides. Reviews map AI applications across documentation, condition assessment, conservation planning, and restoration support [3–6]. Reviews of generative museum technology and AI ethics separately address digital interpretation, data quality, and cultural accountability [2,7]. Colace et al. survey machine-learning applications across cultural-heritage protection and identify challenges in connecting technological development with conservation needs [3]. Li, Lim, and Li map technological developments in heritage conservation, while Kutlu identifies damage detection, structural monitoring, and risk assessment as prominent themes in AI-assisted historical-building research [4,5]. Tan et al. organize architectural conservation into five connected phases—monitoring, damage identification, environmental prediction, deterioration risk assessment, and environmental control—showing how AI supports a conservation workflow [6]. Generative-AI reconstruction research further shows that credible heritage applications require evidence grading, explicit documentation of reconstruction hypotheses, and disclosure of uncertainty rather than visually plausible output alone [8]. Within this wider field, research on generative AI and virtual guides has focused on perceived value and adoption [9,10], virtual-human design and embodiment [11–13], continuance intention [14], offline cultural participation [15], and AI-supported co-creation [16]. The present study addresses a narrower but underdeveloped problem: how visitors perceive a deployed, museum-grounded guide and how those perceptions are associated with heritage stewardship intention. This focus complements conservation and reconstruction research by examining the visitor-facing interpretive and motivational consequences of AI-mediated heritage encounters.

Heritage stewardship represents a different outcome from technology adoption, satisfaction, or continuance intention. These platform-centered outcomes indicate whether visitors accept a system, value its use, or intend to use it again. Heritage stewardship is instead concerned with responsible relationships among people, heritage, places, and communities. It may involve responsible visitation, accurate communication of heritage knowledge, support for conservation and public education, and appropriate responses to heritage-related risks. Conservation-participation research provides context for this broader orientation [17]; the specific activities listed here define the present study's stewardship scope. In this study, pro-heritage stewardship intention (PHSI) refers to visitors' stated willingness to undertake or support such actions. It is treated as a near-term motivational outcome rather than evidence of observed behavior. This outcome is relevant to the cultural sustainability of museums because the continuity and transmission of heritage depend not only on institutional preservation but also on informed and sustained public engagement [18]. The central question is therefore not simply whether a virtual guide is useful or enjoyable, but whether its interpretive and interactive affordances are associated with visitors' perceived heritage-value enhancement, experiential involvement, and capability to act.

Two research gaps motivate the present study. First, many existing studies operationalize virtual-guide qualities through broad constructs such as information quality, interactivity, anthropomorphism, or system quality. Such constructs may not distinguish between different forms of interaction and interpretation. For example, a response that reflects the preceding dialog is not equivalent to a response that is merely fast, while an explanation that discloses its evidential basis is not equivalent to one that is simply fluent or persuasive. An affordance-based perspective is useful here because it treats affordances as

possibilities for action that arise from the relationship between technological arrangements and users, rather than as isolated technical features [19]. This distinction is particularly important in museums, where cultural interpretation is closely related to institutional credibility, historical accountability, and the traceability of knowledge [1,2,20]. The present study, therefore, treats dialogic contingency, source-grounded epistemic transparency, embodied social presence, perceived cultural authenticity, adaptive narrative personalization, and participatory agency as distinct perceived affordances.

Second, existing research on museum AI has commonly examined adoption, satisfaction, continuance, or general cultural participation. The psychological appraisals associated with heritage stewardship intention remain less developed. The present model brings together perceived heritage-value enhancement, flow experience, and heritage stewardship self-efficacy. PHV captures visitors' retrospective perception that the experience deepened their recognition of heritage significance. It does not measure their absolute level of heritage-value recognition or an objectively observed change. FLOW concerns focused attention and absorption, while HSSE concerns confidence in feasible, responsible heritage-supporting action. Examining these appraisals together distinguishes their relationships with a heritage-centered motivational outcome.

To address these gaps, this study develops an SOR model with 24 theoretically specified relationships and evaluates it through four complementary analytical lenses. PLS-SEM estimates conditional associations. Artificial neural networks describe fitted nonlinear input importance. An additive Basis-KAN model is compared with ridge regression and boosted stumps under identical repeated folds to test out-of-sample predictive performance. Necessary-condition analysis identifies conditions that constrain high outcome levels. Keeping these roles separate prevents path coefficients, ANN importance, predictive accuracy, and necessity effects from being treated as interchangeable evidence. The analyses share one questionnaire dataset and are not independent replications.

The empirical setting is Jinghui, a generative-AI virtual guide deployed at the Capital Museum in Beijing. The study draws on 519 adults who affirmed prior Jinghui experience at the questionnaire entry and completed a post-interaction survey. Eligibility therefore rests on self-reported use rather than session-level verification. The single-museum and single-guide setting bounds the interpretation to these respondents and this deployment context.

The study addresses five research questions:

RQ1. How are Jinghui's six perceived affordances associated with perceived heritage-value enhancement, flow experience, and heritage stewardship self-efficacy?

RQ2. How are perceived heritage-value enhancement, flow experience, and heritage stewardship self-efficacy associated with visitors' pro-heritage stewardship intention?

RQ3. Which affordances and organism states have the greatest nonlinear predictive importance in ten-run artificial neural-network models?

RQ4. Does an additive Basis-KAN model improve out-of-sample prediction relative to ridge regression and boosted decision stumps under identical repeated folds?

RQ5. Which perceived affordances and organism states exhibit empirical necessity patterns or bottleneck relationships for higher endogenous outcomes?

Rather than treating museum AI as a single measure of system quality, this study distinguishes six affordances that visitors can experience during an interaction. It gives particular attention to source-grounded epistemic transparency, which concerns visitors' perception of the evidential basis of AI explanation. The analysis also shifts the outcome from technology acceptance to pro-heritage stewardship intention and separates the roles of heritage value, flow, and stewardship self-efficacy. Evidence from 519 respondents reporting use of a deployed guide is examined through structural modeling, out-of-sample

prediction, and necessary-condition analysis, allowing association, predictive performance, and bottleneck patterns to be interpreted as different forms of evidence.

2. Literature Review and Hypothesis Development

2.1. Museum-Grounded Generative-AI Virtual Guides

A museum-grounded generative-AI virtual guide is defined in this study as an embodied conversational system that dynamically generates context-sensitive interpretive responses while anchoring those responses in a museum-curated knowledge base. It differs from three adjacent forms of digital interpretation: scripted digital characters that deliver predetermined content, general-purpose conversational agents that operate without institution-specific heritage grounding, and embodied avatars that present fixed narration without generative dialog. Three properties therefore characterize the focal system. Generativity refers to the production of context-dependent responses during interaction; grounding refers to the use of institutionally curated heritage knowledge to constrain and support those responses; and embodiment refers to the socially perceptible voice, appearance, and behavioral cues through which the guide communicates.

Grounding should, however, be distinguished from epistemic transparency. A system may technically retrieve information from an authoritative museum knowledge base without making the evidential status of its responses perceptible to visitors. Conversely, visitors' judgments of whether an explanation is well supported depend on how clearly the system communicates the basis and limits of its heritage claims. This distinction is especially important in generative-AI interpretation, where fluent responses may otherwise obscure differences among established knowledge, interpretive inference, and uncertain information. Research comparing large language models with structured heritage knowledge likewise suggests that flexible natural-language interaction and domain-specific knowledge structures are complementary rather than substitutable capabilities [20].

Existing museum research has examined several aspects of this emerging interaction form, but largely in isolation. Studies of generative AI in digital museums show that semantic relevance and contextual adaptability can enhance perceived value and adoption, while perceived risk remains consequential [9]. Research on generative-AI virtual assistants in immersive museum environments demonstrates the potential of conversational guidance for visitor engagement while also identifying technical limitations such as latency and voice clarity [12]. Studies of museum digital humans further connect information richness, esthetic experience, flow, cultural identity, and media richness with adoption or continuance intention [10,14]. Experimental evidence comparing interface modalities also suggests that informational richness and embodiment do not produce identical psychological outcomes: voice-plus-text interfaces may enhance perceived information quality, whereas embodied avatars can generate stronger engagement [13]. Taken together, these findings indicate that museum generative AI cannot be adequately represented by a single continuum of technological sophistication, system quality, or human likeness.

Accordingly, this study conceptualizes the museum-grounded generative-AI virtual guide through six visitor-perceived affordance dimensions. Here, affordances are understood relationally as possibilities for interaction and meaning-making that emerge between technological capabilities, visitor perceptions, and the situated museum context, rather than as technical features alone [19]. This distinction is important because affordance research has cautioned against treating system features, user outcomes, and affordances as interchangeable concepts. Dialogic contingency captures whether responses remain semantically connected to the evolving conversation. Source-grounded epistemic transparency concerns whether visitors can understand the evidential basis and epistemic status of heritage explanations. Embodied social presence represents the experienced possibility of socially situated interaction with the

guide. Perceived cultural authenticity reflects whether the interpretation supports culturally coherent and context-sensitive sense-making. Adaptive narrative personalization captures the perceived adjustment of explanatory depth, emphasis, examples, and information pacing to the individual visitor. Participatory agency concerns the visitor's perceived ability to influence the direction and depth of heritage exploration.

These six dimensions are conceptually distinct but potentially complementary. Importantly, they are operationalized as visitor-perceived affordances rather than backend technical specifications. Institutional knowledge grounding, for example, is a system property, whereas source-grounded epistemic transparency concerns whether the evidential basis of the system's claims becomes intelligible to the visitor. Similarly, adaptive algorithms are technical mechanisms, whereas adaptive narrative personalization concerns whether visitors experience the interpretation as responsive to their knowledge, interests, and interaction patterns. This distinction enables the study to examine how different perceived qualities of a museum-grounded generative-AI guide are associated with different psychological processes, rather than attributing visitor outcomes to "AI" as an undifferentiated technological stimulus.

Jinghui is the Capital Museum's AI digital guide within its high-precision smart-guidance platform. The developer reports that the platform entered trial operation in February 2024 and connects 18 on-site touchscreens with smartphone access. Jinghui operates in the intelligent question-and-answer module, combining a customized AI large model with a museum-specific knowledge base to answer natural-language questions about collections, exhibitions, and visitor services [21]. A first-hand developer interview describes the knowledge resources as collection data, exhibition materials, research findings, museum and heritage knowledge, and public-service information, and identifies natural-language processing and personalized interaction as system capabilities [22]. The surrounding platform also supports visit planning, exhibition and collection exploration, resource discovery, and visitor services [21]. Figure 1 organizes these publicly documented components by access channel, knowledge resource, interaction capability, and service scope.

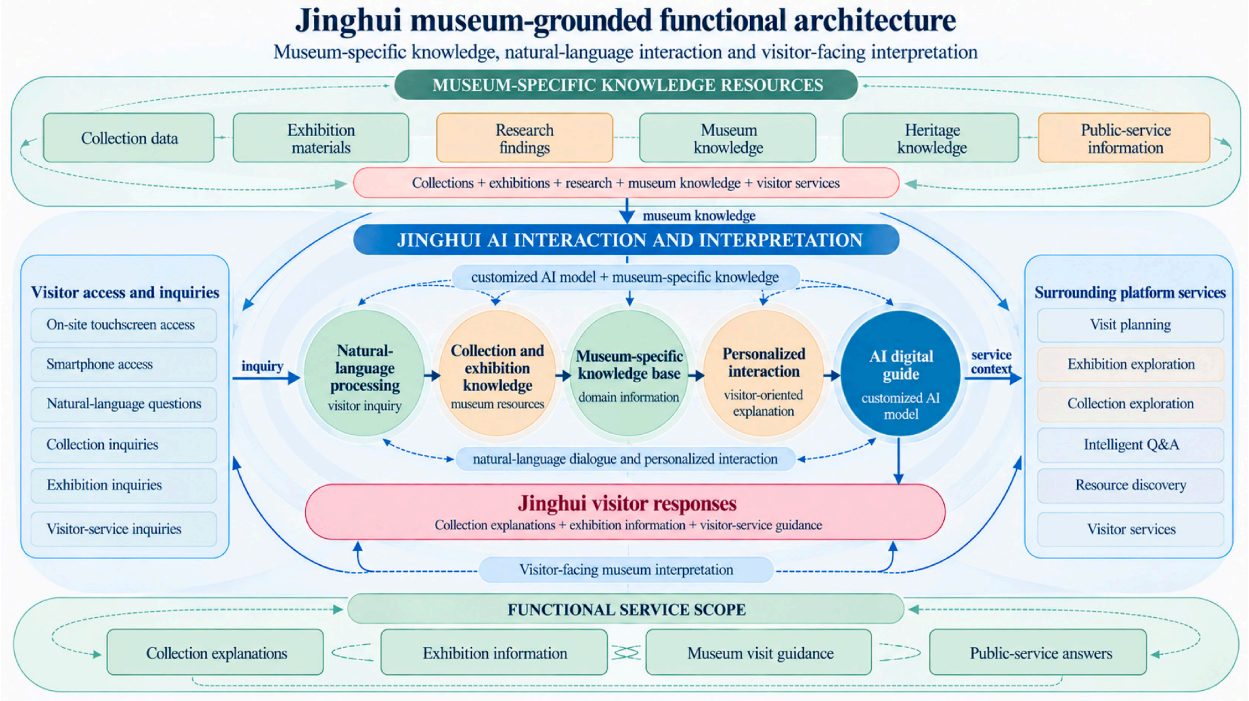


Figure 1. Jinghui's publicly documented functional architecture within the Capital Museum smart-guidance platform. The central field represents the AI question-and-answer guide; the right-hand column identifies services of the surrounding platform, and the lower band summarizes the visitor-facing service scope. Arrows organize functional relationships [21,22].

The public sources document knowledge grounding and functional access, not the presence of visitor-visible citations or uncertainty notices in each survey encounter. SGET therefore measures respondents' perception of evidential intelligibility; neither backend grounding nor an affirmative prior-use response independently verifies exposure to such interface elements. The figure is a functional synthesis, not a reconstruction of session screens or undisclosed retrieval parameters.

2.2. Theoretical Architecture

The stimulus–organism–response (SOR) framework provides the overarching architecture for this study. SOR proposes that environmental stimuli influence internal cognitive and affective states, which subsequently shape approach- or avoidance-oriented responses [23]. In the present context, the six visitor-perceived affordance dimensions of the generative-AI virtual guide constitute the stimuli; perceived heritage-value enhancement (PHV), flow experience (FLOW), and heritage stewardship self-efficacy (HSSE) constitute the organism states; and pro-heritage stewardship intention (PHSI) constitutes the response. This structure shifts evaluation away from whether visitors merely accept or continue using the technology toward how technology-mediated interpretation is associated with heritage-centered motivational outcomes.

SOR specifies the general sequence from environmental perception to internal psychological state and behavioral response, but it does not require different stimuli to affect organism states in identical ways or with equivalent strength. The present study, therefore, adopts a differentiated multi-mechanism structure. Each perceived affordance has a psychologically proximal mechanism through which its influence is expected to be particularly salient, while cross-mechanism effects are also theoretically plausible because cognitive evaluation, experiential absorption, and perceived capability are interrelated during museum interpretation. In this sense, the model does not treat the six affordances as interchangeable indicators of a single "AI quality" construct. Rather, it examines whether each affordance contributes differently to PHV, FLOW, and HSSE while allowing these psychological effects to overlap.

Three complementary mechanisms provide the organizing logic for these relationships. First, an epistemic–cultural mechanism is expected to be particularly salient for source-grounded epistemic transparency and perceived cultural authenticity. Heritage interpretation depends not only on whether information is fluent or engaging but also on whether visitors regard it as evidentially credible and culturally coherent. Perceived authenticity has long been treated as a visitor-level evaluative construct in cultural heritage research, with implications for engagement and behavioral responses rather than merely as an objective property of a site or artifact [24]. Empirical work in heritage tourism likewise shows that authenticity and engagement are closely connected in visitors' evaluation of heritage experiences. SGET and PCA are therefore expected to exert their strongest conceptual influence on PHV, while their effects may also extend to FLOW and HSSE by reducing interpretive uncertainty, sustaining coherent engagement, and providing a more reliable basis for heritage-related judgment.

Second, a dialogic–experiential mechanism is expected to be particularly salient for dialogic contingency, embodied social presence, and adaptive narrative personalization. These affordances concern the continuity, social quality, and personal fit of interaction. Self-determination theory suggests that competence support, personal relevance, and appropriately structured interaction can sustain intrinsic involvement [25]. In a museum setting, coherent dialog, socially legible companionship, and personally adapted interpretation can therefore support concentration and absorption. At the same time, these affordances may also affect PHV and HSSE because responsive

and personalized interpretation can deepen meaning construction, clarify culturally significant information, and provide visitors with more accessible pathways for understanding what responsible heritage engagement entails.

Third, a participatory–agentic mechanism is expected to be particularly salient for participatory agency. Social cognitive theory emphasizes that efficacy beliefs are shaped partly through mastery experiences, perceived control, feedback, and opportunities to exercise agency [26]. When visitors can influence the direction and depth of heritage exploration rather than passively receive fixed information, they may experience themselves as more capable participants. Participatory agency is therefore theoretically closest to HSSE, but its influence need not be confined to efficacy. Active choice and control may also strengthen PHV by encouraging deeper meaning construction and FLOW by increasing autonomy, involvement, and perceived effectance.

The distinction among these mechanisms is therefore one of theoretical proximity rather than exclusivity. SGET and PCA are expected to relate most directly to epistemic and cultural evaluation; DC, ESP, and ANP to experiential absorption; and PA to perceived capability. Nevertheless, all six affordances are hypothesized to affect all three organism states because museum experiences unfold through intertwined processes of interpretation, immersion, and action appraisal. Sections 2.3–2.8 develop these primary and cross-mechanism relationships in detail.

The three organism states are connected through theoretically proposed associations. PHV captures the perceived enhancement of visitors' recognition of historical, cultural, social, and intergenerational significance during the reported experience. FLOW represents focused absorption. HSSE reflects confidence in feasible, responsible heritage-supporting action. The model proposes PHV–FLOW, PHV–HSSE, and FLOW–HSSE relationships, together with a direct association from each state to PHSI. These relationships are developed in Section 2.9.

PHSI is conceptualized as a visitor-level motivational outcome rather than as observed stewardship behavior or a direct measure of cultural sustainability. Nevertheless, it is theoretically relevant to cultural sustainability because the safeguarding, continued interpretation, and intergenerational transmission of cultural heritage depend partly on informed and responsible public participation. Cultural sustainability research has similarly treated heritage, cultural continuity, participation, and transmission as central components of the relationship between culture and sustainable development [27]. The present framework, therefore, connects museum-grounded generative-AI interpretation to cultural sustainability through a visitor-level motivational pathway, rather than treating sustainability as an automatic consequence of technology adoption.

Figure 2 summarizes the proposed model. Each of the six perceived affordances is hypothesized to relate positively to PHV, FLOW, and HSSE, yielding 18 stimulus-to-organism paths. PHV is related to FLOW and HSSE, FLOW to HSSE, and all three states to PHSI. The model contains 24 directional relationships. H1a–H6c identify the affordance-to-state paths; H7–H12 identify the three state-to-state and three state-to-intention paths.

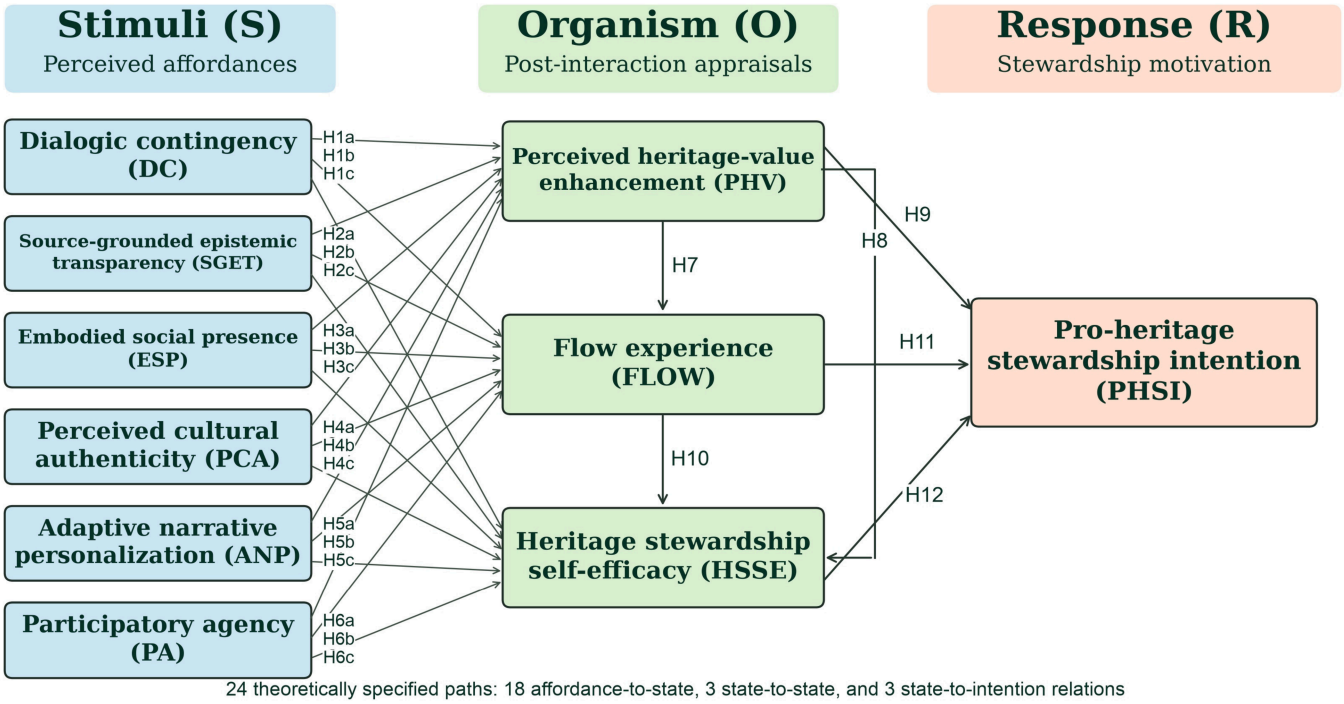


Figure 2. Proposed SOR model linking perceived affordances of a museum-grounded generative-AI virtual guide to pro-heritage stewardship intention.

2.3. Dialogic Contingency

Dialogic contingency refers to the extent to which a guide's responses meaningfully depend on a visitor's preceding and earlier conversational turns. It is narrower than generic interactivity or response speed because contingency requires semantic continuity across exchanges rather than merely rapid or bidirectional communication. A contingent guide recognizes the direction of an inquiry, retains relevant information from prior turns, and adjusts subsequent responses as the conversation develops. Communication research identifies such message interdependence as a defining feature of contingency and shows that interaction history can strengthen perceived dialog and user engagement [28]. Conversational contingency is also distinct from response latency: a reply may be rapid yet insufficiently related to what the user has previously said [29]. More recent human–AI research further indicates that contextual responsiveness is associated with stronger social connection and deeper cognitive and emotional engagement [30].

Among its psychological consequences, dialogic contingency is expected to relate most directly to flow experience. Contextually relevant responses reduce the need to repeat questions, repair misunderstandings, or redirect attention away from the interpretive task, thereby preserving interactional continuity. Prior experimental research has shown that perceived contingency plays an important role in translating message interactivity into stronger user engagement [28]. Although engagement and flow are conceptually distinct, both depend on sustained involvement; therefore, a dialog that remains responsive to a visitor's evolving inquiry should reduce interactional disruption and facilitate the concentration and absorption central to flow.

Dialogic contingency may also influence perceived heritage-value enhancement and heritage stewardship self-efficacy through related cross-mechanism processes. Museum experience is inherently interpretive and relational, with meaning emerging through interactions among visitors, heritage objects, interpretive resources, and place [31]. By linking explanations to visitors' questions and prior understanding, contingent dialog can make the historical and cultural significance of heritage more personally intelligible, thereby

strengthening perceived heritage-value enhancement. At the same time, context-specific clarification and feedback can reduce uncertainty about how heritage information should be understood, communicated, or acted upon. Consistent with social cognitive theory [26], such task-relevant guidance can strengthen visitors' confidence in making responsible heritage-related judgments and undertaking feasible stewardship actions. Accordingly:

H1a. *Dialogic contingency is positively associated with perceived heritage-value enhancement.*

H1b. *Dialogic contingency is positively associated with flow experience.*

H1c. *Dialogic contingency is positively associated with heritage stewardship self-efficacy.*

2.4. Source-Grounded Epistemic Transparency

Source-grounded epistemic transparency refers to the extent to which visitors can perceive the evidential basis, epistemic status, and limits of the heritage claims communicated by a generative-AI guide. It differs from general information quality. An answer may be fluent, relevant, and apparently accurate while providing little indication of what evidence supports it or whether a claim represents established knowledge, interpretation, or uncertainty. Research on human–AI interaction shows that meaningful transparency can reduce uncertainty and strengthen trust by providing users with explanatory knowledge about AI-generated judgments [32]. However, transparency is not uniformly beneficial: its effects depend on the accessibility, informativeness, and interpretability of what is disclosed [33,34]. This distinction is particularly important in heritage interpretation, where unsupported synthesis may distort chronology, provenance, cultural meaning, or contested interpretations. The integration of large language models with structured heritage knowledge therefore provides an important basis for traceable and institutionally constrained interpretation [20].

The most direct psychological consequence of SGET is expected to be perceived heritage-value enhancement. Heritage value is not communicated solely through the quantity or attractiveness of information but also through visitors' ability to understand why an interpretation is credible and how it relates to established heritage knowledge. Empirical research on heritage interpretation shows that interpretive information and principles significantly shape tourists' perceptions of heritage value, with visitors responding particularly to the substance and cultural meaning conveyed through interpretation [35]. When visitors can recognize that important claims are supported by museum or scholarly knowledge and can distinguish relatively established information from interpretive or uncertain content, they have a firmer epistemic basis for evaluating the historical, cultural, and social significance of the heritage. SGET should therefore strengthen PHV by connecting interpretive meaning with perceived evidential legitimacy.

SGET may also affect flow experience and heritage stewardship self-efficacy through cross-mechanism processes. Transparency can reduce epistemic uncertainty and the need for visitors to repeatedly question or externally verify the credibility of an AI-generated explanation, thereby lowering interactional friction and helping sustain attention. This does not imply that greater disclosure necessarily produces greater flow; explanations are beneficial only when they are meaningful and usable [32–34]. The hypothesis therefore concerns perceived epistemic transparency rather than the volume of disclosed information. In addition, access to intelligible evidential grounding can provide visitors with a more reliable basis for judging, communicating, and responding to heritage-related information. Consistent with social cognitive theory [26], reducing uncertainty about what is credible and how it can be justified should strengthen confidence in making responsible heritage-related judgments and undertaking feasible stewardship actions. Accordingly:

H2a. *Source-grounded epistemic transparency is positively associated with perceived heritage-value enhancement.*

H2b. *Source-grounded epistemic transparency is positively associated with flow experience.*

H2c. *Source-grounded epistemic transparency is positively associated with heritage stewardship self-efficacy.*

2.5. Embodied Social Presence

Embodied social presence refers to the extent to which visitors experience the virtual guide as a socially present companion rather than as a disembodied information channel. It is related to, but distinct from, anthropomorphism and visual realism. Humanlike appearance may encourage users to attribute intentions, agency, and social qualities to a nonhuman system, but social presence ultimately depends on whether voice, appearance, gesture, and behavior are experienced as coherent social cues rather than on realism alone [36]. Research on algorithmic human likeness likewise shows that perceptions of humanlike agency can shape social presence and downstream social influence [37]. In virtual museum settings, embodied AI avatars have also been found to generate higher engagement than voice-only or voice-plus-text interfaces even when informational content is held constant [13]. At the same time, highly realistic faces with abnormal features can produce uncanniness, indicating that embodiment should be understood in terms of perceived social presence rather than rendering fidelity [38].

Among the three organism states, ESP is expected to relate most directly to flow experience. A socially present guide can provide a stable interpersonal and attentional anchor, making the interaction feel continuous and accompanied rather than fragmented or purely transactional. Recent research on AI tour guides offers direct empirical support for this relationship: dimensions of automated social presence, including social affability, empathy, responsiveness, communication versatility, and competence, significantly enhance tourists' flow experience [39]. In a museum context, the perception of being accompanied by a responsive guide should therefore help sustain attention and involvement, facilitating the absorption characteristic of flow.

ESP may also influence perceived heritage-value enhancement and heritage stewardship self-efficacy through cross-mechanism effects. Socially situated interpretation can make heritage information feel less like an abstract database output and more like a meaningful interpersonal explanation, thereby increasing the salience and experiential value of cultural content. Tourism research similarly indicates that social presence can support participation-related outcomes [40]. In addition, socially framed guidance may make explanations and normative cues more salient and easier to interpret. From a social cognitive perspective [26], such socially legible guidance can reduce uncertainty about appropriate heritage-related judgments and, in turn, support visitors' confidence in undertaking feasible stewardship actions. Accordingly:

H3a. *Embodied social presence is positively associated with perceived heritage-value enhancement.*

H3b. *Embodied social presence is positively associated with flow experience.*

H3c. *Embodied social presence is positively associated with heritage stewardship self-efficacy.*

2.6. Perceived Cultural Authenticity

Perceived cultural authenticity (PCA) refers to visitors' judgment that the guide's representations and interpretations are culturally appropriate, contextually faithful, and

respectful of the meanings and traditions associated with the heritage. It is a visitor-level perception rather than an expert determination of historical truth or evidential accuracy. This distinction also separates PCA from source-grounded epistemic transparency: SGET concerns whether visitors can understand the evidential basis and epistemic status of a claim, whereas PCA concerns whether the resulting interpretation is experienced as culturally coherent and faithful to its heritage context. Cultural heritage research has long treated authenticity as a perceptual and evaluative construct rather than solely as an objective property of an artifact or site [24]. Empirical research further shows that perceived authenticity shapes visitor engagement and subsequent responses in heritage settings [41,42].

Among the three organism states, PCA is expected to have its most direct relationship with perceived heritage-value enhancement. Authentic interpretation situates objects, practices, and narratives within historically and culturally meaningful contexts, enabling visitors to recognize why the heritage matters beyond its immediate visual or technological presentation. Studies across European and Asian heritage contexts show that authenticity is closely connected with visitors' evaluation and engagement [24,41]. In virtual heritage environments, authenticity has likewise been shown to shape cognitive and affective responses, indicating that mediated representation does not eliminate visitors' concern with cultural and historical credibility [42]. Thus, when a generative-AI guide is perceived as faithfully representing cultural meanings and traditions, visitors should be better able to recognize the historical, cultural, social, and intergenerational significance of the heritage.

PCA may also strengthen flow experience and heritage stewardship self-efficacy through cross-mechanism processes. Authenticity can support experiential absorption by reducing representational incongruity and allowing visitors to engage with a culturally coherent interpretive environment. Importantly, prior tourism research provides direct evidence that authenticity can enhance flow through perceived value and involvement [43], while heritage research links authenticity with stronger visitor engagement [41]. Cultural authenticity may also provide a more credible normative frame for understanding appropriate ways of interpreting, communicating, and interacting with heritage. Consistent with social cognitive theory [26], clearer and culturally grounded guidance can reduce uncertainty about appropriate conduct and strengthen visitors' confidence in making responsible heritage-related judgments and undertaking feasible stewardship actions. PCA is therefore treated as a visitor-perceived judgment of cultural authenticity rather than as a substitute for expert authentication or curatorial validation. Accordingly:

H4a. *Perceived cultural authenticity is positively associated with perceived heritage-value enhancement.*

H4b. *Perceived cultural authenticity is positively associated with flow experience.*

H4c. *Perceived cultural authenticity is positively associated with heritage stewardship self-efficacy.*

2.7. Adaptive Narrative Personalization

Adaptive narrative personalization (ANP) refers to visitors' perception that the guide adjusts the depth, thematic emphasis, examples, and pace or amount of heritage interpretation to their prior knowledge, expressed interests, and ongoing interaction. It differs from dialogic contingency and participatory agency. Dialogic contingency concerns whether the system remains semantically responsive to the evolving conversation, whereas ANP concerns whether the content and form of interpretation are adapted to the individual visitor. Participatory agency, by contrast, concerns the visitor's perceived control over what to explore. Cultural-heritage personalization research has long emphasized the adaptation of content selection, presentation, and interaction to visitors' preferences, background

knowledge, and context, while retaining curatorial control over the underlying interpretive resources [44,45]. In conversational AI, longitudinal evidence likewise shows that personalization can alter perceptions of dialog quality, information relevance, credibility, trust, and privacy risk across repeated interactions [46].

Among the three organism states, ANP is expected to relate most directly to flow experience. Personalized interpretation can reduce mismatch between the amount or complexity of information provided and the visitor's available knowledge and attention. Research on adaptive digital guidance shows that dynamically tailored instructions can reduce cognitive load and improve learning outcomes relative to static guidance [47]. Flow research similarly emphasizes that compelling interactive experiences depend on an appropriate relationship among users' skills, perceived control, challenge, and focused attention [48]. In a museum context, adapting explanatory depth, examples, and information pacing to the visitor should therefore reduce overload or under-stimulation and help sustain attention and absorption during the encounter.

ANP may also strengthen perceived heritage-value enhancement and heritage stewardship self-efficacy through cross-mechanism processes. When heritage significance is explained through examples and levels of detail that are relevant to a visitor's prior knowledge and expressed interests, otherwise unfamiliar historical or cultural meanings may become more personally intelligible, strengthening perceived heritage-value enhancement. Personalization research in cultural heritage specifically argues that adaptive systems can recombine a common body of curated content into experiences that are meaningful to different visitors without relinquishing professional control over interpretation [44,45]. At the same time, adaptive explanation can make heritage-related guidance more comprehensible and realistically actionable by matching the amount and complexity of information to the visitor. Consistent with social cognitive theory [26], guidance that is appropriately matched to an individual's level of understanding can strengthen perceived capability by making responsible judgments and feasible stewardship actions appear more manageable. Accordingly:

H5a. *Adaptive narrative personalization is positively associated with perceived heritage-value enhancement.*

H5b. *Adaptive narrative personalization is positively associated with flow experience.*

H5c. *Adaptive narrative personalization is positively associated with heritage stewardship self-efficacy.*

2.8. Participatory Agency

Participatory agency (PA) refers to visitors' perceived capacity to meaningfully influence the direction and depth of their heritage exploration through their own choices. It extends beyond the mere availability of multiple options because agency requires visitors to perceive that their decisions have meaningful consequences for how the experience unfolds. PA is also distinct from adaptive narrative personalization: ANP concerns system-driven adaptation of interpretive content to the visitor, whereas PA concerns visitor-driven control over what to explore and how far to pursue it. Recent museum research conceptualizes visitor agency as multifaceted, encompassing the physical environment, social engagement, choice and autonomy, and relevance and empowerment, while emphasizing that agency should not be equated with unrestricted open-endedness [49]. Structured opportunities for meaningful choice may therefore support agency more effectively than either rigidly prescribed routes or undirected freedom.

Among the three organism states, PA is expected to relate most directly to heritage stewardship self-efficacy. Social cognitive theory identifies mastery experiences and perceived control as important foundations of efficacy beliefs [26]. When visitors can actively

determine what to investigate, pursue questions in greater depth, and observe that their choices meaningfully shape the interpretive experience, they move from passive reception toward active heritage-related judgment. Museum agency research similarly conceptualizes meaningful participation in terms of autonomy, relevance, and empowerment [49]. More direct heritage evidence comes from a quasi-experimental field study of community archeological participation, in which participants showed increased self-efficacy following active engagement with local heritage [50]. Although such participation is more intensive than interacting with a virtual guide, it supports the broader proposition that exercising agency in heritage-related activities can strengthen perceived capability.

PA may also strengthen perceived heritage-value enhancement and flow experience through cross-mechanism processes. Active choice encourages visitors to allocate attention according to their interests, compare interpretive possibilities, and construct personally meaningful connections with heritage content. Museum research on experience co-creation shows that visitors' active participation and interaction contribute to more intensive and meaningful experience-value formation [51]. At the experiential level, autonomy and perceived control can sustain intrinsic involvement by allowing visitors to regulate the pace and direction of exploration. Evidence from tourism settings, including a field study in an experiential museum, shows that autonomy facilitates immersion, while recent research in virtual museums similarly identifies autonomy as an important driver of presence [52,53]. Accordingly:

H6a. *Participatory agency is positively associated with perceived heritage-value enhancement.*

H6b. *Participatory agency is positively associated with flow experience.*

H6c. *Participatory agency is positively associated with heritage stewardship self-efficacy.*

2.9. Perceived Heritage-Value Enhancement, Flow Experience, Heritage Stewardship Self-Efficacy, and Pro-Heritage Stewardship Intention

Perceived heritage-value enhancement (PHV) denotes visitors' retrospective appraisal that the guide experience deepened their recognition of heritage significance. The four items ask about enhanced recognition of historical significance, cultural/social/spiritual meaning, contemporary and community relevance, and intergenerational importance. A knowledgeable visitor can report substantial heritage value but little additional enhancement; the scale therefore concerns perceived contribution of the experience, not an absolute value-recognition level. Interpretive value and flow are related in tourism research [54], and heritage interpretation has been linked to environmentally responsible responses [55]. A reported enhancement of meaningful understanding may also accompany confidence in making heritage-related judgments, as anticipated by social cognitive theory [26]. We therefore propose positive associations of PHV with FLOW, HSSE, and PHSI, without treating the post-interaction report as an observed pre–post change. Accordingly:

H7. *Perceived heritage-value enhancement is positively associated with flow experience.*

H8. *Perceived heritage-value enhancement is positively associated with heritage stewardship self-efficacy.*

H9. *Perceived heritage-value enhancement is positively associated with pro-heritage stewardship intention.*

Flow experience (FLOW) refers to a state of focused attention, deep absorption, temporal distortion, and immersion during interaction with the virtual guide. Sustained absorption can create repeated opportunities to explore, understand, and successfully navigate heritage information, which may in turn strengthen visitors' per-

ceptions of their own capability. Longitudinal evidence further indicates that flow can predict subsequent self-efficacy at several measurement occasions, supporting the possibility that absorbing experiences contribute to later capability beliefs [56]. Although reciprocal dynamics between flow and self-efficacy have also been reported, the present model focuses on the FLOW-to-HSSE direction because HSSE is conceptualized here as an action-capability appraisal emerging from the heritage interpretive encounter. Flow may also carry motivational involvement into stewardship intention. Nature-based tourism research has found positive associations between flow and environmentally responsible behavior [54], although other evidence indicates that flow may influence pro-environmental intention indirectly through environmental attitudes rather than always exerting a direct effect [57]. These findings suggest that the behavioral implications of flow may vary across contexts and may operate either directly or through additional psychological processes. In the present study, however, flow occurs within a heritage-centered interpretive encounter in which sustained involvement remains connected to heritage meaning and responsibility. A positive relationship with PHSI is therefore expected. Accordingly:

H10. *Flow experience is positively associated with heritage stewardship self-efficacy.*

H11. *Flow experience is positively associated with pro-heritage stewardship intention.*

Heritage stewardship self-efficacy concerns visitors' confidence that they can undertake feasible heritage-supporting actions. Social cognitive theory connects efficacy beliefs with the initiation and persistence of action [26]. Within the present encounter, understanding an appropriate response and believing oneself capable of carrying it out provide a basis for intending to act. HSSE is therefore expected to be positively associated with PHSI. Accordingly:

H12. *Heritage stewardship self-efficacy is positively associated with pro-heritage stewardship intention.*

3. Materials and Methods

3.1. Research Design and Study Workflow

This cross-sectional study examined self-reported post-interaction perceptions of Jinghui and their associations with heritage stewardship intention. It included no alternative-guide or control condition and estimated no experimentally manipulated treatment effect. The design cannot establish temporal ordering among these appraisals or the superiority of Jinghui over another guide.

The study followed a three-step workflow (Figure 3). Step 1 developed the 24-path SOR model, adapted the bilingual instrument, and assembled the screened survey dataset. Step 2 applied four complementary analyses: PLS-SEM estimated conditional associations, ANN described fitted nonlinear input importance, Basis-KAN compared held-out prediction with matched benchmarks, and NCA examined ceiling effects and bottlenecks. Step 3 interpreted the results with checks on measurement quality, multiplicity, predictive validation, and data provenance.

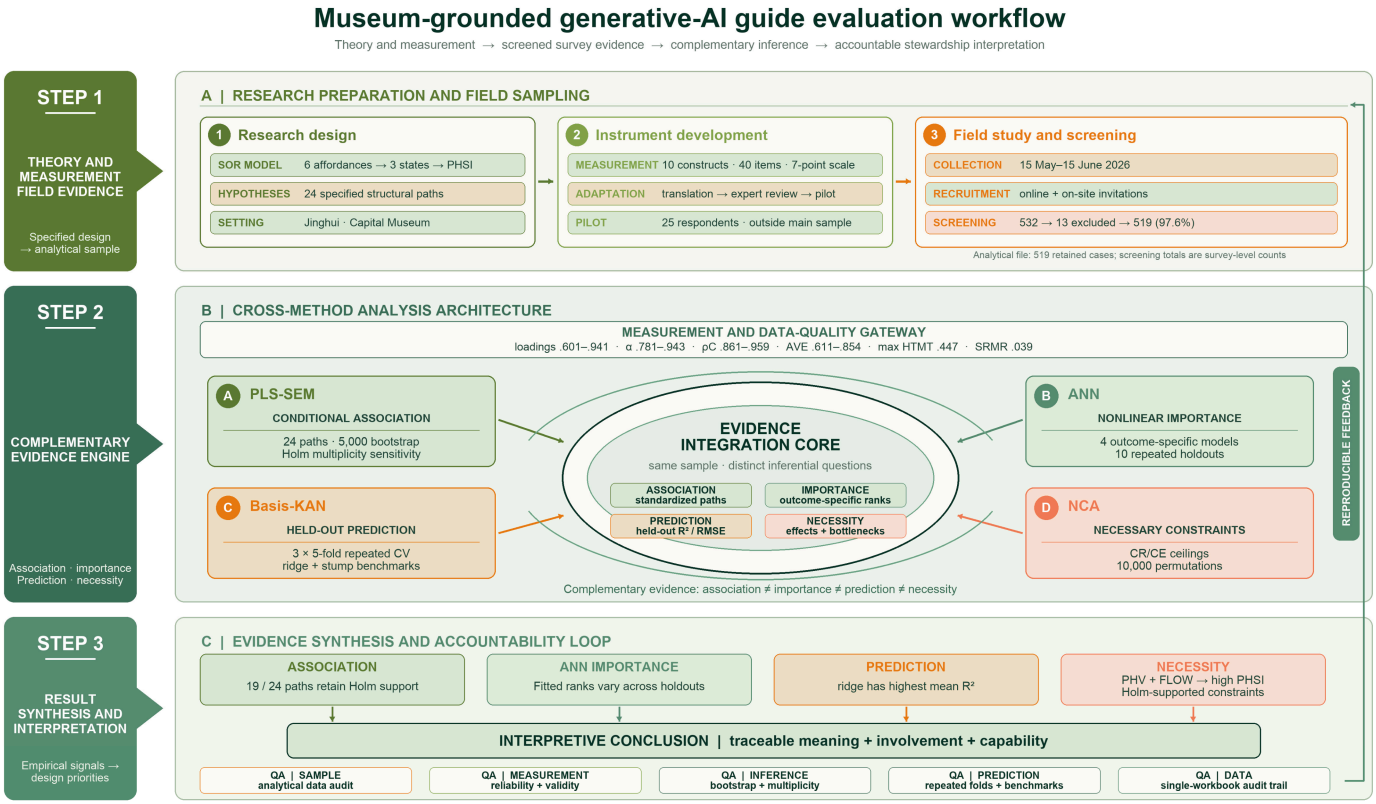


Figure 3. Integrated research workflow linking model development, bilingual measurement, field sampling, PLS-SEM, ANN, Basis-KAN benchmarking, NCA, and evidence synthesis. The synthesis panel distinguishes association, input importance, prediction, and Holm-supported response-stage necessity.

3.2. Measurement Instrument

A structured questionnaire was administered in Chinese to measure visitors' post-interaction perceptions of Jinghui and their pro-heritage stewardship intention. The questionnaire consisted of three sections. The first section provided the study introduction and completion instructions, including the research purpose, the academic use of the data, assurances of anonymity and confidentiality, the estimated completion time, and a statement that there were no right or wrong answers. The second section measured the ten constructs in the research model: six perceived affordances of Jinghui, perceived heritage-value enhancement, flow experience, heritage stewardship self-efficacy, and pro-heritage stewardship intention. Each construct was measured with four items, resulting in 40 measurement items in total. All measurement items were assessed using a seven-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree. The third section collected demographic information, including gender, age, education, occupation, and income. The constructs were specified reflectively: each set of items represents manifestations of a common visitor appraisal or psychological state, so the items are expected to covary. Reliability and convergent and discriminant validity were evaluated to assess this specification.

Items were contextualized or developed for the present museum-guide setting with reference to relevant literature. Table 1 identifies the development status and source basis for each item. DC was contextualized from contingency research [28–30]; SGET items were developed using transparency and explainability research [32–34]; ESP was contextualized from social-presence research [37,39]; PCA from authenticity research [24,41,42]; and ANP from personalization and adaptive-interpretation research [44–47]. PA items were developed from visitor-agency research [49]. PHV was adapted to a perceived-enhancement formulation using heritage-interpretation research [35,41,42], and FLOW was contextualized from flow research [48,54]. HSSE was developed following domain-specific efficacy principles [26,58].

PHSI items were developed for responsible heritage-supporting intentions, informed by conservation-participation and pro-social intention research [17,59]. These references identify measurement or conceptual foundations; the study-specific wording is not presented as an unchanged, previously validated scale.

Table 1. Measurement items, development status, and source basis. Contextual adaptations use study-specific wording; study-developed items are informed by the cited conceptual literature.

Construct	Code	Item	Development and Source
Dialogic contingency	DC1	The guide's replies directly addressed the questions I had just asked.	Contextual adaptation [28–30]
	DC2	The guide's follow-up replies reflected relevant information from earlier parts of our conversation.	Contextual adaptation [28–30]
	DC3	The guide changed its responses appropriately when I refined or redirected my questions.	Contextual adaptation [28–30]
	DC4	The conversation remained coherent across successive exchanges.	Contextual adaptation [28–30]
Source-grounded epistemic transparency	SGET1	The guide made the evidential basis of important heritage claims clear.	Study-developed [32–34]
	SGET2	I could tell which information in the guide's explanations was grounded in museum or scholarly sources.	Study-developed [32–34]
	SGET3	The guide made clear when information was uncertain, interpretive, or not fully established.	Study-developed [32–34]
	SGET4	I could judge how well supported the guide's heritage explanations were.	Study-developed [32–34]
Embodied social presence	ESP1	I felt as though a socially present guide was accompanying me during the visit.	Contextual adaptation [37,39]
	ESP2	The guide felt more like a social partner than a simple information-delivery system.	Contextual adaptation [37,39]
	ESP3	The guide's voice, appearance, and behavior created a sense of interpersonal interaction.	Contextual adaptation [37,39]
	ESP4	I experienced a clear sense of social presence while interacting with the guide.	Contextual adaptation [37,39]
Perceived cultural authenticity	PCA1	The guide's interpretations felt faithful to the historical and cultural context of the heritage.	Contextual adaptation [24,41,42]
	PCA2	The language, examples, and presentation used by the guide felt culturally appropriate.	Contextual adaptation [24,41,42]
	PCA3	The guide represented the meanings and traditions associated with the heritage in a respectful way.	Contextual adaptation [24,41,42]
	PCA4	Overall, the guide's interpretation felt culturally authentic.	Contextual adaptation [24,41,42]
Adaptive narrative personalization	ANP1	The depth of explanation adapted to my level of prior knowledge.	Contextual adaptation [44–47]
	ANP2	The guide emphasized topics that matched the interests I expressed.	Contextual adaptation [44–47]
	ANP3	The examples and comparisons used by the guide were tailored to help me understand the heritage.	Contextual adaptation [44–47]
	ANP4	The pace and amount of information were adapted to the way I interacted with the guide.	Contextual adaptation [44–47]
Participatory agency	PA1	I could decide which heritage topics or objects to explore in greater depth.	Study-developed [49]
	PA2	I could decide when to explore a topic further and when to move on.	Study-developed [49]
	PA3	My choices meaningfully influenced the direction of the experience.	Study-developed [49]
	PA4	I felt that I had meaningful control over how the experience unfolded.	Study-developed [49]

Table 1. *Cont.*

Construct	Code	Item	Development and Source
Perceived heritage-value enhancement	PHV1	The experience deepened my recognition of the heritage's historical significance.	Enhancement adaptation [35,41,42]
	PHV2	The experience deepened my understanding of the heritage's cultural, social, or spiritual significance.	Enhancement adaptation [35,41,42]
	PHV3	The experience helped me recognize the relevance of this heritage to contemporary society and communities.	Enhancement adaptation [35,41,42]
	PHV4	The experience strengthened my recognition of the heritage's significance for future generations.	Enhancement adaptation [35,41,42]
Flow experience	FLOW1	My attention remained fully focused during the interaction.	Contextual adaptation [48,54]
	FLOW2	I became deeply absorbed in exploring the heritage content.	Contextual adaptation [48,54]
	FLOW3	I temporarily lost track of time while interacting with the guide.	Contextual adaptation [48,54]
	FLOW4	I felt fully immersed in the experience throughout the interaction.	Contextual adaptation [48,54]
Heritage Stewardship Self-Efficacy	HSSE1	I am confident that I can recognize situations or behaviors that may threaten cultural heritage.	Study-developed [26,58]
	HSSE2	I am confident that I can make responsible choices when interacting with cultural heritage.	Study-developed [26,58]
	HSSE3	I am confident that I can take feasible actions that contribute to heritage safeguarding.	Study-developed [26,58]
	HSSE4	I am confident that I can respond appropriately when I become aware of a heritage-related risk.	Study-developed [26,58]
Pro-heritage stewardship intention	PHSI1	I intend to behave responsibly when visiting or interacting with cultural heritage.	Study-developed [17,59]
	PHSI2	I intend to communicate heritage information responsibly and avoid spreading unverified claims.	Study-developed [17,59]
	PHSI3	I intend to support heritage safeguarding or public-education activities when opportunities are available.	Study-developed [17,59]
	PHSI4	If I become aware of a threat to heritage, I intend to respond through an appropriate channel.	Study-developed [17,59]

Where English-language source items were used, they were translated into Chinese and back-translated following established procedures for cross-cultural research [60]. The translated questionnaire was reviewed by six experts, including design professors and doctoral researchers, who assessed the conceptual relevance, clarity, and contextual appropriateness of the items. This procedure followed established recommendations for expert review and preliminary scale development [61]. Their feedback was used to refine the item wording and improve the applicability of the instrument to the museum and generative-AI context. A preliminary pilot test was then conducted with 25 respondents to evaluate the clarity, readability, and administration feasibility of the questionnaire. The pilot respondents were excluded from the main survey and the final analytical sample. All final measurement items were positively worded, and no reverse-coded items were included.

3.3. Participants and Data Collection

The study targeted adults reporting prior Jinghui use. Recruitment used Xiaohongshu and WeChat invitations and on-site invitations at the Capital Museum from 15 May to 15 June 2026. At questionnaire entry, respondents were asked whether they had experienced Jinghui; affirmative self-report formed the use-eligibility criterion. Respondents then evaluated their reported experience and provided consent after receiving information about the study purpose, voluntary participation, anonymity, and academic use. Use eligibility was based on self-report rather than session-record verification.

The survey yielded 532 questionnaires. Thirteen questionnaires were excluded during screening because of duplicate submissions ($n = 6$), incomplete responses ($n = 2$), failure to meet participant eligibility criteria ($n = 3$), and clearly patterned responding indicative of inattentive completion ($n = 2$). The final analytical sample comprised 519 cases, corresponding to a retention rate of 97.6%. All retained cases contained complete responses to the 40 measurement items and the corresponding ten construct means; no item-level missing-data imputation was required. Table 2 summarizes the demographic characteristics of the final sample.

Table 2. Demographic characteristics of the respondents.

Variable	Category	Frequency (*n*)	Percentage (%)
Gender	Male	247	47.59
	Female	272	52.41
Age	18–29 years	298	57.42
	30–40 years	165	31.79
	≥41 years	56	10.79
Education	High school or below	49	9.44
	Associate degree	119	22.93
	Bachelor's degree	257	49.52
	Master's degree or above	94	18.11
Occupation	Student	64	12.33
	Government or public-institution employee	138	26.59
	Company employee	183	35.26
	Freelancer or self-employed	108	20.81
	Other	26	5.01
Monthly income or living expenses (RMB)	≤3000	54	10.40
	3001–6000	149	28.71
	6001–9000	168	32.37
	9001–12,000	103	19.85
	≥12,001	45	8.67
Total		519	100.00

3.4. Data Analysis

PLS-SEM used reflective Mode A blocks, path inner weighting, and standardized item data in a reproducible Python 3.13.5/NumPy 2.2.6 implementation of the iterative PLS algorithm [62]. All outer weights began at +1; normalized scores and weights were updated until the largest absolute weight change was below 10^{-7}, with a maximum of 3000 iterations. The full model comprised 24 paths. Participant-level bootstrap resampling used 5000 draws and seed 20261016; item standardization and the complete weight-estimation algorithm were repeated within every draw. Signs were oriented to the positively keyed indicators. We report bootstrap standard errors, 95% percentile intervals, two-sided normal-approximation *p*-values from coefficient/standard-error ratios, and Holm-adjusted *p*-values across all 24 tests. The latter are not empirical bootstrap tail probabilities. A matched 23-path model omitting only PHV → HSSE provides a specification-sensitivity comparison on the same data. Measurement assessment includes standardized and raw alpha, rho_A, composite reliability, AVE, loadings, construct correlations, and HTMT with bootstrap intervals. SRMR is calculated from the observed and model-implied item correlations. Paired bootstrap coefficient contrasts assess uncertainty in close rankings. The separate 5000-resample HTMT procedure retains seed 20261010 from the submitted analysis. The

equations test direct associations; formal indirect effects and sequential mediation were not tested.

All substantive variables came from one post-interaction questionnaire. Procedural separation of predictor and outcome blocks, construct-specific wording, and anonymous participation reduced evaluation apprehension and demand cues. The first unrotated component accounted for 17.7% of item variance, and full-collinearity VIFs based on arithmetic-mean scores ranged from 1.10 to 1.48.

ANN analysis used four outcome-specific feed-forward networks: the six affordances predicted PHV; the six affordances plus PHV predicted FLOW; the six affordances plus PHV and FLOW predicted HSSE; and PHV, FLOW, and HSSE predicted PHSI. Each network had one sigmoid hidden layer with four, four, four, or two nodes, respectively, and a linear output. Ten repeated holdouts used seeds 2026–2035. Each split assigned 363 cases to development and 156 to testing; within development, 291 cases optimized the weights and 72 supplied validation for early stopping. Input z-scores and target min–max scaling were fitted exclusively on the optimization subset. Adam used a learning rate of 0.001, batches of 32, an L2 weight penalty of 0.0001, and a maximum of 500 epochs. Training stopped after 30 epochs without a validation-MSE improvement of at least 0.0000001, and the best validation weights were restored. Testing data were used only for final evaluation. RMSE was calculated on the optimization-derived target scale. After restoring the best validation weights, training RMSE was evaluated on all 363 development cases, including the 72 validation cases. The constant baseline predicted the outcome mean of these same 363 development cases and was evaluated on the 156 test cases, alongside testing R^2. Input importance was the absolute signed sum of input–hidden–output weight products, normalized to sum to one within each run, averaged across runs, and then expressed relative to the largest mean importance within each outcome [63,64]. Predictors and outcomes were arithmetic means of their four original-scale items, not PLS-weighted scores. Per-run importance ranks, their observed ranges, and first-place counts summarize split-and-training variability; they do not estimate population rank uncertainty.

The additive Basis-KAN implementation was compared with ridge regression and boosted stumps using three repeats of five-fold cross-validation (seed 2026). Identical folds and predictor sets were used for all three models within each outcome. PHV used the six affordances; FLOW used the six affordances plus PHV; HSSE used the six affordances plus PHV and FLOW; and PHSI used all six affordances and three organism states. The Basis-KAN implementation summed fitted univariate expansions, with five triangular piecewise-linear basis functions per predictor, knots spanning the training-fold range, and a ridge penalty of 1. It was a restricted additive basis model, without learned cross-predictor interactions or a multilayer KAN architecture. Predictor scaling and basis construction were fitted within the training folds. Ridge used alpha = 1; boosted stumps used 100 depth-one learners and a learning rate of 0.05. These settings were held fixed across folds. The cleaned analytical data and analysis outputs are consolidated in Reviewer_Data_and_Analysis_Package.xlsx; the bilingual questionnaire and comprehensive reproducibility archive are supplied separately.

Necessary-condition analysis used the four-item arithmetic means on the original questionnaire scale, with CE-FDH and CR-FDH ceiling procedures [65,66]. The six affordances were examined as conditions for PHV, FLOW, and HSSE, and the three organism states as conditions for PHSI, giving 21 relationships per ceiling procedure. CE-FDH uses the upper-left free-disposal frontier at unique coordinate pairs; CR-FDH fits an unweighted regression through those unique frontier vertices. The empirical scope is the product of the observed condition and outcome ranges. Each relationship used 10,000 outcome permutations, seed 20261016, and p = (1 + number of permuted effects at least as large

as the observed effect)/(10,001). Holm adjustment was applied separately to each 21-test ceiling family. CR-FDH bottlenecks are expressed as percentages of the observed ranges. Raw score ranges, frontier coordinates, ceiling equations, and scatterplots accompany the numerical results. No fuzzy-set calibration was used.

4. Results

4.1. Measurement Model Assessment

Table 3 reports measurement estimates recomputed from the same 519-case item dataset as the structural model. Standardized alpha ranged from 0.781 to 0.943, rho_A from 0.800 to 0.948, composite reliability from 0.861 to 0.959, and AVE from 0.611 to 0.854. Loadings ranged from 0.601 to 0.941. HSSE1 had the lowest loading (0.601); the four-item HSSE block nevertheless had composite reliability of 0.861 and AVE of 0.611. Maximum HTMT was 0.447, and the largest upper 95% percentile-bootstrap bound was 0.523. Each square root of AVE exceeded the corresponding absolute between-construct score correlations. Complete loadings, weights, correlations, and bootstrap results are supplied in the supplementary workbook and analysis archive. HSSE's comparatively weak loading and reliability indicate a need to refine action-specific coverage, including confidence in identifying risks, choosing an appropriate response, and carrying it out with museum support.

Table 3. Measurement-model assessment.

Construct	Standardized α	rho_A/rho_C	AVE	Loading Range
DC	0.943	0.946/0.959	0.854	0.901–0.941
SGET	0.921	0.948/0.943	0.807	0.841–0.927
ESP	0.906	0.909/0.934	0.780	0.850–0.906
PCA	0.931	0.933/0.951	0.830	0.894–0.928
ANP	0.899	0.901/0.930	0.767	0.850–0.900
PA	0.892	0.895/0.925	0.755	0.832–0.902
PHV	0.886	0.898/0.921	0.745	0.800–0.904
FLOW	0.857	0.860/0.903	0.699	0.808–0.871
HSSE	0.781	0.800/0.861	0.611	0.601–0.852
PHSI	0.883	0.889/0.919	0.740	0.834–0.884

4.2. Structural Model Results

Table 4 and Figure 4 report the complete 24-path model. The algorithm converged in six iterations; estimated-model SRMR was 0.039. Twenty-three coefficients were positive and met nominal $p < 0.05$, but only 19 of the 24 tests retained significance after Holm adjustment. PHV → HSSE was near zero and unsupported (β = −0.010, 95% CI −0.101 to 0.077, $p = 0.823$). ANP → HSSE, PA → HSSE, PHV → FLOW, and FLOW → HSSE were positive at the nominal level but did not retain Holm significance. Thus, the data do not support every proposed positive relationship.

For PHV, fitted coefficients were DC 0.236, ANP 0.208, SGET 0.206, PCA 0.186, ESP 0.178, and PA 0.123. DC had the largest point estimate, while the close leading values identify several jointly relevant candidates rather than a demonstrated superiority of one feature. For FLOW, DC and PA had coefficients of 0.240 and 0.226, followed by SGET 0.194, ESP 0.187, PCA 0.155, ANP 0.138, and PHV 0.108. The paired PA-minus-DC difference was −0.014 (95% bootstrap CI −0.120 to 0.095), providing no evidence that their conditional associations differ.

Table 4. Structural-model results.

H	Path	β	95% CI	t	*p*	Holm *p*
H1a	DC → PHV	0.236	0.165, 0.309	6.435	<0.001	<0.001
H1b	DC → FLOW	0.240	0.166, 0.314	6.382	<0.001	<0.001
H1c	DC → HSSE	0.139	0.047, 0.229	3.001	0.003	0.019
H2a	SGET → PHV	0.206	0.131, 0.285	5.208	<0.001	<0.001
H2b	SGET → FLOW	0.194	0.115, 0.271	4.891	<0.001	<0.001
H2c	SGET → HSSE	0.192	0.112, 0.271	4.709	<0.001	<0.001
H3a	ESP → PHV	0.178	0.103, 0.253	4.716	<0.001	<0.001
H3b	ESP → FLOW	0.187	0.112, 0.262	4.885	<0.001	<0.001
H3c	ESP → HSSE	0.127	0.043, 0.206	3.043	0.002	0.019
H4a	PCA → PHV	0.186	0.114, 0.255	5.111	<0.001	<0.001
H4b	PCA → FLOW	0.155	0.080, 0.229	4.128	<0.001	<0.001
H4c	PCA → HSSE	0.191	0.111, 0.274	4.578	<0.001	<0.001
H5a	ANP → PHV	0.208	0.140, 0.280	5.806	<0.001	<0.001
H5b	ANP → FLOW	0.138	0.060, 0.214	3.499	<0.001	0.004
H5c	ANP → HSSE	0.098	0.014, 0.184	2.265	0.023	0.070
H6a	PA → PHV	0.123	0.041, 0.200	3.039	0.002	0.019
H6b	PA → FLOW	0.226	0.148, 0.309	5.380	<0.001	<0.001
H6c	PA → HSSE	0.083	0.007, 0.163	2.112	0.035	0.070
H7	PHV → FLOW	0.108	0.023, 0.189	2.547	0.011	0.054
H8	PHV → HSSE	−0.010	−0.101, 0.077	−0.224	0.823	0.823
H9	PHV → PHSI	0.267	0.192, 0.343	6.958	<0.001	<0.001
H10	FLOW → HSSE	0.116	0.023, 0.208	2.441	0.015	0.059
H11	FLOW → PHSI	0.249	0.173, 0.324	6.410	<0.001	<0.001
H12	HSSE → PHSI	0.215	0.135, 0.293	5.302	<0.001	<0.001

Note: β = standardized coefficient. Intervals are 95% participant-bootstrap percentiles from 5000 complete PLS refits. *p*-values use two-sided normal approximations with bootstrap standard errors; Holm *p*-values adjust the full family of 24 path tests. A positive hypothesis is supported only when its direction and the stated significance criterion are both met.

In the HSSE equation, SGET and PCA had closely matched coefficients of 0.192 and 0.191. DC, ESP, FLOW, ANP, and PA had coefficients of 0.139, 0.127, 0.116, 0.098, and 0.083; PHV had β = −0.010. The paired PCA-minus-SGET contrast was −0.001 (95% CI −0.110 to 0.111). Epistemic transparency and cultural authenticity are therefore jointly relevant correlates of capability; their relative superiority is not established. PHV, FLOW, and HSSE were positively associated with PHSI (β = 0.267, 0.249, and 0.215), all retaining

Holm significance. Paired intervals for differences among these response-stage coefficients included zero, so their numerical ordering does not establish a population hierarchy.

Explained variance was 21.7% for PHV, 27.6% for FLOW, 17.0% for HSSE, and 27.3% for PHSI. A matched reduced specification omitting PHV → HSSE had corresponding R^2 values of 0.217, 0.276, 0.170, and 0.273, and its HSSE → PHSI coefficient was 0.215. Full-model HSSE → PHSI was also 0.215. The response-stage interpretation was therefore stable under inclusion or exclusion of the unsupported PHV → HSSE relationship. Between 72.4% and 83.0% of organism-state variance and 72.7% of intention variance remained unexplained. The results identify candidates for service evaluation rather than accurate individual-level stewardship predictions.

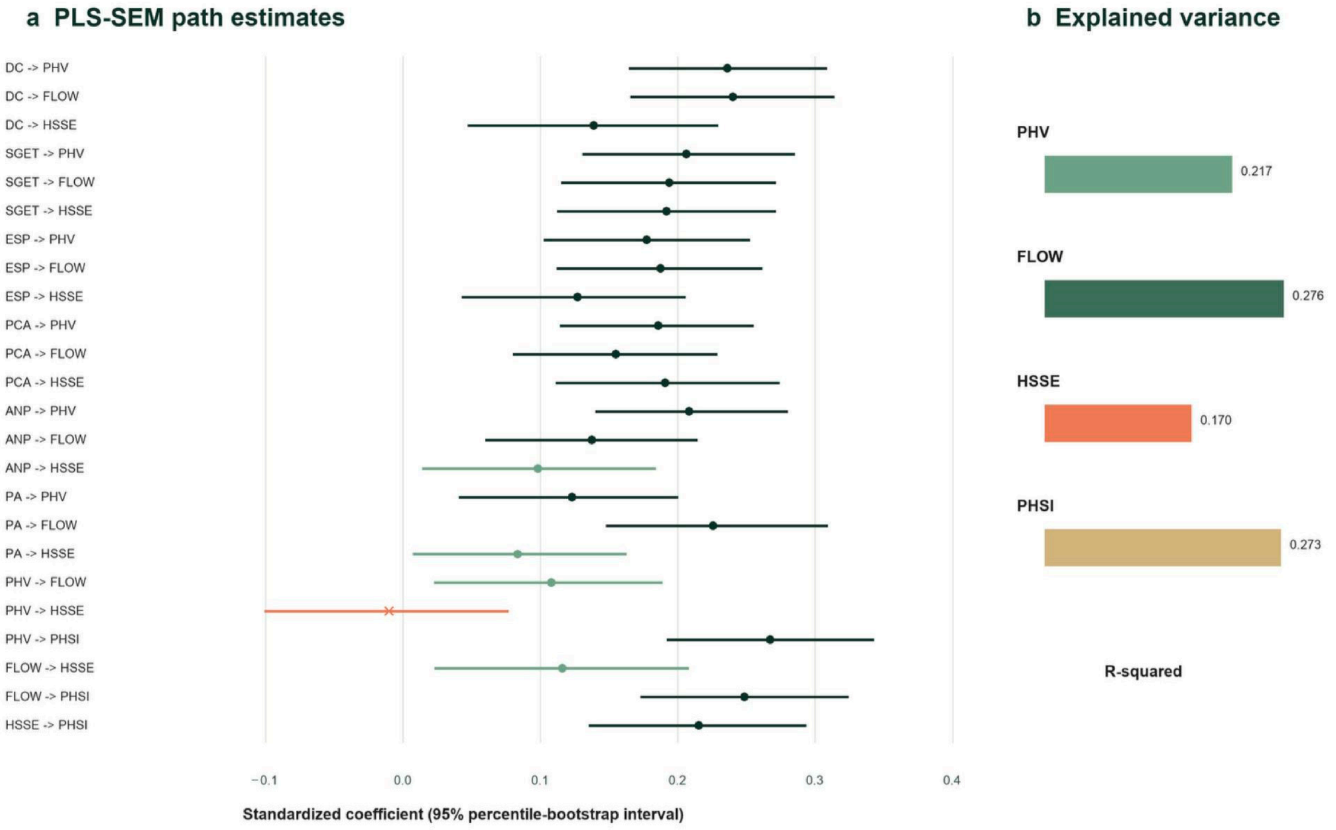


Figure 4. Complete 24-path PLS-SEM estimates and explained variance. Points show standardized coefficients and lines show 95% percentile intervals from 5000 participant bootstrap samples. Dark green identifies Holm $p < 0.05$, sage identifies nominal $p < 0.05$ only, and the coral cross marks the unsupported PHV → HSSE hypothesis.

4.3. Artificial Neural-Network Analysis

Table 5 summarizes ANN performance across ten repeated holdouts. Mean training/testing RMSE was 0.169/0.171 for PHV, 0.171/0.181 for FLOW, 0.172/0.182 for HSSE, and 0.148/0.156 for PHSI. Testing RMSE was lower on average than the constant development-mean baseline for each outcome; corresponding baseline errors were 0.184, 0.207, 0.189, and 0.176. Mean testing R^2 was 0.118, 0.225, 0.066, and 0.208. HSSE had the weakest held-out performance. Its ANN includes PHV and FLOW alongside the six affordances. Small average training–testing gaps do not remove the substantial split variation, particularly for PHV.

Table 5. Ten-run artificial neural-network performance.

Outcome	Inputs	Training RMSE	Testing RMSE	Baseline RMSE	Testing R^2	Top Input
PHV	6	0.169 ± 0.016	0.171 ± 0.013	0.184 ± 0.016	0.118 ± 0.106	DC
FLOW	7	0.171 ± 0.011	0.181 ± 0.011	0.207 ± 0.015	0.225 ± 0.069	PA
HSSE	8	0.172 ± 0.004	0.182 ± 0.008	0.189 ± 0.011	0.066 ± 0.071	SGET
PHSI	3	0.148 ± 0.013	0.156 ± 0.018	0.176 ± 0.019	0.208 ± 0.058	PHV

Note: Mean ± SD across ten repeated holdouts. RMSE is on the optimization-derived min–max target scale. Training RMSE is evaluated on all 363 development cases (291 optimization + 72 validation) after restoring the best validation weights. The baseline predicts the mean outcome of all 363 development cases and is evaluated on the 156 test cases. Testing R^2 is dimensionless.

The fitted mean-importance profiles differ by outcome (Figure 5). For PHV, DC was 100%, ESP 89.4%, ANP 89.3%, SGET 80.9%, PCA 63.4%, and PA 37.8%. For FLOW, PA was 100%, DC 91.1%, ESP 83.9%, SGET 74.2%, PCA 64.3%, ANP 58.9%, and PHV 54.3%. For the eight-input HSSE model, SGET was 100%, PCA 98.3%, FLOW 81.0%, DC 74.9%, PA 47.6%, ESP 46.7%, ANP 43.3%, and PHV 29.5%. SGET and PCA are closely ranked in this fitted capability model and are treated as jointly relevant inputs.

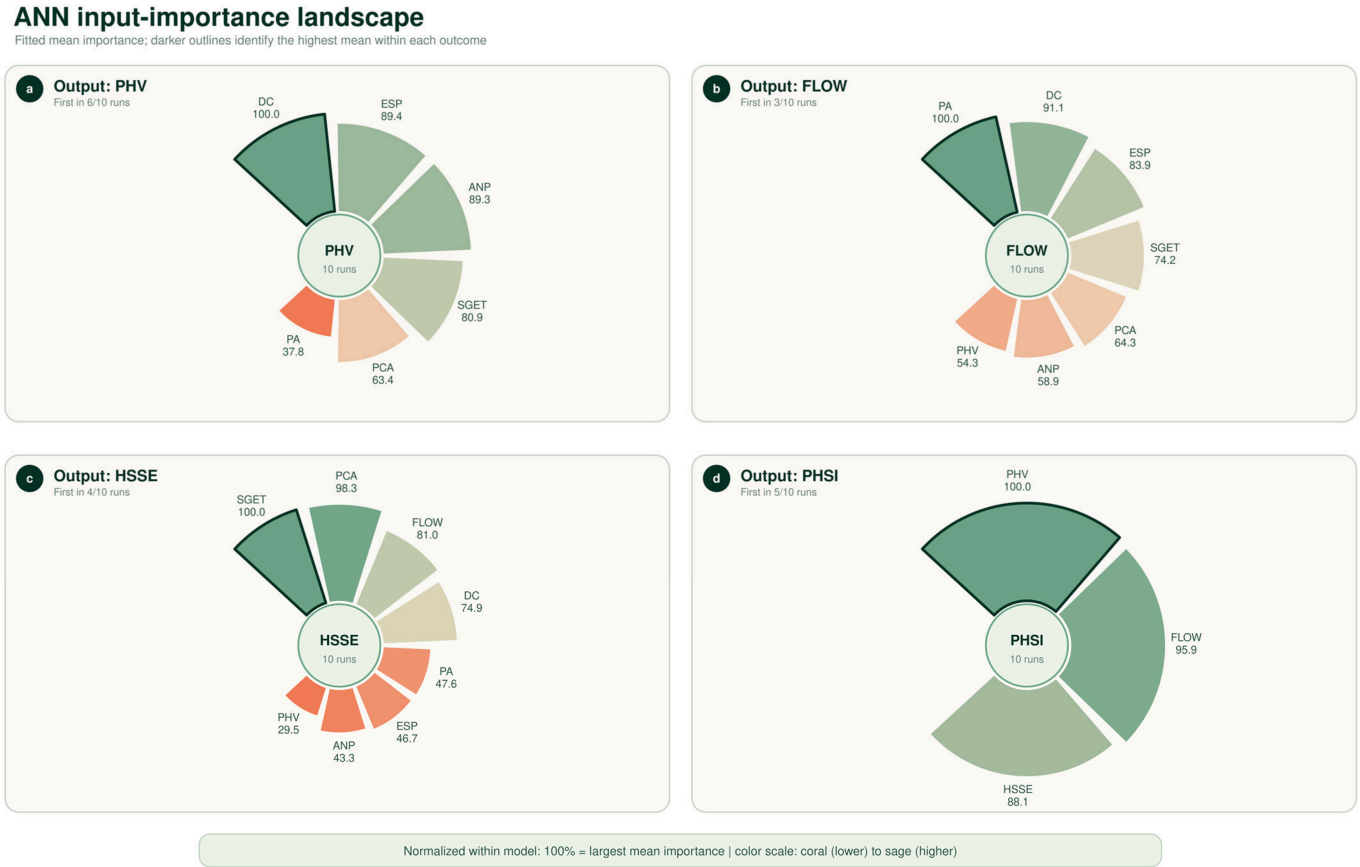


Figure 5. ANN mean normalized connection-weight importance across ten repeated holdouts. Each outcome is normalized to 100% for its largest mean input importance. Panel subtitles report how often that input ranked first; the supplementary rank table gives all per-run ranks and observed rank ranges. These summaries represent split-and-training variability within the same 519 cases.

For PHSI, the mean normalized importances were PHV 100%, FLOW 95.9%, and HSSE 88.1%. All three occupied each of the first three positions across individual runs. PHV ranked first in five runs, FLOW in three, and HSSE in two. The mean ordering describes these networks, not a statistically established difference between the underlying constructs.

Importance is normalized within each outcome; it is neither a share of explained variance nor an effect size comparable across outcomes.

PLS-SEM and ANN give conditional descriptions of the same sample. DC has the largest fitted mean or coefficient for PHV, but ranked first in only six of ten ANN runs. For FLOW, PA has the largest mean ANN importance and DC the largest structural coefficient; each ranked first in three ANN runs. For HSSE, SGET and PCA have near-identical structural coefficients and mean ANN importances, and ranked first in four and three runs. These results nominate responsive dialog, visitor agency, epistemic transparency, and cultural authenticity for subsequent comparative testing. Agreement between analyses of the same questionnaire is not independent validation or evidence that altering a feature changes an outcome.

4.4. Basis-KAN Model Comparison

Figure 6 compares the additive Basis-KAN implementation with ridge regression and boosted stumps under the same three-repeat, five-fold partitions. Each outcome and model was evaluated on 15 held-out folds. Ridge had the highest mean held-out R^2 for every outcome: 0.164 ± 0.054 for PHV, 0.237 ± 0.084 for FLOW, 0.125 ± 0.052 for HSSE, and 0.238 ± 0.068 for PHSI. The corresponding Basis-KAN values were 0.122 ± 0.055, 0.175 ± 0.087, 0.085 ± 0.078, and 0.199 ± 0.084. Basis-KAN's mean R^2 was therefore lower by 0.042, 0.062, 0.040, and 0.039, respectively. These comparisons establish the ordering of mean performance under the evaluated specifications; fold-level SD describes dispersion and is not a confidence interval for a model difference.

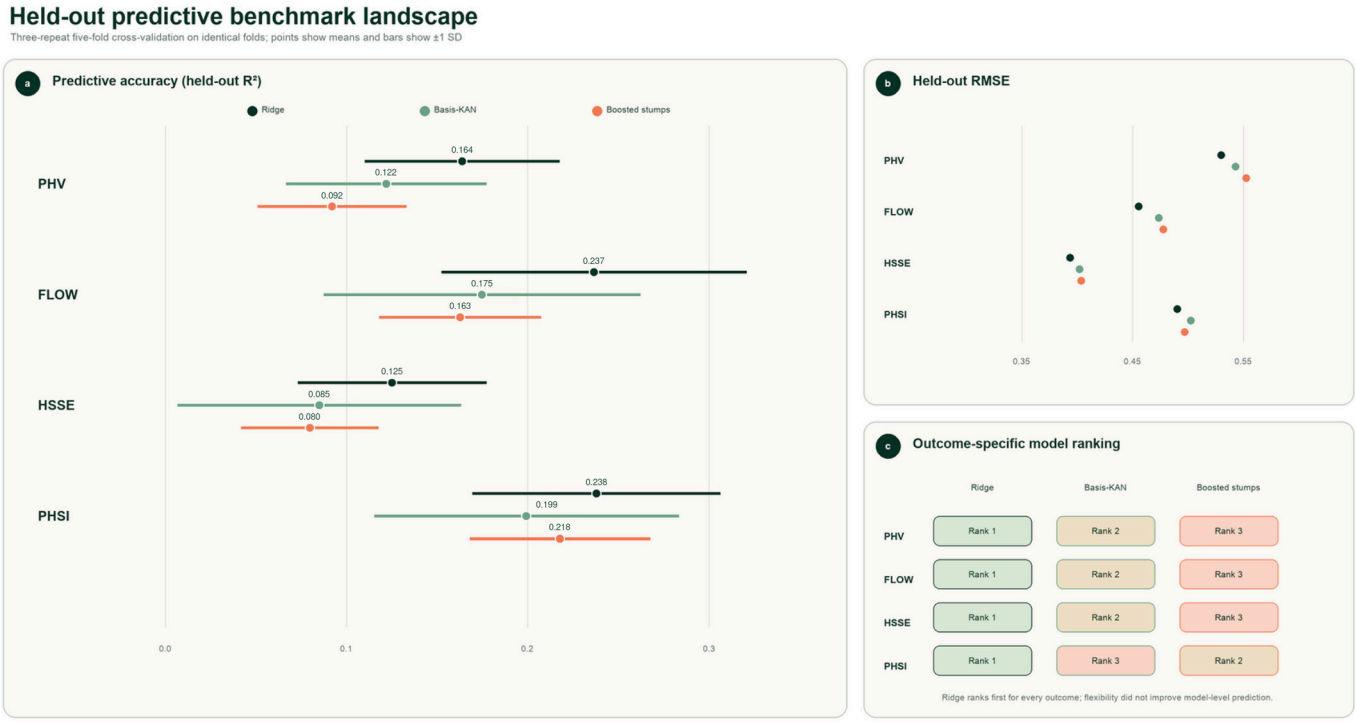


Figure 6. Matched repeated-cross-validation performance of ridge regression, additive Basis-KAN, and boosted stumps. Points and intervals show mean ± SD across 15 held-out folds; ranking uses held-out R^2. RMSE is in original 1–7 arithmetic-mean questionnaire-score units. SD describes fold dispersion, not a confidence interval for a between-model difference.

The error metrics gave the same practical conclusion. Mean RMSE for ridge versus Basis-KAN was 0.530 versus 0.543 for PHV, 0.456 versus 0.474 for FLOW, 0.394 versus 0.402 for HSSE, and 0.491 versus 0.503 for PHSI. Boosted stumps yielded mean R^2 values of 0.092,

0.163, 0.080, and 0.218. Basis-KAN exceeded boosted stumps for PHV, FLOW, and HSSE, but boosted stumps ranked above Basis-KAN for PHSI. The flexible additive basis model thus occupied an intermediate position for the organism states and the lowest position for the response. The results support ridge as the strongest tested benchmark for mean predictive accuracy in this dataset.

The benchmark evaluates the predictive value of a specific additive basis representation. Five piecewise-linear functions per predictor allow each input to have a flexible univariate relationship with the outcome, but the fitted model does not include cross-predictor interaction terms. Its lower held-out performance shows that this additional univariate flexibility did not improve prediction over the ridge specification. ANN importance addresses a separate question about input relevance. Its RMSE values are not directly comparable with the benchmark because the analyses use different validation schemes, outcome scales, and, for PHSI, predictor sets. The absence of a Basis-KAN gain, therefore, does not contradict the outcome-specific ANN rankings.

4.5. Necessary-Condition Analysis and Bottlenecks

Figure 7 distinguishes ceiling effect sizes from observed-range bottlenecks. At nominal $p < 0.05$, CR-FDH identified eight relationships. No individual affordance met this criterion for PHV. For FLOW, ANP had d = 0.145 ($p = 0.015$) and PA d = 0.132 ($p = 0.011$). For HSSE, the nominal relationships were DC d = 0.092 ($p = 0.016$), SGET d = 0.130 ($p = 0.044$), PCA d = 0.059 ($p = 0.016$), and PA d = 0.074 ($p = 0.006$). For PHSI, PHV had d = 0.225 and FLOW d = 0.154 (both $p = 0.00010$); HSSE had d = 0.112 ($p = 0.071$) and did not meet the nominal criterion. The supplied numerical outputs and scatterplots use the same arithmetic-mean item scores.

CE-FDH identified the same eight nominally significant relationships. Its effects were 0.162 and 0.139 for ANP and PA in the FLOW equation; 0.117, 0.194, 0.078, and 0.089 for DC, SGET, PCA, and PA in the HSSE equation; and 0.264 and 0.182 for PHV and FLOW in the PHSI equation. Agreement between the two ceiling procedures preserves the main descriptive ordering, including the largest necessity effect for PHV in relation to PHSI. Both procedures use the same observations, so their agreement is a sensitivity check on ceiling specification rather than an independent replication.

Holm adjustment across the 21 tests in each ceiling family retained only PHV and FLOW for PHSI, with adjusted $p = 0.0021$ for both relationships under CE-FDH and CR-FDH. The other six nominal relationships did not retain familywise significance and remain exploratory. Permutation p-values use a declared seed and a finite-permutation correction with a lower bound of 1/10,001.

The CR-FDH bottlenecks describe minimum condition levels along the estimated ceilings. At the 70% PHSI target, the levels were 36.7% for PHV and 18.3% for FLOW; at 90%, they rose to 55.4% and 58.6%; and at 100%, to 64.7% and 78.8%. FLOW's estimated requirement increased more steeply toward the upper end of PHSI, while PHV had the larger overall necessity effect. These metrics capture different features of the ceiling: effect size summarizes the excluded region, whereas a bottleneck identifies a condition level at a particular outcome target. HSSE's estimated levels of 49.5% at the 90% target and 73.7% at 100% are descriptive.

The exploratory FLOW and HSSE bottlenecks are retained in the supplement, including the observed score ranges and six ceiling scatterplots. For the two Holm-retained response-stage relations, Figure 8 plots all 519 observations and both ceiling procedures. PHV ranged from 3.75 to 7, FLOW from 4.25 to 7, and PHSI from 3.5 to 7. The 70%, 90%, and 100% PHSI targets correspond to raw mean scores of 5.95, 6.65, and 7.00. A 90% target required estimated CR-FDH PHV and FLOW levels of 55.4% and 58.6% of their observed

ranges. These percentages are positions in a sample-specific score range, not percentages of visitors or universal questionnaire cut-points.

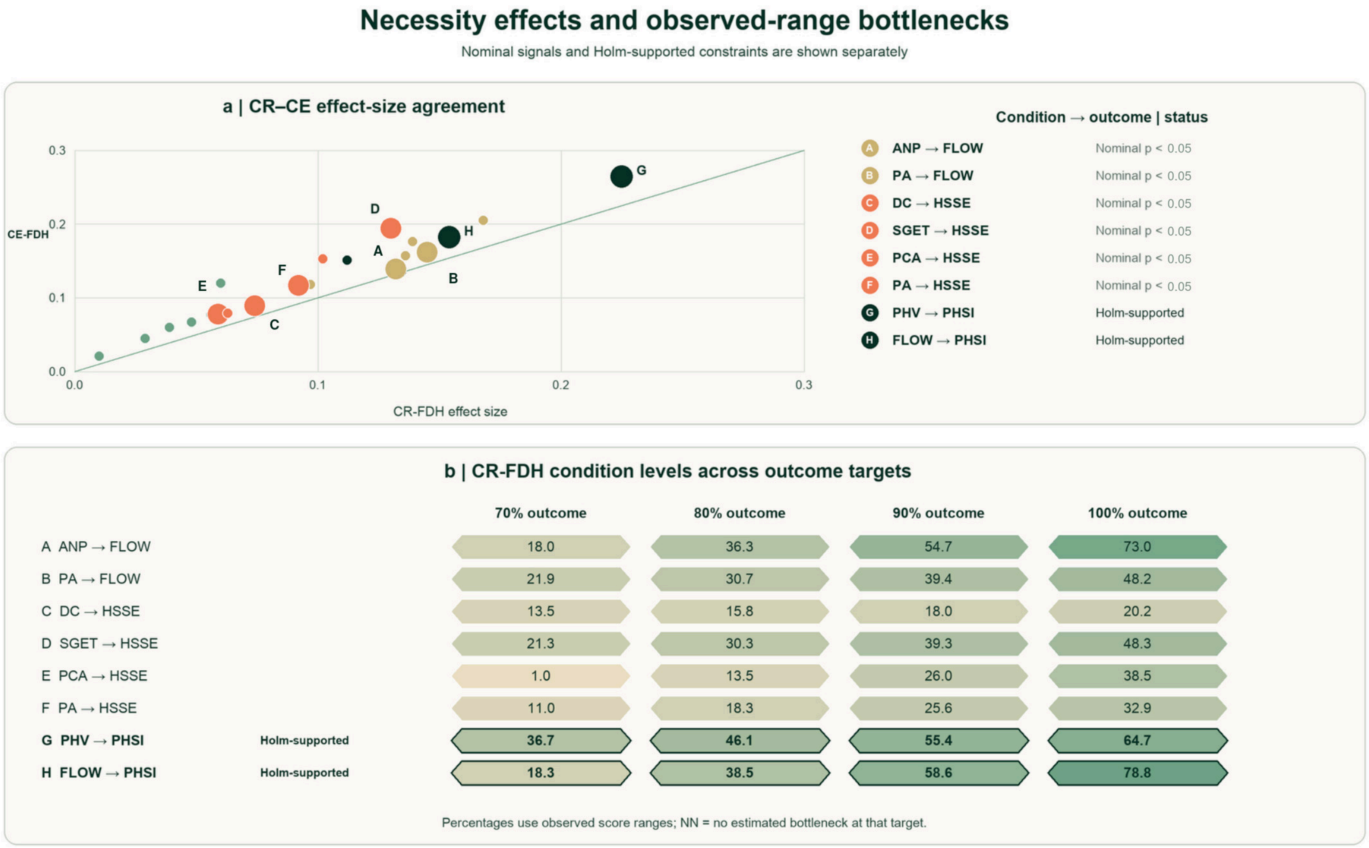


Figure 7. CR-FDH and CE-FDH necessity effects and CR-FDH bottlenecks. Letters identify eight relationships with nominal $p < 0.05$. Dark outlines mark PHV and FLOW for PHSI, which remain significant after Holm adjustment. Bottleneck percentages use observed condition and outcome ranges; NN indicates no estimated bottleneck at the target.

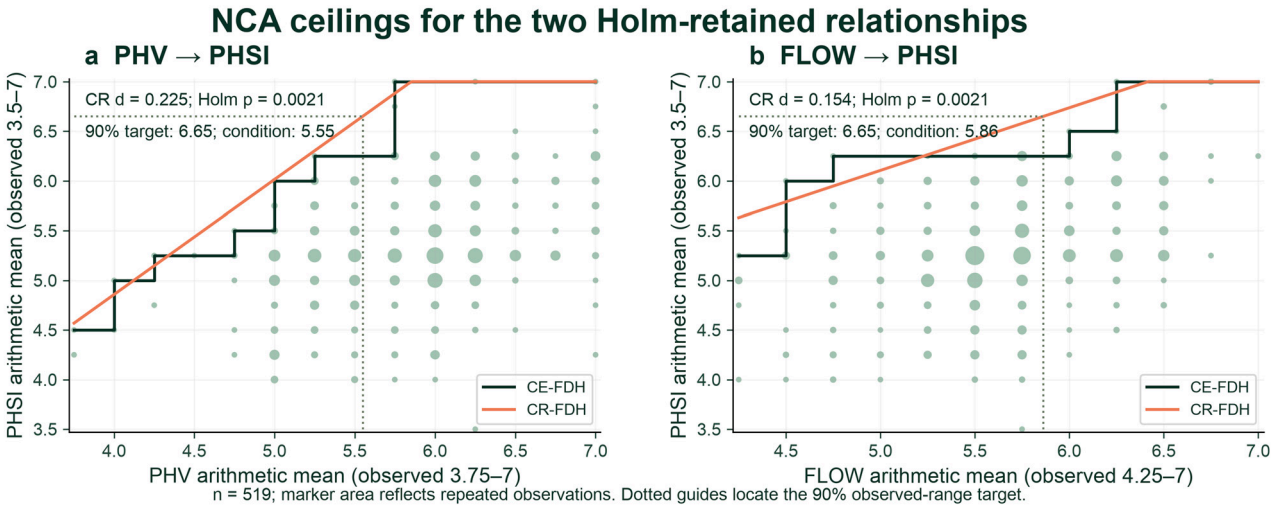


Figure 8. Observed-range NCA scatterplots for PHV → PHSI and FLOW → PHSI, the two relationships retained after Holm adjustment. All 519 respondents are represented; point area encodes repeated score pairs. Solid lines identify CE-FDH and CR-FDH ceilings. Dotted guides locate the 90% PHSI target and its CR-FDH condition level. PHV measures perceived experience-related enhancement of value recognition.

The necessity interpretation follows the item content. PHV bottlenecks concern reported experience-related enhancement of heritage-value recognition, not an absolute level of heritage value. They do not imply that a knowledgeable visitor who reports little additional enhancement lacks stewardship motivation. PHV and FLOW retained empirical ceiling constraints within this sample, while HSSE was positively associated with intention without meeting the necessity criterion. Conditional association, fitted importance, held-out accuracy, and ceiling constraints remain distinct forms of evidence.

5. Discussion

5.1. Main Findings and Interpretation

5.1.1. Extending the S–O–R Framework from Product Evaluation to Stewardship Intention

The complete model provides a differentiated rather than uniformly supported account of post-interaction appraisals. Nineteen of the 24 relationships retained Holm significance, four additional positive paths met only the nominal criterion, and PHV → HSSE was unsupported. The strongest response-stage evidence combines positive associations of PHV, FLOW, and HSSE with intention and sample-specific necessity evidence for PHV and FLOW. This distinguishes complementary motivational appraisals without validating every theoretical connection.

Compared with earlier SOR research on GenAI-enabled museum cultural and creative products [67], the present study examines a different encounter and outcome. The earlier work centered on product novelty, originality, cultural congruence, and purchase intention; here, visitors evaluated guide affordances after interacting with a deployed system, and the outcome was stewardship intention. This extends the use of SOR from product appraisal to museum interpretation. The cross-sectional design supports associations among constructs, not a temporal or causal sequence.

5.1.2. Differentiated Roles of Museum-Guide Affordances

DC has the largest point estimate and mean ANN importance for PHV, while ESP and ANP occupy closely spaced subsequent ANN positions. SGET ranks fourth in the PHV ANN model and is positively associated with PHV in PLS-SEM; these results do not identify it as the dominant PHV input. The profile is consistent with responsive, socially legible, and adaptive interpretation accompanying perceived enhancement of heritage-value recognition. PCA is also positively associated with PHV. Visitor-perceived authenticity remains distinct from independently assessed historical accuracy.

For FLOW, DC and PA are jointly relevant candidates. Their point estimates and fitted ANN ordering differ, but the paired coefficient-difference interval includes zero and each ranks first in three of ten ANN runs. ANP and PA have nominal NCA constraints that do not survive multiplicity adjustment. These results justify evaluating conversational continuity and meaningful visitor control together, rather than declaring one the superior design intervention.

For HSSE, SGET and PCA have closely matched structural coefficients and ANN mean importances. Their paired coefficient contrast gives no evidence of a difference. Both are candidates for capability-oriented design, while FLOW and responsive dialog also remain relevant fitted inputs. PHV's unsupported direct coefficient indicates that perceiving enhanced heritage understanding does not, by itself, establish confidence in feasible action after the other predictors are considered.

5.1.3. From Visitor Experience to Stewardship Intention

PHV, FLOW, and HSSE are positively associated with PHSI, with coefficients of 0.267, 0.249, and 0.215. The mean ANN ordering is the same, but per-run ranks vary and paired

coefficient differences include zero. PHV and FLOW retain necessity evidence under both ceilings and Holm adjustment; HSSE does not. The evidence therefore distinguishes the presence of a positive capability–intention association from an individual necessity constraint, rather than demonstrating that capability is less valuable as an intervention target.

This result is consistent with social cognitive theory, which emphasizes perceived capability as an important basis for action [26,58]. Related heritage-tourism work examines pro-social compliance with health guidelines through norms and anticipated emotions [59]; that context supplies conceptual comparison, not validation of a conservation-stewardship scale. A broader meta-analysis of tourism and hospitality research similarly shows that sustainable behavioral intentions are associated with multiple psychological antecedents, although their relative importance varies across contexts [68]. The present study adds a capability-based perspective by examining whether visitors believe that they can identify heritage-related risks and take feasible action.

PA is positively associated with the organism states at the nominal level and has the largest mean ANN importance for FLOW; its HSSE path does not retain Holm significance. Its nominal necessity effects also remain exploratory. Meaningful visitor choice is therefore a testable candidate for absorption-oriented design, rather than an established causal priority. Subsequent interface studies can compare choices that alter interpretive depth, direction, or evidence while recording visitors' actual use.

5.1.4. How the Complementary Analyses Refine the Findings

ANN and Basis-KAN address different predictive questions. ANN ranks relative input importance, whereas the additive Basis-KAN benchmark tests held-out accuracy against ridge and boosted stumps. Agreement between PLS-SEM and ANN on the same questionnaire does not independently validate a design priority. Ridge achieved a higher mean held-out R^2 than the evaluated Basis-KAN implementation for every outcome. Thus, distinct importance profiles do not establish that a flexible basis model improves predictive accuracy.

Under identical repeated folds, ridge regression achieved the highest mean held-out R-squared for PHV (0.164), FLOW (0.237), HSSE (0.125), and PHSI (0.238), compared with Basis-KAN values of 0.122, 0.175, 0.085, and 0.199. The ten-run ANN rankings remain informative about relative input relevance, but the Basis-KAN comparison favors the simpler linear benchmark for prediction in this sample.

NCA further distinguishes robust response-stage constraints from exploratory affordance-stage patterns. PHV and FLOW retained necessity evidence for PHSI under both ceilings and multiplicity adjustment. ANP and PA for FLOW, and several affordances for HSSE, met only the nominal significance criterion. These ceilings describe sample-specific restrictions on high outcomes.

5.2. Theoretical Contributions

The affordance-based model distinguishes interpretive qualities that a single system-quality score would combine. Dialogic responsiveness had the largest fitted coefficients for perceived heritage-value enhancement and flow; epistemic transparency and cultural authenticity were closely ranked for capability. Paired contrasts do not establish superiority among the leading coefficients. Visitor agency had the highest mean ANN importance for flow. These findings refine the psychological content of the SOR framework by differentiating outcome-specific associations and fitted importance profiles. The unsupported PHV $\rightarrow$ HSSE path limits the proposed organism sequence. The necessary-condition results add robust ceiling constraints at the stewardship-intention stage, while the predictive benchmark gives no evidence that the evaluated additive basis representation improves accuracy over ridge.

The study also extends the focal outcome beyond adoption and continued use. PHSI captures intended responsible visitation, accurate communication, support for safeguarding, and appropriate responses to risk. PHV, FLOW, and HSSE were each associated with intention, while PHV and FLOW also retained necessity constraints. The contribution is an account of stewardship motivation that distinguishes perceived capability from perceived experience-related enhancement of heritage-value recognition and experiential absorption. The data support these relationships within the specified model.

Source-grounded epistemic transparency is the study's most heritage-specific conceptual addition. In this setting, transparency concerns whether visitors can recognize the museum or scholarly basis of a heritage claim, distinguish established information from interpretation, and notice uncertainty. The measure captures visitor perception and does not independently validate factual accuracy or cultural legitimacy. The four analyses separate association, nonlinear importance, held-out predictive accuracy, and necessary constraints.

5.3. Practical Implications

Museums can use the outcome-specific profiles to formulate comparative interface tests. Responsive dialog and visitor control are closely ranked candidates for absorption, while evidential transparency and cultural coherence are closely ranked candidates for perceived capability. PHV, FLOW, and HSSE are each positively associated with stewardship intention; PHV and FLOW also retain sample-specific ceiling constraints. These results support evaluating how interpretation makes heritage significance intelligible, sustains involvement, and connects visitors with feasible actions. They do not establish which single interface change produces the greatest benefit.

Source-grounded epistemic transparency requires usable evidence at the point of interpretation. Museums can attach a concise source card to a consequential factual claim, identify the relevant object or collection record, and offer access to the supporting catalog entry or scholarly source. Established information, curatorial interpretation, and contested accounts should be clearly distinguished. When approved sources do not support an answer, the guide should identify the gap and direct the visitor to an appropriate museum resource. Curators should maintain versioned source records and audit whether generated claims remain linked to their evidence. SGET and PCA have closely matched HSSE coefficients and ANN mean importances. Evidential transparency and cultural coherence should therefore be evaluated jointly in capability-oriented interface tests. Perceived transparency and authenticity should be assessed alongside expert checks of factual accuracy and cultural appropriateness. Altaf's bibliometric analysis places human oversight and accountable deployment within the broader development of human-in-the-loop AI [69]. For museum interpretation, we propose retaining curatorial responsibility for approving sources, reviewing generated claims, and correcting errors.

Conversational design should preserve the visitor's line of inquiry across turns. DC has the largest fitted PHV point estimates and is closely ranked with PA for FLOW. The guide should retain the current object, question, and interpretive context; ask targeted clarification questions when an inquiry is ambiguous; and repair a misunderstanding without restarting the encounter. Visitors should also be able to request the reasoning behind an explanation or compare alternative interpretations. QA can test multi-turn continuity, successful repair, and consistency between a response and its cited evidence. Because ESP ranked third in ANN importance for FLOW, voice, text, and embodied cues should remain synchronized and intelligible. Coherent interaction is a candidate indicated by these sample-specific results; the evaluation should focus on continuity and comprehension.

Personalization should be directed by consequential visitor choices. PA had the highest mean importance in the fitted FLOW ANN profile, while ANP and PA showed exploratory

ceiling constraints for high FLOW. Interfaces can offer topic and depth controls, a clear route back to the museum's core explanation, and the option to compare perspectives supported by approved sources. A visitor's selection should change the content or direction of exploration, and those changes should remain visible. Adaptation should preserve evidence and curatorial boundaries across different narrative versions. The exploratory NCA bottlenecks can guide hypotheses for subsequent interface testing, but the observed percentages do not provide universal deployment thresholds. Usability tests should establish whether the choices are understandable, reversible, and effective for visitors with different prior knowledge and digital skills.

Stewardship prompts should connect heritage meaning to an accessible action. PHV, FLOW, and HSSE were positively associated with PHSI, while PHV and FLOW retained sample-specific necessity evidence. PHV here concerns perceived experience-related enhancement of value recognition. A subsequent interface test can examine an interpretive sequence that explains why an object or practice matters, invites sustained exploration of that significance, and then offers an appropriate next step. Examples include museum-approved advice on responsible visitation, a verified resource for sharing heritage information, an education or volunteer program, or a channel for reporting a specific risk. HSSE's positive relationship with PHSI supports making the requested action concrete: explain what the visitor can do, how to do it, and where institutional support is available. These are design proposals derived from the intention findings; their effects on actual stewardship require behavioral evaluation.

Evaluation should combine interpretive quality with evidence of use and action. Museums can audit claim accuracy and source coverage, assess whether visitors understand and control the guide, and track voluntary uptake of the verified stewardship options. Follow-up evaluation should distinguish stated intention, opening a resource, enrolling in an activity, and completing an action. Interaction records should be retained only with an appropriate purpose, consent, and data minimization. For prediction, ridge is the first benchmark warranted by the present comparison; added model complexity should demonstrate an improvement on independent visitors or another museum before adoption. The modest explained variance and held-out R^2 values favor evaluation at the service level rather than treating a model score as an individual visitor's stewardship propensity.

5.4. Limitations of the Study

The main limitation concerns the research design. The study used cross-sectional self-reports obtained after reported Jinghui use, and no alternative guide or control condition was included. The findings, therefore, describe associations among visitors' perceptions and intentions, but cannot establish that Jinghui caused changes in those responses or performed better than another form of guidance. They also do not establish the temporal ordering implied by the S–O–R framework or provide a formal test of sequential mediation.

The generalizability of the findings is limited by the sampling and research setting. Participants were recruited through convenience sampling from one museum through online and on-site channels. The sample was concentrated among visitors aged 18–40 and participants with a bachelor's degree or above. The findings should therefore not be assumed to represent all museum visitors, particularly those with different levels of education, museum experience, digital literacy, or familiarity with generative AI. Reported prior use was not verified against session records, and the data cannot establish individual exposure duration or timing.

A further limitation concerns the measurement approach. All constructs were measured through one positively worded self-report questionnaire at one time point. Common method variance, social desirability, respondents' immediate affect, and a general positive

evaluation of Jinghui may therefore contribute to the predominantly positive path pattern. A statistical diagnostic can indicate whether method variance is severe under a particular model, but it cannot remove this design limitation. Several constructs were adapted for this context and require replication with independent samples and alternative measurement sources. Perceived cultural authenticity is a visitor judgment of cultural coherence, not evidence of historical accuracy or curatorial legitimacy; those properties require independent expert and community assessment. PHSI measures stated intention, not observed stewardship behavior. PHV records retrospective perceived enhancement and does not establish an objective pre–post change. Future HSSE measurement should use concrete action scenarios and external task criteria to determine whether the present four items adequately capture feasible, responsible action in a museum-AI setting.

Finally, all predictive and necessity findings require independent validation. Ten ANN runs and repeated Basis-KAN folds characterize variation within one sample. The rank tables describe split-and-training variability, not population confidence intervals. Paired PLS coefficient contrasts do not establish superiority among closely ranked inputs. The benchmark evaluates one fixed additive implementation and differs from ANN in its PHSI predictor set, validation design, and outcome scaling. NCA bottlenecks depend on the observed range, ceiling specification, and tie convention; they should not be deployed as universal intervention thresholds.

5.5. Directions for Future Research

For future prospective studies, we propose a standardized protocol to verify participants' use of Jinghui before questionnaire completion. Consenting adults would interact with the system for at least six minutes, completing four question–answer exchanges involving an exhibit query, a follow-up question, an explanation adjustment, and a request for sources or clarification of uncertainty. Access route, device, interaction duration, and task completion would be recorded using anonymous identifiers. Participants would complete the 40-item questionnaire within ten minutes after interaction, with predefined eligibility, duplication, completeness, and attention checks. Future research should establish the causal and behavioral implications of the proposed relationships. Preregistered field experiments could compare Jinghui with a conventional guide, a less embodied interface, or alternative levels of source transparency and personalization. Pre- and post-interaction measures would help clarify temporal changes, while follow-up studies could examine responsible information sharing, participation in educational or volunteer activities, or reporting of heritage-related risks. Where appropriate, behavioral measures could be combined with interaction records collected with participants' informed consent.

The model should also be replicated across museums, heritage types, visitor groups, and cultural settings. Future studies could examine whether cultural authenticity, personalization, social presence, and agency operate differently according to visitors' age, prior knowledge, digital literacy, language, and the cultural sensitivity of the heritage being interpreted. Cross-group measurement invariance and independent assessments by curators or heritage communities would help determine whether the constructs remain stable beyond the present sample.

Future research should validate the ANN rankings, Basis-KAN performance, and NCA bottlenecks in independent museums and preregister architectures, data partitions, hyperparameters, and calibration choices. Cross-site studies can test whether the same affordances remain structurally important, predictively dominant, and necessary across heritage types and visitor populations.

6. Conclusions

This study extends the evaluation of museum-grounded generative-AI guides to heritage stewardship motivation. Analysis of 519 questionnaires from adults reporting prior use linked six perceived guide affordances to perceived heritage-value enhancement, flow, and stewardship self-efficacy, and linked these organism states to stewardship intention. Dialogic contingency had the largest fitted coefficients for value enhancement and flow, whereas epistemic transparency and perceived cultural authenticity were closely ranked for capability. The restored direct value-enhancement–self-efficacy path was unsupported. The findings identify distinct interpretive and experiential profiles within the proposed SOR model.

ANN, Basis-KAN, and NCA clarify the analytical roles of these relationships. ANN mean importance followed the PHV, FLOW, and HSSE ordering for stewardship intention, with rank variation across runs. The matched benchmark favored ridge in mean predictive accuracy over the evaluated additive Basis-KAN implementation. NCA retained constraints for PHV and FLOW in relation to high intention under both ceiling procedures and multiplicity adjustment, while affordance-level bottlenecks remained exploratory. The combined evidence separates conditional association, input importance, held-out accuracy, and sample-specific necessity.

For museum practice, the findings inform tests of traceable heritage claims, sustained contingent dialog, and consequential visitor control within curatorial boundaries. These functions connect interpretation to perceived enhancement of meaning, involvement, and capability associated with stewardship intention. Their implementation should be evaluated through independent checks of interpretive quality and subsequent visitor actions. The study contributes an integrated evaluation of generative-AI interpretation in which cultural accountability and visitor experience are linked to stewardship motivation without equating stated intention with observed protection behavior.

Author Contributions: Conceptualization, Q.W. and Q.Z.; methodology, Q.W. and M.S.; software, formal analysis, data curation, and visualization, Q.W.; validation, Y.Z.; writing, original draft preparation, Q.W.; writing, review, and editing, M.S. and Y.Z.; supervision, Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Humanities and Social Science Fund of the Ministry of Education of China (grant 25XJC760009); the China Postdoctoral Science Foundation (grant 2026M794021); the 2026 National Postdoctoral Researcher Funding Program (grant GZC20261369); the 2025 Shaanxi Provincial Department of Education General Special Scientific Research Programme (grant 25JK0046); and the China National Social Science Fund Art Project (grant 2021AG02652).

Institutional Review Board Statement: Ethical review and approval were waived by the Ethics Review Committee of the School of Design and Arts, Shaanxi University of Science and Technology, because the study involved an anonymous, non-interventional questionnaire administered to adult participants and collected no sensitive or personally identifiable information (date of exemption: 14 April 2026). The study was conducted in accordance with the Declaration of Helsinki and applicable institutional requirements.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study. Before beginning the questionnaire, participants were informed of the purpose of the study, the voluntary nature of participation, and the anonymized use of their responses. Completion and submission of the questionnaire constituted informed consent to participate.

Data Availability Statement: The data and supporting materials are available from the first author or the corresponding author upon reasonable request.

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

References

1. UNESCO. *Recommendation on the Ethics of Artificial Intelligence*; UNESCO: Paris, France, 2021. Available online: https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence (accessed on 4 October 2026).
2. Tiribelli, S.; Pansoni, S.; Frontoni, E.; Giovanola, B. Ethics of artificial intelligence for cultural heritage: Opportunities and challenges. *IEEE Trans. Technol. Soc.* **2024**, *5*, 293–305. [CrossRef]
3. Colace, F.; Gaeta, R.; Lorusso, A.; Pellegrino, M.; Santaniello, D. New AI challenges for cultural heritage protection: A general overview. *J. Cult. Herit.* **2025**, *75*, 168–193. [CrossRef]
4. Li, B.; Lim, Y.L.; Li, W. Systematic review: A bibliometric analysis of building technology and its potential applications to artificial intelligence in the field of cultural heritage conservation from 2013 to 2023. *J. Asian Archit. Build. Eng.* **2026**, *25*, 1021–1043. [CrossRef]
5. Kutlu, İ. Scientific mapping of artificial intelligence (AI) assisted applications in historical building conservation. *J. Asian Archit. Build. Eng.* **2026**, *25*, 1258–1278. [CrossRef]
6. Tan, X.; Xie, H.; Xia, C.; Ying, Y.; Hokoi, S.; Li, Y. Artificial intelligence in five key phases of architectural heritage conservation: A systematic review. *J. Build. Eng.* **2026**, *117*, 114862. [CrossRef]
7. Xu, J.; Yan, L.; Zhang, R.; Zhou, M. A review of the development and application of generative technology in digital museums. *npj Herit. Sci.* **2025**, *13*, 589. [CrossRef]
8. Park, S.W.; Park, N.H.; Park, J.J.; Jun, H.J. A systematic and procedural framework for digital reconstruction of architectural cultural heritage using generative AI. *npj Herit. Sci.* **2026**, *14*, 549. [CrossRef]
9. Hao, X.; Xu, J.; Wang, Y. How generative AI shapes user perceived value and adoption intention in digital museum experiences. *npj Herit. Sci.* **2025**, *13*, 608. [CrossRef]
10. Mo, J.; Chen, H.; Ye, C.; Wang, Z.; Chen, C. Exploring the drivers of users' adoption of museum digital humans. *npj Herit. Sci.* **2026**, *14*, 43. [CrossRef]
11. Li, J. The impact of hyperrealistic virtual humans on cultural heritage dissemination. *npj Herit. Sci.* **2025**, *13*, 515. [CrossRef]
12. Ariya, P.; Khanchai, S.; Intawong, K.; Puritat, K. Enhancing textile heritage engagement through generative AI-based virtual assistants in virtual reality museums. *Comput. Educ. X Real.* **2025**, *7*, 100112. [CrossRef]
13. Ariya, P.; Worragin, P.; Khanchai, S.; Poollapalin, D.; Julrode, P. Voice, text, or embodied AI avatar? Effects of generative AI interface modalities in VR museums. *Informatics* **2026**, *13*, 42. [CrossRef]
14. Sun, X.; Wang, F.; Jin, W. Continuance intention of cultural museum virtual human based on PLS-SEM analysis of MRT and UGT. *npj Herit. Sci.* **2026**, *14*, 139. [CrossRef]
15. Lai, S.; Tian, Y.; Zhang, Q. The impact of AI-generated technologies-driven digital cultural heritage platforms on users' offline cultural participation intentions. *npj Herit. Sci.* **2025**, *13*, 574. [CrossRef]
16. Kong, X.; Liu, Y.; Shi, Y.; Xu, J.; Liu, M. Mechanism of public behavioral intention to use generative AI for folk story image co-creation. *npj Herit. Sci.* **2026**, *14*, 164. [CrossRef]
17. Xie, J.; Li, H.; Furuya, K.; Chen, J.; Luo, S. Participatory intention and behavior in green cultural heritage conservation: An application of the extended theory of planned behavior. *Herit. Sci.* **2024**, *12*, 299. [CrossRef]
18. Hansson, P.; Öhman, J. Museum education and sustainable development: A public pedagogy. *Eur. Educ. Res. J.* **2022**, *21*, 469–483. [CrossRef]
19. Volkoff, O.; Strong, D.M. Critical realism and affordances: Theorizing IT-associated organizational change processes. *MIS Q.* **2013**, *37*, 819–834. [CrossRef]
20. Vasic, I.; Fill, H.-G.; Quattrini, R.; Pierdicca, R. Knowledge graphs vs. large language models: Competitors or partners in supporting virtual museums. *ACM J. Comput. Cult. Herit.* **2025**, *18*, 61. [CrossRef]
21. Frontop Digital Creative. Capital Museum High-Precision AI Smart-Guidance System Selected as a Benchmark Case. 13 May 2025. Available online: https://www.efrontop.cn/news-detail/2000790763716358144.html (accessed on 8 October 2026).
22. Yang, Y.T. Digital Creativity Empowers Multiple Industries: Frontop's Integrated Digital-Intelligence Solutions. *Economic Daily*, 15 April 2025. Available online: https://static.jingjiribao.cn/static/jjrbrss/rsshtml/20250415/580972.html (accessed on 8 October 2026).
23. Mehrabian, A.; Russell, J.A. *An Approach to Environmental Psychology*; MIT Press: Cambridge, MA, USA, 1974.
24. Kolar, T.; Zabkar, V. A consumer-based model of authenticity: An oxymoron or the foundation of cultural heritage marketing? *Tour. Manag.* **2010**, *31*, 652–664. [CrossRef]
25. Deci, E.L.; Ryan, R.M. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychol. Inq.* **2000**, *11*, 227–268. [CrossRef]
26. Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1986.
27. Soini, K.; Birkeland, I. Exploring the scientific discourse on cultural sustainability. *Geoforum* **2014**, *51*, 213–223. [CrossRef]
28. Sundar, S.S.; Bellur, S.; Oh, J.; Jia, H.; Kim, H.-S. Theoretical importance of contingency in human–computer interaction: Effects of message interactivity on user engagement. *Commun. Res.* **2016**, *43*, 595–625. [CrossRef]

29. Lew, Z.; Walther, J.B.; Pang, A.; Shin, W. Interactivity in online chat: Conversational contingency and response latency in computer-mediated communication. *J. Comput.-Mediat. Commun.* **2018**, *23*, 201–221. [CrossRef]
30. Cao, B.; Wen, C.; Zhang, R. How conversational contingency shapes trust in human–AI interaction: Evidence from a two-wave mediation study. *J. Broadcast. Electron. Media* **2025**, *69*, 398–417. [CrossRef]
31. McCarthy, J.; Ciolfi, L. Place as dialogue: Understanding and supporting the museum experience. *Int. J. Herit. Stud.* **2008**, *14*, 247–267. [CrossRef]
32. Shin, D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *Int. J. Hum.-Comput. Stud.* **2021**, *146*, 102551. [CrossRef]
33. Grimmelikhuijsen, S. Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Adm. Rev.* **2023**, *83*, 241–262. [CrossRef]
34. Wang, P.; Ding, H. The rationality of explanation or human capacity? Understanding the impact of explainable artificial intelligence on human–AI trust and decision performance. *Inf. Process. Manag.* **2024**, *61*, 103732. [CrossRef]
35. Weng, L.; Liang, Z.; Bao, J. The effect of tour interpretation on perceived heritage values: A comparison of tourists with and without tour guiding interpretation at a heritage destination. *J. Destin. Mark. Manag.* **2020**, *16*, 100431. [CrossRef]
36. Epley, N.; Waytz, A.; Cacioppo, J.T. On seeing human: A three-factor theory of anthropomorphism. *Psychol. Rev.* **2007**, *114*, 864–886. [CrossRef]
37. Liu, B.; Wei, L. Machine gaze in online behavioral targeting: The effects of algorithmic human likeness on social presence and social influence. *Comput. Hum. Behav.* **2021**, *124*, 106926. [CrossRef]
38. Seyama, J.; Nagayama, R.S. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence Teleoperators Virtual Environ.* **2007**, *16*, 337–351. [CrossRef]
39. Liu, F.; Ren, L.; Hu, Y. Exploring how AI tour guides' automated social presence influences tourist flow experience and behavioral intentions. *Int. J. Tour. Res.* **2026**, *28*, e70326. [CrossRef]
40. Shen, Y.; Du, J.; Liu, Y.; Peng, J. The effect of social presence on user's participation willingness under tourism live streaming. *Int. J. Tour. Res.* **2024**, *26*, e2777. [CrossRef]
41. Bryce, D.; Curran, R.; O'Gorman, K.; Taheri, B. Visitors' engagement and authenticity: Japanese heritage consumption. *Tour. Manag.* **2015**, *46*, 571–581. [CrossRef]
42. Atzeni, M.; Del Chiappa, G.; Mei Pung, J. Enhancing visit intention in heritage tourism: The role of object-based and existential authenticity in non-immersive virtual reality heritage experiences. *Int. J. Tour. Res.* **2022**, *24*, 240–255. [CrossRef]
43. Zhang, S.-N.; Li, Y.-Q.; Liu, C.-H.; Ruan, W.-Q. How does authenticity enhance flow experience through perceived value and involvement: The moderating roles of innovation and cultural identity. *J. Travel Tour. Mark.* **2019**, *36*, 710–728. [CrossRef]
44. Not, E.; Petrelli, D. Blending customisation, context-awareness and adaptivity for personalised tangible interaction in cultural heritage. *Int. J. Hum.–Comput. Stud.* **2018**, *114*, 3–19. [CrossRef]
45. Not, E.; Petrelli, D. Empowering cultural heritage professionals with tools for authoring and deploying personalised visitor experiences. *User Model. User-Adapt. Interact.* **2019**, *29*, 67–120. [CrossRef]
46. Araujo, T.; Bol, N. From speaking like a person to being personal: The effects of personalized, regular interactions with conversational agents. *Comput. Hum. Behav. Artif. Hum.* **2024**, *2*, 100030. [CrossRef]
47. Ferguson, C.; van den Broek, E.L.; van Oostendorp, H. AI-induced guidance: Preserving the optimal zone of proximal development. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100089. [CrossRef]
48. Novak, T.P.; Hoffman, D.L.; Yung, Y.-F. Measuring the customer experience in online environments: A structural modeling approach. *Mark. Sci.* **2000**, *19*, 22–42. [CrossRef]
49. Letourneau, S.M.; Schloss, D.; Perez, S.; Aucapina, F.; Baburyan, M.; Curto, K.; Gomez, E.; McDonald, T.; Multani, S.; Taylor, T.; et al. What do we mean by "agency"? A framework and tools for supporting visitors' agency in museums and science centers. *Curator Mus. J.* **2025**, *68*, 417–433. [CrossRef]
50. Brizi, A.; Rabinovich, A.; Lewis, C. Psychological outcomes of local heritage engagement: Participation in community archeological excavations increases well-being, self-efficacy, and perceived community support. *J. Appl. Soc. Psychol.* **2023**, *53*, 850–861. [CrossRef]
51. Antón, C.; Camarero, C.; Garrido, M.-J. Exploring the experience value of museum visitors as a co-creation process. *Curr. Issues Tour.* **2018**, *21*, 1406–1425. [CrossRef]
52. Lunardo, R.; Ponsignon, F. Achieving immersion in the tourism experience: The role of autonomy, temporal dissociation, and reactance. *J. Travel Res.* **2020**, *59*, 1151–1167. [CrossRef]
53. Vega, E.; Camarero, C.; Hamari, J.; Xi, N. What leads to better immersive experiences? The role of guidance and mobility in virtual reality museums. *Int. J. Hum.–Comput. Stud.* **2026**, *209*, 103739. [CrossRef]
54. Kim, M.; Thapa, B. Perceived value and flow experience: Application in a nature-based tourism context. *J. Destin. Mark. Manag.* **2018**, *8*, 373–384. [CrossRef]

55. Wang, H.; Tong, Y.; Zeng, H. Driving tourists' pro-environmental behaviour through heritage interpretation messages. *npj Herit. Sci.* **2025**, *13*, 182. [CrossRef]
56. Jia, F.; Meng, J.; Ma, Y.; Mao, Y. Flow experience and self-efficacy in undergraduates' English learning: A weekly diary investigation with cross-lagged panel modeling approach. *System* **2024**, *123*, 103312. [CrossRef]
57. Han, G.-S. Relationships between outdoor recreation-associated flow, pro-environmental attitude, and pro-environmental behavioral intention. *Sustainability* **2023**, *15*, 10581. [CrossRef]
58. Bandura, A. Guide for constructing self-efficacy scales. In *Self-Efficacy Beliefs of Adolescents*; Pajares, F., Urdan, T., Eds.; Information Age Publishing: Greenwich, CT, USA, 2006; pp. 307–337.
59. Zhu, P.; Chi, X.; Ryu, H.B.; Ariza-Montes, A.; Han, H. Traveler pro-social behaviors at heritage tourism sites. *Front. Psychol.* **2022**, *13*, 901530. [CrossRef]
60. Brislin, R.W. Back-translation for cross-cultural research. *J. Cross-Cult. Psychol.* **1970**, *1*, 185–216. [CrossRef]
61. Boateng, G.O.; Neilands, T.B.; Frongillo, E.A.; Melgar-Quiñonez, H.R.; Young, S.L. Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front. Public Health* **2018**, *6*, 149. [CrossRef]
62. SmartPLS. PLS-SEM Algorithm. Available online: https://www.smartpls.com/documentation/algorithms-and-techniques/core-algorithm/pls/ (accessed on 16 October 2026).
63. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson: Upper Saddle River, NJ, USA, 2009.
64. Olden, J.D.; Joy, M.K.; Death, R.G. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecol. Model.* **2004**, *178*, 389–397. [CrossRef]
65. Dul, J. Necessary condition analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organ. Res. Methods* **2016**, *19*, 10–52. [CrossRef]
66. Dul, J.; van der Laan, E.; Kuik, R. A statistical significance test for necessary condition analysis. *Organ. Res. Methods* **2020**, *23*, 385–395. [CrossRef]
67. Shi, M.; Guo, Q.; Li, H.; Lee, K. Understanding purchase intention for genAI-enabled museum cultural and creative products using a SOR model. *Sci. Rep.* **2026**, *16*, 5858. [CrossRef]
68. Lin, M.-T.; Zhu, D.; Liu, C.; Kim, P.B. A meta-analysis of antecedents of pro-environmental behavioral intention of tourists and hospitality consumers. *Tour. Manag.* **2022**, *93*, 104566. [CrossRef]
69. Altaf, A. Mapping the evolution of human-in-the-loop artificial intelligence: A bibliometric analysis. *J. Sch. Res. Metr.* **2026**, *1*, 27–44. [CrossRef]

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.