

## Article

# Designing an Accountable Generative AI Co-Creation System for Museum Visitors: A Field Study at the Lanting Calligraphy Museum

Sheng Jin <sup>1</sup>, Min Fan <sup>2,\*</sup> and Zhiyun Lei <sup>3,\*</sup><sup>1</sup> School of Animation and Digital Arts, Communication University of Zhejiang, Hangzhou 310018, China; sjin@cuz.edu.cn<sup>2</sup> School of Animation and Digital Arts, Communication University of China, Beijing 100024, China<sup>3</sup> Faculty of Humanities and Arts, Macau University of Science and Technology, Macao 999078, China

\* Correspondence: mfan@cuc.edu.cn (M.F.); 3220005350@student.must.edu.mo (Z.L.)

## Abstract

Heritage museums increasingly seek to support visitors as active meaning-makers during the visit and to extend that experience beyond the museum through personally meaningful takeaways. Generative AI (GenAI) may support this goal by helping visitors transform museum source material into personalized artifacts, but open-ended generation may weaken heritage linkage and obscure curatorial responsibility. This paper proposes two paired design principles for heritage GenAI: bounded creativity, which defines what visitors may change, and curatorial accountability, which defines how cultural limits are set, explained, and revised. We implemented these principles at the Lanting Calligraphy Museum in Shaoxing, China, through a visit-linked co-creation system. Visitors collected source images during the museum tour and later created a personalized 3D-printed takeaway through either a rule-based editor with predefined templates and options or a GenAI-mediated dialogue interface. A between-subjects field study with adult visitors ( $N = 60$ ; 20 per condition) compared conventional consumption (a standard tour and souvenir), rule-based co-creation, and GenAI-mediated co-creation. Both co-creation pathways improved cultural engagement, perceived connection to heritage sources, personal meaning and self-expression, and psychological ownership compared with conventional consumption. The GenAI-mediated pathway produced further gains in cultural engagement, personal meaning, self-expression, and psychological ownership. Still, it did not reliably increase perceived connection to heritage sources beyond that of the rule-based pathway. These findings suggest that the GenAI pathway added value mainly on personal authenticity and psychological ownership, by helping visitors articulate and revise intentions within accountable cultural boundaries; its incremental effect on heritage-linked authenticity remains undetermined at this sample size.

**Keywords:** generative AI; cultural heritage; bounded creativity; curatorial accountability; human–AI co-creation; cultural boundaries; authenticity; psychological ownership; museum experience

Academic Editor: George Pavlidis

Received: 26 May 2026

Revised: 19 July 2026

Accepted: 23 July 2026

Published: 30 July 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

Heritage museums increasingly seek to support visitors as active meaning-makers during the visit and to extend that experience beyond the museum through personally meaningful takeaways [1–3]. Interactive museum technologies, including mobile guides, augmented or virtual reality, kiosks, and material-digital interfaces, have made on-site interpretation more situated and participatory [4–12]. The connection between on-site attention and the post-visit takeaway has received less direct attention. Visitors most often leave with a generic souvenir that retains little trace of the specific objects or choices that shaped their encounter. The design problem is therefore not only how to enrich on-site interaction, but also how to sustain a traceable relation between the museum encounter and the object that remains afterward.

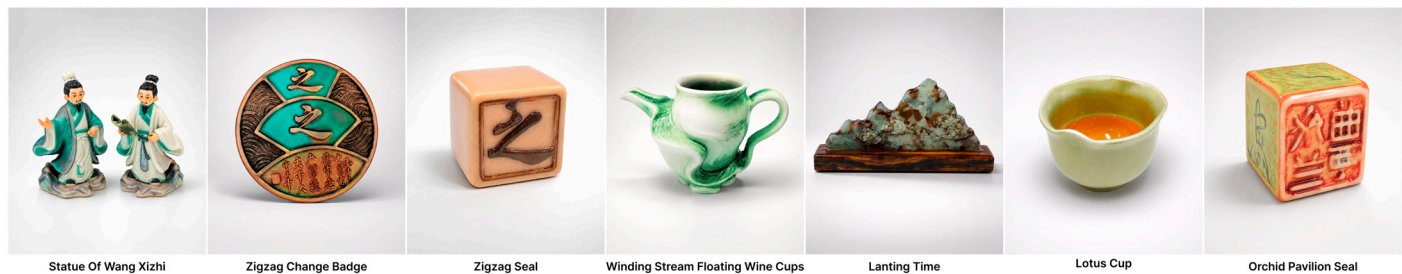
Generative AI (GenAI) may extend this design space by making museum systems productive as well as interpretive. Large language models, image generators, multimodal agents, and text-to-3D systems may help a visitor select a calligraphic detail, motif, or exhibit photograph and use ordinary language to develop it into a personalized output [13–19]. Open-ended generation also raises heritage-specific concerns. Generated outputs may weaken perceived connection to source material, distort symbolic meaning, or shift attention away from the cultural references that support museum learning [13,14,19–21]. The rules that shape acceptable outputs are often hidden inside model defaults, prompts, datasets, or interface logic, making it difficult to identify who defined a cultural boundary and why [22–27].

Recent heritage-GenAI work has begun to address parts of this space. Studies have explored AI museum guides, chatbot-based artwork engagement, automated heritage content generation, workshop-based co-creation of heritage narratives [28], generative re-interpretation of heritage records [29], and broader ethical questions around generative systems in heritage practice [13,14,17–21]. Most of this work addresses one part of the experience at a time, such as on-site explanation, guidance, or generated illustration. Less attention has been paid to visit-linked digital–physical co-creation, in which visitors collect source material during the museum visit, author a personalized output after the tour, receive cultural-boundary feedback, and obtain a physical takeaway that remains connected to that source. Heritage GenAI of this form requires a design structure that supports visitor authorship while keeping source linkage and curatorial responsibility visible.

This paper addresses these gaps through two research questions. RQ1 asks how visitor authorship and curatorial responsibility may be operationalized in a visit-linked digital–physical co-creation system for heritage museums. RQ2 asks whether source-linked co-creation improves visitors’ heritage and personal experience compared with conventional consumption, and whether a GenAI-mediated authoring pathway adds value beyond a rule-based co-creation pathway.

To address these questions, we propose two paired design principles for heritage GenAI. Bounded creativity defines what visitors may change in a heritage transformation. Curatorial accountability defines how cultural limits are set, explained, and revised. Together, the two principles answer RQ1 by operationalizing visitor authorship and curatorial responsibility within a visit-linked digital–physical co-creation system. Section 3 develops both principles in detail. We implemented them at the Lanting Calligraphy Museum in Shaoxing, China. The site is associated with Wang Xizhi’s Preface to the Orchid Pavilion Gathering (Lantingji Xu, 353 CE), a central work in the history of Chinese running-script calligraphy [30–33]. During the visit, visitors collect calligraphic strokes, motifs, and exhibit details with a mobile companion program. After the tour, a tablet workstation transforms selected source images into personalized designs through one of two authoring pathways. The rule-based pathway provides curatorially authored templates and options. The GenAI-mediated pathway translates natural-language requests into

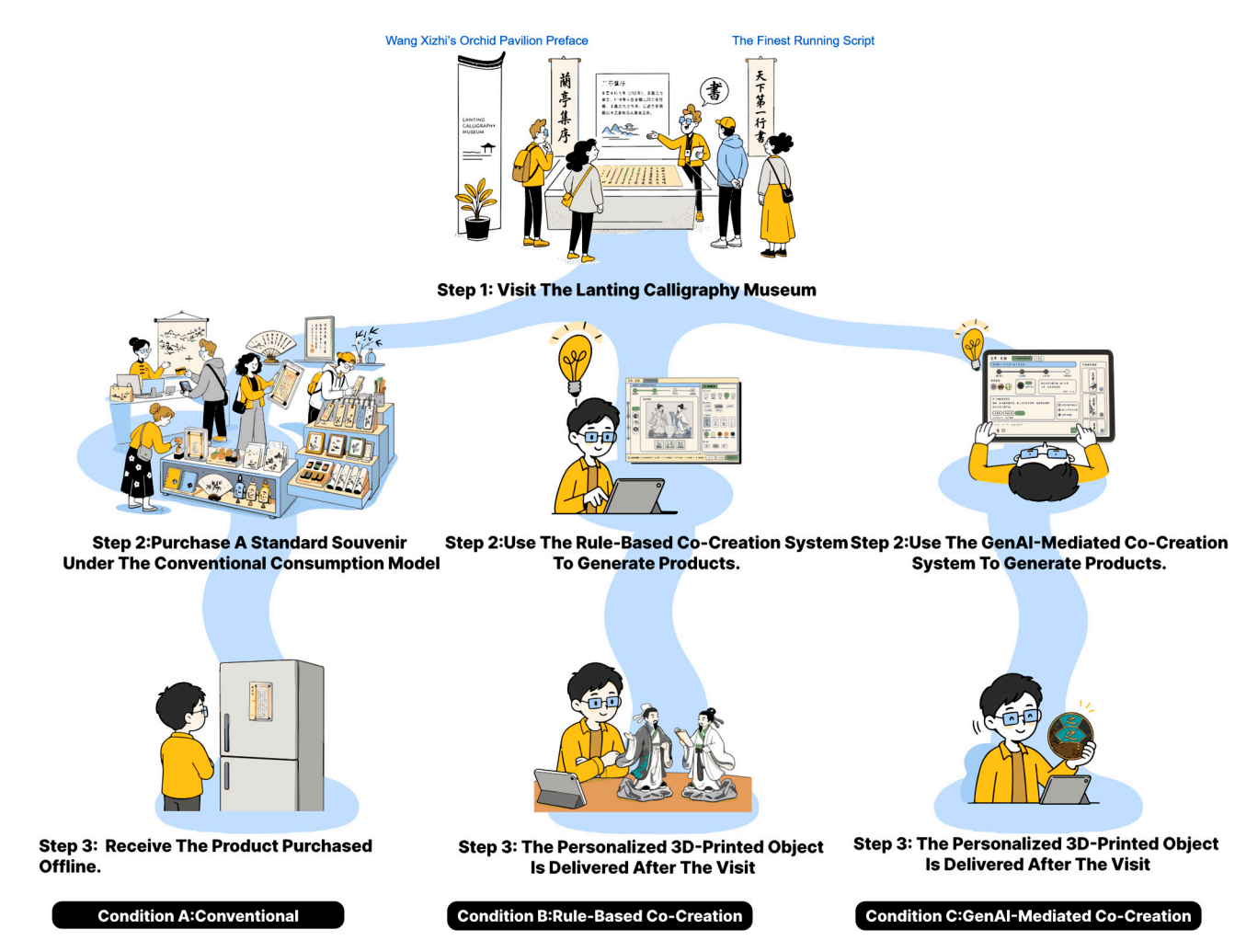
design operations under documented cultural boundaries. Approved designs are checked for manufacturability before confirmation. Visitors receive a printed provenance card showing the selected source and the approved design, and the personalized 3D-printed object is delivered after the visit (see Figure 1).



**Figure 1.** Examples of personalized 3D-printed takeaways from the field deployment at the Lanting Calligraphy Museum. Each object began from calligraphic source material that visitors collected during the museum visit and was authored under documented cultural boundaries through either the rule-based or GenAI-mediated pathway.

We evaluated the system in a between-subjects field study with adult visitors ( $N = 60$ ; 20 per condition). Participants were assigned to one of three conditions (see Figure 2). In conventional consumption, visitors completed the standard guided tour and received a standard visit-related souvenir. In rule-based co-creation, visitors collected source material during the visit and used predefined templates and options to author their object after the tour. In GenAI-mediated co-creation, visitors used the same source-linked structure but expressed and revised design intentions through dialogue. The rule-based condition served as a matched active control, allowing the study to separate the effect of source-linked co-creation in general from the additional effect of GenAI-mediated dialogue. Both co-creation pathways produced higher cultural engagement, perceived connection to heritage sources, personal meaning and self-expression, and psychological ownership than conventional consumption. The GenAI-mediated pathway produced further gains in cultural engagement, personal meaning and self-expression, and psychological ownership, but did not reliably increase perceived connection to heritage sources beyond the rule-based pathway.

This paper offers three contributions. First, it proposes bounded creativity and curatorial accountability as paired design-and-governance principles for heritage GenAI. The pair links creative freedom to explicit, reviewable cultural boundaries. Second, it presents a visit-linked digital–physical co-creation system deployed in a real museum. The system connects on-site source collection, post-tour authoring, culturally bounded transformation, fabrication-aware validation, provenance documentation, and post-visit delivery of a personalized 3D-printed takeaway. Third, it provides a field study with a matched rule-based active control, which separates the broader effects of source-linked co-creation from the additional effects of GenAI-mediated dialogue. Taken together, the findings locate the observed value of heritage GenAI in accountable, source-linked co-creation that connects a visitor’s museum encounter, authoring process, cultural boundaries, and final physical takeaway. The gains over the rule-based pathway were concentrated in the personal outcomes. Whether unrestricted generation would produce different outcomes was not tested, and the null result for Heritage-Linked Authenticity cannot separate no effect from a smaller effect that this sample could not detect (Section 5.5).



**Figure 2.** Process flows for the three between-subjects conditions. Conventional consumption shows a standard guided tour followed by receipt of a standard visit-related souvenir. Both co-creation conditions add an on-site source image collection step and an after-tour authoring step that produces a personalized 3D-printed takeaway. The two co-creation conditions differ in whether the authoring step uses predefined templates and options or natural-language dialogue under documented cultural boundaries.

## 2. Related Work

### 2.1. Heritage Experience: Meaning-Making, Authenticity, Ownership, and Authority

Research on heritage experience has been shaped by several conceptual traditions that are directly relevant to visitor-facing museum technologies. Four are especially important for the present study: constructivist museum learning, authenticity in heritage experience, psychological ownership, and the politics of heritage interpretation. Together, these traditions suggest that heritage systems should not be evaluated only by information quality or interface usability. They should also be understood in terms of how they support visitor meaning-making, sustain perceived authenticity, foster personal connection and ownership, and distribute interpretive authority.

Constructivist museum learning provides the first foundation. Hein [1] argues that museum visitors actively construct meaning rather than passively absorb fixed knowledge, while Falk and Dierking [2] show that museum learning is shaped by the interaction of personal, sociocultural, and physical contexts. Visitor-experience research further treats the museum encounter as a multifaceted process involving attention, emotion, prior knowledge, social interaction, and environmental context [3]. This body of work

suggests that museum technologies should not only deliver pre-packaged interpretation. They may also create conditions in which visitors connect institutional narratives with their own observations, choices, and experiences.

A second foundation is the authenticity literature in tourism and heritage. Wang [34] distinguishes among objective, constructive, and existential authenticity, shifting attention from authenticity as a property of objects alone to authenticity as a visitor experience. Later work has further shown that perceived authenticity in heritage and tourism is shaped by both the perceived connection to cultural sources and the visitor's own sense of meaningful involvement, emotional fit, and self-related experience [35–37]. This distinction is important for visitor-facing co-creation systems. A personalized heritage object may be experienced as culturally grounded, personally meaningful, or both. Conversely, a visitor may perceive an object as strongly connected to heritage sources but emotionally distant, or as personally expressive while only loosely connected to the source tradition. This literature therefore motivates treating perceived heritage connection and personal meaning as related but distinct dimensions of visitor experience.

Psychological ownership provides a third foundation. Psychological ownership refers to the feeling that something is “mine” even without legal ownership [38,39]. Prior work suggests that this feeling can emerge through several routes, including investment of effort, perceived control, intimate knowledge of the target, and self-association [38–40]. These mechanisms are especially relevant to co-creation because a visitor-designed object may not only be received as a museum souvenir, but also experienced as the outcome of one's own choices, revisions, and attention. Psychological ownership therefore offers a useful lens for understanding why a personalized takeaway may differ from a standard museum gift, and why different authoring pathways may produce different levels of felt attachment.

The fourth foundation concerns authority. Critical heritage studies show that heritage is not simply inherited or transmitted; it is selected, interpreted, authorized, and governed. Smith's critique of Authorized Heritage Discourse shows how heritage value is often stabilized through institutional expertise [25]. Simon's participatory museum framework argues for more active public contribution to museum meaning-making [26], while broader critical heritage scholarship emphasizes that heritage interpretation is shaped by social, institutional, and political processes [27]. Generative AI complicates this relationship. When interpretation or creative transformation is mediated by models, authority does not disappear. It may be redistributed into training data, prompts, interface vocabulary, ranking logic, and default constraints. Visitor-facing heritage AI therefore raises not only a participation question, but also an accountability question: who defines cultural limits, how those limits are documented, how they are explained to visitors, and how they may be revised.

Together, these traditions inform the conceptual framework developed in Section 3 and clarify why the field study examines both heritage-related and personally meaningful aspects of visitor experience.

## *2.2. Interactive Technologies in Museums: From Information Delivery to Visit-Linked Digital–Physical Co-Creation*

Research on interactive museum technology has expanded how visitors access, explore, and engage with heritage material. Mobile guides and augmented-reality tours provide contextual information in the gallery [7,8], virtual and mixed-reality systems support immersive access to otherwise inaccessible sites or past environments [7], and material-digital interaction has explored tangible, personalized, and context-aware forms of visitor engagement [9–12]. This work has moved museum interaction beyond the static display-and-label model by making interpretation more situated, multimodal, and participatory.

A related body of work highlights the role of tangible objects in extending museum and tourism experiences beyond the moment of the visit. Data-souvenir research shows that material objects can encode traces of an individual museum visit and support later reflection across physical and digital settings [4]. Work on tangible and augmented museum artifacts similarly suggests that physical objects can carry stories, visitor activity, and social memory beyond the gallery space [9–12]. Tourism research further shows that souvenirs operate through locality, memory, display, and everyday re-encounter, rather than through purchase alone [5,6]. Across these studies, the post-visit object is not simply a commodity. It can become a material anchor for remembering, displaying, discussing, and reinterpreting the experience after the visit.

These strands of work point toward a digital–physical understanding of museum experience. On-site technologies can guide attention and enrich interpretation, while tangible objects can sustain memory and affect after the museum exit. However, these elements are often treated separately. Mobile, AR, and VR systems mainly enrich the on-site encounter; tangible and souvenir systems often support post-visit continuity but provide limited visitor authorship or limited traceability to specific source materials encountered during the visit. The design opportunity is therefore not only to personalize a souvenir, but to connect on-site attention, visitor authoring, and post-visit materialization into a more continuous process.

Generative AI may extend this design space by making museum systems productive rather than only interpretive. Recent heritage-GenAI work has explored large-language-model tour guidance, AI chatbots for artwork engagement, automated heritage content generation, and broader ethical questions around generative systems in heritage practice [13,14,17–21]. These studies show that GenAI can support explanation, access, dialogue, and narrative interaction. At the same time, many such systems still focus on interpretation or generated content as an on-screen experience. Less attention has been paid to how generative systems might support material takeaways that remain traceable to what a visitor noticed, selected, and transformed during a particular museum visit.

This paper uses the term visit-linked digital–physical co-creation to describe such a continuous process. Visitors collect source material in the gallery, transform it through a digital authoring system, and receive a physical object that remains connected to the museum encounter.

### *2.3. Generative AI as a Co-Creative Partner: Heritage-Specific Tensions*

The digital–physical perspective above raises a further question: how should authorship be organized when generative systems are used to transform heritage source material into personalized outputs? HCI research increasingly frames generative models as co-creative partners rather than as automatic content-generation tools [15,16]. From this perspective, the central design issue is not only what the model can generate, but how initiative, communication, and contribution are distributed between human and system. Frameworks such as COFI are useful here because they treat co-creation as an interactional process involving turn-taking, feedback, shared control, and evolving contribution [15].

This perspective is especially relevant in museum contexts, where many visitors are domain novices rather than trained designers or heritage specialists. Natural-language interaction may lower the barrier to participation by allowing visitors to express aesthetic intentions, personal meanings, or revision requests in ordinary terms. Dialogue may also support iterative authoring: the system can ask for clarification, generate alternatives, respond to feedback, and help visitors refine an output without requiring technical design knowledge. In this sense, GenAI may support not only personalization, but also a stronger feeling that the visitor has actively shaped the final artifact. This emphasis on guided rather than unrestricted authoring is also consistent with creativity-support and constraint

research, which suggests that constraints may support novice creativity when they structure rather than replace decision-making [16,41–43].

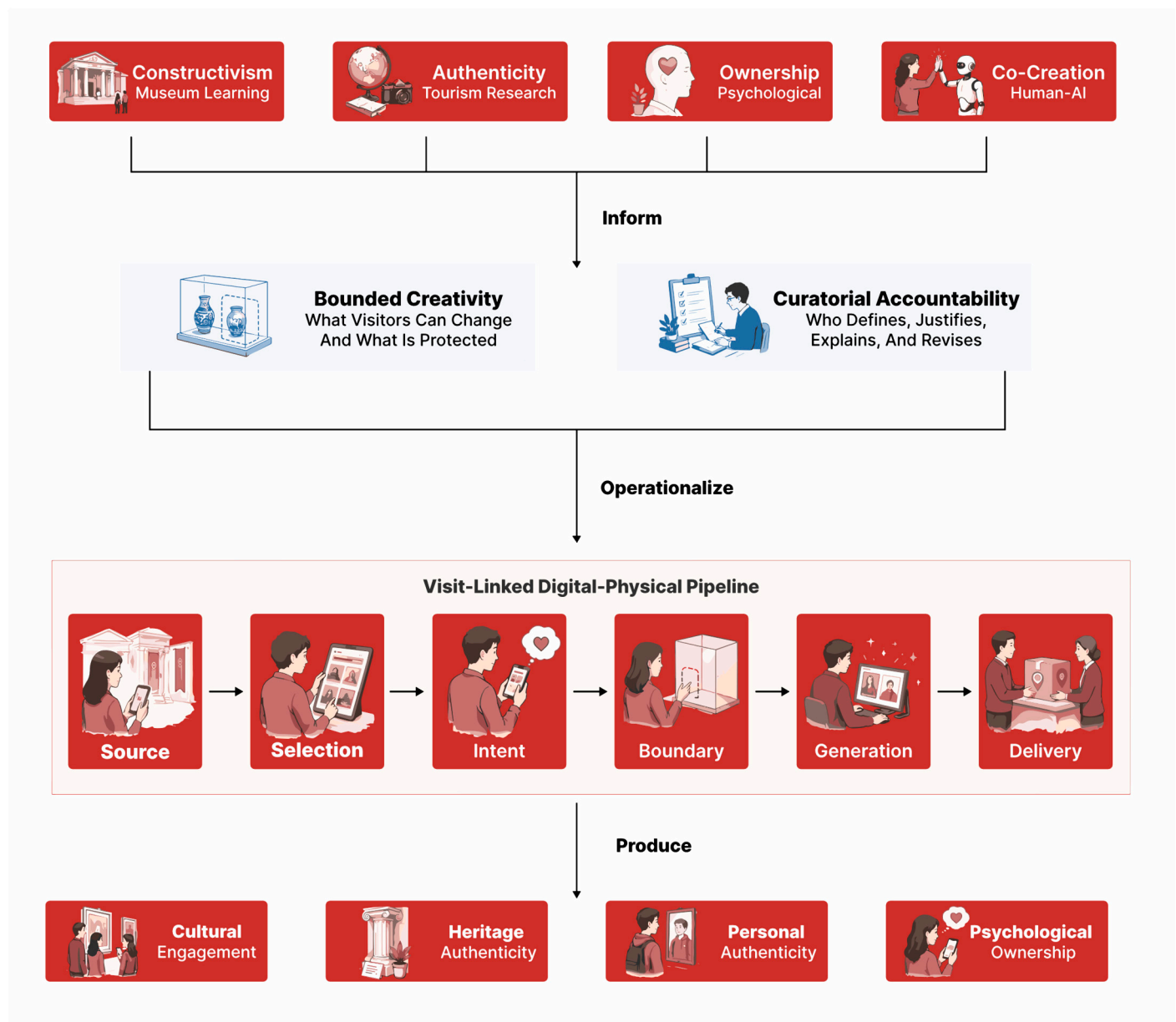
In heritage settings, however, dialogue and generation also carry interpretive consequences. A request is not only a design instruction; it may involve cultural symbols, historical forms, or source traditions that require care. A culturally inappropriate or unsupported request should therefore not be handled only as a technical failure or a generic refusal. It may become a moment of guided interpretation if the system explains why a request has been redirected and offers culturally appropriate alternatives. This makes dialogue a potential site for both creative participation and cultural boundary communication.

Generative AI also introduces a heritage-specific tension with three related faces. The content face concerns outputs that may appear visually attractive but weaken heritage authenticity, distort symbolic meaning, or sever the connection to the source tradition [13,19–21,44]. The process face concerns the possibility that automatic generation may reduce the close attention to source material that supports museum learning [1,2]. The governance face concerns the rules that shape acceptable outputs. In generative systems, these rules may be hidden in training data, prompts, model defaults, ranking mechanisms, or interface constraints, making it difficult to identify who defined a cultural boundary and why [22–27]. Museum-sector governance work responds to this concern: practitioner-grounded frameworks call for preserving curatorial vision and keeping human oversight in visitor-facing AI [45]. What remains underspecified is how such oversight attaches to the concrete cultural limits inside a co-creation system.

These tensions suggest that heritage GenAI requires more than powerful generation or smooth interaction. It needs a design structure that supports visitor authorship while keeping source linkage, cultural limits, and institutional responsibility visible. Recent systems share this concern. A field-deployed system for the Kaiping Diaolou, published in 2026, encodes non-negotiable heritage invariants that generation may not alter [44]. Such convergence indicates an active design space. It does not yet supply a governance account of who sets these invariants, how they are documented and explained, and how they can be revised.

### **3. Conceptual Framework: Bounded Creativity and Curatorial Accountability**

Building on these three bodies of work, the conceptual framework treats GenAI as one authoring mechanism within a visit-linked, culturally bounded co-creation process. It responds to the central tension identified above: visitors need room to author personally meaningful outputs, while the system must preserve source linkage, post-visit continuity, and visible curatorial responsibility. We capture this response through two paired principles: bounded creativity, which defines what visitors may change and what should remain protected, and curatorial accountability, which defines how cultural limits are set, implemented, explained, and revised. Figure 3 summarizes this relationship.



**Figure 3.** Design-research trajectory of bounded creativity and curatorial accountability. Three research streams reviewed in Section 2 motivate the conceptual proposition. The proposition is justified through constraint-creativity research, heritage-authority scholarship, and AI accountability practices; extracted into four design guidelines through a formative focus group; implemented as a visit-linked co-creation system with two authoring pathways under one documented rule set; and validated in a three-condition field study with four primary outcomes. The bottom row states what the framework inherits from participatory heritage and human–AI co-creation research and what it adds.

### 3.1. Bounded Creativity

Bounded creativity addresses how visitor authorship can be supported without turning heritage transformation into unrestricted generation. The need for this principle follows from the tensions reviewed in Section 2. Visitors need opportunities to notice, choose, express, revise, and connect cultural material with their own experience; these activities are important for meaning-making, personal connection, and psychological ownership [1–4,38–40]. At the same time, a heritage-linked output should remain visibly connected to the source material that the visitor encountered, because perceived authenticity depends not only on personal expression but also on cultural and source-based connection [34–37]. GenAI-mediated authoring intensifies this balance: dialogue and

generation may help visitors express intentions, but open-ended transformation may also weaken source linkage or distort culturally significant forms [13–21].

We define bounded creativity as a design principle in which the authoring space is deliberately open in some dimensions and deliberately restricted in others. In a heritage co-creation system, visitors may vary aspects such as product type, layout, color, texture, surface treatment, degree of abstraction, and personal expression. However, culturally significant features, such as the recognizability of a calligraphic character, the structure of key brushstrokes, or the symbolic associations of a heritage motif, should remain protected when their distortion would weaken the connection to the source tradition.

The term extends an established line of creativity research. Studies of creative work under constraints show that well-chosen restrictions can support creative outcomes, because restrictions focus search, structure variation, and make departures from convention meaningful [41,42]. Bounded creativity carries this position into heritage co-creation and adds a domain-specific assignment: the restricted dimensions protect culturally significant features, and the open dimensions carry visitor expression. The principle therefore has two components: a division of the authoring space into open and restricted dimensions, and an assignment rule that ties each restriction to a documented cultural rationale.

This principle does not treat visitor creativity as a problem to be minimized. Rather, it treats creativity as something that needs to be organized around source material. The goal is to give visitors meaningful room for authorship while keeping the final output connected to what they encountered in the museum. In this sense, bounded creativity responds to both the content and process faces of the heritage-GenAI tension: it limits culturally inappropriate transformation, and it keeps visitor attention anchored to selected source material rather than to arbitrary generation.

Three considerations justify cultural boundaries in this setting. First, heritage institutions hold stewardship obligations. An output produced from museum source material within a museum system carries institutional endorsement, so cultural misrepresentation is an institutional risk rather than a private one [25–27]. Second, source linkage is part of the value a heritage takeaway holds for visitors, and a transformation that severs the link removes the quality the object is meant to carry [34–37]. Third, the constraint-creativity evidence cited above indicates that structured limits can support rather than suppress creative outcomes [41,42]. These considerations justify boundaries as a design premise. Whether bounded generation produces better heritage outcomes than unrestricted generation remains an empirical question, and this study does not test it, because the design includes no unrestricted-generation condition. The study instead examines whether co-creation under documented boundaries improves visitor outcomes relative to conventional consumption, and whether a GenAI-mediated pathway adds value relative to a rule-based pathway under the same boundaries. Section 7.4 returns to this scope limit.

### 3.2. Curatorial Accountability

Curatorial accountability addresses how cultural limits are defined and made visible. In this paper, cultural limits refer to the abstract rules that curators set, explain, and revise. Cultural boundaries refer to the concrete instances through which those rules act on a specific visitor request. The need for this principle follows from the authority and governance concerns reviewed in Section 2. Heritage interpretation is not neutral: cultural value is selected, interpreted, authorized, and governed through institutional and social processes [25–27]. In GenAI systems, these processes may become harder to inspect because cultural limits can be embedded in training data, prompts, model defaults, interface options, ranking logic, or hidden validation rules [22–24]. If a system redirects a visitor's request or blocks a generated output, the decision should not appear as an unexplained technical constraint. It should be connected to an identifiable cultural rationale.

We define curatorial accountability as the requirement that culturally consequential boundaries be attributable, documented, implemented, explained, and revisable. They should be attributable, so it is clear who defined or approved a rule. They should be documented, so the rule and its rationale can be inspected. They should be implemented, so the rule is connected to concrete interface and technical behavior. They should be explained, so visitors receive a clear reason when a boundary is triggered. They should be revisable, so the institution can update the boundary when curatorial knowledge, stakeholder expectations, or deployment experience changes.

These five properties join two established bodies of practice. Museum scholarship treats heritage interpretation as an exercise of authority that requires justification and openness to challenge [25–27]. Documentation practices in AI governance require that consequential system behavior be recorded, attributable, and reviewable [22–24]. Curatorial accountability applies the second body of practice to the cultural limits identified by the first: the limits that curators set in a co-creation system receive the same documentation-and-review treatment that model cards and datasheets give to model behavior and training data.

This principle treats cultural boundaries as interpretive decisions rather than as neutral technical settings. Curatorial accountability does not mean that curators simply close down visitor interpretation. It means that culturally consequential limits are made visible, justified, and open to review. In a GenAI co-creation system, this is especially important because authority may otherwise be displaced into model behavior or interface defaults without being recognizable as curatorial judgment.

The two principles are paired. Bounded creativity without curatorial accountability may support source-linked variation, but the reasons for its limits may remain hidden. Curatorial accountability without bounded creativity may make rules visible, but it does not by itself create a meaningful space for visitor authorship. Together, the two principles define a design stance for heritage GenAI: visitors should be able to author personalized outputs from source material they encountered, while cultural boundaries remain explicit, justified, and revisable. This framework guides the system design described in Section 4 and motivates the comparison between conventional consumption, rule-based co-creation, and GenAI-mediated co-creation in the field study.

The paired principles form a design-and-governance framework rather than a new theory of heritage or of creativity [46,47]. From participatory heritage research, the framework inherits the position that visitors can hold interpretive agency in heritage settings [25–27]. From human–AI co-creation research, it inherits mixed-initiative authoring, in which system scaffolding and user intent shape the output together [13–19]. The framework adds two requirements that these traditions do not combine. Creative variation remains anchored to source material the visitor collected during the visit, and the cultural limits on that variation satisfy the five accountability properties defined above. Figure 3 places the two principles in a design-research trajectory from conceptual proposition to validation protocol.

## 4. System and Experience Design

### 4.1. From Principles to Practice: Formative Focus Group and Design Guidelines

We conducted a formative focus group [48] to translate the two high-level principles into design guidelines specific to the Lanting Calligraphy Museum. The focus group included nine participants: two museum curators, two museum educators, two cultural product designers, one digital fabrication specialist, and two HCI or AI-system researchers. Participants were selected because they could comment on at least one of three issues: the cultural boundaries of Lanting calligraphy, visitor interpretation and learning, or the technical feasibility of personalized heritage products. All participants provided written informed consent, and the study was approved by the university ethics board.

The session lasted approximately four hours, was conducted in Mandarin Chinese, and followed a structured protocol. The research team first introduced the two design principles of bounded creativity and curatorial accountability. The team then presented a preliminary design scenario in which a visitor collects museum source material during the visit and later uses it to create a personalized physical object. The scenario was used to support discussion rather than to present a finished system specification. Participants discussed four topics: which features of Lanting calligraphy and related material should be protected during reinterpretation; which aspects could remain open to visitor choice; where visitors would need guidance or curatorial intervention; and what requirements would apply when digital designs were turned into physical objects. The research team collected discussion notes, annotated materials, and a transcript prepared from the audio recording.

The materials were analyzed through pragmatic thematic synthesis [49–51]. Two team members reviewed the transcript, notes, and annotated sketches, identified recurrent concerns, grouped them into higher-level design inputs, and translated these inputs into design guidelines. Because the data came from a single formative focus group, the results are reported as design inputs for this deployment rather than as generalizable qualitative findings.

The synthesis produced four main findings. Participants treated cultural protection and visitor personalization as interdependent: character recognizability, brushstroke structure, and Lanting-related symbolic associations should be protected, while product type, layout, color, texture, and degree of abstraction could remain open to visitor choice. They also stressed that co-creation should begin from source material encountered during the museum visit, that visitors would need scaffolding to translate ordinary-language preferences into design operations, and that physical production requirements should be considered before final confirmation.

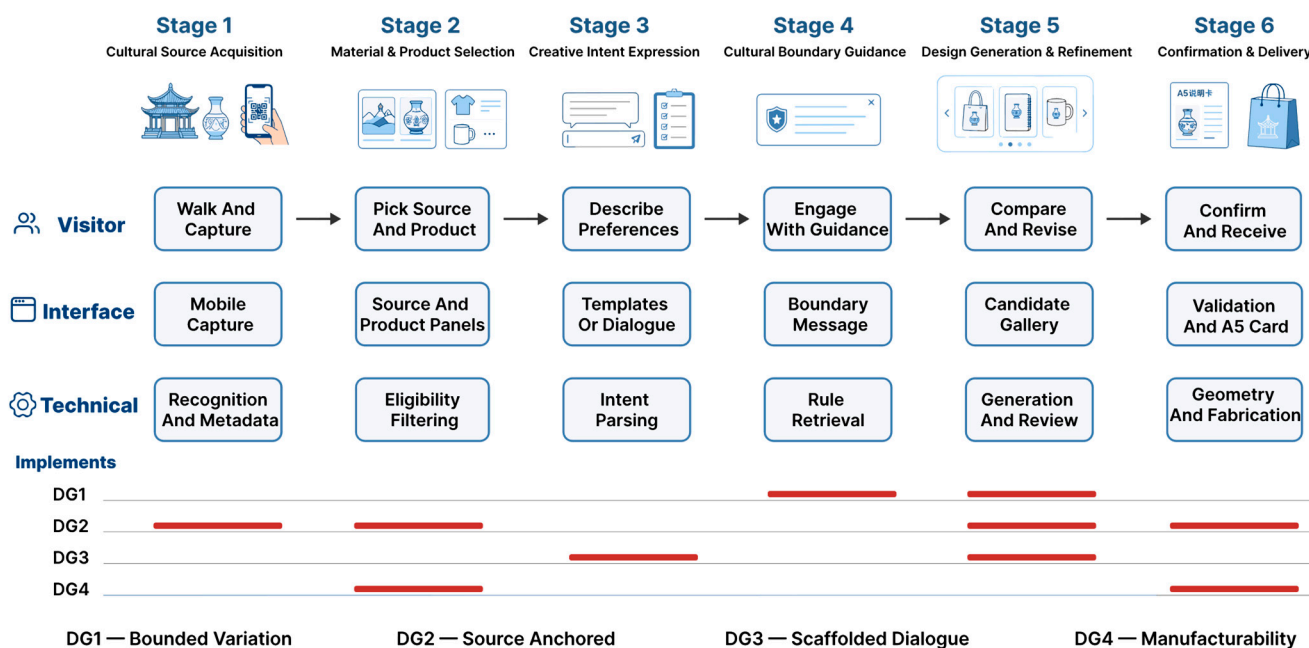
These inputs did not map onto the two principles in a one-to-one way. Rather, they specified how the principles should be operationalized in the Lanting context. DG1 directly translates bounded creativity and curatorial accountability into cultural-boundary handling. DG2 anchors creative variation to visitor-selected source material. DG3 scaffolds visitor authoring within those boundaries. DG4 extends the principles to the physical-production stage by ensuring that approved designs remain feasible and traceable. Table 1 summarizes how the formative focus-group inputs were translated into Lanting-specific design guidelines.

**Table 1.** Translating high-level principles into Lanting-specific design guidelines.

| Lanting-Specific Focus-Group Input                                                                                   | Principle Emphasis and Role                                                                                                         | Design Guideline                                                        |
|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Protect key calligraphic features while allowing variation in product form, layout, color, texture, and abstraction. | Both principles, direct. Defines the protected/variable boundary and makes it curatorially reviewable.                              | DG1. Define curatorially accountable boundaries for creative variation. |
| Begin co-creation from material visitors personally encountered during the museum visit.                             | Mainly bounded creativity; supports accountability. Keeps authoring source-linked and traceable.                                    | DG2. Visitors begin from source materials collected during the visit.   |
| Help visitors translate ordinary-language preferences into design operations                                         | Mainly bounded creativity; supports accountability through explanation. Enables guided authoring without unrestricted generation.   | DG3. Scaffold authoring through guided controls or dialogue.            |
| Check production feasibility before visitors confirm the final design.                                               | Operational support for both principles. Prevents fabrication constraints from silently altering the approved source-linked design. | DG4. Check manufacturability before final confirmation.                 |

#### 4.2. System Architecture and Experience Flow

The deployed system instantiated DG1–DG4 as a visit-linked digital–physical co-creation pipeline (see Figure 4). The pipeline connected six stages: source collection, source and product selection, intent expression, cultural-boundary checking, generation and revision, and validation and delivery. These stages were supported by four linked components: a mobile companion program for source capture, a cloud backend for session and source metadata, a tablet-based co-creation workstation for post-tour authoring, and a fabrication management layer for validation and delivery.



**Figure 4.** Overall workflow of the visit-linked digital–physical co-creation system. It shows how the design guidelines informed both visitor-facing interaction and backend processing: source capture and metadata support DG2; guided controls and dialogue support DG3; cultural-boundary checking supports DG1; and manufacturability validation and delivery tracking support DG4.

The purpose of the architecture was to keep the visitor experience simple while preserving source linkage and accountable cultural boundaries. Visitors collected source material during the tour, created from that material after the tour, and received a physical object linked to the same source record. Internally, a session token connected source images, design choices, generated candidates, boundary events, validation results, the provenance card, and the final object.

#### 4.2.1. Contextual Exploration and Heritage Source Collection

Stage 1 occurred during the guided museum tour. Visitors used the mobile companion program to capture calligraphic details, motifs, architectural details, or exhibit-related visual material that they wanted to use later. The interface used a single-action capture function to keep attention on the exhibits rather than on screen management. Each capture was linked to the session token and basic exhibit metadata, so later outputs could be traced to visitor-selected museum source material.

#### 4.2.2. AI-Mediated Workstation Co-Creation and Cultural-Boundary Checking

After the tour, visitors used a museum-provided iPad as the co-creation workstation. The workstation retrieved the visitor's Heritage Photo Library and presented a source-linked creation workspace. Visitors selected one or more source images, chose a product category, and used the Lanting Co-Creation Agent to describe desired changes in ordinary

language. The selected source image remained visible beside the current candidate, and quick revision actions allowed visitors to adjust the design without prompt-engineering knowledge.

Cultural-boundary checking operationalized bounded creativity and curatorial accountability. The system did not ask the AI model to decide cultural appropriateness on its own. Instead, it applied documented cultural-boundary rules defined and reviewed with curatorial input. Supplementary Table S4 reproduces a representative excerpt from the rule register, including each rule's protected feature, permitted variation, cultural rationale, approver, and revision status. These rules protected culturally significant features, including character recognizability, brushstroke structure, and Lanting-related symbolic associations. They allowed variation in palette, texture, layout, scale, and product application. When a request exceeded the permitted range, the interface explained the issue and offered alternatives. For example, a request to turn a calligraphic character into a cartoon mascot was redirected because it could make the character structure unrecognizable.

#### 4.2.3. Generation, Review, Confirmation, and Physicalization

In the AI-mediated authoring path, the agent translated the visitor's request into structured design variables. These variables were combined with source metadata, product constraints, and applicable cultural-boundary rules. The system then generated a small set of candidate outputs per round and reviewed each candidate before it was shown to the visitor. This design reduced choice overload and kept each candidate connected to the selected museum source.

After a design was approved, the system converted it into a product-specific fabrication file. Final confirmation occurred only after manufacturability validation. The validation checked scale, minimum thickness, surface continuity, base stability, and product-category constraints. If automatic repair would substantially change the design, the visitor returned to revision. Before leaving the museum, the visitor received a two-sided A5 provenance card showing the captured source image and the approved design preview. The final 3D-printed object was fabricated off-site and delivered within five to seven business days.

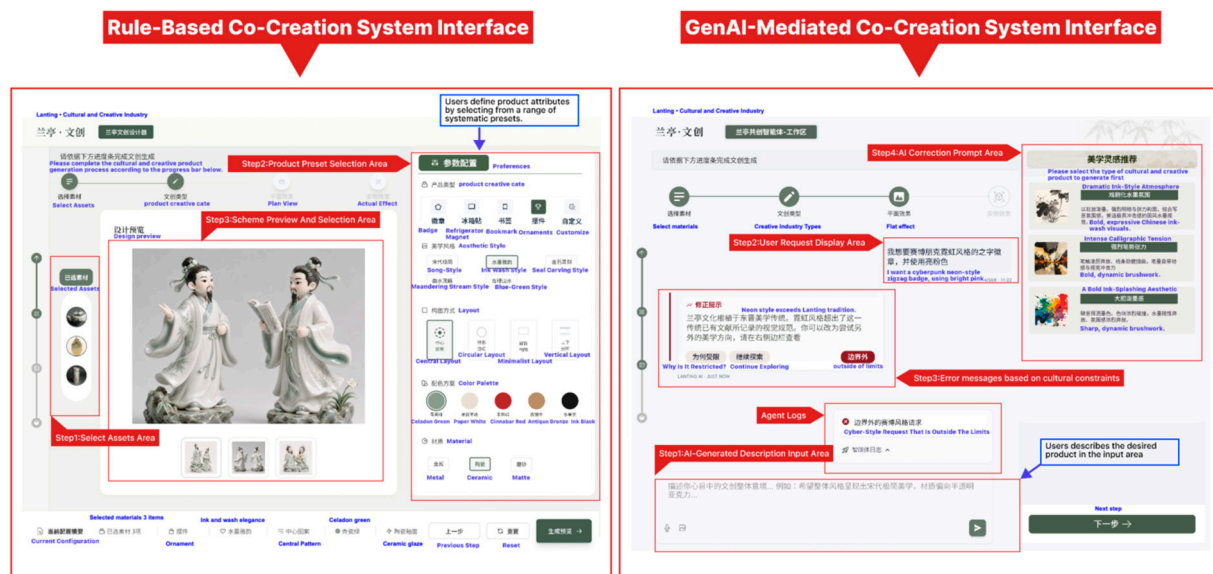
#### 4.2.4. Rule-Based Authoring Variant (For Evaluation)

To support the later field comparison, we also implemented a rule-based authoring variant using the same source-collection, product-selection, provenance, validation, and delivery infrastructure. This variant replaced dialogue-based generation with curatorially authored templates, layout options, color choices, and product-specific design controls (see Figure 5). Cultural boundaries were embedded in the available templates and option ranges rather than explained through dialogue.

The rule-based variant was designed as a matched active-control pathway rather than as a separate museum system. It preserved the core source-linked co-creation structure while removing GenAI-mediated dialogue and generative output production. The experimental use of this variant, together with the conventional-consumption baseline, is described in Section 5.

The two authoring pathways were built as variants of one system rather than as two systems. Table 2 maps the shared and differing modules. Five of seven modules were identical across pathways: source collection, source and product selection, cultural-boundary rule content, manufacturability validation, and provenance and delivery. The pathways differed in two respects. The authoring mechanism differed: the rule-based pathway composed designs from curatorially authored templates, layout options, color choices, and product-specific controls, while the GenAI-mediated pathway translated natural-language requests into structured design variables and generated candidate outputs. The enforcement point of cultural boundaries also differed: the rule-based pathway applied the documented rules in advance, because curators encoded them into the template

and option space, while the GenAI-mediated pathway applied the same rules at request time and explained redirections to the visitor. The templates themselves instantiated the same design guidelines: DG1 through curatorial authorship of the option space, DG2 through mandatory use of visitor-collected source images, DG3 through structured controls in place of dialogue, and DG4 through the shared validation pipeline. This module-level correspondence supports the structural comparability of the two pathways. The C-versus-B contrast therefore isolates the authoring mechanism and its enforcement point while holding the remaining modules constant.



**Figure 5.** Tablet workstation interfaces for the two authoring pathways. The left panel shows the rule-based interface, in which visitors select from predefined templates and options to compose a personalized design from their collected source images. The right panel shows the GenAI-mediated interface, in which visitors describe and revise design choices in natural language, and the system generates, redirects, and explains outputs under documented cultural boundaries.

**Table 2.** Module-level mapping between the rule-based and GenAI-mediated authoring pathways.

| Module                       | Rule-Based Pathway (Condition B)                                                          | GenAI-Mediated Pathway (Condition C)                                                | Status                                    |
|------------------------------|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------|
| Source collection            | Mobile companion program during the guided tour                                           | Same                                                                                | Identical                                 |
| Source and product selection | Heritage Photo Library; shared product catalog                                            | Same                                                                                | Identical                                 |
| Intent expression            | Curatorially authored templates, layout options, color choices, product-specific controls | Natural-language dialogue parsed into structured design variables                   | Differs                                   |
| Cultural-boundary rules      | Same documented rule set, encoded in advance into the template and option space           | Same documented rule set, applied at request time with visitor-facing explanation   | Same content; different enforcement point |
| Output production            | Deterministic composition from the selected template and options                          | Generated candidates screened by automated source-consistency review before display | Differs                                   |
| Manufacturability validation | Shared validation pipeline                                                                | Same                                                                                | Identical                                 |

|                         |                                                     |      |           |
|-------------------------|-----------------------------------------------------|------|-----------|
| Provenance and delivery | Provenance card, session token, post-visit delivery | Same | Identical |
|-------------------------|-----------------------------------------------------|------|-----------|

#### 4.2.5. Technical Implementation

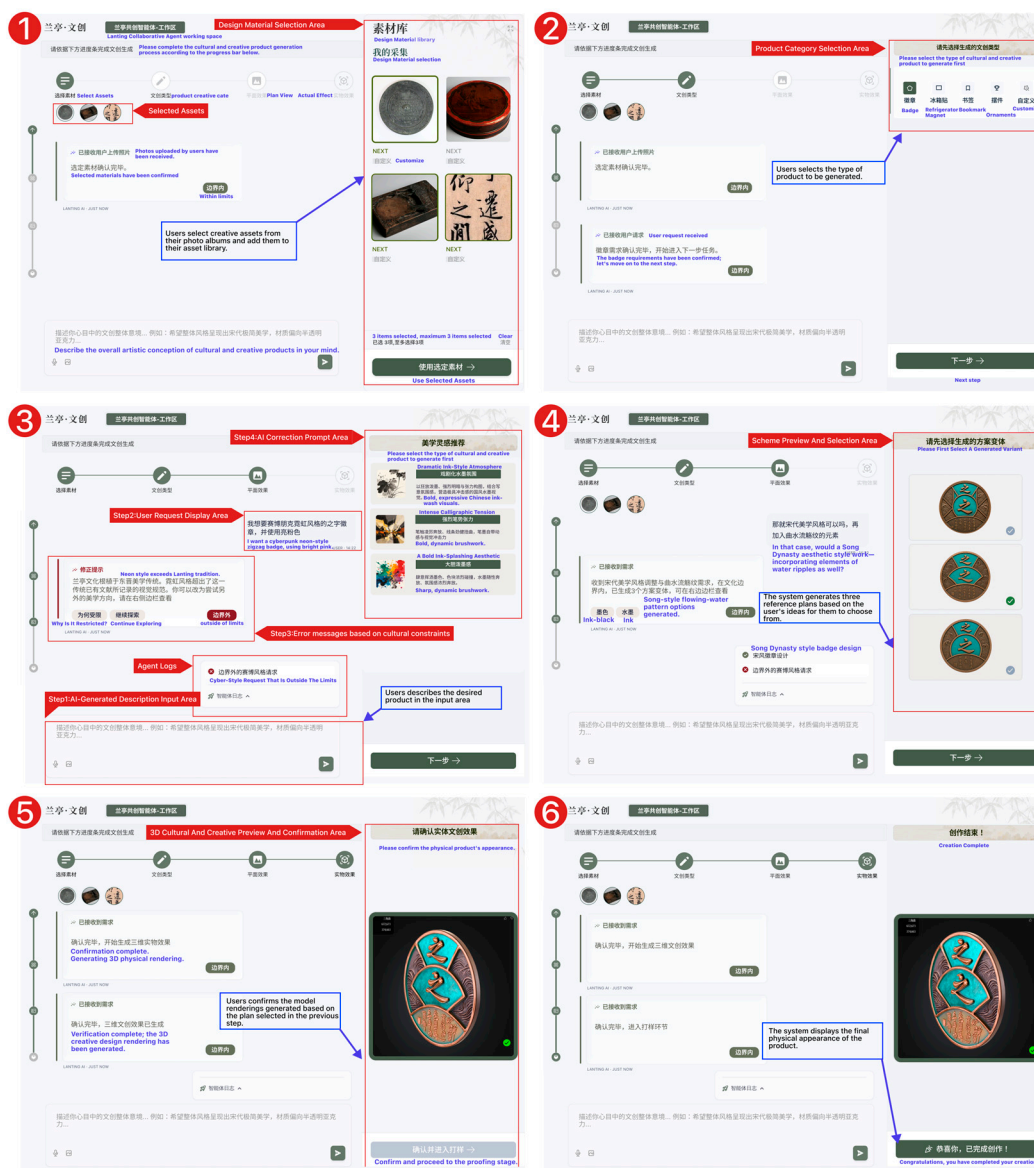
The implementation used separate components for dialogue management, image generation, source recognition, output review, image-to-3D conversion, and manufacturability validation. The study evaluates a concept-guided system rather than the standalone capability of any single model. Table 3 itemizes each component, its role in the pipeline, and its key configuration to support reproducibility. Dialogue management used GPT-4o [52] accessed through its public API, operating under a structured prompt framework that defined the agent tasks described below. Image generation used Stable Diffusion XL [53] fine-tuned with low-rank adaptation [54] on 2400 curated images. The corpus combined 1600 high-resolution images from the museum digital archive, covering calligraphy works, stele rubbings, Eastern Jin artifacts, and celadon ceramics under a research license agreement, with 800 open-access images of Wei-Jin decorative arts. The museum curatorial team reviewed the corpus before training to confirm cultural and historical suitability. Fine-tuning used rank  $r = 8$ , a learning rate of  $10^{-4}$ , and 3000 steps. This adaptation shifts the learned distribution toward Lanting-specific visual patterns and reduces generic or anachronistic elements. Source recognition used MobileNetV3 [55] to link captured photographs to exhibit metadata, with approximately 96% recognition accuracy on an evaluation set of 50 exhibit photographs. This evaluation set is small; the reported accuracy characterizes recognition in this deployment and does not establish general performance. Output review used CLIP [56] to screen each candidate against the selected source image before display, with a minimum similarity threshold of 0.72. The threshold was set empirically during development: the team inspected similarity scores for outputs judged consistent or inconsistent with their source images and selected the value that excluded the inconsistent cases. In internal testing, approximately 8% of initial generations fell below the threshold and were regenerated, and all second attempts passed. The value is specific to this deployment rather than a general recommendation. Image-to-3D conversion used TripoSR [57] to produce a mesh from the approved design, followed by product-template constraints. Manufacturability validation used a pipeline built on the Trimesh geometry library [58] with custom scripts. The pipeline checked scale, minimum wall thickness of 1.2 mm for resin printing and 1.5 mm for fused-deposition printing, surface continuity, base stability, and a maximum overhang angle of 45 degrees. In internal testing, 94% of models passed automatic validation without manual repair. Prompt templates, the cultural-boundary rule schema, and the fine-tuned adapter weights are available from the corresponding author on reasonable request. Supplementary Figures S1–S3 provide step-by-step photographic documentation of the three condition flows in the field deployment, including a cultural-boundary redirection.

**Table 3.** Model and pipeline components of the deployed system.

| Function            | Component                                    | Role in the Pipeline                                                                                    | Key Configuration                                                                 |
|---------------------|----------------------------------------------|---------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Dialogue management | GPT-4o [58]                                  | Parses visitor requests, assembles structured design variables, explains cultural-boundary redirections | Public API; structured prompt framework                                           |
| Image generation    | Stable Diffusion XL, LoRA fine-tuned [59,60] | Generates 2D candidate designs from source image and design variables                                   | 2400 curated images; rank $r = 8$ ; learning rate $1 \times 10^{-4}$ ; 3000 steps |
| Source recognition  | MobileNetV3 [61]                             | Links captured photographs to exhibit metadata                                                          | Approximately 96% accuracy on a 50-photograph exhibit evaluation set              |

|                              |                             |                                                                            |                                                                                            |
|------------------------------|-----------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| Output review                | CLIP [62]                   | Screens candidates for consistency with the selected source image          | Minimum similarity threshold 0.72                                                          |
| Image-to-3D conversion       | TripoSr [63]                | Converts the approved 2D design into a 3D mesh                             | Product-template constraints applied after conversion                                      |
| Manufacturability validation | Trimesh-based pipeline [64] | Checks scale, wall thickness, surface continuity, base stability, overhang | Wall $\geq 1.2$ mm resin, $\geq 1.5$ mm FDM; overhang $\leq 45^\circ$ ; 94% automatic pass |

The GenAI-mediated version used the Lanting Co-Creation Agent to parse visitor requests, assemble structured design variables, generate and revise candidate outputs, and explain cultural-boundary redirections. Figure 6 shows the six interface steps a visitor moves through in this pathway, from source image selection to the final 3D preview. The rule-based variant used the same source metadata, product catalog, provenance card, validation workflow, and delivery process, but replaced GenAI-mediated generation with deterministic template controls.



**Figure 6.** Six-step walkthrough of the GenAI-mediated authoring interface. Visitors first select source images from their on-site collection, then describe design intentions in natural language. The system generates candidate outputs, supports iterative revision through dialogue, applies cultural-

boundary checks with visitor-facing explanations when a request is redirected, and produces a 3D preview before fabrication. Each panel shows what the visitor sees at one step of the authoring sequence.

## 5. Evaluation Methods

### 5.1. Study Design

To address RQ2, we designed a three-condition between-subjects field study. The independent variable was the type of heritage experience. Condition A was a conventional-consumption baseline: participants completed the standard guided museum tour and received a standard visit-related souvenir. Condition B was a rule-based co-creation condition: participants collected source material during the visit and used a template editor with curatorially authored layouts and option sets. Condition C was an AI-mediated co-creation condition: participants followed the same source-linked co-creation structure but used the Lanting Co-Creation Agent to express and revise design intentions through dialogue under documented cultural-boundary rules.

The design supported two planned comparisons. The first compared Condition A with the two co-creation conditions combined, estimating the effect of source-linked co-creation relative to conventional consumption. The second compared Condition C with Condition B, estimating the additional value of the AI-mediated authoring pathway relative to a matched rule-based co-creation pathway. Because the two active conditions shared source collection, post-tour authoring, cultural boundaries, product validation, and personalized delivery, the C-versus-B comparison should be interpreted as a pathway-level comparison rather than as an isolated test of a single model feature.

### 5.2. Participants

Sixty adult visitors were recruited through the museum's online channels and on-site invitation on non-study days. Eligibility required participants to be at least 18 years old, available for an approximately half-day on-site session during museum operating hours, willing to receive a mailed takeaway within five to seven business days, and not previously exposed to the Lanting co-creation system. All sixty enrolled participants completed the full session and the post-session measures, with no withdrawals.

Participants ranged in age from 18 to 64 years ( $M = 31.27$ ,  $SD = 11.98$ ). Gender was evenly balanced, with 28 women and 32 men. Prior generative-AI experience and prior museum-visit history were collected as baseline covariates.

Random assignment was stratified by age group and prior generative-AI experience, because both variables could plausibly affect comfort with the tablet interface and confidence in creative tasks. The allocation sequence was generated before recruitment began and concealed from facilitators until participants entered the session. This procedure assigned 20 participants to each of the three conditions ( $n = 20$  per condition).

All participants provided written informed consent, and the study was approved by the university ethics board. Each participant received a completion gift valued at approximately RMB 100 (about USD 14). Participants in Conditions B and C also received the personalized object as the outcome of the co-creation task, not as additional compensation.

### 5.3. Instruments and Apparatus

The study used two types of instruments and apparatus: a self-report questionnaire battery and the system apparatus used during the visit and post-tour activity.

The self-report questionnaire battery was administered using a 7-point Likert response format, from 1 = strongly disagree to 7 = strongly agree. Composite scores were

calculated as item means. All questionnaire items were administered in Mandarin Chinese, the working language of the museum and of all participants.

Four primary-outcome scales were administered. Cultural Engagement measured attention, interest, absorption, and involvement with Lanting heritage during the visit, with items adapted from the User Engagement Scale Short Form [59]. Heritage-Linked Authenticity measured whether the visit-associated takeaway felt genuinely connected to Lanting heritage, with items adapted from authenticity scales in tourism and heritage research [37–40]. Personal Authenticity measured self-expression, personal meaning, and emotional fit, also adapted from the same authenticity literature [34–37]. Psychological Ownership measured the extent to which participants felt that the visit-associated takeaway was their own, with items adapted from psychological-ownership measures [38–40].

Two manipulation-check scales measured Perceived Creative Participation and Perceived AI Mediation. Three diagnostic scales measured Perceived Source Traceability, Perceived Boundary Transparency, and Process Clarity and Usability. These diagnostic scales supported interpretation but were not treated as primary outcomes.

Psychometric evidence for the four primary-outcome scales was examined before hypothesis testing. Cronbach's alpha ranged from 0.820 to 0.851, and composite reliability computed as McDonald's omega ranged from 0.823 to 0.853. Corrected item-total correlations for all 31 items were 0.35 or higher. Kaiser–Meyer–Olkin values ranged from 0.78 to 0.84, and Bartlett's test of sphericity was significant for each scale, all  $p < 0.001$ . These values support factor analysis at the scale level. Single-factor principal-axis solutions produced loadings between 0.38 and 0.85. The first factor explained 59.4% of item variance for Heritage-Linked Authenticity, 45.6% for Personal Authenticity, 40.8% for Psychological Ownership, and 27.2% for Cultural Engagement. Average variance extracted was 0.59 for Heritage-Linked Authenticity and ranged from 0.27 to 0.46 for the other three scales. Values below the 0.50 convention remain acceptable when composite reliability exceeds 0.60 [60]. The lower value for Cultural Engagement reflects the breadth of the 14-item scale, which covers attention, interest, absorption, and involvement and functions as a broad composite index rather than a narrow unidimensional measure. Discriminant validity met the Fornell–Larcker criterion: within-condition correlations between the four scales ranged from 0.37 to 0.55 and did not exceed the square root of the corresponding average variance extracted, which ranged from 0.52 to 0.77. We did not estimate a confirmatory factor analysis because a four-factor measurement model for 31 items requires more parameters than a sample of 60 supports [61]. Item-level statistics and the full correlation matrix are reported in Supplementary Tables S2 and S3.

The system apparatus included the mobile companion program and the tablet-based co-creation workstation. During the guided tour, participants accessed the mobile companion program by scanning a QR code with their own smartphones. This bring-your-own-device arrangement reflected normal museum use and avoided assigning unfamiliar phones during the gallery tour. After the tour, participants in Conditions B and C used a museum-provided iPad tablet to complete the co-creation task. The same tablet setup, workstation location, product catalog, and confirmation procedure were used in both active conditions. The intended difference between Conditions B and C was the authoring mechanism: the rule-based template editor in Condition B and the AI-mediated dialogue interface in Condition C.

#### 5.4. Procedure

Each session followed a standardized procedure across all three conditions and occupied an approximately half-day block in the museum. A between-subjects design was used because repeated exposure to the same museum visit, source material, and product workflow could alter novelty, first-encounter meaning-making, and perceived

authenticity. Each participant therefore completed only one condition. The study was conducted in the operational museum environment to preserve ecological validity for visitor-facing heritage AI deployment.

The procedure included four stages. Stage 1 was arrival and pre-tour preparation. Participants completed informed consent, demographic questions, prior generative-AI experience, and prior museum-visit history measures. Stage 2 was the guided museum tour. All participants followed the same route with the same docent script and museum-education content. In Conditions B and C, participants used the mobile companion program to capture source material during the tour. In Condition A, the mobile companion program was used in tour-only mode, with capture and creation functions disabled.

Stage 3 was the condition-specific post-tour activity. Condition A participants completed a matched post-tour rest and reflection period and received a standard visit-related souvenir. Condition B participants used the rule-based template editor. Condition C participants used the AI-mediated dialogue interface. Active-condition participants worked until they confirmed a design or reached the maximum workstation time. Stage 4 consisted of the post-session questionnaire battery and a brief semi-structured interview.

The session sequence was fixed by design: source collection occurred during the guided tour, and authoring occurred afterward at the workstation. The Heritage Photo Library available at the workstation contained the images captured during the tour, and participants did not return to the galleries to capture additional material during the authoring stage. This separation matches the intended use case, in which a visitor authors a takeaway after the visit, and it holds the source-collection stage constant across the two active conditions. Iteration was supported within the authoring stage rather than across stages: participants could reselect source images from their library, revise designs, switch templates in Condition B, and request regeneration in Condition C. Behavioral logs recorded this iteration, and photo counts, workstation duration, and design-revision counts were comparable across the two active conditions (Section 6.1). Because the collection-then-authoring sequence was identical for every participant, sequence is a constant of the design rather than a between-condition confound. Scheduling across days and times is addressed in Section 7.4.

The active co-creation sessions for Conditions B and C were run in the same physical workstation area to avoid setting confounds. Participants in B and C received the same orientation length, used the same product catalog, and followed the same delivery schedule. The only intended difference after source collection was the authoring mechanism. Condition A did not use the tablet-based co-creation workstation. This structural difference is part of the A-versus-(B+C) contrast and is interpreted as the effect of the source-linked co-creation bundle rather than as an isolated interface effect.

A trained facilitator team handled all three conditions using a written script and a non-intervention rule on creative content. Facilitators could answer technical questions but were instructed not to suggest source images, aesthetic choices, product categories, prompt wording, or revision directions. Condition C participants received a brief standard caution about possible AI misinterpretation. Condition B participants received a parallel statement that choices could be revised before final confirmation. Full facilitator script rules and caution wording are reported in Appendix A. Figure 7 documents the deployment: source collection during the guided tour, the workstation in use, the boundary-explanation interface with the provenance card, and finished objects beside their source material.



**Figure 7.** Field deployment at the Lanting Calligraphy Museum: (a) source collection during the guided tour; (b) tablet workstation co-creation; (c) cultural-boundary explanation and provenance card; (d) delivered 3D-printed takeaways. Faces and identifying details are blurred in line with participant consent.

### 5.5. Data Collection and Analysis

Data collected per participant included the pre-tour questionnaire, the post-session questionnaire battery, system-side behavioral logs for Conditions B and C, and the post-session semi-structured interview audio recording. All data were de-identified before analysis using the participant ID issued at consent.

Behavioral logs were collected only in the two active co-creation conditions because Condition A did not use the tablet-based co-creation workstation. Shared logs for Conditions B and C included the number of heritage photos captured, workstation duration, and design revisions. Condition B additionally recorded template switches. Condition C additionally recorded dialogue turns and cultural-boundary redirections.

A brief post-session semi-structured interview was conducted with all participants. The interview covered the visit experience, perceived authorship, and the takeaway object. In Conditions B and C, the prompt set also covered the creation tool. Interviews were audio-recorded with consent and transcribed verbatim. Interview quotes presented in this paper were translated by a bilingual member of the research team.

The confirmatory analysis was specified before data collection and maps onto the two parts of RQ2. For each of the four primary outcomes, a one-way between-subjects ANOVA with condition as a three-level factor was followed by two planned contrasts. The A-versus-(B+C) contrast estimated the effect of source-linked co-creation relative to conventional consumption and addresses the first part of RQ2. The C-versus-B contrast estimated the additional effect of the GenAI-mediated pathway relative to rule-based co-creation and addresses the second part. Pairwise comparisons of each active condition against the conventional condition (B versus A; C versus A) were computed as supplementary checks outside the confirmatory family. Their only role was to confirm that the pooled contrast was not driven by a single active condition, and no other pairwise tests were run.

The confirmatory family consisted of the eight planned contrasts defined above, two for each of the four primary outcomes. Family-wise Type I error was controlled at  $\alpha = 0.05$  using Holm's sequential procedure [62]. Effect sizes are reported as  $\eta^2$  for omnibus ANOVA and Hedges'  $g$  for planned contrasts, following reporting recommendations for cumulative quantitative research [63]. Manipulation checks and diagnostic measures were reported separately and were not included in the confirmatory family.

Because the study did not include an a priori power analysis, we conducted a post hoc sensitivity analysis. With  $N = 60$ , a two-tailed alpha of 0.05, and power of 0.80, the design could detect omnibus effects of  $f \geq 0.41$ , which corresponds to  $\eta^2 \geq 0.145$ . The A-versus-(B+C) contrast could detect effects of  $d \geq 0.78$ , and the C-versus-B contrast could detect effects of  $d \geq 0.91$ . Under the most stringent step of the Holm procedure, the detectable

effect for the C-versus-B contrast rises to  $d \geq 1.19$ . The study was therefore adequately powered for large effects only. Section 7.4 discusses the implications for non-significant results.

We ran a sensitivity check using Welch correction for Psychological Ownership, where homogeneity of variance was less clearly satisfied. To show effect-size uncertainty, we report approximate 95% confidence intervals for Hedges'  $g$  for the eight planned contrasts. As a further robustness check, we ran a non-parametric bias-corrected and accelerated bootstrap sensitivity check with 10,000 resamples drawn within each condition [64]. This robustness check did not change the substantive interpretation of the planned contrasts. Item-level missing data on the post-session questionnaire was minimal, and complete-case analysis was used.

Interviews were analyzed using a codebook thematic approach [49–51]. Codes were generated and refined iteratively. Formal inter-coder reliability was not computed for participant interviews. Interview material is therefore reported as illustrative evidence used to interpret the quantitative pattern, not as an independent qualitative finding.

## 6. Results

This section reports the quantitative and qualitative results among the three conditions: conventional-consumption baseline (A), rule-based co-creation (B), and GenAI-mediated co-creation (C).

### 6.1. Primary Quantitative Outcomes

The primary outcome analysis followed the confirmatory structure specified in Section 5.5 and addressed RQ2 through two planned contrasts. The four primary outcomes use scale names from Section 5.3 that correspond to the descriptive terms in the abstract: Cultural Engagement to cultural engagement, Heritage-Linked Authenticity to perceived connection to heritage sources, Personal Authenticity to personal meaning and self-expression, and Psychological Ownership to psychological ownership. The A-versus-(B+C) contrast tested whether the source-linked co-creation bundle differed from conventional consumption. The C-versus-B contrast tested whether the GenAI-mediated pathway differed from the matched rule-based co-creation pathway. Across all four outcomes, both co-creation conditions showed higher means than the conventional condition. The C-versus-B pattern was outcome-specific: the GenAI-mediated pathway showed higher scores than the rule-based pathway for Cultural Engagement, Personal Authenticity, and Psychological Ownership, while the difference for Heritage-Linked Authenticity was smaller and not statistically significant; effects below the detectable threshold reported in Section 5.5 cannot be ruled out. Table 4 reports descriptive statistics and reliability for the four primary outcome measures, Table 5 reports the omnibus ANOVA and planned contrast results, and Figure 8 summarizes condition means.

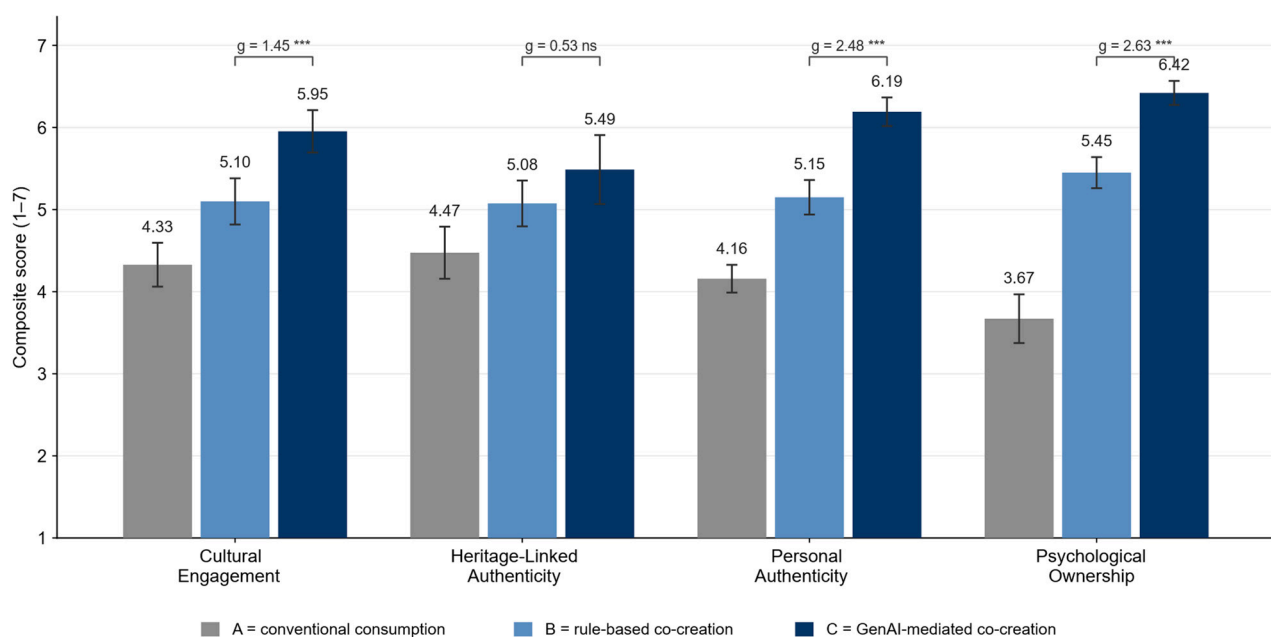
**Table 4.** Descriptive statistics and reliability for the four primary outcome measures.

| Measure                      | Items | Alpha | A: M (SD)   | B: M (SD)   | C: M (SD)   |
|------------------------------|-------|-------|-------------|-------------|-------------|
| Cultural Engagement          | 14    | 0.832 | 4.33 (0.57) | 5.10 (0.60) | 5.95 (0.55) |
| Heritage-Linked Authenticity | 4     | 0.851 | 4.47 (0.68) | 5.08 (0.60) | 5.49 (0.90) |
| Personal Authenticity        | 6     | 0.829 | 4.16 (0.36) | 5.15 (0.45) | 6.19 (0.37) |
| Psychological Ownership      | 7     | 0.820 | 3.67 (0.63) | 5.45 (0.40) | 6.42 (0.31) |

**Table 5.** Omnibus and planned contrast results for the four primary outcomes, with approximate 95% confidence intervals for Hedges' *g*.

| Outcome                      | Omnibus ANOVA                                  | A vs. (B + C)                                                       | C vs. B                                                             |
|------------------------------|------------------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------------|
| Cultural Engagement          | $F(2, 57) = 39.98, p < 0.001, \eta^2 = 0.584$  | $t(58) = 6.52, p < 0.001, g = 1.76, 95\% \text{ CI } [1.13, 2.39]$  | $t(38) = 4.68, p < 0.001, g = 1.45, 95\% \text{ CI } [0.75, 2.15]$  |
| Heritage-Linked Authenticity | $F(2, 57) = 9.59, p < 0.001, \eta^2 = 0.252$   | $t(58) = 3.93, p < 0.001, g = 1.06, 95\% \text{ CI } [0.49, 1.63]$  | $t(38) = 1.71, p = 0.095, g = 0.53, 95\% \text{ CI } [-0.10, 1.16]$ |
| Personal Authenticity        | $F(2, 57) = 132.05, p < 0.001, \eta^2 = 0.822$ | $t(58) = 9.46, p < 0.001, g = 2.56, 95\% \text{ CI } [1.85, 3.27]$  | $t(38) = 7.99, p < 0.001, g = 2.48, 95\% \text{ CI } [1.65, 3.31]$  |
| Psychological Ownership      | $F(2, 57) = 175.88, p < 0.001, \eta^2 = 0.861$ | $t(58) = 13.41, p < 0.001, g = 3.63, 95\% \text{ CI } [2.78, 4.48]$ | $t(38) = 8.49, p < 0.001, g = 2.63, 95\% \text{ CI } [1.77, 3.49]$  |

Note. Confidence intervals are approximate large-sample intervals for Hedges' *g* and are reported to show effect-size uncertainty. The planned contrasts remain the main basis for inference.

**Figure 8.** Condition means with 95% confidence intervals for Cultural Engagement, Heritage-Linked Authenticity, Personal Authenticity, and Psychological Ownership. Brackets show the C-versus-B planned contrast (Hedges' *g*; \*\*\*  $p < 0.001$ ; ns = not significant).

Across all four primary outcomes, the omnibus ANOVA was significant, and the A-versus-(B+C) contrast was significant, indicating a broad co-creation effect relative to conventional consumption. The C-versus-B contrast was significant for Cultural Engagement, Personal Authenticity, and Psychological Ownership, but not for Heritage-Linked Authenticity. The largest C-versus-B effects were on Psychological Ownership and Personal Authenticity, indicating that the GenAI-mediated pathway was most differentiated from the rule-based pathway on personal outcomes. Holm correction left the same significance pattern: all four A-versus-(B+C) contrasts remained significant, and the C-versus-B contrasts remained significant for Cultural Engagement, Personal Authenticity, and Psychological Ownership, but not for Heritage-Linked Authenticity.

Several checks supported interpretation of these planned contrasts. Baseline checks provided no evidence of meaningful imbalance across conditions on age, cultural-activity frequency, calligraphy familiarity, gender, prior generative-AI experience, or prior museum-visit history. Supplementary Table S1 reports full demographic and baseline statistics with balance tests; no test reached significance, all  $p \geq 0.545$ . Reliability values for the

four primary scales are reported in Table 4, and item-level psychometric statistics and inter-scale correlations are reported in Supplementary Tables S2 and S3. Distributional diagnostics did not change the planned analysis, and the Welch sensitivity check for Psychological Ownership yielded the same substantive conclusion as the standard ANOVA. The planned contrasts therefore remain the main basis for inference.

Manipulation checks confirmed that the conditions were experienced as intended. Perceived Creative Participation differed strongly by condition,  $F(2, 57) = 106.16, p < 0.001, \eta^2 = 0.788$ , with scores increasing from conventional consumption to rule-based co-creation and GenAI-mediated co-creation. Perceived AI Mediation also differed strongly,  $F(2, 57) = 492.92, p < 0.001, \eta^2 = 0.945$ , with Condition C much higher than Condition B. Diagnostic measures further showed that Condition C was higher than Condition B on Perceived Source Traceability,  $t(38) = 5.44, p < 0.001, g = 1.69$ , and Perceived Boundary Transparency,  $t(38) = 2.27, p = 0.029, g = 0.70$ . Process Clarity and Usability did not differ between B and C,  $t(38) = -0.63, p = 0.535$ , providing no evidence of a simple usability advantage for the GenAI-mediated pathway. Behavioral logs also showed comparable source collection, workstation duration, and revision counts across the two active conditions.

## 6.2. Illustrative Interview Themes

Interview material was used to interpret the quantitative pattern rather than to establish a separate qualitative theory. The analysis identified five recurring themes and one qualification. A structured codebook excerpt with theme definitions, evidence patterns, example quotes, and interpretive links is provided in Table A1. First, co-creation changed the takeaway from a received object into an authored object. A Condition B participant said, “I chose the photo, the frame, and the color. It feels more like something I put together, not something I just bought” (B07). A Condition C participant similarly said, “I kept changing the direction until it felt like mine” (C12). By contrast, Condition A participants more often described the takeaway as a standard museum gift.

The interview findings help interpret why the conventional condition scored lower. Participants in Condition A described the visit as informative and pleasant, but they generally described the takeaway as something provided by the museum. In contrast, participants in Conditions B and C used verbs such as choose, change, compare, and decide when describing the object. This language difference was not treated as a separate outcome, but it is consistent with the strong co-creation effect on ownership.

Second, source-photo collection supported attention and memory. Participants in Conditions B and C described the Heritage Photo Library as part of the visit rather than only a technical input. Third, GenAI dialogue increased felt expression and authorship. Condition C participants emphasized the value of describing preferences in ordinary language and seeing the system respond. Fourth, rule-based co-creation was clear and controlled, but sometimes felt bounded by the option set. Fifth, boundary explanations sometimes became cultural guidance. One Condition C participant explained that a redirection helped them understand why the character shape should remain recognizable (C02).

The qualification is important. Not all boundary encounters were experienced positively. A small number of participants described redirections as interrupting creative flow when explanations were brief or alternatives felt narrow. These cases indicate that visible constraints do not automatically become acceptable constraints. Explanation quality and alternative design paths remain important design issues.

Taken together, the post-session results show a clear distinction between co-creation effects and GenAI-mediated pathway effects. Source-linked co-creation, whether rule-based or GenAI-mediated, improved all four primary outcomes relative to conventional consumption. The GenAI-mediated pathway further increased Perceived Creative Participation and produced reliable incremental gains in Cultural Engagement, Personal Authenticity, and

Psychological Ownership, while its additional effect on Heritage-Linked Authenticity was smaller and not statistically significant. These themes should be read as explanatory and illustrative. They clarify how participants made sense of the co-creation pathways, but they do not replace the planned quantitative contrasts as the primary basis for inference.

## 7. Discussion

### 7.1. Principal Findings: Source-Linked Co-Creation and GenAI-Mediated Authoring

The findings distinguish between two effects. The first is the effect of source-linked co-creation as a whole. Both the rule-based and GenAI-mediated pathways produced higher scores than conventional consumption across Cultural Engagement, Heritage-Linked Authenticity, Personal Authenticity, and Psychological Ownership. This pattern is consistent with prior work showing that museum visitors do not simply receive information, but construct meaning through personal context, attention, choice, and interaction [1–3]. In this study, visitors reported stronger outcomes when they selected museum source material, made design choices, revised an output, and received a takeaway connected to their own visit. The benefit was therefore not simply a matter of adding AI. It came first from turning the visit into a source-linked authoring process.

This finding also extends work on tangible and material-digital museum experiences. Prior research suggests that physical objects can carry traces of a visit, support memory, and extend engagement beyond the museum site [4,9–12]. The present study adds that the post-visit object may become more meaningful when it is not only a souvenir, but the endpoint of a visitor's own source selection and authoring process. In this sense, the physical takeaway functioned as part of the experience design rather than as an external reward.

The second effect is more specific to the GenAI-mediated pathway. Compared with the rule-based pathway, GenAI-mediated co-creation produced stronger gains in Cultural Engagement, Personal Authenticity, and Psychological Ownership, with the largest differences on the two personal outcomes. This pattern is consistent with human–AI co-creation work that treats generative systems not only as output generators, but as interaction partners through which initiative, feedback, and contribution are distributed between human and system [15,16]. In the present study, dialogue-based authoring allowed visitors to express intentions in ordinary language, request revisions, and see the system respond to their input. These features made personal contribution more visible, which may explain why the final object felt more expressive and more “owned.” This interpretation also fits prior work on psychological ownership, which emphasizes effort, control, knowledge, and self-association as routes through which people come to feel that something is theirs [38–40].

The result for Heritage-Linked Authenticity is especially important. Although the GenAI-mediated pathway produced higher perceived source traceability than the rule-based pathway, it did not reliably increase Heritage-Linked Authenticity beyond the rule-based condition. This suggests that heritage linkage is not automatically strengthened by more flexible generation. A generated output may be more personally expressive while not necessarily feeling more culturally grounded. This pattern resonates with authenticity research, which distinguishes source-based or object-related authenticity from personally meaningful or existential experience [34–37]. It also echoes concerns in heritage-GenAI work that visually appealing generation may weaken heritage linkage if the connection to source material and cultural meaning is not carefully maintained [13,19–21].

Taken together, the findings suggest that GenAI's value in this setting lies in accountable, source-linked authoring rather than unrestricted generation. A rule-based pathway can already support meaningful heritage co-creation when it begins from visitor-selected source material and leads to a personalized physical object. The GenAI-mediated pathway adds value when it helps visitors articulate, revise, and recognize their own intentions within that bounded process.

### 7.2. Bounded Creativity and Curatorial Accountability in Practice

The field deployment clarifies how the two principles work together. Bounded creativity defined the space for visitor authorship: visitors could vary product form, layout, color, texture, abstraction, and personal expression, but they could not freely distort culturally diagnostic calligraphic features or create unsupported symbolic combinations. Curatorial accountability defined the responsibility behind those limits: cultural boundaries were documented, connected to interface and technical behavior, explained when triggered, and subject to curatorial review. This design responds to heritage scholarship showing that cultural value is selected, interpreted, authorized, and governed through institutional and social processes [25–27].

The two pathways implemented the same rule set at different points, and the diagnostic data show that this difference was visible to participants. Boundaries in the rule-based pathway were embedded in templates and option ranges in advance. Boundaries in the GenAI-mediated pathway acted at request time through request checking, output review, and dialogue redirection with visitor-facing explanations. Participants in the GenAI-mediated condition reported higher Perceived Boundary Transparency,  $t(38) = 2.27$ ,  $p = 0.029$ ,  $g = 0.70$ , and higher Perceived Source Traceability,  $t(38) = 5.44$ ,  $p < 0.001$ ,  $g = 1.69$ . These governance perceptions did not convert into a reliable advantage in Heritage-Linked Authenticity, where the C-versus-B difference was small and not significant ( $g = 0.53$ ). Making boundaries visible therefore appears to be a governance outcome in its own right rather than an automatic route to stronger heritage linkage. This distinction is consistent with responsible-AI work that treats documentation, traceability, and user-facing communication as accountability conditions rather than as experience guarantees [19–21,57].

The outcome pattern also connects to constraint research in creativity support [16,25–27]. The GenAI-mediated pathway produced large additional gains in Personal Authenticity ( $g = 2.48$ ) and Psychological Ownership ( $g = 2.63$ ) over the rule-based pathway. Three checks indicate that these gains did not come from a simple usability or effort difference: Process Clarity and Usability did not differ between the two pathways ( $p = 0.535$ ), and workstation duration and design-revision counts were comparable (Section 6.1). The added value therefore appears to lie in how dialogue structured the authoring process, helping participants articulate and revise intentions within protected limits. In this deployment, constraints carried a dual role: they scaffolded novice authoring, and they protected culturally significant forms. This dual role separates bounded creativity in heritage GenAI from constraint use in general creative-tool design.

Boundary encounters were not uniformly positive. Cultural-boundary redirections were logged in the GenAI-mediated condition, and the interview material indicates that some participants experienced redirections as interruptions when explanations were brief, or alternatives felt narrow (Section 6.2). Visible constraints are not automatically acceptable constraints. They require clear rationales and meaningful alternative paths.

The transferable contribution is a design pattern rather than a validated general principle. Heritage GenAI systems can make cultural limits attributable, explainable, and revisable instead of leaving them hidden in prompts, model behavior, or interface defaults. Whether this pattern holds outside single-institution visual heritage remains to be tested. In living, contested, sacred, or community-held heritage, accountability may need to involve artists, educators, tradition bearers, descendant communities, and visitors as well as museum curators [25–27].

### 7.3. Methodological and Practical Implications

The study shows why matched active controls matter in evaluations of generative AI. If the study had compared only conventional consumption with GenAI-mediated co-creation, all observed gains might have been attributed to AI. The rule-based pathway shows that much of the benefit came from the broader co-creation structure: source selection, choice-making, revision, and a personalized object. The C-versus-B comparison then identifies where the GenAI-mediated pathway added value beyond those shared elements. This is especially relevant to museum GenAI studies that compare an AI condition to a conventional baseline and leave open whether gains come from AI itself, from richer interaction, or from personalization more broadly [17,18].

The manipulation and diagnostic checks are important for the same reason. Participants recognized the GenAI-mediated condition as GenAI-mediated, and they experienced it as more creatively participatory. At the same time, Process Clarity and Usability did not differ between the two active conditions, providing no evidence that the GenAI-mediated gains were simply a usability effect. However, the C-versus-B comparison should still be interpreted as a pathway-level comparison, not as an isolated test of a model component. The GenAI-mediated pathway combined dialogue, generation, revision, source visibility, and boundary explanation.

For museums, the practical implication is that AI adoption should not be the first design question. A more useful starting point is whether visitors can begin from source material they personally encountered, whether they can see how their choices shape the output, whether cultural limits are understandable, and whether the final object preserves a traceable relation to the visit. AI is most useful when it supports capabilities that templates cannot easily provide: ordinary-language intent expression, adaptive revision, richer variation, and explainable redirection.

For interaction designers, error states and boundary messages should be treated as part of interpretation, not merely as system status. A redirected request, a failed output review, or a manufacturability warning can shape how visitors understand cultural limits, source linkage, and institutional reliability. For curators and museum educators, cultural-boundary work should be prepared before deployment and revised after use, rather than handled only through ad hoc review.

### 7.4. Limitations and Future Work

Several limitations bound the conclusions. First, the study was conducted in one museum and one heritage tradition. Lanting calligraphy has a specific visual language, institutional context, and visitor profile. Other domains, including sacred, contested, community-held, or non-visual heritage, may require different forms of source collection, cultural-boundary governance, and product design. The findings therefore support claims about one class of settings: single-institution visual heritage with clearly identifiable source objects and an on-site curatorial team. Application to history, art, or archaeological museums requires new boundary work and new validation, and we make no empirical claim beyond the studied setting.

Second, the field setting involved methodological trade-offs. The study did not include an a priori power analysis, full participant blinding, full counterbalancing across all day-time combinations, a formal external pilot, or pre-registration. The post hoc sensitivity analysis in Section 5.5 shows that the C-versus-B contrast could reliably detect effects of  $d \geq 0.91$  only. The observed C-versus-B effect for Heritage-Linked Authenticity was  $g = 0.53$  and falls below this threshold. The non-significant result should therefore be read as no reliable evidence of a large increment at this sample size, not as evidence that no smaller increment exists.

Third, the design comparisons estimate bundled experiences. The A-versus-(B+C) comparison captures a source-linked co-creation bundle rather than isolating source capture, post-tour authoring, personalization, physical delivery, or time-on-task. The C-versus-B comparison is also not a clean test of AI as a single technical feature, because the GenAI-mediated pathway increased perceived creative participation and made boundary explanations more visible. Future studies could separately manipulate source visibility, boundary explanation, degree of generative freedom, and physical-output format. An unrestricted-generation condition would directly test whether cultural boundaries improve heritage outcomes, which this study treats as a justified design premise rather than a tested claim.

Fourth, the outcome measures were adapted for this museum-takeaway context. The reliability and factor evidence in Section 5.3 indicates within-study coherence and satisfies the Fornell–Larcker criterion for discriminant validity. This evidence does not constitute full scale validation. Confirmatory factor analysis and measurement-invariance testing require larger samples and remain future work. Interview material was used illustratively, and formal inter-coder reliability was not computed. These limitations matter because the study measured visitor perceptions of engagement, authenticity, and ownership using adapted scales and interpretive interview evidence rather than fully validated heritage-specific instruments [34–40,49–51,59].

Finally, the main outcomes were measured immediately after the session. Longer-term work is needed to examine whether visitors keep, display, gift, discuss, or reinterpret the object over time. The provenance structure also supports output forms beyond 3D-printed objects. The same source-linked record could drive personalized video content, interpretive posters, digital collages, or condensed catalogues that document a visitor's selected works, their creators, and their historical context. These forms could extend post-visit engagement while keeping source traceability and accountable interpretation. Future work should also examine data governance, environmental cost, and multi-stakeholder rule ownership in deployments where cultural authority cannot rest with a museum team alone.

## 8. Conclusions

Heritage museums increasingly need ways to support visitors as active meaning-makers during the visit and to extend that experience through takeaways that remain connected to what visitors encountered. This paper examined how GenAI may support this goal without separating visitor authorship from cultural responsibility. We proposed bounded creativity and curatorial accountability as paired principles for heritage GenAI. Bounded creativity defines which aspects of a heritage-linked output visitors may vary and which culturally significant features should remain protected. Curatorial accountability defines how those limits are attributed, documented, implemented, explained, and revised.

We implemented these principles in a visit-linked digital–physical co-creation system at the Lanting Calligraphy Museum. The system connected on-site source collection, post-tour authoring, cultural-boundary checking, manufacturability validation, provenance documentation, and delivery of a personalized 3D-printed takeaway. In the field study, both rule-based and GenAI-mediated co-creation were associated with higher cultural engagement, perceived connection to heritage sources, personal meaning and self-expression, and psychological ownership than conventional consumption. The GenAI-mediated pathway was associated with higher cultural engagement, personal meaning and self-expression, and psychological ownership than the rule-based pathway, while its additional effect on perceived connection to heritage sources was smaller and not statistically reliable.

The broader conclusion, within the scope of this setting, is that GenAI delivered its value through accountable, source-linked co-creation: beginning from source material visitors personally encounter, giving them meaningful room to author and revise outputs,

keeping cultural boundaries visible and reviewable, and preserving a traceable relation between the museum encounter and the final physical takeaway. Whether unconstrained generation would produce similar gains was not tested. Heritage GenAI systems that share this visit-linked structure can treat boundary explanation, provenance, and institutional revision of cultural limits as part of the visitor experience rather than as secondary safeguards, and testing this design stance in other heritage domains is the next step.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/heritage9080295/s1>, Table S1: demographic and baseline characteristics by condition with balance tests; Table S2: item-level descriptive statistics, corrected item-total correlations, and single-factor loadings; Table S3: inter-scale correlations for the four primary outcomes; Table S4: representative excerpt from the cultural-boundary rule register; Figures S1–S3: step-by-step photographic documentation of the three condition flows in the field deployment, including the delivered takeaways.

**Author Contributions:** Conceptualization, S.J. and Z.L.; methodology, S.J., M.F. and Z.L.; software, S.J. and Z.L.; validation, S.J., M.F. and Z.L.; formal analysis, S.J.; investigation, S.J.; resources, S.J.; data curation, S.J.; writing—original draft preparation, S.J.; writing—review and editing, M.F. and Z.L.; visualization, S.J.; supervision, M.F. and Z.L.; funding acquisition, M.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Fundamental Research Funds for the Central Universities under Grant [CUC26XZ06].

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## Appendix A. Operational Protocol Details for the Field Study

Facilitator protocol. The same trained facilitator team handled all three conditions, following a written script for orientation, condition-specific instructions, and procedural steps. During the active creation phase, facilitators could answer technical questions but were instructed not to suggest source images, aesthetic choices, product categories, prompt wording, or revision directions. This non-intervention rule was used to keep authorship with participants and avoid asymmetric facilitator help across conditions.

AI-error caution. Because Condition C used a generative model, the facilitator script included a brief standard caution delivered to every Condition C participant before the workstation phase. The wording informed visitors that the dialogue agent could occasionally misinterpret a request, produce a candidate that did not match the intended description, or apply a cultural-boundary rule in a way the visitor would want to revise. Visitors were told that they were free to ask the system to try again, switch to a different source image, ignore an alternative, or ask the facilitator to intervene if the workstation appeared to be stuck. Condition B participants received a parallel statement that choices could be revised at any point before final confirmation.

## Appendix B. Codebook Excerpt for Illustrative Interview Themes

Table A1 provides a codebook excerpt for the interview material used in Section 6.2. The table reports theme definitions, evidence patterns, illustrative quotes, and links to interpretation. The interview material was not analyzed as a standalone qualitative study and formal inter-coder reliability was not computed.

**Table A1.** Codebook excerpt for illustrative interview themes.

| Theme                                                          | Definition and Evidence                                                                                                                                                                                                                                                              | Interpretive Link                                                               |
|----------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Co-created takeaways were authored objects                     | Participants described the object as something they helped make rather than simply received. B and C participants described selection, revision, and decision-making. Example: “I chose the photo, the frame, and the color” (B07).                                                  | Explains strong gains in Psychological Ownership.                               |
| Source collection supported attention and memory               | Visitor-captured source images helped connect the object to the museum encounter. Participants linked outputs to photos taken during the tour and described the Heritage Photo Library as part of the visit. Example: “I keep coming back to the photo I chose at the museum” (B14). | Supports B/C gains over A in Heritage-Linked Authenticity.                      |
| AI dialogue increased felt expression                          | Ordinary-language dialogue made personal intention more visible. C participants emphasized refinement, response, and iterative change. Example: “I kept changing the direction until it felt like mine” (C12).                                                                       | Explains C > B in Personal Authenticity and Ownership.                          |
| Templates were clear but more bounded                          | Rule-based co-creation supported control but offered fewer expressive routes. B participants valued clarity but noted limited expressiveness. Example: “It was easy to use, but the choices were already set” (B05).                                                                 | Explains why B improved over A but remained lower than C for personal outcomes. |
| Boundary explanations became guidance, with some interruptions | Boundary messages sometimes supported cultural learning but could interrupt flow. C participants described redirections as informative, while some wanted fuller alternatives. Example: “It helped me understand why the character shape should remain recognizable” (C02).          | Supports curatorial accountability while showing a design challenge.            |

## References

1. Hein, G.E. *Learning in the Museum*; Routledge: London, UK, 1998. <https://doi.org/10.4324/9780203028322>.
2. Falk, J.H.; Dierking, L.D. *Learning from Museums: Visitor Experiences and the Making of Meaning*; AltaMira Press: Walnut Creek, CA, USA, 2000.
3. Packer, J.; Ballantyne, R. Conceptualizing the Visitor Experience: A Review of Literature and Development of a Multifaceted Model. *Visit. Stud.* **2016**, *19*, 128–143. <https://doi.org/10.1080/10645578.2016.1144023>.
4. Petrelli, D.; Marshall, M.T.; O'Brien, S.; McEntaggart, P.; Gwilt, I. Tangible Data Souvenirs as a Bridge between a Physical Museum Visit and Online Digital Experience. *Pers. Ubiquitous Comput.* **2017**, *21*, 281–295. <https://doi.org/10.1007/s00779-016-0993-x>.
5. Liu, Y.; Chen, Y.; Font, X.; Zhang, Y.; Hu, J. Echoes of a Destination: The Representation of Locality in Souvenirs and Their Role in Sustaining Post-Trip Tourist Well-Being. *Tour. Manag.* **2025**, *108*, 105125. <https://doi.org/10.1016/j.tourman.2024.105125>.
6. Byrom, J.; Light, D.; Medway, D.; Parker, C.; Zenker, S. Post-Holiday Memory Work: Everyday Encounters with Fridge Magnets. *Ann. Tour. Res.* **2024**, *105*, 103724. <https://doi.org/10.1016/j.annals.2024.103724>.
7. Bekele, M.K.; Pierdicca, R.; Frontoni, E.; Malinverni, E.S.; Gain, J. A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. *ACM J. Comput. Cult. Herit.* **2018**, *11*, 7. <https://doi.org/10.1145/3145534>.
8. Tom Dieck, M.C.; Jung, T. A Theoretical Model of Mobile Augmented Reality Acceptance in Urban Heritage Tourism. *Curr. Issues Tour.* **2018**, *21*, 154–174. <https://doi.org/10.1080/13683500.2015.1070801>.
9. Petrelli, D.; Ciolfi, L.; van Dijk, D.; Hornecker, E.; Not, E.; Schmidt, A. Integrating Material and Digital: A New Way for Cultural Heritage. *Interactions* **2013**, *20*, 58–63. <https://doi.org/10.1145/2486227.2486239>.
10. 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. <https://doi.org/10.1016/j.ijhcs.2018.01.001>.
11. Petrelli, D.; O'Brien, S. Phone vs. Tangible in Museums: A Comparative Study. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2018; pp. 1–12. <https://doi.org/10.1145/3173574.3173686>.
12. Claisse, C.; Ciolfi, L.; Petrelli, D. Containers of Stories: Using Co-Design and Digital Augmentation to Empower the Museum Community and Create Novel Experiences of Heritage at a House Museum. *Des. J.* **2017**, *20*, S2906–S2918. <https://doi.org/10.1080/14606925.2017.1352801>.
13. Spennemann, D.H.R. Generative Artificial Intelligence, Human Agency and the Future of Cultural Heritage. *Heritage* **2024**, *7*, 3597–3609. <https://doi.org/10.3390/heritage7070170>.

14. Witte, J.; Lee, E.; Brausem, L.; Shillabeer, V.; Bonacchi, C. Generative AI in Heritage Practice: Improving the Accessibility of Heritage Guidance. *Heritage* **2025**, *8*, 513. <https://doi.org/10.3390/heritage8120513>.
15. Rezwana, J.; Maher, M.L. Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *ACM Trans. Comput.-Hum. Interact.* **2023**, *30*, 45. <https://doi.org/10.1145/3519026>.
16. Shneiderman, B. Creativity Support Tools: Accelerating Discovery and Innovation. *Commun. ACM* **2007**, *50*, 20–32. <https://doi.org/10.1145/1323688.1323689>.
17. Wang, H.; Matviienko, A. Experiencing Art Museum with a Generative Artificial Intelligence Chatbot. In *Proceedings of the 2025 ACM International Conference on Interactive Media Experiences*; ACM: New York, NY, USA, 2025; pp. 430–436. <https://doi.org/10.1145/3706370.3731650>.
18. Wang, Z.; Yuan, L.P.; Wang, L.; Jiang, B.; Zeng, W. VirtuWander: Enhancing Multi-Modal Interaction for Virtual Tour Guidance through Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2024; p. 612. <https://doi.org/10.1145/3613904.3642235>.
19. 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. <https://doi.org/10.1038/s40494-025-02164-1>.
20. Foka, A.; Griffin, G. AI, Cultural Heritage, and Bias: Some Key Queries That Arise from the Use of GenAI. *Heritage* **2024**, *7*, 6125–6136. <https://doi.org/10.3390/heritage7110287>.
21. Menotti, G. The Model Is the Museum: Generative AI and the Expropriation of Cultural Heritage. *AI Soc.* **2025**, *40*, 5593–5597. <https://doi.org/10.1007/s00146-025-02290-1>.
22. National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*; NIST AI 100-1; NIST: Gaithersburg, MD, USA, 2023. <https://doi.org/10.6028/NIST.AI.100-1>.
23. Mitchell, M.; Wu, S.; Zaldivar, A.; Barnes, P.; Vasserman, L.; Hutchinson, B.; Spitzer, E.; Raji, I.D.; Gebru, T. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*; ACM: New York, NY, USA, 2019; pp. 220–229. <https://doi.org/10.1145/3287560.3287596>.
24. Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J.W.; Wallach, H.; Daume, H., III; Crawford, K. Datasheets for Datasets. *Commun. ACM* **2021**, *64*, 86–92. <https://doi.org/10.1145/3458723>.
25. Smith, L. *Uses of Heritage*; Routledge: London, UK, 2006. <https://doi.org/10.4324/9780203602263>.
26. Simon, N. *The Participatory Museum*; Museum 2.0: Santa Cruz, CA, USA, 2010.
27. Harrison, R. *Heritage: Critical Approaches*; Routledge: London, UK, 2013. <https://doi.org/10.4324/9780203108857>.
28. He, Z.; Su, J.; Chen, L.; Wang, T.; Lc, R. “I Recall the Past”: Exploring How People Collaborate with Generative AI to Create Cultural Heritage Narratives. *Proc. ACM Hum.-Comput. Interact.* **2025**, *9*, CSCW108. <https://doi.org/10.1145/3711006>.
29. Moldovan, S.; Moldovan, I.; Rice, T. Speculative Memory and Machine Augmentation: A Polyvocal Rendering of Brutalist Architecture Through AI and Photogrammetry. *Heritage* **2025**, *8*, 401. <https://doi.org/10.3390/heritage8100401>.
30. Harist, R.E.; Fong, W.C. *The Embodied Image: Chinese Calligraphy from the John B. Elliott Collection*; Princeton University Art Museum: Princeton, NJ, USA, 1999.
31. Ledderose, L. *Mi Fu and the Classical Tradition of Chinese Calligraphy*; Princeton University Press: Princeton, NJ, USA, 1979.
32. Nakata, Y.; Hunter, J. *Chinese Calligraphy*; Weatherhill: New York, NY, USA; Tankosha: Tokyo, Kyoto, 1983.
33. Li, W. *Chinese Writing and Calligraphy*; University of Hawai'i Press: Honolulu, HI, USA, 2009. <https://doi.org/10.1515/9780824860691>.
34. Wang, N. Rethinking Authenticity in Tourism Experience. *Ann. Tour. Res.* **1999**, *26*, 349–370. [https://doi.org/10.1016/S0160-7383\(98\)00103-0](https://doi.org/10.1016/S0160-7383(98)00103-0).
35. 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. <https://doi.org/10.1016/j.tourman.2009.07.010>.
36. Rickly-Boyd, J.M. Authenticity and Aura: A Benjaminian Approach to Tourism. *Ann. Tour. Res.* **2012**, *39*, 269–289. <https://doi.org/10.1016/j.annals.2011.05.003>.
37. Fu, X. Existential Authenticity and Destination Loyalty: Evidence from Heritage Tourists. *J. Destin. Mark. Manag.* **2019**, *12*, 84–94. <https://doi.org/10.1016/j.jdmm.2019.03.008>.
38. Pierce, J.L.; Kostova, T.; Dirks, K.T. Toward a Theory of Psychological Ownership in Organizations. *Acad. Manag. Rev.* **2001**, *26*, 298–310. <https://doi.org/10.5465/AMR.2001.4378028>.
39. Van Dyne, L.; Pierce, J.L. Psychological Ownership and Feelings of Possession: Three Field Studies Predicting Employee Attitudes and Organizational Citizenship Behavior. *J. Organ. Behav.* **2004**, *25*, 439–459. <https://doi.org/10.1002/job.249>.

40. Peck, J.; Shu, S.B. The Effect of Mere Touch on Perceived Ownership. *J. Consum. Res.* **2009**, *36*, 434–447. <https://doi.org/10.1086/598614>.
41. Stokes, P.D. *Creativity from Constraints: The Psychology of Breakthrough*; Springer Publishing Company: New York, NY, USA, 2005.
42. Acar, O.A.; Tarakci, M.; van Knippenberg, D. Creativity and Innovation under Constraints: A Cross-Disciplinary Integrative Review. *J. Manag.* **2019**, *45*, 96–121. <https://doi.org/10.1177/0149206318805832>.
43. Cummiskey, B.M.; Baer, M. Bounded Creativity: A Theory of the Effects of Constraints on the Production of New Ideas. *Acad. Manag. Proc.* **2018**, *2018*, 15243. <https://doi.org/10.5465/AMBPP.2018.257>.
44. Han, L.; Gao, Y.; Lu, X.; Wang, B.; Zhang, L.; Wang, Z.; Yip, D. Gen-Diaolou: An Integrated AI-Assisted Interactive System for Diachronic Understanding and Preservation of the Kaiping Diaolou. In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), Barcelona, Spain, 13–17 April 2026. <https://doi.org/10.1145/3772318.3790720>.
45. Derda, I.; Predescu, D. Towards Human-Centric AI in Museums: Practitioners' Perspectives and Technology Acceptance of Visitor-Centered AI for Value (Co-)Creation. *Mus. Manag. Curatorship* **2025**, *40*, 532–554. <https://doi.org/10.1080/09647775.2025.2467703>.
46. Zimmerman, J.; Forlizzi, J.; Evenson, S. Research through Design as a Method for Interaction Design Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2007; pp. 493–502. <https://doi.org/10.1145/1240624.1240704>.
47. Baskerville, R.; Kaul, M.; Storey, V. Genres of Inquiry in Design-Science Research: Justification and Evaluation of Knowledge Production. *MIS Q.* **2015**, *39*, 541–564. <https://doi.org/10.25300/MISQ/2015/39.3.02>.
48. Krueger, R.A.; Casey, M.A. *Focus Groups: A Practical Guide for Applied Research*, 5th ed.; SAGE: Thousand Oaks, CA, USA, 2015.
49. Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
50. Braun, V.; Clarke, V.; Hayfield, N.; Terry, G. Thematic Analysis. In *Handbook of Research Methods in Health Social Sciences*; Liamputtong, P., Ed.; Springer: Singapore, 2019; pp. 843–860. [https://doi.org/10.1007/978-981-10-5251-4\\_103](https://doi.org/10.1007/978-981-10-5251-4_103).
51. King, N. Doing Template Analysis. In *Qualitative Organizational Research: Core Methods and Current Challenges*; Symon, G., Cassell, C., Eds.; SAGE: London, UK, 2012; pp. 426–450. <https://doi.org/10.4135/9781526435620.n24>.
52. OpenAI. GPT-4o System Card. *arXiv* **2024**, arXiv:2410.21276. <https://doi.org/10.48550/arXiv.2410.21276>.
53. Podell, D.; English, Z.; Lacey, K.; Blattmann, A.; Dockhorn, T.; Müller, J.; Penna, J.; Rombach, R. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis. *arXiv* **2023**, arXiv:2307.01952. <https://doi.org/10.48550/arXiv.2307.01952>.
54. Hu, E.J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual, 25–29 April 2022.
55. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
56. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning Transferable Visual Models from Natural Language Supervision. In Proceedings of the 38th International Conference on Machine Learning (ICML), Virtual, 18–24 July 2021; pp. 8748–8763.
57. Tochilkin, D.; Pankratz, D.; Liu, Z.; Huang, Z.; Letts, A.; Li, Y.; Liang, D.; Laforte, C.; Jampani, V.; Cao, Y.-P. TripoSR: Fast 3D Object Reconstruction from a Single Image. *arXiv* **2024**, arXiv:2403.02151. <https://doi.org/10.48550/arXiv.2403.02151>.
58. Dawson-Haggerty, M. trimesh: Python Library for Loading and Using Triangular Meshes. Available online: <https://trimsh.org> (accessed on 6 July 2026).
59. O'Brien, H.L.; Cairns, P.; Hall, M. A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. *Int. J. Hum.-Comput. Stud.* **2018**, *112*, 28–39. <https://doi.org/10.1016/j.ijhcs.2018.01.004>.
60. Fornell, C.; Larcker, D.F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* **1981**, *18*, 39–50. <https://doi.org/10.2307/3151312>.
61. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 4th ed.; Guilford Press: New York, NY, USA, 2016.
62. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
63. Lakens, D. Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs. *Front. Psychol.* **2013**, *4*, 863. <https://doi.org/10.3389/fpsyg.2013.00863>.
64. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall/CRC: New York, NY, USA, 1993. <https://doi.org/10.1201/9780429246593>.

---

**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.