

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/foarSOUTHEAST
UNIVERSITY

RESEARCH ARTICLE

ArchiBuilder: A virtual reality-based gamified experience design for architectural images in Dunhuang murals

Zhizheng Liu, Xiran Ji, Yizhu Cui, Tianzi Zhang, Jian Yao*



School of Design and Arts, Beijing Institute of Technology, Beijing 102488, China

KEYWORDS

Cultural heritage;
Historical
architectural images;
Dunhuang murals;
Virtual reality;
Serious games

Abstract Historical architectural images serve as vital carriers of architectural heritage, preserving visual information about structural forms. However, due to the complexity of architectural knowledge, the general public often struggles to intuitively comprehend the spatial and structural characteristics embedded in these images. With the advancement in digital technology, virtual reality serious games (VR SG) offer new opportunities to enhance user engagement and understanding. This study presents ArchiBuilder, a VR SG system developed based on architectural images in Dunhuang murals, aimed at fostering immersive learning and exploring its potential in public education. A between-subjects experiment ($N = 30$) using a mixed-methods approach was conducted to compare ArchiBuilder with a conventional VR image-text information mode in terms of learning performance, immersion, and cognitive load. The results indicate that ArchiBuilder significantly reduced users' cognitive load and improved immersion, although no statistically significant difference was observed in learning performance. These findings demonstrate the effectiveness of gamified VR experiences in enhancing user experience. Through narrative-driven mechanisms and staged reconstruction tasks, ArchiBuilder reshapes how users cognitively engage with architectural images. While learning outcomes remained comparable, the interactive scenario design improved learning strategies and sustained user engagement. This study offers a practical framework for the digital communication of historical architectural images and provides insights for public-oriented cultural heritage education.

© 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: yaojian@bit.edu.cn (J. Yao).

Peer review under the responsibility of Southeast University.

1. Introduction

Historical architectural images refer to visual records of historical architectural forms, including plans, elevations, sections, paintings, and photographs (Moyano et al., 2022). These images document components, spatial layouts, and decorative features, typically sourced from historical archives, survey data, or site records (Pietroni and Ferdani, 2021). As part of the UNESCO World Cultural Heritage, Dunhuang Mogao Caves preserve murals that vividly depict the structural forms and spatial configurations of Tang dynasty wooden architecture and earlier. These include detailed renderings of dougong (bracket sets), beams and columns, and roof types (Han, 2022). With most physical wooden structures from that era no longer extant, architectural images in Dunhuang murals serve as vital resources for studying traditional Chinese architecture and construction techniques. They also provide one of the few accessible channels for the public to engage with ancient architectures (Wang and Qiu, 2014).

In recent years, advances in 3D modeling technology have enabled researchers to digitally reconstruct architectural images in Dunhuang murals, aiming to enhance public understanding and cultural dissemination (Wang and Zhang, 2022; Zhang and Wang, 2018). However, most existing efforts focus primarily on structural modeling and visual display, offering limited interactivity and immersion. Users often remain passive recipients of information, hindering deeper engagement with the architectural logic and cultural context embedded in the images. This raises a critical question: how can interactive design be leveraged to foster public interest and cultural resonance in the interpretation of historical architectural images?

Virtual reality serious games (VR SG), which integrate virtual reality (VR) technologies with gamification, provide innovative means for digital heritage dissemination (Ting and Min, 2025). By transforming complex cultural contexts into interactive experiences, VR SG significantly improves user engagement in heritage contexts (Capecchi et al., 2024). Currently, VR SG has been widely applied in areas such as historical event reenactments, cultural knowledge transmission, and architectural reconstruction (DaCosta and Kinsell, 2022; Kara, 2024). However, systematic design and empirical research remain scarce when it comes to applications targeting architectural images in Dunhuang murals.

To address the current research gap, this study focuses on the architectural images of the front hall of the central main halls on the north wall of Mogao Cave 172. Based on this prototype, we designed and developed ArchiBuilder, a gamified experience system grounded in VR SG. Centered on a participatory task mechanism, the system guides users through three game stages—basic component extraction, core structure assembly, and general architecture reconstruction—aiming to deepen their understanding of the spatial logic and cultural meanings embedded in the murals.

Based on this design framework, the study addresses the following research questions:

RQ1: Can an interactive gamified experience effectively enhance users' understanding and cultural cognition of architectural images in Dunhuang murals?

RQ2: Compared to conventional VR image-text information, does ArchiBuilder offer a significantly improved user experience?

To systematically examine the above research questions, the following hypotheses are proposed:

H1: Through scene design, task-driven guidance, and knowledge integration, ArchiBuilder can enhance users' understanding and cultural cognition of architectural images in Dunhuang murals.

H2: Compared to conventional VR image-text experiences, ArchiBuilder significantly improves users' learning performance, immersion, and reduces cognitive load.

Accordingly, this study follows a “design—implementation—evaluation” workflow. It first conducts a systematic review of prior research on Dunhuang architectural images and digital cultural experiences. Based on the architectural images from the north wall of Mogao Cave 172, a structured knowledge framework and interactive task model are developed. The system prototype is then built using 3D modeling tools and Unreal Engine. A between-subjects experiment is conducted, combining quantitative and qualitative methods to evaluate user performance across different experience modes. Finally, design insights and future research directions are proposed, offering methodological support and empirical evidence for the public communication and educational application of historical architectural images.

2. State of the art

2.1. Architectural images in Dunhuang murals

The Mogao Caves in Dunhuang, first constructed during the Sixteen Kingdoms period (366 CE), systematically document the historical development of religion, art, and society in ancient China through their extensive murals. Due to the perishable nature of timber architecture, physical remains from the Tang dynasty and earlier are extremely scarce, with only a few surviving examples such as the main halls of Foguang Temple, Nanchan Temple, and Guangrenwang Temple. As visual records of ancient wooden architecture, the architectural images in Dunhuang murals offer valuable evidence for studying the formal aesthetics, spatial organization, and functional characteristics of early Chinese architecture (Yao and Yu, 2023).

Current research on Dunhuang mural architectural images mainly focuses on three areas: (1) **Cultural and historical value:** Scholars analyze decorative elements and component styles to reveal the symbolic meanings embedded in the architecture—such as religious iconography, social hierarchy, and cultural ideologies. For example, Sun (2020) explored how ridge-end tiles reflect symbolic variations across political systems; Jin (2022) examined imperial architectural motifs in murals to trace the evolution of Ming and Qing palace architecture; Zhang and Pei (2025) applied typological analysis to classify mural architecture based on both explicit and implicit visual features. (2) **Structural analysis and reconstruction:** This stream emphasizes identifying architectural components, restoring construction methods, and analyzing spatial forms. Meng et al. (2021) outlined the key interface elements in

corridor structures; Geng and Tang (2024) used statistical analysis and image tracing to extract standardized structural patterns from 77 mural depictions of city walls; and Wang (2024) digitally reconstructed Buddhist temple courtyards by integrating multisource images. (3) **Information presentation and education:** Studies in this area explore how digital tools enhance the communication of architectural images. Mu et al. (2024) developed VR tours based on 2D mural slices to present spatial layouts of Cave 061; Xu et al. (2022) used 3D modeling and printing to reconstruct the spatial structure of Cave 172, enabling more tangible user interaction beyond 2D images. An overview of these research directions is shown in Fig. 1.

In summary, while current research has made initial progress in image interpretation, structural modeling, and digital visualization, most efforts remain expert-driven, focusing on visual restoration and information presentation. There is a noticeable lack of participatory experience mechanisms tailored to the general public. Therefore, the next section will further examine the development of digital experience approaches in this field, highlighting their potential for cultural dissemination and public education.

2.2. Digital experiences of historical architectural images

With rapid advances in technology, digital experiences have become essential tools for conserving and disseminating cultural heritage. Compared to static 2D presentations, digital technologies demonstrate significant potential for enhancing public engagement with historical architectural images. Current research primarily focuses on three key areas: **3D reconstruction, information visualization, and interactive display.**

First, in **3D reconstruction**, researchers extract 2D data from historical images to construct 3D models that restore proportions, structures, and spatial layouts. For instance, the Yuanmingyuan digital project reconstructed the Wanfang Anhe Pavilion using historical drawings and survey data, combining modeling and texture mapping to visualize both exterior and interior details (Chen and del Blanco García, 2022). Second, **information visualization**, improves the accessibility of complex architectural data. Zhang et al. (2022) developed a multi-stage visualization system for Shanxi’s Qinglian Temple using knowledge graphs, VR navigation, and mobile platforms to convey temporal, structural, and decorative information. Third, **interactive display technologies** enable users to actively explore architectural knowledge. Dewitz et al. (2019) combined AR with mobile apps and 3D-printed models to let users interact with historical buildings in Dresden, enhancing engagement through spatial exploration.

While these approaches enhance spatial understanding and knowledge acquisition, most remain focused on structural presentation, offering limited interaction. High cognitive barriers persist, especially in grasping architectural logic and techniques. With the emergence of VR SG, research shows promising potential for improving learning outcomes (Liu et al., 2024). The next section explores the role of VR SG in advancing digital engagement with historical architecture.

2.3. Historical architecture and virtual reality serious games

VR SG integrates narrative storytelling with task-based interaction, offering innovative approaches for digital engagement with historical architecture. Current VR SG

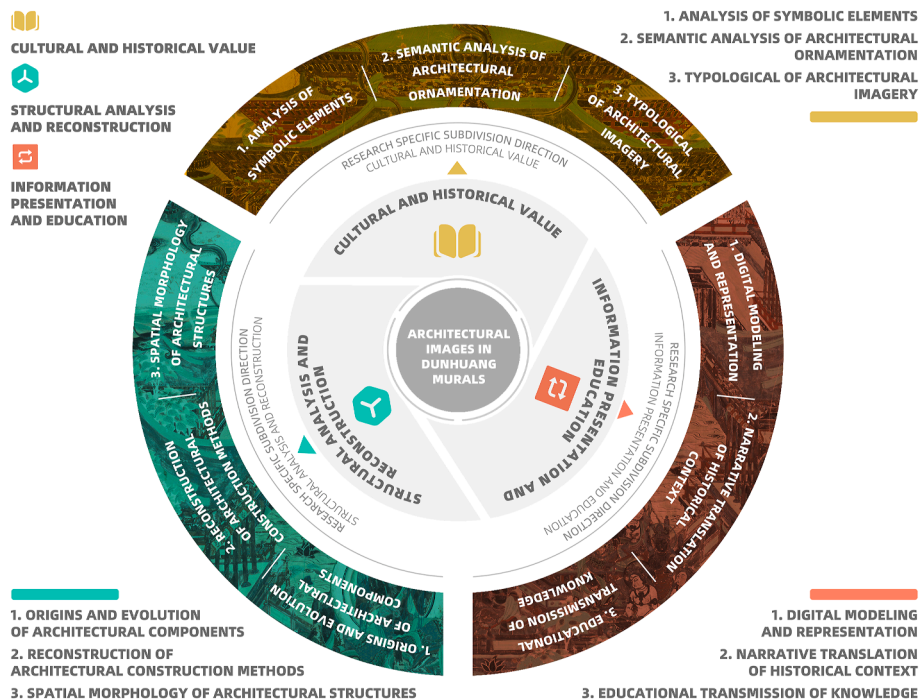


Fig. 1 Overview of research directions on Dunhuang mural architectural images.

research on historical architecture focuses on three aspects: **construction simulation**, **spatial exploration**, and **narrative event experiences**.

First, **construction simulation** enables users to assemble architectural components and perform spatial tasks, supporting intuitive learning. For instance, the Temple Lego project reconstructed an ancient temple complex using CAD data and integrated it into Unreal Engine. The system uses time limits, controller input, and score-based feedback to guide users through sequential assembly tasks, simulating a participatory restoration process (Maji et al., 2024). Second, **spatial exploration** allows users to navigate reconstructed environments and understand architectural logic. Ferdani et al. (2020) employed photogrammetry and BIM techniques to rebuild the Forum of Augustus, using Agisoft Photoscan and Unreal Engine. The experience, optimized with LOD (levels of detail) and mapping techniques, provided users with a “learning-by-doing” journey through Roman architecture. Third, **narrative event experiences** increase emotional involvement through virtual storytelling. In the ArkaeVision project, users followed a virtual guide, “Ariadne,” through ritual scenes in the Temple of Hera II. Built in Unreal Engine with motion capture and character modeling, the design enhanced cultural understanding and engagement (Pagano et al., 2020).

In summary, VR SG in historical architecture typically follows three experiential paths—constructive, spatial, and narrative—and utilizes tools like CAD, BIM, and real-time rendering. Building on this foundation, this study develops ArchiBuilder, a VR SG system based on Dunhuang mural architectural images. It explores how gamified interaction can improve public learning performance, immersion, and cognitive efficiency, offering a novel approach for cultural heritage experiences.

3. ArchiBuilder design process

3.1. Knowledge content

To support knowledge representation in ArchiBuilder, six representative architectural elements were selected for content development and analysis (Zhou et al., 2024).

Platform base (taiji): Serving as the architectural foundation, the platform stabilizes the structure and prevents ground moisture damage (Liu, 2018). Dunhuang murals depict bases in rectangular or square forms, often built from stone or rammed earth and adorned with decorative lines. Its dimensions influence both the structural stability and visual proportions of the architecture.

Column: Columns are key load-bearing components that shape spatial stability and rhythm (Hao et al., 2024). In the murals, columns often appear with beam frames and roofs to form a cohesive support system. Variations in column shape, size, and arrangement affect both the openness and balance of the architecture.

Main structure (wushen): This is the core volume of the architecture, often enclosed by columns and beams (Zou and Bahauddin, 2024). Its form considers function, symbolism, and environmental factors such as lighting and ventilation, reflecting a holistic design approach.

Dougong (bracket set): Positioned between columns and roof eaves, dougong serves a dual role in structure and decoration (Wu et al., 2022). Murals depicted it in multi-layered forms, with single-step or double-step types indicating hierarchy and building function. It exemplifies both mechanical ingenuity and artistic refinement.

Beam frame (liangjia): The beam frame forms the structural skeleton of the roof system, enabling large interior spans without metal fasteners (Zhang et al., 2024). In Dunhuang murals, they are precisely rendered with overlapping elements, showcasing the mortise-and-tenon craftsmanship of traditional timber construction.

Roof: Roofs are the most visually prominent architectural elements. Common types in murals include hip roof (wudian ding), gable-and-hip roof (xie shan ding), and overhanging gable roof (xuan shan ding) (Yuk et al., 2023). Their layered tiles and sweeping curves reflect both structural function and regional aesthetic styles.

These six elements form the core knowledge foundation in ArchiBuilder, guiding the structuring of interactive tasks and educational content. A summary is provided in Table 1.

3.2. Design process

Based on the architectural knowledge system, ArchiBuilder centers on gamified learning, employing task-driven progression and interactive exploration to guide players in gradually mastering architectural elements, structural logic, and cultural context. The knowledge framework is presented in Table 2.

The design follows a progressive learning path—cognition, practice, and immersion—and integrates three gameplay levels: **Knowledge introduction**, **Component assembly**, and **Complete reconstruction**. This structure forms a recursive, incremental model that guides users from basic recognition to hands-on construction and contextual understanding. In the following section, we detail how ArchiBuilder transforms this design process into interactive experiences that enhance users’ architectural comprehension.

4. ArchiBuilder design implementation

4.1. 3D reconstruction of architectural image in Dunhuang murals

Mogao Cave 172, constructed during the High Tang period (705–781 CE), exemplifies Tang dynasty architectural styles as depicted in Dunhuang murals (Zhou and Li, 2024). The north wall murals provide a detailed portrayal of courtyard layouts and wooden structural elements, highlighting the integration of architecture and Tang culture. This study selected the image of the front hall of the central main halls on the north wall as the case study for the implementation of ArchiBuilder.

4.1.1. Image analysis and multi-source information integration

Due to the planar nature of Dunhuang mural images, which lack proportional annotations and detailed structural data, directly reconstructing 3D forms poses significant

Table 1 Architectural elements and knowledge content in Dunhuang mural architectural images. Image source: Authors' redrawing based on images from the Digital Dunhuang website (Dunhuang Academy).

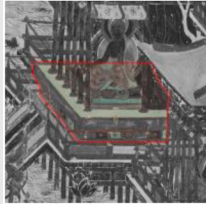

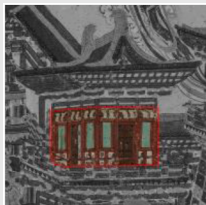
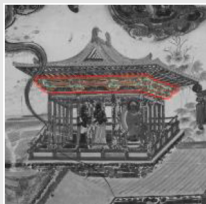
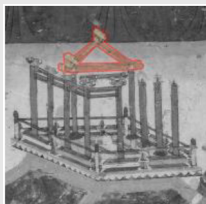

Architectural element	Image	Knowledge content	Knowledge description
Platform base		Foundation support	Learn about the forms and construction process of platform bases.
Column		Spatial division	Learn about column structure and load-bearing function.
Main structure		Spatial enclosure	Learn about the assembly and spatial configuration of the main structure.
Dougong		Load-bearing structure	Learn the composition of dougong and its mortise-and-tenon joints.
Beam frame		Structural connection	Learn how to assemble and construct beam frames.
Roof		Top coverage	Learn about the artistic forms and stylistic features of traditional roofs.

Table 2 ArchiBuilder game framework and design concepts.

Knowledge content	Game framework	Design concept description
Architectural image introduction	Narrative onboarding	Create a story-driven entry to guide users in recognizing architectural images.
Cultural background knowledge	Immersive exploration	Enable free exploration of Dunhuang caves and murals to access historical context.
Assembly logic knowledge	Task-based learning	Guide users from basic parts to complex assemblies to grasp modular design principles.
Construction process knowledge	Step-by-step building	Use lighting, demos, and real-time interaction for hands-on building experience.
General reconstruction knowledge	Integrated restoration tasks	Involve users in full-scale reconstruction to deepen understanding and cultural connection.
Historical and artistic insights	Cultural narrative and free roaming	Allow post-reconstruction exploration to experience architectural artistry immersively.

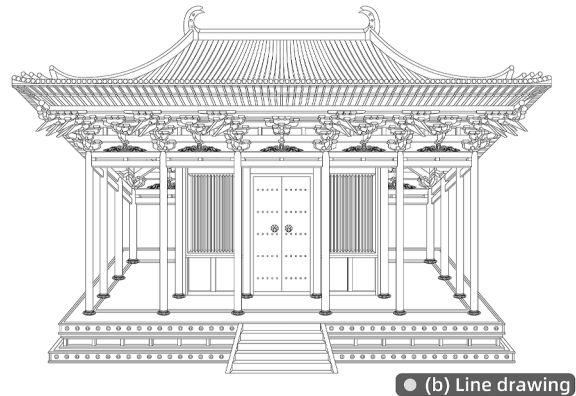


Fig. 2 Architectural image in Dunhuang mural and corresponding line drawing. (a) Architectural image in the mural; (b) Line drawing. Source: Authors' redrawing based on images from the Digital Dunhuang website (Dunhuang Academy).

challenges. To address this, the study employed high-resolution image analysis combined with multi-source information integration.

First, based on the digital image of the north wall in Mogao Cave 172 (Fig. 2(a)), the primary forms and spatial layout of architectural components were identified. Line-drawing techniques were applied to extract and segment key elements from the mural, clarifying their hierarchical relationships and spatial organization (Fig. 2(b)).

Second, architectural practices in the Tang and Song dynasties followed consistent construction principles. By referencing historical texts such as the *Yingzao Fashi* (Song dynasty) and empirical data from surviving Tang structures, the study supplemented and verified the mural's architectural forms and proportions (Shi et al., 2024). Using the modular system documented in the *Yingzao Fashi*, along with Tang-era construction dimensions, the research team inferred key measurements for reconstruction.

The "modular system based on cai, qi, and fen" (材分模数制) outlines precise standards for timber architecture. In this system, cai (材) and qi (槩) refer to cross-sectional units, while fen (分) is a linear measurement (Bao, 2024). The section "Timber Specifications for Major Constructions (Damuzuo Zhidu Yi · Cai)" records eight size levels, as shown

in Table 3 (Wells and Xue, 2024). Based on the line drawing analysis of the north wall's front hall, the structure was determined to have five bays across and three in depth, corresponding to the third-class timber level in the standard.

To reconstruct the spatial layout and component dimensions of Tang dynasty timber architecture, we derived scale data using third-class timber standards and Tang-era fen° values. Historical sources indicate that one chi (尺) equaled approximately 300 mm, and one fen° measured about 21 mm (Zhang et al., 2007). The modular system defines cai height as 15 fen° (315 mm), cai width as 210 mm, qi as 6 fen° (126 mm) and 84 mm wide, with fen° as the base unit.

These values serve as the dimensional basis for reconstructing the front hall's proportions and component sizes on the north wall of Mogao Cave 172.

4.1.2. Dimensional derivation and orthographic drawing

Based on the modular standards outlined above, we conducted dimensional derivation for the architectural components, referencing empirical data from preserved Tang dynasty architectures such as the main hall of Foguang Temple and Nanchan Temple (Xiao, 2017). The derivation covered overall width and depth (number of bays), vertical

Table 3 Modular system of Cai, Qi, and Fen° in the Yingzao Fashi.

Timber class	Cai (chi) Height × width	Qi (chi) Height × width	Fen° (chi) 1/10 width	Typical application
First class	0.9 × 0.6	0.36 × 0.24	0.06	Main halls with 9–11 bays
Second class	0.825 × 0.55	0.33 × 0.22	0.055	Halls with 5–7 bays
Third class	0.75 × 0.5	0.3 × 0.2	0.05	Halls with 3–5 bays; large halls with 7 bays
Fourth class	0.72 × 0.48	0.288 × 0.192	0.048	Halls with 3 bays; medium halls with 5 bays
Fifth class	0.66 × 0.44	0.264 × 0.176	0.044	Small 3-bay halls; large 3-bay pavilions
Sixth class	0.6 × 0.4	0.24 × 0.16	0.04	Pavilions and small halls
Seventh class	0.525 × 0.35	0.21 × 0.14	0.035	Small pavilions and temples
Eighth class	0.45 × 0.3	0.18 × 0.12	0.03	Small pavilions and caisson ceilings

Table 4 Dimensional derivation of architectural images in Dunhuang murals.

Element name	Modular rule description	Derived size (fen°)	Approximate size (mm)
Total width	5 bays: $2 \times 210 + 3 \times 240$ fen° (Yu et al., 2024)	1140	23940
Central bay width	Referencing Foguang temple east hall: 240 fen° (Yu et al., 2024)	240	5040
Secondary bay width	Referencing Foguang temple east hall: 240 fen° (Yu et al., 2024)	240	5040
End bay width	Referencing Foguang temple east hall: 210 fen° (Yu et al., 2024)	210	4410
Total depth	3 bays: 3×210 fen° (Yu et al., 2024)	630	13230
Central depth bay	Referencing Foguang temple east hall: 210 fen° (Yu et al., 2024)	210	4410
End depth bay	Referencing Foguang temple east hall: 210 fen° (Yu et al., 2024)	210	4410
Column height	Not exceeding bay width: 240 fen° (Wang, 2017)	240	5040
Column diameter	2 cai + 2 qi to 3 cai: 42–45 fen° (Wang, 2017)	45	945
Dougong layer height	Referencing Foguang temple east hall: 117 fen° (Zhang et al., 2018)	117	2457
Beam frame height	$11 \times$ cai (21 fen°): 231 fen° (Yu et al., 2024)	231	4851
Dougong spacing	Center spacing: 125 ± 25 fen° (Wang, 2017)	120	2520
Purlin length	Equal to bay width: 240 fen° (Du and Chen, 1977)	240	5040
Purlin diameter	1 cai + 1 qi or more: 21–30 fen° (Du and Chen, 1977)	30	630
Rafter length	Not exceeding 1/2 purlin: 120 fen° (Du and Chen 1977)	120	2520
Rafter diameter	In hall architectures: 9–10 fen° (Du and Chen, 1977)	10	210
...
Column rise (shengqi)	Rise of 5 fen° per bay from center: 10 fen° (Chen and Wang, 2016)	10	210

Note: 1 fen° ≈ 21 mm; all metric values in parentheses are approximate.

section heights (columns, dougong, beam frame), spacing of dougong sets, and sizes of purlins and rafters.

By aligning mural proportions with known construction ratios, we iteratively refined and confirmed component dimensions. These values were translated into detailed AutoCAD drawings, integrating modular rules, historical architecture data, and prior research (Xu, 2015; Zhou, 2025). The finalized dimensions are presented in Table 4.

Based on these results, we created the plan (Fig. 3(a)) and elevation (Fig. 3(b)) drawings. The plan outlines spatial layout and proportions of the front hall, while the elevation shows component hierarchies and external form. Together, these drawings provide a precise geometric basis for accurate 3D reconstruction.

4.1.3. 3D digital modeling and texture mapping

Based on the derived dimensions and geometric relationships, modular 3D modeling was completed in SketchUp

2020. The component was assembled following structural logic to form a full 3D model of the front hall.

Architectural textures and decorative details were extracted from high-resolution mural images using Photoshop 2020. To address weathering and partial loss in the original murals, missing color and material features were digitally restored to reflect the artistic style and symbolism of High Tang architecture.

The final 3D model of the front hall in Mogao Cave 172 is shown in Fig. 4, including a front view (Fig. 4(a)) and an axial view (Fig. 4(b)).

4.2. ArchiBuilder scene, narrative, and task design

4.2.1. Scene design

ArchiBuilder includes two main scenes: the reconstructed Mogao Cave 172 from the High Tang period (Fig. 5(a)) and a stylized Buddhist world based on its mural (Fig. 5(b)). These

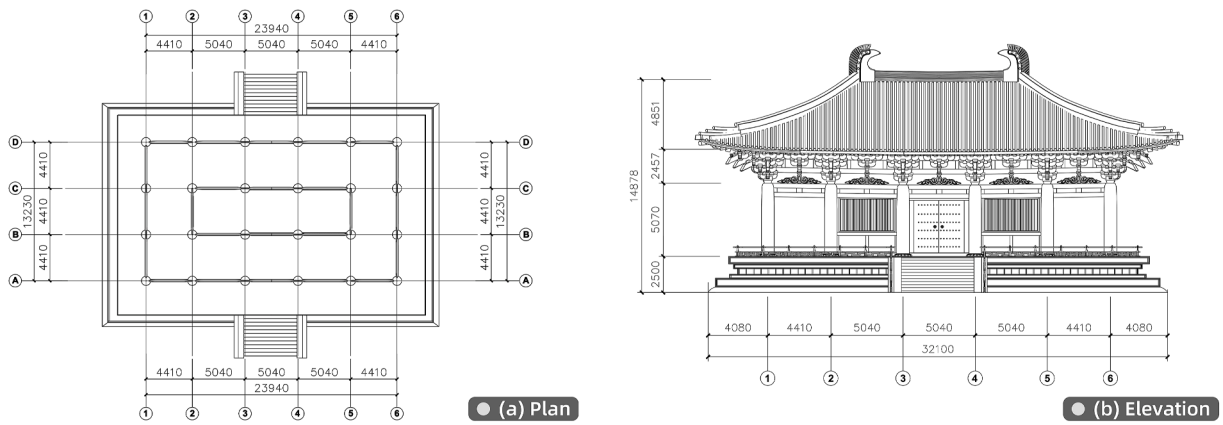


Fig. 3 CAD drawings of architectural images in Dunhuang murals: (a) plan; (b) elevation.

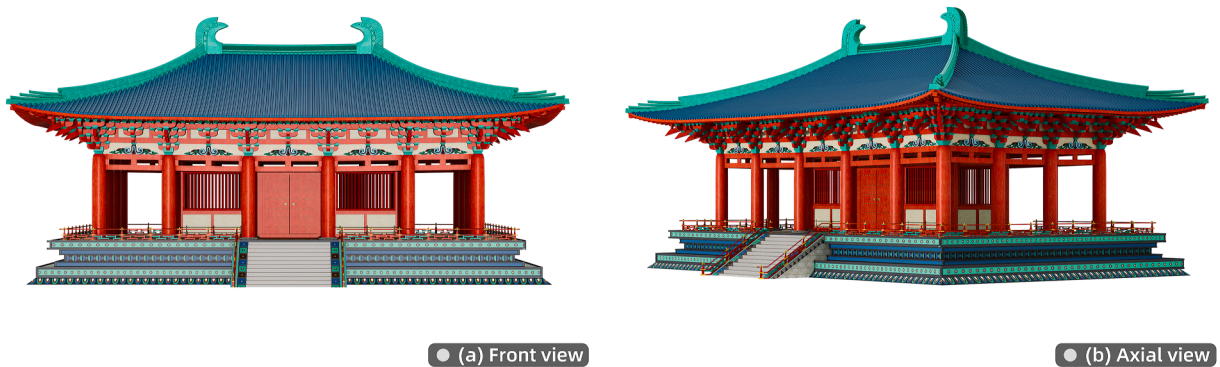


Fig. 4 3D reconstruction model of architectural images in Dunhuang murals: (a) front view; (b) axial view.

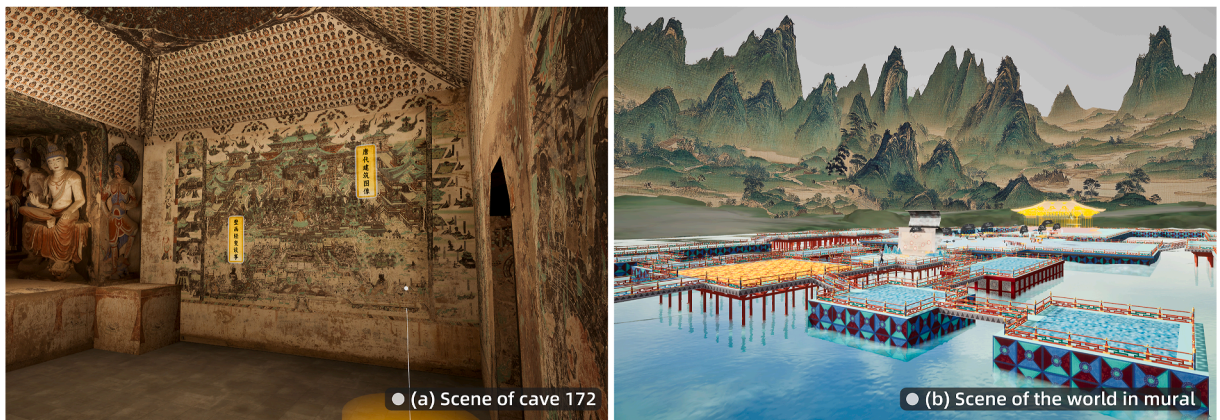


Fig. 5 Scene design in ArchiBuilder: (a) scene of cave 172; (b) scene of the world in mural.

scenes aim to blend architectural culture with immersive interaction.

Mogao Cave 172: Built using high-resolution mural images and 3D spatial data from the Digital Dunhuang project (Yu et al., 2020), this scene faithfully reproduces the cave's wall structures, mural details, and lighting. Users can explore the murals up close and interact with the interface to access historical and cultural information.

Buddhist world scene: Inspired by the north wall mural of Cave 172, this abstract environment integrates architectural

elements into a symbolic space featuring platforms, water reflections, and sky backdrops. Here, users progress through three task-based construction stages.

4.2.2. Narrative design

ArchiBuilder adopts a role-playing narrative structure. Players assume the role of a university student who reconstructs architecture while uncovering its cultural meaning. The storyline is shown in Fig. 6.

Plot 1: The player arrives at Mogao Cave 172. In the dim light, the north wall mural glows faintly. Drawn by its

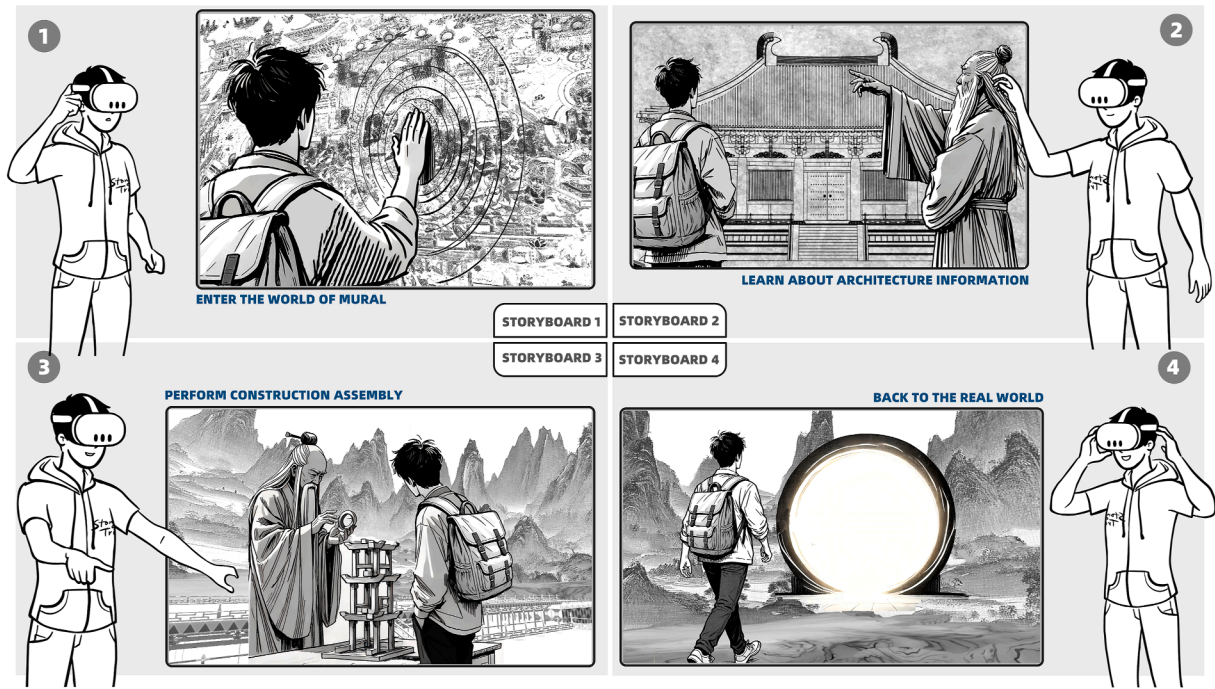


Fig. 6 Narrative design in ArchiBuilder.

architectural image, the player touches the mural and is transported into its world.

Plot 2: Inside the unfinished Buddhist world, a robed craftsman appears. He tells the player that reconstructing the architecture will reveal its secrets and the way home.

Plot 3: The craftsman unrolls a scroll and guides the player through the reconstruction process, explaining component functions and sharing cultural stories.

Plot 4: Upon completion, the world glows. The craftsman opens a portal, and the player returns to the cave with deeper cultural insight.

Building on this story framework, the following section introduces the task design that supports interactive learning and progressive knowledge construction.

4.2.3. Task design

The task design of ArchiBuilder is based on the structural logic and component knowledge of Dunhuang mural architectural images, progressing through five stages: **Prologue**, **Basic Component Extraction**, **Core Structure Assembly**, **General Architecture Reconstruction**, and **Epilogue**, as shown in Fig. 7. This structure aligns with the learning path



Fig. 7 Task design in ArchiBuilder.

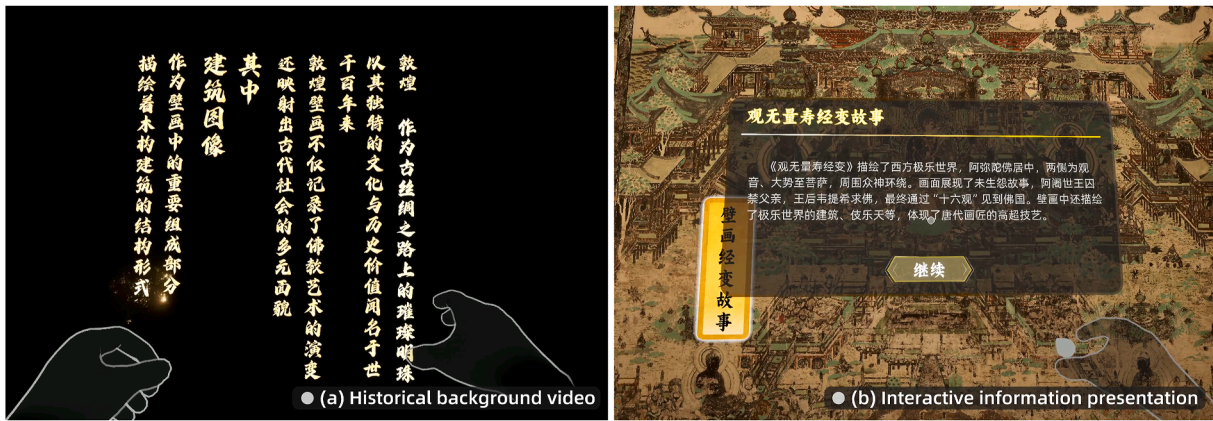


Fig. 8 Detail design of the prologue stage: (a) historical background video; (b) interactive information presentation.

of knowledge acquisition—interactive learning—cultural experience, guiding players from basic perception to holistic understanding.

Tasks revolve around six key components—platform base, column, main structure, dougong, beam frame, and roof—each linked to specific knowledge points. Players complete tasks such as identifying, assembling, and exploring components, supported by NPC (non-player character) guidance, visual cues, and animations. This “part-to-whole” assembly approach provides a unified interaction framework for evaluating learning performance, immersion, and cognitive load. The next section details the core interactions and user experience of each stage.

4.3. ArchiBuilder game detail design

4.3.1. Prologue design

The prologue begins with the player clicking “Start Game”, launching a video that introduces the architectural and cultural significance of Dunhuang murals, as well as the game’s storyline and objectives (Fig. 8(a)). After the video, the player enters Mogao Cave 172 and interacts with clickable points on the mural to access explanations of key architectural elements and their cultural meanings (Fig. 8(b)). Once all points are explored, a transition transports

the player into the mural’s Buddhist world, officially starting the reconstruction journey.

4.3.2. Basic component extraction design

In this stage, the player interacts with a large scroll displaying architectural line drawings to identify and extract key components. Upon reaching the first water platform, an NPC introduces the task and explains the six core architectural components. The player locates each component on the scroll based on these descriptions, aided by dynamic lighting cues (Fig. 9(a)).

Clicking a correct component triggers a pop-up with detailed information on its structure, function, and cultural meaning (Fig. 9(b)). This stage transitions players from recognizing architectural images to understanding individual components, essential knowledge for later tasks.

4.3.3. Core structure assembly design

In this stage, the player manually assembles architectural structures using components identified earlier. Upon reaching the second water platform, the player accesses assembly zones featuring component parts, instructional videos, and interactive 3D models (Fig. 10(a)). The player can rotate translucent models and click hotspots for detailed knowledge.



Fig. 9 Detail design of the basic component extraction stage: (a) dynamic effects contact; (b) architectural knowledge presentation.

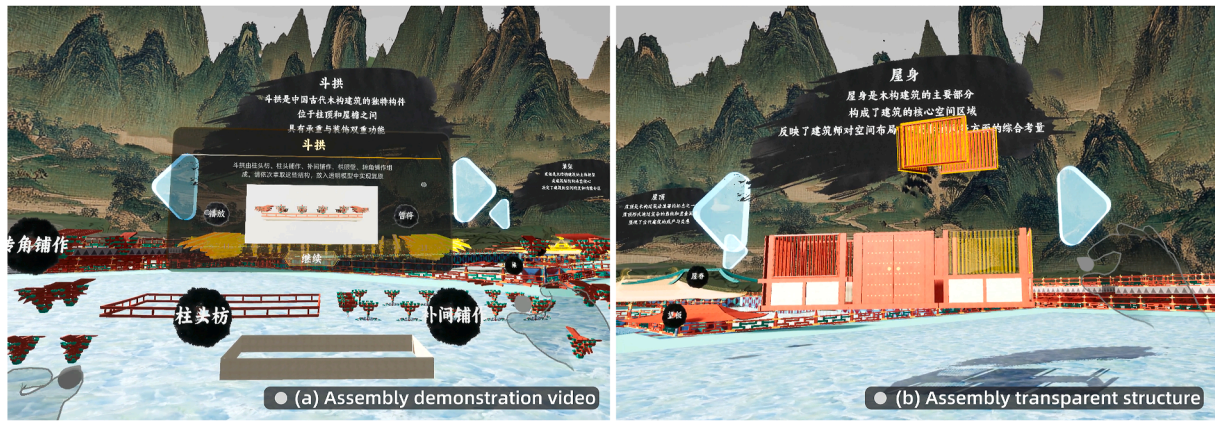


Fig. 10 Detail design of the core structure assembly stage: (a) assembly demonstration video; (b) assembly transparent structure.

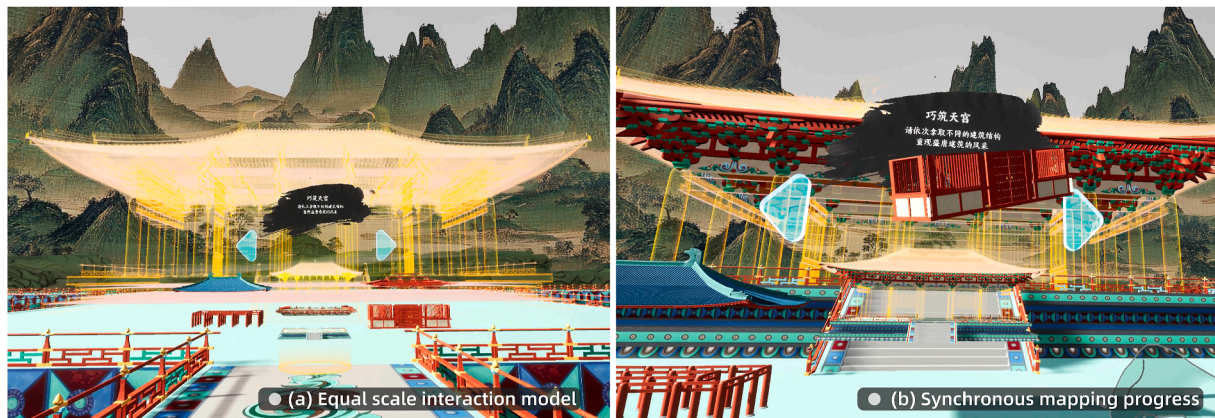


Fig. 11 Detail design of the general architecture reconstruction stage: (a) equal scale interaction model; (b) synchronous mapping progress.

Guided by the NPC, the player selects and places components step-by-step into floating substructures using visual cues (Fig. 10(b)). The goal is to complete six partial modules. After assembly, a summary screen reviews each component's structural function and cultural meaning, deepening architectural understanding.

4.3.4. General architecture reconstruction design

In this stage, the player integrates all partial modules to reconstruct the full structure of the front hall, achieving a complete 3D restoration of Dunhuang architecture. A transparent scaled miniature reference model assists intuitive interaction (Fig. 11(a)).

At the third water platform, the player follows a guided sequence to place components, with each step updating the reconstruction progress on the central model (Fig. 11(b)). Once completed, the entire architecture is revealed, allowing exploration of structural and decorative details from multiple angles. This stage reinforces the player's spatial understanding and completes the architectural assembly.

4.3.5. Epilogue design

In the epilogue, after completing the reconstruction, the player enters a free-roaming mode to explore the interior

and decorative details of the restored architecture in 3D (Fig. 12(a)). The NPC offers a final summary, highlighting the cultural value of the experience. When ready, the player follows a glowing portal to exit the game (Fig. 12(b)). A knowledge review and extended resources are provided, encouraging continued exploration of Dunhuang's architectural heritage beyond gameplay.

4.4. Implementation

The experimental setup includes a 13th Gen Intel® Core™ i9-13900HX (2.20 GHz), NVIDIA GeForce RTX 4060 GPU, and Windows 11 OS. The VR experience was delivered using the Meta Quest 3 headset. Scene modeling for ArchiBuilder was completed in SketchUp 2020, exported via Unreal DataSmith, and imported into Unreal Engine 5.3 for interaction and optimization.

To ensure spatial and structural accuracy, all 3D components—platform base, columns, main structure, dougong, etc.—were modeled based on proportional analysis of architectural images from the north wall of Mogao Cave 172. These were further calibrated using empirical data from Tang dynasty architectures such as Foguang Temple and Nanchan Temple. An expert in Dunhuang architectural images reviewed the models to ensure academic accuracy.

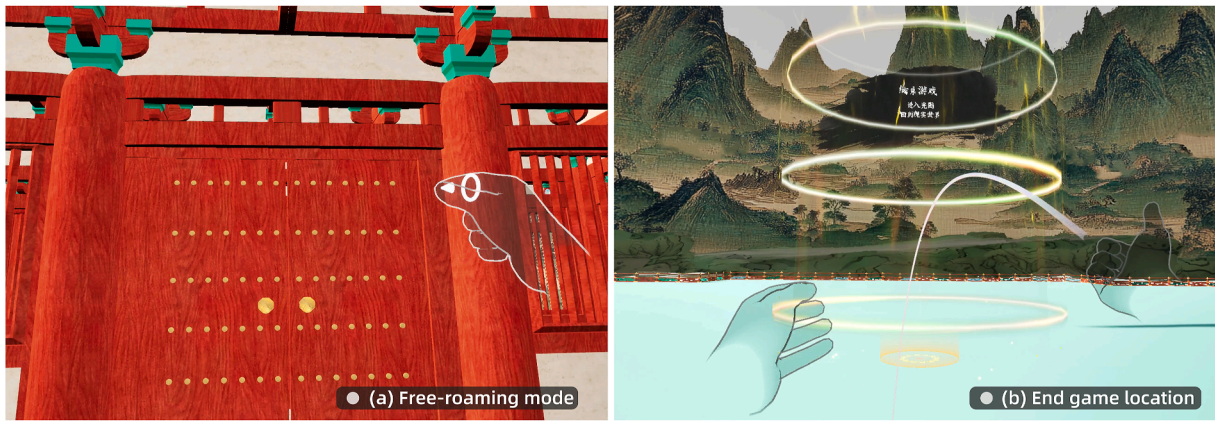


Fig. 12 Detail design of the epilogue stage: (a) free-roaming mode; (b) end game location.

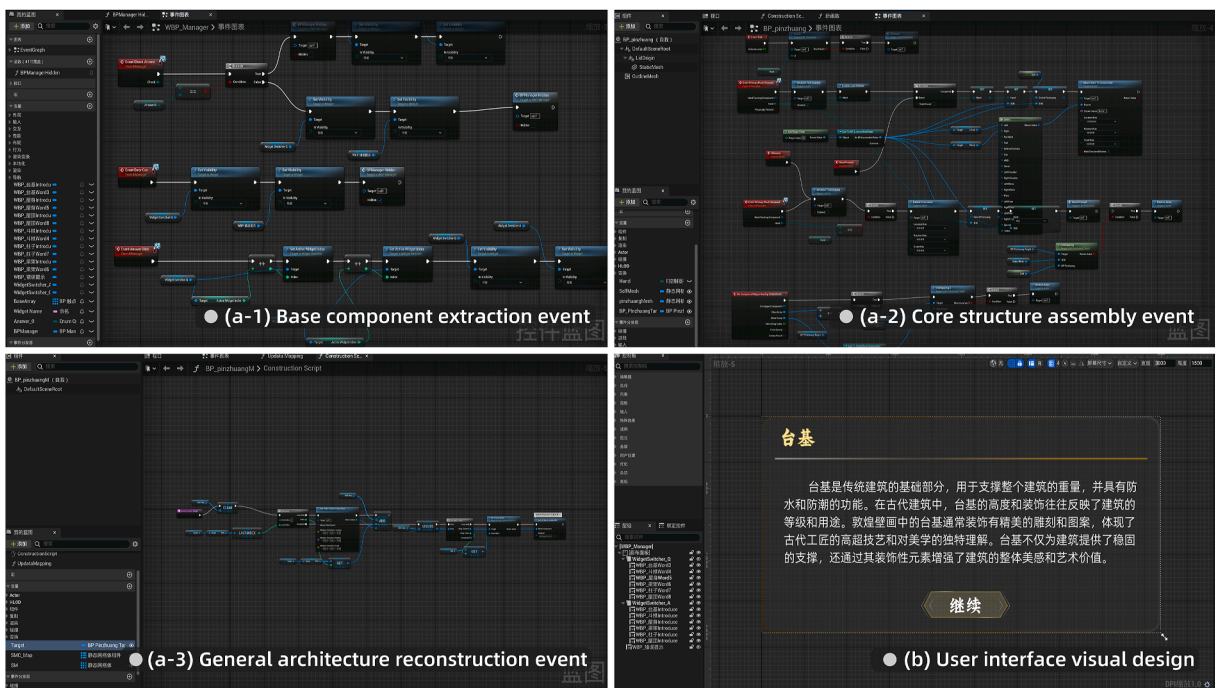


Fig. 13 Development blueprints of ArchiBuilder: (a-1) base component extraction event; (a-2) core structure assembly event; (a-3) general architecture reconstruction event; (b) user interface visual design.

The game’s core interaction logic was implemented using Unreal Engine’s Blueprint visual scripting system. Fig. 13(a-1) shows the blueprint for basic component extraction, while Fig. 13(a-2) and Fig. 13(a-3) show core structure assembly and general architecture reconstruction. Fig. 13(b) shows the user interface design.

5. Evaluation

5.1. Evaluation objectives and variable definitions

To assess the user experience of ArchiBuilder, a between-subjects experiment was conducted. Participants were randomly assigned to either a control group using a VR image-text information or an experimental group using

ArchiBuilder. Evaluation metrics were based on the system’s design goals and common indicators in virtual learning and gamified experiences. Three core dimensions were measured:

Learning Performance: Users’ knowledge acquisition and retention of Dunhuang mural architectural images.

Immersion: Users’ attention, emotional engagement, and sense of presence during interaction.

Cognitive Load: The mental effort perceived during tasks, reflecting system usability.

These metrics align with the study’s two hypotheses (H1, H2). The evaluation aimed to:

- 1) Assess whether ArchiBuilder improves learning performance more than the VR image-text information.

- 2) Assess whether ArchiBuilder offers more immersion than the VR image-text information.
- 3) Assess whether ArchiBuilder reduces cognitive load more than the VR image-text information.

5.2. Participants

A total of 35 participants (18 males, 17 females) aged 18–35 ($M = 25.09$, $SD = 4.09$) were recruited. All had normal or corrected vision and hearing, provided informed consent, and received monetary compensation. 19 participants had prior experience with digital historical architecture, and 26 had used VR devices before. Most were familiar with basic VR interactions and could complete tasks after a brief tutorial.

After excluding outliers and invalid responses, valid data from 30 participants were retained, with 15 in each group. The experimental group included 9 males and 6 females ($M = 26.07$, $SD = 3.52$), and the control group included 7 males and 8 females ($M = 24.33$, $SD = 4.22$).

5.3. Materials

The experiment was conducted in a university lab (Fig. 14), providing a quiet, spacious, and distraction-free environment to support focused interaction.

In the control group, participants experienced VR image-text information in observational mode. They navigated the virtual space and accessed knowledge through visual-text interfaces, following guided instructions to learn about the historical and cultural context of Dunhuang mural architectural images.

In contrast, the experimental group engaged in an immersive, gamified experience using ArchiBuilder.

Participants completed interactive reconstruction tasks embedded in a narrative structure. Both groups received the same knowledge content but differed in user engagement and interaction design, ensuring structural comparability for between-group evaluation.

5.4. Experimental procedure

As shown in Fig. 15, participants first signed in and completed a pre-experience questionnaire, including demographic data and a basic knowledge pre-test to assess prior knowledge. Researchers then introduced the equipment and guided participants through the interaction methods to ensure familiarity with the devices and tasks.

Participants were randomly assigned to either the ArchiBuilder group or the VR image-text group, without being informed of the study's purpose. Each session lasted 30 min under consistent conditions.

After the VR experience, participants completed a post-experience questionnaire and the knowledge instant post-test questionnaire on a lab computer. Semi-structured interviews followed to collect user feedback. One week later, participants completed an online knowledge delay post-test questionnaire to assess knowledge retention.

5.5. Measures

To evaluate the research objectives, we adopted a mixed-methods approach combining quantitative and qualitative data, including knowledge tests, standardized questionnaires, and interviews. All questionnaires used were validated in prior VR and digital heritage studies and followed standardized procedures and scoring guidelines. This ensured reliable measurement of the



Fig. 14 Experimental environment.

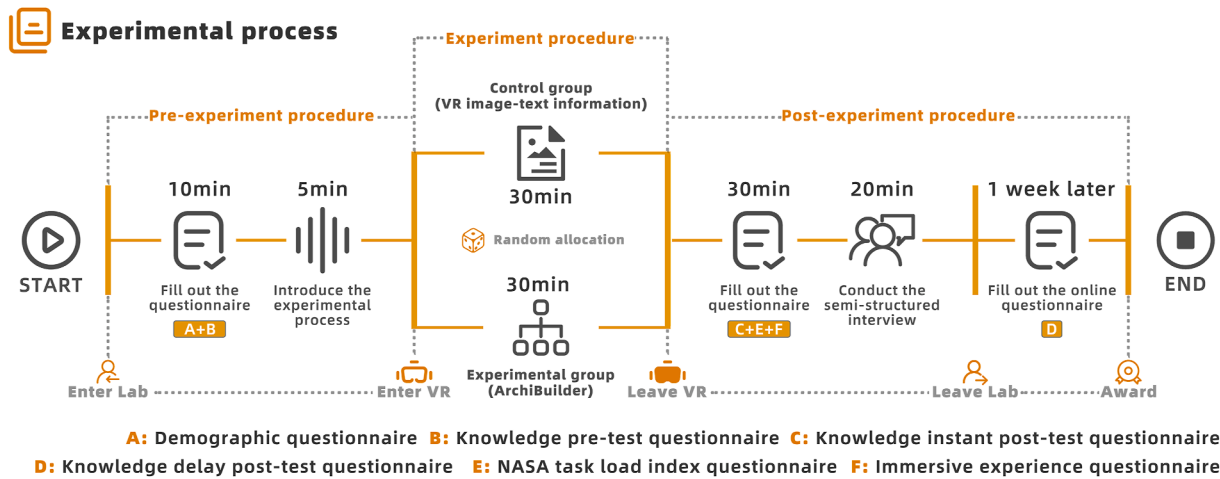


Fig. 15 User evaluation procedure.

core experiential dimensions aligned with the experimental design.

5.5.1. Knowledge test

To assess learning performance, two key indicators were measured: **knowledge instant retention** and **knowledge delay retention** (Makransky and Mayer, 2022). The test was developed based on the content of ArchiBuilder and VR image-text experience, drawing on prior research on Dunhuang mural architectural images. An expert reviewed and validated the items.

The testing included three stages: a knowledge pre-test questionnaire (before the experiment), a knowledge instant post-test questionnaire (immediately after), and a knowledge delay post-test questionnaire (one week later). It comprised single-choice (1 point), multiple-choice (2 points), and true/false (1 point) questions, with a total score of 24 points. Sample items are shown in Table 5. To reduce memory bias (Mancuso et al., 2023), the delayed post-test used the same questions but randomized item and option order.

The final learning performance score was calculated as the average of the knowledge instant retention and knowledge delay retention scores.

5.5.2. Scales

Immersion was measured using the Immersive Experience Questionnaire (IEQ) (Luria et al., 2025), a widely adopted tool for assessing immersion in virtual environments. The IEQ includes 31 items across five dimensions—Cognitive Involvement (9), Real-World Dissociation (7), Challenge (4), Emotional Involvement (6), and Control (5)—rated on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Participants rated each item, and dimension scores were averaged to compute the final immersion score.

Cognitive load was measured using the NASA Task Load Index (NASA-TLX) (Ercolani et al., 2024), a standard tool for evaluating perceived workload. It comprises six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration, each rated on a 0–100 scale. Participants also completed 15 pairwise

Table 5 ArchiBuilder learning performance test design.

Item type	Sample question	Sample options
Single choice	Q1. In which part of an architecture is the xumizuo commonly located?	Dougong/Beam frame/Platform base/Main structure

Multiple choice	Q6. In which part of an architecture is the chiwei commonly located?	Column/Roof/Platform base/Main structure
	Q7. Which two of the following components belong to the platform base?	Balustrade/Chiwei/Steps/Wall

True/False	Q12. Which two of the following components belong to the roof?	Ceiling/Rafter/Architrave/Chiwei
	Q13. The platform base includes xumizuo, balustrade, and steps.	True/False

	Q18. The roof includes columns, column bases, and architraves.	True/False

comparisons to determine weighting. The final score was a weighted average of the six dimensions.

Standardized procedures were followed for both tools to ensure valid, reliable measurement and enable consistent comparison between the ArchiBuilder and VR image-text information groups.

5.5.3. Interviews

Following the questionnaires, participants completed a semi-structured interview focused on three key topics:

- 1) What are your primary needs and expectations for a digital experience of Dunhuang mural architectural images?
- 2) How effective was ArchiBuilder or VR image-text information experience in supporting your learning, and why?
- 3) How appealing did you find ArchiBuilder or VR image-text information overall?

Experimental group participants were also invited to offer feedback and suggestions for improving ArchiBuilder to enhance its design and user experience.

5.6. Data analysis

5.6.1. Quantitative data

Quantitative analysis was conducted using IBM SPSS Statistics 24. The Shapiro-Wilk test was employed to examine normality. A $p > 0.05$ indicated a normal distribution, while a $p \leq 0.05$ indicated a non-normal distribution (Fiandini et al., 2024). For normally distributed data, Welch's t -test was used for between-group comparisons (Curtis, 2024); otherwise, the Mann-Whitney U test was applied (Nguyen et al., 2025).

The significance level of $\alpha=0.05$ was adopted. Results with $p < 0.05$ indicated significant differences and rejection of the null hypothesis; $p \geq 0.05$ indicated no significant difference (Latini et al., 2023). These analyses determined whether ArchiBuilder outperformed the VR image-text information condition across the measured variables.

5.6.2. Qualitative data

For qualitative data, thematic analysis was used to systematically analyze the interview transcripts (Braun and Clarke, 2023). Three researchers independently performed open coding to ensure comprehensive coverage.

The codes were then reviewed and discussed to refine a unified coding framework. Key themes were identified that best addressed the core research questions, ensuring analytical rigor and consistency.

6. Results

6.1. Quantitative results

6.1.1. Enhancing learning performance

The Shapiro-Wilk test was used to assess normality, and Welch's t -test or Mann-Whitney U test was applied accordingly. The results of learning performance are presented in Table 6.

Within-group comparison: Both groups showed significant improvement post-intervention: Control ($t = -15.602$, $p = 0.000$), Experimental ($Z = -4.678$, $p = 0.000$). The score rose from 6.733 to 18.033 in the experimental group and from 7.533 to 17.233 in the control group.

Between-group comparison: No significant differences were found in Prior knowledge ($Z = -1.453$, $p = 0.146$), Immediate post-test ($t = -0.543$, $p = 0.592$), Delayed post-test ($Z = -1.903$, $p = 0.057$), or overall Learning performance ($t = 1.218$, $p = 0.234$).

In summary, both ArchiBuilder and VR image-text information effectively enhanced learning performance. Although ArchiBuilder users had slightly higher average scores, differences were not statistically significant. Nonetheless, the task-driven and immersive approach of ArchiBuilder may have promoted better knowledge organization and active learning behaviors, as suggested by stronger within-group improvements.

6.1.2. Enhancing user immersion

Cronbach's alpha confirmed strong internal consistency for the immersion scale: Immersion (0.870), Cognitive involvement (0.847), Real-world dissociation (0.759), Challenge (0.710), Emotional involvement (0.713), and Control (0.717), all above the 0.7 threshold. The results of immersion are presented in Table 7.

Immersion: Immersion was assessed across five dimensions: Cognitive involvement ($t = 2.257$, $p = 0.032$); Real-world dissociation ($t = 1.156$, $p = 0.261$); Challenge ($Z = -3.092$, $p = 0.002$); Emotional involvement ($t = 2.409$, $p = 0.027$); Control ($t = 2.448$, $p = 0.021$). ArchiBuilder demonstrated significantly higher levels of Immersion ($t =$

Table 6 Normality test p -values (p), mean (M), standard deviation (SD), and p -values for learning performance.

Dependent variable	Control group		Experimental group		p -value
	Normality p	M (SD)	Normality p	M (SD)	
Prior knowledge	0.479	7.533 (1.685)	0.023	6.733 (2.865)	0.146
Learning performance	0.879	17.233 (1.720)	0.116	18.033 (1.875)	0.234
Within-group difference (p)	—	0.000***	—	0.000***	
Immediate post-test	0.147	19.267 (2.154)	0.191	18.800 (2.541)	0.592
Delayed post-test	0.002	15.200 (2.651)	0.924	17.267 (2.915)	0.057

Note: * indicates a significant difference ($0.01 \leq p < 0.05$); ** indicates a highly significant difference ($0.001 \leq p < 0.01$); *** indicates an extremely significant difference ($p < 0.001$).

Table 7 Normality test p -values (p), mean (M), standard deviation (SD), and p -values for immersion.

Dimension	Control group		Experimental group		p -value
	Normality p	M (SD)	Normality p	M (SD)	
Immersion (+)	0.325	4.888 (0.399)	0.076	5.469 (0.441)	0.001**
Cognitive involvement	0.222	4.859 (0.671)	0.218	5.400 (0.641)	0.032*
Real-world dissociation (+)	0.783	4.990 (0.457)	0.407	5.295 (0.914)	0.261
Challenge (+)	0.048	4.383 (0.906)	0.157	5.500 (0.675)	0.002**
Emotional involvement	0.254	4.978 (0.872)	0.238	5.556 (0.319)	0.027*
Control (+)	0.421	5.093 (0.792)	0.257	5.707 (0.560)	0.021*

Note: * indicates a significant difference ($0.01 \leq p < 0.05$); ** indicates a highly significant difference ($0.001 \leq p < 0.01$); *** indicates an extremely significant difference ($p < 0.001$).

3.782, $p = 0.001$) than the VR image-text information condition. Notably, cognitive involvement, challenge, emotional involvement, and control were all significantly higher in the experimental group.

In summary, ArchiBuilder significantly enhanced user immersion compared to the VR image-text information. Its gamified, context-driven design promoted greater engagement, emotional involvement, and a stronger sense of control. While real-world dissociation showed no significant difference, this suggests ArchiBuilder maintained an effective balance between immersion and cognitive awareness, helping users stay focused and process information more efficiently.

6.1.3. Reducing cognitive load

Cronbach's alpha for the cognitive load scale was 0.888, indicating high internal consistency and sufficient reliability for further analysis. The results of participants' cognitive load assessment are shown in Table 8.

Cognitive load: Cognitive load was assessed across six dimensions: Mental demand ($t = -2.600$, $p = 0.017$); Physical demand ($Z = -1.456$, $p = 0.145$); Temporal demand ($Z = -0.640$, $p = 0.522$); Effort ($Z = -2.701$, $p = 0.007$); Performance ($t = -2.129$, $p = 0.044$); Frustration ($t = -6.011$, $p = 0.000$). ArchiBuilder resulted in significantly lower Cognitive load ($t = -3.562$, $p = 0.001$) than the VR image-text information condition. Significant improvements were observed in the dimensions of mental demand, effort, performance, and frustration.

In summary, ArchiBuilder significantly reduced users' cognitive load compared to VR image-text information. Its interactive, task-driven design likely improved information processing efficiency and eased mental strain in acquiring architectural knowledge. Significant reductions were observed in mental demand, effort, performance pressure, and frustration. Physical and temporal demands showed no significant difference, likely due to the active operations—dragging, rotating, and placing—required in virtual tasks, which increased physical engagement and time usage.

6.2. Qualitative results

Thematic analysis of user interviews identified four key themes: learning performance, immersion, cognitive load, and opportunities and challenges. This section focuses on how the core mechanisms and design strategies in VR SG contributed to enhancing the user experience.

6.2.1. Interactive scenes enhance learning performance

Participants widely agreed that the interactive scene design helped them intuitively understand architectural components and reinforced memory through hands-on assembly. Compared to the VR image-text information condition, ArchiBuilder's real-time feedback and visual cues encouraged active exploration. As participant P10 noted: "When assembling the components, I could see how different parts were put together. It was much easier to

Table 8 Normality test p -values (p), mean (M), standard deviation (SD), and p -values for cognitive load.

Dimension	Control group		Experimental group		p -value
	Normality p	M (SD)	Normality p	M (SD)	
Cognitive load (+)	0.190	65.378 (9.517)	0.814	47.778 (16.601)	0.001**
Mental demand	0.155	69.333 (12.228)	0.758	52.000 (22.741)	0.017*
Physical demand	0.065	57.333 (14.864)	0.028	48.000 (14.243)	0.145
Temporal demand	0.000	52.000 (9.411)	0.068	47.333 (18.310)	0.522
Effort	0.034	74.000 (12.421)	0.467	52.667 (21.536)	0.007**
Performance (+)	0.575	63.333 (15.887)	0.462	46.667 (25.820)	0.044*
Frustration	0.210	69.333 (15.337)	0.220	38.667 (12.459)	0.000***

Note: * indicates a significant difference ($0.01 \leq p < 0.05$); ** indicates a highly significant difference ($0.001 \leq p < 0.01$); *** indicates an extremely significant difference ($p < 0.001$).

understand the architecture than just looking at static images and text”.

The visualization of task progress further strengthened the learning effect. As participant P21 noted: “Each time I completed a structure, I better understood the connections between components”. This process made learning more logical and cohesive. In contrast, participants in the control group found the experience less engaging, with participant P15 remarked: “There was a lot of information, but it was easy to lose focus—especially with complex structures”. These insights suggest that interactive scenes promoted more efficient, structured, and engaging learning through visual and operational guidance.

6.2.2. Narrative mechanisms enhance immersion

Participants generally agreed that the integration of narrative-driven gameplay enhanced their immersion, helping them become emotionally and cognitively engaged in the reconstruction process. As participant P18 described: “The moment I entered the virtual world, I truly felt like I had walked into the mural itself”. This indicates that the narrative elements shifted users from passive observers to active participants, creating a stronger sense of presence.

Contextualized tasks also improved understanding. Many participants noted that the NPC guidance was effective in blending knowledge with actions. Participant P5 remarked: “The NPC not only taught me how to assemble the parts but also explained their functions and history. It felt like having a tour guide”. This narrative support added coherence to learning.

Additionally, story progression helped reduce cognitive isolation. Some users felt they were “advancing the story” with each task, enhancing emotional connection. These results indicate that narrative design not only boosts immersion but also facilitates deeper, more meaningful cultural learning.

6.2.3. Staged tasks reduce cognitive load

Staged tasks were widely recognized for easing cognitive load. Participants noted that breaking the reconstruction process into smaller, sequential steps made complex architectural knowledge easier to grasp without feeling overwhelmed. As participant P12 shared: “The reconstruction was divided into several stages, so it didn’t feel too difficult”. This hierarchical structure made learning more gradual and accessible.

The progression from simple to complex tasks also supported better understanding. Participant P7 commented: “When assembling the dougong, it started simple and got harder later. That helped me understand how it’s built”. This approach reduced mental strain while reinforcing learning continuity.

In contrast, control group participants found the VR image-text information condition more linear and harder to retain. Participant P15 noted: “There was a lot of useful information, but it was harder to remember”. These findings show that staged, interactive tasks can reduce overload and improve the internalization and transfer of knowledge.

6.2.4. Opportunities and challenges

Opportunities: ArchiBuilder offers a promising digital approach for cultural heritage education, especially in

visualizing historical architecture and reducing cognitive load. Through dynamic assembly and narrative-driven tasks, users gained intuitive insights into the functions and cultural meanings of architectural components. Many participants noted that interactive reconstruction helped them better understand each element’s historical role, enhancing their grasp of the architectural system.

The ability to revisit and freely explore the virtual environment further reinforced learning and memory. This experiential model provides a more engaging alternative for cultural education, showcasing the potential of digital tools in heritage communication and learning practices.

Challenges: Despite its strengths, participants identified areas needing improvement. Some found the fixed task sequence restrictive for self-paced learning. As participant P6 noted: “Sometimes I wanted to go at my own pace rather than follow the system’s steps”. Others mentioned occasional delays or mismatches in visual cues during complex assembly tasks, disrupting the interaction flow.

These findings suggest a need to refine visual guidance and incorporate adaptive learning features. Future versions should aim to enhance flexibility and interaction accuracy to further support user engagement and knowledge retention.

7. Discussion

This study introduced ArchiBuilder, a VR-based gamified system aimed at enhancing users’ understanding of architectural images in Dunhuang murals by improving learning performance, immersion, and reducing cognitive load. A between-subjects experiment and user interviews were conducted to evaluate its effectiveness in cultural heritage education and explore the influence of its core design strategies on user experience.

First, although no significant difference in learning performance between ArchiBuilder and the VR image-text information was found, participants in the ArchiBuilder group showed stronger engagement and developed a deeper grasp of architectural components through interactive tasks. This indicates that gamified learning may not always boost scores but can enhance the learning process by supporting active knowledge construction (Feng et al., 2022). Moreover, the interactive feedback helped reduce the cognitive strain of passive learning and encouraged deeper understanding through exploration (Capatina et al., 2024).

Second, ArchiBuilder significantly outperformed the VR image-text information in immersion—especially in cognitive involvement, challenge, emotional engagement, and control. Participants attributed this to the narrative-driven task structure, which made learning more coherent and engaging (Li et al., 2025). This aligns with prior research showing that immersive interaction enhances learning motivation (Chang and Suh, 2025). While no significant difference was found in real-world dissociation, users noted that structured tasks and role-play elements effectively supported immersion (Lee and Wang, 2025).

Finally, NASA-TLX results showed that ArchiBuilder significantly reduced cognitive load, particularly in mental

Table 9 Key advancements of ArchiBuilder compared to prior research.

Research dimension	Previous studies	This study (ArchiBuilder)
Interaction design	Relied mainly on text and visuals with limited user interaction (Agnello et al., 2019)	Combines 2D extraction, 3D assembly, and real-time interaction to foster active exploration and agency
Narrative immersion	Fragmented scene logic and weak narrative immersion (Fanani and Syarif, 2023)	Integrates 2D-to-3D progression with NPC-guided storytelling for cognitive coherence and cultural immersion
Architectural image representation	Often used VR videos or free exploration, lacking spatial clarity (De Fino et al., 2022)	Translates mural images into structured 3D reconstruction, enhancing spatial understanding and learning depth
Evaluation metrics	Focused on expert knowledge or subjective feedback; rarely evaluated interaction quality (Chan et al., 2022)	Applies a multi-dimensional framework (learning, immersion, cognitive load) for comprehensive user experience assessment

demand, effort, performance, and frustration. This suggests that task-based gamification lightened the burden of processing complex knowledge (Friebs et al., 2020). The lack of significant differences in physical and temporal demand may relate to the inherent interaction costs of VR systems (Fan et al., 2023). While users still invested time to complete tasks, the main benefit of cognitive load reduction appeared in smoother processing and better task execution (Greenberg and Zheng, 2022).

As summarized in Table 9, our research differs from previous work in the following four key dimensions:

In summary, ArchiBuilder demonstrates the potential of gamified VR in cultural heritage education. By integrating interactive tasks, narrative immersion, and staged cognitive load management, it offers an innovative approach to the digital representation of historical architectural images. Experimental results confirm that gamified interaction enhances user engagement and facilitates more effective, accessible learning of complex architectural knowledge.

7.1. Design implications

This study offers key insights to guide future research and development in VR-based gamified cultural heritage education:

3D Reconstruction Enhances Learning: ArchiBuilder translated historical architectural images into interactive 3D models, enabling users to intuitively understand structural logic and cultural significance. The immersive 3D environment fostered engagement and improved memory retention. Future work may enhance model accuracy and incorporate advanced visualization to further support cultural learning.

Task-Driven Gamification Supports Active Exploration: The task-based design with real-time feedback encouraged users to engage actively and learn by doing. Compared to passive image-text learning, gamified tasks offered a more interactive path. Future studies could examine the impact

of different game mechanics and how to tailor task difficulty to diverse user backgrounds.

Narrative Enhances Emotional Engagement: The narrative structure in ArchiBuilder enriched immersion and motivation by embedding learning within culturally meaningful scenarios. Role-based guidance helped users internalize architectural knowledge while fostering cultural identity. Future designs may expand storylines and character roles to deepen user connection.

Interactive Scene Enables Multisensory Learning: By combining dynamic assembly, visual cues, and real-time feedback, ArchiBuilder supported a multisensory learning experience that boosted engagement and eased cognitive load. Future research could explore broader multimodal interactions to improve situational immersion and user involvement.

Progressive Task Design Manages Cognitive Load: Staged tasks with incremental complexity helped users process architectural knowledge more effectively. ArchiBuilder's design facilitated gradual learning and reduced overload. Future systems could incorporate adaptive mechanisms that adjust task complexity based on real-time user performance for more personalized experiences.

7.2. Limitations and future work

Despite the promising results, this study has several limitations that inform future directions:

First, the experiment was limited to a single session in a controlled laboratory setting, which may have introduced novelty effects and restricted insight into long-term learning outcomes. Future research should adopt longitudinal approaches to evaluate sustained learning and cultural understanding.

Second, the study relied primarily on self-reported data, which may be subject to bias. Incorporating objective measures such as behavioral tracking and physiological responses could yield more robust insights into user experience and cognitive processes.

Third, the fixed task flow limited opportunities for personalized learning. Future developments should explore adaptive systems with multimodal and collaborative interactions to enhance flexibility, engagement, and user autonomy in cultural heritage education.

8. Conclusion

This study explored interactive learning through historical architectural images and proposed a gamified experience model based on VR SG. Centered on Dunhuang mural architectural images, the prototype system ArchiBuilder was developed and evaluated through a between-subjects experiment, assessing its impact on learning performance, immersion, and cognitive load.

For RQ1, ArchiBuilder employed a three-stage task flow—basic component extraction, core structure assembly, and general architecture reconstruction—combined with NPC storytelling and progressive interaction to enhance users' understanding of structural features and cultural context. This confirms the effectiveness of gamified approaches in architectural and cultural learning.

For RQ2, compared to traditional VR image-text experiences, ArchiBuilder significantly improved immersion and reduced cognitive load. Although learning performance was not statistically higher, the interactive structure facilitated more efficient cognitive processing.

Theoretically, the study advances an integrated design paradigm of knowledge construction, task-driven learning, and immersive interaction for cultural heritage education. It expands the application of VR SG in architectural interpretation and enriches theoretical models of cultural learning. Practically, ArchiBuilder presents a transferable framework combining spatial cognition and staged tasks, offering methodological insights for digital heritage display and public education.

Future work will explore adaptive learning paths, multimodal interaction, and long-term user engagement to further integrate VR SG into cultural education and build a sustainable digital dissemination model.

Author contributions

Zhizheng Liu: design concept, game concept and plot design, data research and analysis, writing (first draft). Xiran Ji: scene modeling, game environment design. Yizhu Cui: game interaction design, UI and visual design. Tianzi Zhang: game interaction design, UI and visual design. Jian Yao: project administration, writing, review, and editing. All authors reviewed the manuscript.

Ethics statement

The authors declare that their Institutional Ethics Committee confirmed that no ethical review was required for this study. Written informed consent for participation was not required because all participants' data were anonymized before the statistical analyses were done.

Availability of data and materials

The data underlying this article cannot be shared publicly due to the privacy of the individuals that participated in the study. The data will be shared, on reasonable request, with the corresponding author.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We are grateful to all survey participants for their contributions to this research. We also appreciate the anonymous reviewers for their thoughtful comments and suggestions, which greatly enhanced the rigor and clarity of this manuscript.

This study was supported by the National Social Science Foundation of China (Special Program for Unpopular Disciplines) Academic Team Project "Research on the Genealogical Construction and Transmission Path of Dunhuang mural architectural images Art" (24VJXT010).

References

- Agnello, F., Avella, F., Agnello, S., 2019. Virtual reality for historical architecture. *Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci.* 42, 9–16.
- Bao, Y., 2024. Research on the yingzao chi restoration of palace buildings sans dougong in Ming and Qing Dynasties: a case study of Chongqing. *Herit. Sci.* 12 (1), 306.
- Braun, V., Clarke, V., 2023. Toward good practice in thematic analysis: avoiding common problems and being a knowing researcher. *Int. J. Transgender Health* 24 (1), 1–6.
- Capatina, A., Juarez-Varon, D., Micu, A., Micu, A.E., 2024. Leveling up in corporate training: unveiling the power of gamification to enhance knowledge retention, knowledge sharing, and job performance. *J. Innovation Knowl.* 9 (3), 100530.
- Capecchi, I., Bernetti, I., Borghini, T., Caporali, A., Saragosa, C., 2024. Augmented reality and serious game to engage the alpha generation in urban cultural heritage. *J. Cult. Herit.* 66, 523–535.
- Chan, C.S., Bogdanovic, J., Kalivarapu, V., 2022. Applying immersive virtual reality for remote teaching architectural history. *Educ. Inf. Technol.* 27 (3), 4365–4397.
- Chang, S., Suh, J., 2025. The impact of digital storytelling on presence, immersion, enjoyment, and continued usage intention in VR-Based museum exhibitions. *Sensors* 25 (9), 2914.
- Chen, X., Wang, X., 2016. Parameterization of *Yingzao Fashi*: generating the palace style structural carpentry system based on algorithmic language. *Huazhong Archit.* 34 (12), 41–43 (*in Chinese*).
- Chen, Y., del Blanco García, F.L., 2022. Constructive analysis and digital 3D reconstruction of the Yuanmingyuan ruins: Wanfanganhe pavilion (China). *Virtual Archaeol. Rev.* 13 (27), 1–16.
- Curtis, D., 2024. Welch's test is more sensitive to real world violations of distributional assumptions than student's t test but

- logistic regression is more robust than either. *Stat. Pap.* 65 (6), 3981–3989.
- DaCosta, B., Kinsell, C., 2022. Serious games in cultural heritage: a review of practices and considerations in the design of location-based games. *Educ. Sci.* 13 (1), 47.
- De Fino, M., Bruno, S., Fatiguso, F., 2022. Dissemination, assessment and management of historic buildings by thematic virtual tours and 3D models. *Virtual Archaeol. Rev.* 13 (26), 88–102.
- Dewitz, L., Kröber, C., Messemer, H., Maiwald, F., Münster, S., Brusckke, J., Niebling, F., 2019. Historical photos and visualizations: potential for research. *Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci.* 42, 405–412.
- Du, G., Chen, M., 1977. Achievements in statics of the Northern Sung dynasty as seen from "Codes of Construction". *Architect. J.* 1 (42–46), 36–52 (*in Chinese*).
- Ercolani, J., Cure, L., Misasi, P., 2024. Identifying and validating perceived workload metrics for emergency medical services. *Appl. Ergon.* 118, 104270.
- Fan, L., Wang, J., Li, Q., Song, Z., Dong, J., Bao, F., Wang, X., 2023. Eye movement characteristics and visual fatigue assessment of virtual reality games with different interaction modes. *Front. Neurosci.* 17, 1173127.
- Fanani, A.Z., Syarif, A.M., 2023. Historical building 3D reconstruction for a virtual reality-based documentation. *Int. J. Adv. Comput. Sci. Appl.* 14 (9).
- Feng, Y., Yi, Z., Yang, C., Chen, R., Feng, Y., 2022. How do gamification mechanics drive solvers' Knowledge contribution? A study of collaborative knowledge crowdsourcing. *Technol. Forecast. Soc. Change* 177, 121520.
- Ferdani, D., Fanini, B., Piccioli, M.C., Carboni, F., Vigliarolo, P., 2020. 3D reconstruction and validation of historical background for immersive VR applications and games: the case study of the Forum of Augustus in Rome. *J. Cult. Herit.* 43, 129–143.
- Fiandini, M., Nandiyanto, A.B.D., Al Husaeni, D.F., Al Husaeni, D. N., Mushiban, M., 2024. How to calculate statistics for significant difference test using SPSS: understanding students comprehension on the concept of steam engines as power plant. *Indones. J. Sci. Technol.* 9 (1), 45–108.
- Friebs, M.A., Dechant, M., Vedress, S., Frings, C., Mandryk, R.L., 2020. Effective gamification of the stop-signal task: two controlled laboratory experiments. *JMIR Serious Games* 8 (3), e17810.
- Geng, M., Tang, X., 2024. Study in the evolution of Dunhuang frescoes city wall paradigm based on the interpretation of historical prospect information. *J. Human Settlements West China* 40 (1), 218–224 (*in Chinese*).
- Greenberg, K., Zheng, R., 2022. Cognitive load theory and its measurement: a study of secondary tasks in relation to working memory. *J. Cognit. Psychol.* 34 (4), 497–515.
- Han, Q., 2022. Architectural history research and the Dunhuang Caves—On the correspondence between Liang Sicheng and Paul Pelliot. *Dunhuang Res.* 1, 143–149 (*in Chinese*).
- Hao, Y., Yao, Z., Wu, R., Bao, Y., 2024. Damage and restoration technology of historic buildings of brick and wood structures: a review. *Herit. Sci.* 12 (1), 301.
- Jin, H., 2022. Cases Study of palace architectural elements in Dunhuang Grottoes murals. *Palace Mus. J.* 11 (4–14), 142.
- Kara, N., 2024. A mixed-methods study of cultural heritage learning through playing a serious game. *Int. J. Hum. Comput. Interact.* 40 (6), 1397–1408.
- Latini, A., Di Giuseppe, E., D'Orazio, M., 2023. Immersive virtual vs real office environments: a validation study for productivity, comfort and behavioural research. *Build. Environ.* 230, 109996.
- Lee, I.J., Wang, W.Z., 2025. Applying theatre-based role-playing combined with AR and VR game strategies to enhance imagination and symbolic play skills in children with autism spectrum disorder. *Univers. Access Inf. Soc.* 1–22.
- Li, Q., Chen, Z., Wu, T., Shen, C., Yang, Y., Zhao, X., 2025. Exploring the effects of narrative augmented reality in cultural heritage: a study on emotional arousal, intrinsic motivation, cognitive load, and learning. *Int. J. Hum. Comput. Interact.* 1–11.
- Liu, Y., 2018. The earthwork, trabeated system, and the roof: three architectural elements of Foguangsi from a view of *Sachlichkeit*. *Architect. J.* 9, 11–18 (*in Chinese*).
- Liu, Z., Chen, D., Zhang, C., Yao, J., 2024. Design of a virtual reality serious game for experiencing the colors of Dunhuang frescoes. *Herit. Sci.* 12, 370.
- Luria, E., Gross, Z., Rothgangel, M., 2025. Unveiling the power of educational media: enhancing immersion and affective engagement in holocaust education through virtual reality. *Educ. Media Int.* 62 (3), 380–404.
- Maji, S., Saha, A., Pal, H., 2024. Exploring enjoyment while playing 'temple Lego': a virtual reality cultural heritage serious game for architecture students. *Discov. Educ.* 3 (1), 1–23.
- Makransky, G., Mayer, R.E., 2022. Benefits of taking a virtual field trip in immersive virtual reality: evidence for the immersion principle in multimedia learning. *Educ. Psychol. Rev.* 34 (3), 1771–1798.
- Mancuso, V., Bruni, F., Stramba-Badiale, C., Riva, G., Cipresso, P., Pedroli, E., 2023. How do emotions elicited in virtual reality affect our memory? A systematic review. *Comput. Hum. Behav.* 146, 107812.
- Meng, X., Zhang, Q., Pei, Q., Ye, M., 2021. A periodizational study on the historical evolution of veranda and corridor buildings on Dunhuang frescoes. *Architect. J.* 2, 74–80 (*in Chinese*).
- Moyano, J., Carreño, E., Nieto-Julián, J.E., Gil-Arízón, I., Bruno, S., 2022. Systematic approach to generate Historical Building Information Modelling (HBIM) in architectural restoration project. *Autom. Construct.* 143, 104551.
- Mu, R., Nie, Y., Cao, K., You, R., Wei, Y., Tong, X., 2024. Pilgrimage to Pureland: art, perception and the Wutai mural VR reconstruction. *Int. J. Hum. Comput. Interact.* 40 (8), 2002–2018.
- Nguyen, T.K., Ahmad, Z., Nguyen, D.T., Kim, J.M., 2025. A remaining useful lifetime prediction model for concrete structures using Mann-Whitney U test state indicator and deep learning. *Mech. Syst. Signal Process.* 222, 111795.
- Pagano, A., Palombini, A., Bozzelli, G., De Nino, M., Cerato, I., Ricciardi, S., 2020. *ArkaeVision* VR game: user experience research between real and virtual Paestum. *Appl. Sci.* 10 (9), 3182.
- Pietroni, E., Ferdani, D., 2021. Virtual restoration and virtual reconstruction in cultural heritage: terminology, methodologies, visual representation techniques and cognitive models. *Information* 12 (4), 167.
- Shi, L., Chen, X., Xu, Y., Gao, X., Lai, J., Wang, S., 2024. Towards an effective architectural form: the composition of squareness and roundness based on Scale Proportion—Evidence from the yingxian wooden pagoda. *Buildings* 14 (5), 1472.
- Sun, Y., 2020. Evolution of two kinds of vertical ridge-end tile decoration styles seen in the Tang and Western Xia murals of Dunhuang. *J. Architect. Hist.* 1 (1), 125–135 (*in Chinese*).
- Ting, M.P., Min, C.D., 2025. What drives user churn in serious games? An empirical examination of the TAM, SOR theory, and game quality in Chinese cultural heritage games. *Entertain. Comput.* 52, 100758.
- Wang, J., Qiu, Z., 2014. An analysis on the Dunhuang Mogao Northern Dynasties architectural images. *J. Asian Architect. Build Eng.* 13 (1), 1–8.
- Wang, N., 2017. Circles, squares, and the place where Buddha lives: an explorative study on compositional proportion of the main hall of Foguangsi. *Architect. J.* 6, 29–36 (*in Chinese*).
- Wang, Q., 2024. Digital reconstruction of Buddhist compounds in Dunhuang murals using BIM Technology-A case Study of the north Wall of the main chamber in Cave 172, mogao grottoes. *J. Dunhuang Stud.* 3, 65–75 (*in Chinese*).
- Wang, Q., Zhang, J., 2022. Digital construction of the architecture in Dunhuang murals—With a case study of the Buddhist temple

- architecture in a painting in Mogao Cave 361. *Dunhuang Res.* 2, 125–135 (*in Chinese*).
- Wells, M., Xue, J., 2024. Survey as method: field trips of the society for research in Chinese architecture, 1930–1937. *J. Architect.* 29 (4), 556–583.
- Wu, C., Xue, J., Song, D., Liang, X., 2022. Analysis on the mechanical performance of Dougong bracket sets under eccentric vertical load. *Constr. Build. Mater.* 314, 125652.
- Xiao, M., 2017. Measurement regulations of the main hall of Foguangsi. *Architect. J.* 6, 37–42 (*in Chinese*).
- Xu, L., Wu, J., Yu, T., Ding, X., 2022. 3D digital rendering of architectural images in Cave 172 of Dunhuang Mogao grottoes. *Sci. Conserv. Archaeol.* 34 (1), 2002–2018.
- Xu, T., 2015. Study on the regularity and control for giant timber construction scale included in *Treatise on Architectural Methods (Yingzao Fashi)*. *Palace Mus. J.* 6 (36–44), 157–158 (*in Chinese*).
- Yao, J., Yu, Q., 2023. The form and decoration of balustrade in Tang-dynasty murals of the Mogao Grottoes in Dunhuang. *Art-magazine* 11, 96–105 (*in Chinese*).
- Yu, M., Yang, C., Chen, S., 2024. A maze of data: doubts on the multiple solution of the scale design in the East Hall of Foguangsi Temple. *Architect. J.* 2, 117–123 (*in Chinese*).
- Yu, T., Wu, J., Zhao, L., Ding, X., Ye, Q., 2020. The design and implementation of the framework for the database of "Digital Dunhuang". *Dunhuang Res.* 2, 120–130 (*in Chinese*).
- Yuk, H., Jo, H.H., Choi, J.Y., Nam, J., Chang, S.J., Kim, S., 2023. Evaluation of suitability for passive retrofit of wooden roof considering the specificity of historic buildings. *Build. Environ.* 242, 110608.
- Zhang, H., Gao, W., Wang, Y., 2024. A wooden pin reinforcement of Ancient Chinese wooden temple: a case of Daxiong hall. *Adv. Civ. Eng.* 2024 (1), 8824310.
- Zhang, J., Wang, Q., 2018. Digital representation and reconstruction of architecture in Dunhuang frescoes. *J. Dunhuang Stud.* 3, 187–196 (*in Chinese*).
- Zhang, Q., Pei, Q., 2025. Research on the interpretation of architectural image information in the "Mount Wutai Map" in Cave 61 of Dunhuang Mogao grottoes. *J. Northwest Normal Univ. (Nat. Sci.)* 61 (3), 86–96 (*in Chinese*).
- Zhang, R., Liu, C., Zang, C., 2007. Explanation of the numerical and digital data resulting from the on-site survey of the Large Eastern Hall of Foguang Si (Buddha Light Temple). *Palace Mus. J.* 02 (28–51), 155–156 (*in Chinese*).
- Zhang, X., Zhi, Y., Xu, J., Han, L., 2022. Digital protection and utilization of architectural heritage using knowledge visualization. *Buildings* 12 (10), 1604.
- Zhang, Y., Han, X., Ding, T., Ma, M., Luo, G., Hao, Z., 2018. Reconstructive study on length of carpentry ruler of a timber structure of Chinese Tang dynasty named Main Hall of Foguangsi Temple from perspective of traditional architecture design method named Integer Scale Method. *J. Human Settlements West China* 33 (2), 85–92 (*in Chinese*).
- Zhou, C., Dong, Y., Hou, M., 2024. DGPCD: a benchmark for typical official-style Dougong in ancient Chinese wooden architecture. *Herit. Sci.* 12 (1), 201.
- Zhou, Z., Li, L., 2024. Vision and site: revisiting a pure Land cave of Dunhuang. *Religions* 15 (3), 329.
- Zhou, Z., 2025. Dunhuang architectural studies, 1926–2024. *Heritage* 8 (3), 101.
- Zou, M., Bahauddin, A., 2024. The creation of "Sacred Place" through the "Sense of Place" of the Daci'en Wooden Buddhist Temple, Xi'an, China. *Buildings* 14 (2), 481.