ACCSCIENCE PUBLISHING

Arts & Communication

ARTICLE

Augmented reality in Chinese museums: A case study on visitor experience

Guangyi Wu* and **Nuno Bernardo**

Department of Industrial Design, Design School, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, China

Abstract

This study examines how visitor satisfaction co-varies with augmented reality (AR)-related evaluations in a Chinese museum context, using Suzhou Museum as a case study. We collected on-site survey data from 420 visitors. Because several questionnaire blocks were path-dependent, we conducted exploratory factor analysis and confirmatory factor analysis using the complete 12-item AR-evaluation block (Q24; $n = 169$), whereas regression analyses included cases with complete Q24 and six-item satisfaction block responses (Q28; $n = 135$). The Q24 items supported a two-dimensional structure consisting of positive AR experience and technical concerns. In the conventional ordinary least squares model, positive AR experience showed the dominant positive association with core satisfaction (standardized $\beta = 0.747$, $p < 0.001$), whereas technical concerns showed a smaller positive coefficient under raw scoring (standardized $\beta = 0.121$, $p = 0.032$). However, HC3-robust standard errors and response-pattern sensitivity analyses attenuated the technical concerns estimate, so that it was no longer statistically distinguishable from zero. The study is cross-sectional and supports associations rather than causal effects. The main defensible conclusion is therefore that visitor satisfaction co-varies primarily with a coherent positive AR experience dimension, whereas the concern-related dimension remains measurable but is subject to greater inferential uncertainty in this sample.

Keywords: Visitor satisfaction; Museum augmented reality experience; On-site survey; Factor analysis; Empirical study

***Corresponding author:**
Guangyi Wu
(Guangyi.Wu22@student.xjtlu.edu.cn)

Citation: Wu G, Bernardo N. Augmented reality in Chinese museums: A case study on visitor experience. *Arts & Communication*. doi: 10.36922/AC026300057

Received: July 21, 2026

Revised: August 19, 2026

Accepted: September 14, 2026

Published online: October 8, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons AttributionNoncommercial License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Augmented reality (AR) has become an integral component in contemporary museum exhibitions. Adopting the technology has modernized museums and is associated with increased visitor engagement.[1] Digital overlays enable interactive learning experiences by bringing contextual information to life through visual animations. AR technology enables layering virtual elements onto physical artifacts, which is associated with greater visitor involvement through physical–digital reciprocal interactions within exhibition spaces. Often playful and interactive, content also became better aligned with the expectations of younger generations, contributing to broader visitor demographics. However, research on its application and on visitor interaction and engagement remains young and not broadly applicable when considering factors such as local context and culture, technology acceptance, or existing infrastructure.

Museums are non-profit institutions serving society through the collection, preservation, interpretation, and exhibition of tangible and intangible heritage.[2] In recent years, museums have undergone extensive digital transformation, integrating information and communication technologies such as interactive kiosks, projection mapping, virtual reality (VR), and AR. These technologies go beyond conventional exhibition formats by introducing sensory-driven and interactive systems that can enhance visitor engagement and reshape the museum experience.

Digital technologies have transformed visitors from passive observers into active participants within semi-immersive environments. AR and VR enable visitors to interact with digital content in real time while staying aware of their physical surroundings, fostering curiosity, improving information retention, and supporting personalized learning experiences.[3] Mixed reality (MR) further strengthens the social dimension of museum visits by reducing visitor isolation and encouraging interaction among visitors and museum staff.[4] In disciplines such as archaeology, integrating three-dimensional and extended reality technologies has enabled more active engagement with historical artifacts and spatial data, enhancing interpretative depth.[5]

The adoption of digital technologies has accelerated with the emergence of digital-smart museums, a trend further intensified by the COVID-19 pandemic. Mobile applications and virtual platforms have allowed museums to maintain educational functions during periods of restricted physical access.[6] Digital content is increasingly tailored to reflect site-specific historical and cultural contexts, reinforcing locality and narrative authenticity.[7] Successful implementations include AR and VR for remote and on-site engagement at the National Museum of Singapore[8], as well as digital game-based learning approaches in children's museums, which promote learning through play and enhance educational outcomes.

From a research perspective, AR is commonly situated within the reality–virtuality continuum, occupying an intermediate, semi-immersive position that bridges physical and digital environments.[9] The development of open-source frameworks such as ARToolKit and the later commercialization of wearable devices, including Google Glass and Microsoft HoloLens, have driven sustained academic interest in AR-based museum applications.[10] Bibliometric analyses reveal a steady increase in AR–museum research, with a notable rise after 2020, reflecting growing demand for digital and remote cultural experiences.[11]

Empirical studies consistently demonstrate that AR enhances visitor engagement and supports informal learning by embedding virtual content into physical exhibition spaces. AR-based reconstructions and storytelling approaches enable visitors to better understand historical artifacts and contexts.[12,13] Evaluative research further suggests that AR strengthens narrative connection, emotional engagement, and knowledge acquisition when appropriately designed.[14,15]

Current AR research in museums broadly follows two directions. The first focuses on usability and human–computer interaction, examining how interface design, navigability, and feedback influence user experience and acceptance.[11,16,17] The second investigates visitors' emotional and psychological responses, including memory, comfort, and satisfaction, using qualitative interviews and quantitative methods such as eye-tracking and usability scales.[18,19,20] Theoretical models, such as the MR visitor experience model, emphasize usability and interaction quality as central drivers of engagement[21], while broader immersive technology studies indicate that AR and VR enhance enjoyment and well-being.[22,23]

While previous studies have demonstrated the educational and experiential value of AR in museums, most empirical evidence comes from Europe, North America, and other technologically mature museum environments. Comparatively fewer studies have examined visitor responses to AR applications in Chinese museums, where cultural contexts, exhibition practices, and visitor expectations may differ. Cultural differences, varying levels of digital literacy, and distinct museum visitation patterns may influence how visitors evaluate AR technologies, making findings from other contexts difficult to generalize directly to Chinese museums. This highlights the need for further empirical investigation in Chinese museum settings.

Although AR in museums is increasingly regarded as a mature and rapidly expanding research field[10], several gaps remain. Many studies document overall benefits of AR for engagement or learning, yet fewer report how visitors' on-site evaluations empirically organize into measurement dimensions and how those dimensions relate to satisfaction. In particular, theoretically motivated facets such as ease of use, informational value, and interactivity are often discussed as distinct drivers, but they may collapse into a broader positive-experience factor in field data. At the same time, design discourse frequently notes technical frictions and authenticity-related concerns, which warrant explicit measurement alongside enabling evaluations. Evidence from Chinese museum settings also remains comparatively limited relative to studies conducted in Europe, North America, and other technologically mature

museum environments.

This study examines how experiential and technical evaluations relate to visitor satisfaction in AR-enhanced museum exhibitions within a Chinese museum context. Drawing on the technology acceptance model (TAM) and the information systems success model as design heuristics for item coverage, the questionnaire included statements related to clarity and usability, informational value, engagement and endorsement of AR, and technical or authenticity-related concerns. Rather than treating the four a priori facets as confirmed latent predictors, the analysis first estimates the empirical dimensionality of the AR evaluation items and then examines associations between the supported composites and satisfaction. This sequence keeps the measurement and regression models aligned and avoids overclaiming a four-factor structure not supported by the present sample.

Specifically, this study aimed to:

(i) Identify the key dimensions of visitors' AR-related evaluations using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), and construct composite scores that are consistent with the supported measurement structure.
(ii) Estimate the associations between the supported AR evaluation dimensions and visitor satisfaction, using a primary satisfaction composite that reduces conceptual overlap with predictor content, together with sensitivity checks.
(iii) Describe the participant flow from the full on-site sample to the analytical subsamples, compare demographic profiles across samples, and report covariate-adjusted associations where theoretically relevant covariates are available.

Based on the measurement strategy above, two primary hypotheses were specified for the regression stage:

H1: Positive AR experience is positively associated with visitor satisfaction in AR-enhanced museum exhibitions.

H2: Technical concerns are associated with visitor satisfaction in AR-enhanced museum exhibitions after accounting for positive AR experience. The direction of this association is treated as an empirical question under consistent raw scoring (higher scores = greater endorsement of concern statements), rather than as a pre-fixed negative effect.

The earlier four-facet hypothesis list (ease of use, information effectiveness, interactivity, and technical concerns as four simultaneous predictors) is not retained as the primary confirmatory framework, because the EFA did not support four separable factors and ease of use was represented by a single item.

Content related to information and interactivity is retained interpretively as part of positive AR experience rather than as separate regression predictors.

We selected Suzhou Museum as the study site because it is a national first-grade museum in China with active implementation of augmented reality exhibitions and a diverse visitor population, providing a representative context for this study.

We seek to address the following research questions (RQ):

(i) RQ1: What empirical dimensions organize visitors' evaluations of AR-enhanced museum exhibitions in this on-site sample?
(ii) RQ2: How are positive AR experience and technical concerns associated with visitor satisfaction, after mutual adjustment and selected covariates?
(iii) RQ3: How do the factor-analysis and regression subsamples compare with the full on-site sample in demographic profile, and what does the participant flow imply for interpretation?

The remainder of this paper is organized as follows. **Section 2** describes the study design, questionnaire development, measurement approach, and statistical analyses. **Section 3** reports sample characteristics, factor-analytic results, and regression findings. **Section 4** discusses the results, limitations, and implications. **Section 5** concludes.

2. Methods

This study used a cross-sectional, on-site intercept survey design. Visitors interacted with the museum AR content under ordinary exhibition conditions, completed the bilingual questionnaire on site, and then entered the measurement and regression pipeline summarized in **Section 3**. The statistical workflow comprised descriptive screening, factor analysis of the Q24 AR-evaluation block, construction of measurement-aligned composites, regression modeling of visitor satisfaction, and robustness checks using HC3-robust standard errors and response-pattern sensitivity analyses.

2.1. Data collection

To address the research questions, this study used a structured questionnaire and non-probability convenience sampling through an on-site intercept approach (**Table S1**). Unlike many AR-in-museum studies that rely on online surveys or post-visit recall, data collection was conducted in situ at the moment of visitor interaction with AR-integrated exhibits, enabling a more context-sensitive and experience-proximal assessment of AR user experience. The study was reviewed and approved

by the Ethics Committee of the Design School at Xi'an Jiaotong-Liverpool University (approval number: 00100000897202404151129 56). Data collection was conducted at Suzhou Museum, China, during the first week of June 2024. Researchers approached participants within museum spaces after they had interacted with AR-integrated exhibits and invited them to complete the anonymous questionnaire on site.

Sampling was conducted across public areas of the main Suzhou Museum, with intercept locations near AR exhibits and visitor rest areas. Data collection was performed between June 1 and June 7, 2024, during regular museum opening hours and covered both weekdays and weekends. Visitors were invited to participate through a brief verbal introduction outlining the study purpose, the voluntary nature of participation, and the expected completion time of approximately 5–10 minutes. Eligibility was limited to on-site museum visitors who had interacted with AR-related exhibition content. The survey was anonymous and collected no direct personal identifiers. Eight respondents subsequently selected the under-18 age category. Because the current archival file set does not preserve separate documentation of guardian consent, this study transparently reports this age composition and acknowledges the absence of age-specific consent records in the working archive as a limitation that warrants caution in interpretation and should be addressed in future replication planning.

2.2. Augmented reality exhibits and protocol

The AR-enabled exhibits used in this study were documented to enhance transparency and replicability. Visitors primarily accessed AR content using their own smartphones, which represented a range of device models and operating system versions. When compatible devices were unavailable, researchers provided tablets to ensure participation. The AR system employed a marker-based design in which visitors scanned designated exhibit markers to access digital overlays. These overlays combined explanatory text, video content, and interactive visual elements intended to enrich visitors' understanding of the exhibits.

After interacting with the AR exhibits, participants completed the questionnaire on site. After launching the AR application, visitors scanned exhibit markers and engaged with the associated digital content before completing the questionnaire.

A typical interaction involved scanning an artifact, viewing the AR overlays, and exploring available interactive features, including explanatory text, video materials, and basic user interactions. Visitors generally spent several minutes engaging with each AR exhibit prior to completing the survey. The interaction process reflected typical museum visitation behavior and was not experimentally manipulated. This naturalistic, cross-sectional observational approach captured associations between visitors' evaluations and satisfaction under ordinary visit conditions.

2.3. Questionnaire development and measures

The questionnaire was developed for this on-site study of Suzhou Museum's mobile AR application rather than adopted as a single pre-existing multi-item battery with published scoring keys. Item coverage was guided by three design heuristics commonly used in technology-experience research: the TAM (perceived ease of use and usefulness), information-systems success ideas concerning information and system quality, and museum AR user-experience themes emphasized in prior work (engagement and interpretive value, technical friction, authenticity or realism concerns, distraction, and learning-related appraisal).[14,15,17,19,22] Within that framework, the research team drafted context-specific statements referring to the museum app's description and content clarity, onboarding support, expected usefulness, additional insights into exhibits, engagement, navigation ambiguity, authenticity and realism concerns, distraction, future-oriented endorsement, recommendation, and willingness to try new features (Q24), together with multi-attribute satisfaction ratings of overall experience, technical quality, information, interactivity, and learning contribution (Q28). An item-to-facet map is provided in the supplementary materials (item source map). Because several theoretically labeled facets were represented by few items—and ease-of-use content by a single statement—the study treated these facets as design heuristics for item writing, not as a priori confirmed multi-item scales.

The instrument was prepared in a bilingual Chinese–English format and administered on site through an online survey platform, with Chinese as the primary language for most visitors and English displayed in parallel for semantic support. The research team reviewed the wording for clarity and appropriateness to the museum context. The available study records did not document a formal external expert content-validity panel with quantitative content-validity indices, and a separate cognitive interview or pilot sample with an item-revision log. Accordingly, this study does not claim full content validation of the questionnaire. Instead, it reports transparent item coverage, the empirical factor structure of the 12-item AR-evaluation block (Q24), and internal-consistency evidence for the composites used in regression. Demographic, AR-familiarity, and visit-use items preceded the evaluation blocks; conditional display

logic routed respondents according to prior AR awareness and app-use pathways, which is why complete-case counts differ across analyses.

2.4. Data preprocessing, response-quality screening, and composite construction

We conducted data completeness checks before analysis. For descriptive analyses, we used the full retained sample (n = 420). Questionnaire items in the AR-evaluation block (Q24) were then examined using EFA and CFA to evaluate empirical dimensionality. Following the supported two-factor solution, we calculated composite scores as arithmetic means of the corresponding items: positive AR experience (Q24_01, Q24_03, Q24_04, Q24_05, Q24_06, Q24_10, Q24_11, Q24_12) and technical concerns (Q24_02, Q24_07, Q24_08, Q24_09). Technical concerns were analyzed in the raw item direction, so that higher values indicate greater endorsement of concern-related statements. Visitor satisfaction items (Q28) were summarized as (i) a primary core composite of overall experience, satisfaction appraisal, and technical quality (Q28_01–Q28_03), chosen to reduce conceptual overlap with informational and interactive predictor content, and (ii) a full six-item composite (Q28_01–Q28_06) used in sensitivity analysis (Table 1). All focal items used five-point scales, so composite means remain on the original metric.

Since several questionnaire blocks were conditionally displayed based on visitors' prior AR awareness and app-use pathways, valid case counts differed across analyses. Descriptive analyses used the full retained sample (n = 420; **Table S2**). Factor analyses required complete Q24 responses (n = 169). Regression analyses required complete Q24 and Q28 responses (n = 135); 34 respondents completed Q24 but not Q28. Within the regression sample, two response-pattern indicators were computed for sensitivity analyses: the longest within-block run of identical Q24 responses and an indicator for respondents who selected the same response category for all 12 Q24 items. These indicators were not used as post hoc primary exclusion criteria, but they were used to test whether the technical concerns coefficient was robust to plausible low-information response screens. The findings support associations only and should not be interpreted as evidence that specific AR features cause satisfaction.

In the complete Q24–Q28 regression sample, the longest within-block run of identical Q24 responses ranged from 2 to 12, and 35 respondents (25.9%) selected the same category for all 12 Q24 items. Because uniformly positive or uniformly negative patterns can be genuine in a museum satisfaction context, these cases were not automatically deleted from the primary model. Instead, they were retained in the complete-case analysis and then examined through sensitivity analyses that excluded fully invariant Q24 patterns or restricted the sample to respondents with a longest run of 10 or fewer identical Q24 responses.

2.5. Diagnostic factor analysis

We conducted EFA in SPSS to examine the dimensionality of the 12 Q24 items. Sampling adequacy and sphericity were evaluated using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity. Principal axis factoring with Promax rotation was used because related evaluative dimensions were expected. Interpretation of dimensionality followed common multivariate-analysis guidance, with emphasis on eigenvalues, the scree pattern, and the substantive interpretability of the rotated loading matrix.[24] The EFA was treated as the primary basis for deciding how many composites to carry into regression, rather than forcing a four-predictor scheme that was not recovered empirically.

2.6. Confirmatory factor analysis

Confirmatory factor analysis was conducted in R using the lavaan package and the weighted least squares mean-and-variance adjusted estimator, which is appropriate for ordinal questionnaire items.[25] The CFA evaluated the two-factor structure suggested by the EFA on the same complete-Q24 subsample. Because the CFA used the same

Table 1. Variable definitions and analytical roles

Variable	Role in analysis	Items
Visitor satisfaction (core, primary)	Dependent variable	Q28_01, Q28_02, Q28_03
Visitor satisfaction (full, sensitivity)	Dependent variable (sensitivity)	Q28_01–Q28_06
Positive augmented reality experience	Predictor (Factor 1 composite)	Q24_01, Q24_03, Q24_04, Q24_05, Q24_06, Q24_10, Q24_11, Q24_12
Technical concerns (raw scoring)	Predictor (Factor 2 composite)	Q24_02, Q24_07, Q24_08, Q24_09 (higher = greater concern endorsement)

Notes: Technical concerns were analyzed in their original raw direction throughout exploratory factor analysis, confirmatory factor analysis, reliability, and regression; higher scores indicate greater endorsement of concern statements. No numerical reverse transformation was applied. Item labels therefore omit any "R" suffix; negatively worded item stems are retained only as wording, not as reversed scores.

respondents as the EFA, and because residual covariances suggested by modification indices were considered during model refinement, the CFA is reported as same-sample evaluation and exploratory refinement of the measurement structure, not as independent cross-validation of the scale. In addition, all positive AR experience indicators are positively worded, and all technical concerns indicators are negatively worded, so wording/valence is perfectly confounded with factor membership in the present item set.

Model fit was assessed using the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Standardized factor loadings were examined for each item. Composite reliability (CR) and average variance extracted (AVE) were computed for the two factors. We interpreted global fit using widely used guideline ranges, acknowledging that such cut-offs are heuristic rather than absolute decision rules.[25,26] We also reported Cronbach's alpha for the regression composites, including the satisfaction measures.

Reliability and convergent validity for the two-factor solution were summarized using CR and AVE. Discriminant evidence was inspected via factor correlations and Fornell–Larcker-style comparisons where applicable. These indices describe the two empirically supported factors; they do not support claims of a fully validated multi-scale battery beyond the present sample.

2.7. Reliability analysis

Cronbach's alpha coefficients were calculated for positive AR experience, technical concerns, the primary core satisfaction composite (Q28_01–Q28_03), and the full satisfaction composite (Q28_01–Q28_06). Reliability was reported for the factor subsample and/or the regression subsample as appropriate to each analysis.

2.8. Multiple regression analysis

Multiple linear regression was used to estimate associations between the two AR-evaluation composites and visitor satisfaction. The primary model specified core satisfaction as the outcome and positive AR experience and technical concerns as focal predictors. An adjusted model additionally included gender (female indicator), age group (18–24 versus other), and prior AR awareness. Further models entered on-site AR use frequency or average session duration where available. A sensitivity model replaced the outcome with the full six-item satisfaction mean. For each model, we report unstandardized coefficients, standard errors, standardized coefficients, *t*-statistics, *p*-values, 95% confidence intervals (CIs), R^2 values, and variance inflation factors. Diagnostic checks included residual distribution, an approximate Breusch–Pagan assessment of heteroscedasticity, and Cook's distance for influential observations. Because residual non-normality and response-pattern concentration remained concerns, we additionally report HC3-robust standard errors for the primary model[27] and compare the technical concerns coefficient after excluding respondents with fully invariant Q24 response patterns. All regression results are interpreted as associations in a cross-sectional observational design.

2.9. Analytical rationale

The analytical strategy prioritized aligning the measurement model with the regression model. EFA/CFA informed a two-composite specification; regression then tested how those composites relate to satisfaction. Theoretically motivated facets (e.g., informational value and interactivity) remain useful for item design and qualitative interpretation of positive AR experience, but they are not treated as four separately validated predictors in the primary quantitative model. This choice responds directly to the need for coherence between factor-analytic evidence and explanatory modeling in the present dataset.

3. Results

3.1. Demographic information

Table 2 summarizes the demographic characteristics of the 420 participants. Among respondents, 125 (29.8%) were male, 282 (67.1%) were female, and 13 (3.1%) preferred not to say. Most participants were aged 18–24 (n = 361, 86.0%), and 378 (90.0%) reported prior awareness of AR. The factor-analysis subsample (complete Q24, n = 169) and the regression subsample (complete Q24 and Q28, n = 135) remained predominantly young and AR-aware, with somewhat higher female shares than the full sample (factor: 72.2% female; regression: 74.1% female). Chi-square (χ^2) tests comparing full-sample membership versus regression-subsample membership indicated demographic differences for gender ($\chi^2(2) = 6.39$, $p = 0.041$), age group ($\chi^2(5) = 20.32$, $p = 0.001$), and AR awareness ($\chi^2(1) = 5.94$, $p = 0.015$). These contrasts are expected under path-dependent completion of app-use and evaluation blocks and are treated as scope conditions for generalization rather than as post hoc exclusions.

The participant-flow logic implies that the EFA sample includes visitors who completed the Q24 evaluation block after routed exposure to the AR content, regardless of whether they later completed the full satisfaction block. Thus, the EFA sample is broader than the regression sample but still anchored in respondents who progressed to the AR-evaluation section of the routed questionnaire.

Table 2. Descriptive statistics for demographic characteristics

Category	Frequency	Percentage (%)
Gender		
Male	125	29.8
Female	282	67.1
Prefer not to say	13	3.1
Age		
Under 18	8	1.9
18–24	361	86.0
25–34	15	3.6
35–44	13	3.1
45–54	8	1.9
55–64	15	3.6
65 or above	0	0.0
Education		
Primary education	2	0.5
Middle school	1	0.2
High school	109	26.0
Bachelor's degree	286	68.1
Master's degree	22	5.2
PhD or equivalent	0	0.0
Augmented reality awareness		
Yes	378	90.0
No	42	10.0

3.2. Exploratory factor analysis

Exploratory factor analysis examined the structure of the 12 Q24 items among respondents with complete Q24 data (n = 169). The KMO value was 0.888, and Bartlett's test of sphericity was significant, indicating adequate factorability.

The eigenvalue analysis indicated that only the first two factors exceeded the commonly accepted threshold of 1.0, with eigenvalues of 6.491 and 2.333, respectively. The scree plot (Figure 1) also showed a clear inflection after the second factor, supporting a two-factor solution.

Principal axis factoring with Promax rotation yielded a clear two-factor pattern (Table 3). Factor 1 (positive AR experience) comprised high loadings for Q24_01, Q24_03, Q24_04, Q24_05, Q24_06, Q24_10, Q24_11, and Q24_12, covering clarity/usability-related appraisal, informational and engagement evaluations, and forward-looking endorsement of museum AR. Factor 2 (technical concerns) comprised Q24_02, Q24_07, Q24_08, and Q24_09, covering navigation ambiguity, authenticity disruption, reduced realism, and distraction. Cross-loadings were low relative to primary loadings. This pattern does not recover four separate factors corresponding to ease of use, information effectiveness, interactivity, and technical concerns; instead, the first three content areas cohere within a single positive-experience dimension in this sample.

The two retained factors jointly explained 73.7% of the total variance, with Factor 1 accounting for 50.2% and Factor 2 for 23.5%. These results provide provisional same-sample support for a two-factor pattern and motivate the subsequent confirmatory evaluation on the same subsample.

3.3. Confirmatory factor analysis (same-sample evaluation)

Confirmatory factor analysis evaluated the two-factor structure on the same complete-Q24 subsample (n = 169; Figure 2). Standardized loadings were substantial for both factors (Table 4). The retained two-factor model showed close global fit (CFI = 0.999, TLI = 0.999, RMSEA = 0.049, 90% CI: 0.016–0.075, SRMR = 0.042; Table 5). CR and AVE supported internal consistency and convergent evidence for positive AR experience (CR = 0.969, AVE = 0.796) and technical concerns (CR = 0.861, AVE = 0.614) (Table 6 and Figure 3). Figure 4 reports the corresponding standardized loadings, while Figure 5 summarizes the Fornell–Larcker comparison. Because the CFA used the same subsample as the EFA, and modification indices suggested residual covariances by entered model refinement, these results are interpreted as preliminary same-sample evidence rather than independent confirmation. Cronbach's alpha for the regression composites was high (positive AR experience α = 0.957; technical concerns α = 0.853; core satisfaction α = 0.955; full satisfaction α = 0.974; n = 135).

3.4. Regression analysis results

Multiple linear regression examined associations between the two measurement-aligned composites and visitor satisfaction in the complete Q24–Q28 subsample (n = 135). Technical concerns were entered in the raw metric (higher = greater concern endorsement). The primary outcome was core satisfaction, defined as the mean of Q28_01–Q28_03.

The primary two-predictor model was statistically significant and explained a large share of variance in core satisfaction (R^2 = 0.633, adjusted R^2 = 0.628, $F(2, 132) = 113.93$, $p < 0.001$; Table 7; Figure 6). Positive AR experience was strongly and positively associated with core satisfaction (B = 0.827, standardized β = 0.747, $p < 0.001$). Technical concerns showed a smaller positive coefficient under conventional ordinary least squares estimation

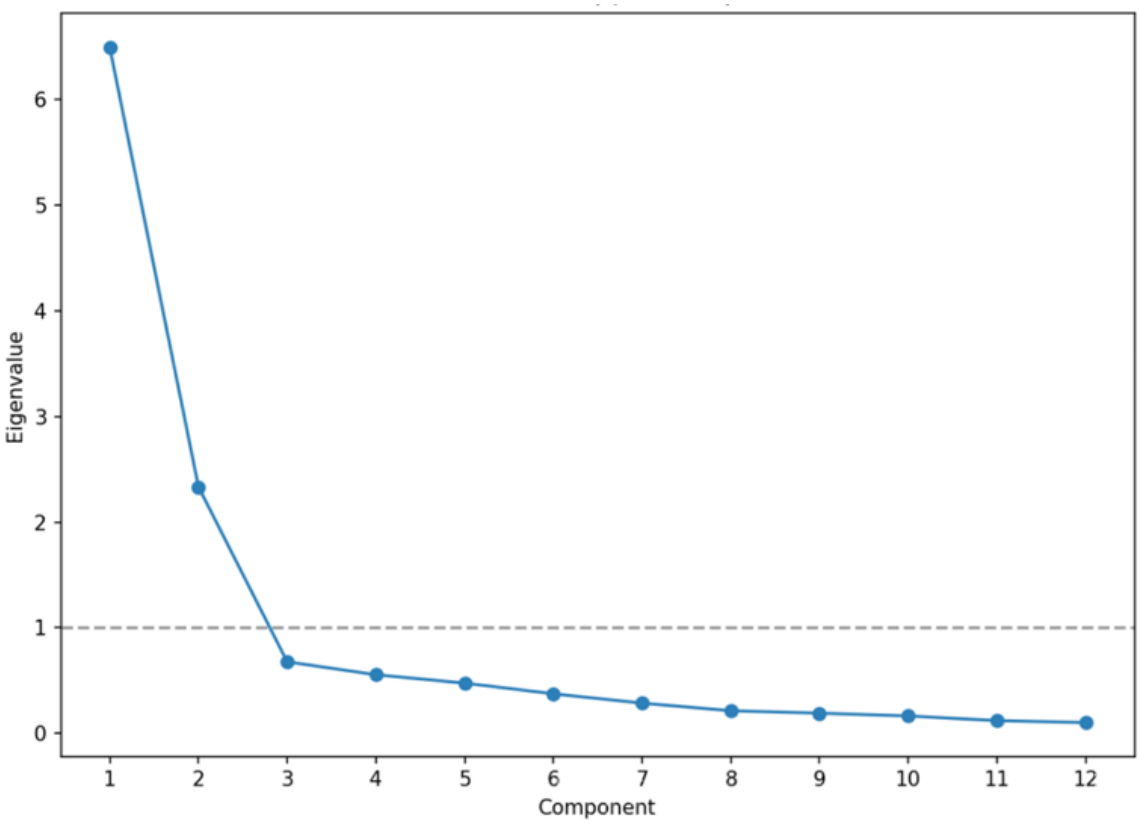


Figure 1. Scree plot of eigenvalues for the Q24 augmented reality-evaluation block

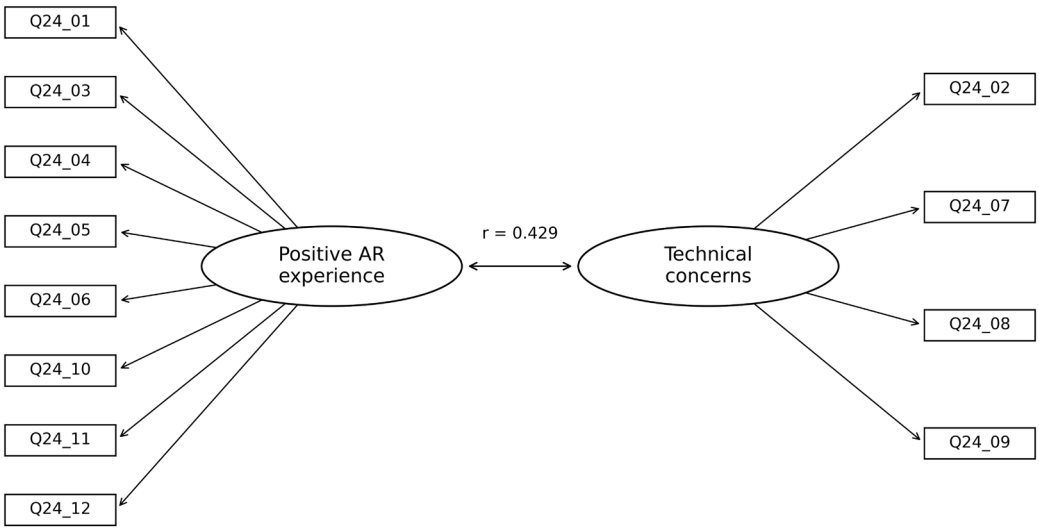


Figure 2. Two-factor measurement model for the Q24 block using reader-friendly construct labels
Abbreviation: AR: Augmented reality

Table 3. Rotated factor loadings and communalities from the exploratory factor analysis

Item	Factor 1	Factor 2	h2
Q24_01	0.820	−0.075	0.678
Q24_02	0.102	0.711	0.516
Q24_03	0.877	0.018	0.769
Q24_04	0.887	−0.023	0.787
Q24_05	0.850	0.022	0.723
Q24_06	0.897	−0.029	0.806
Q24_07	0.068	0.820	0.677
Q24_08	−0.098	0.925	0.865
Q24_09	−0.077	0.879	0.779
Q24_10	0.781	0.082	0.617
Q24_11	0.896	−0.012	0.803
Q24_12	0.907	0.022	0.823
Rotated sums of squared loadings: Factor 1 = 6.022, Factor 2 = 2.821 (initial eigenvalues were 6.491 and 2.333; see scree plot)	Rotated sums of squared loadings: Factor 1 = 6.022, Factor 2 = 2.821 (initial eigenvalues were 6.491 and 2.333; see scree plot)	Rotated sums of squared loadings: Factor 1 = 6.022, Factor 2 = 2.821 (initial eigenvalues were 6.491 and 2.333; see scree plot)	Rotated sums of squared loadings: Factor 1 = 6.022, Factor 2 = 2.821 (initial eigenvalues were 6.491 and 2.333; see scree plot)
Variance explained: Factor 1 = 50.2%, Factor 2 = 23.5%	Variance explained: Factor 1 = 50.2%, Factor 2 = 23.5%	Variance explained: Factor 1 = 50.2%, Factor 2 = 23.5%	Variance explained: Factor 1 = 50.2%, Factor 2 = 23.5%
Cumulative = 73.7%	Cumulative = 73.7%	Cumulative = 73.7%	Cumulative = 73.7%

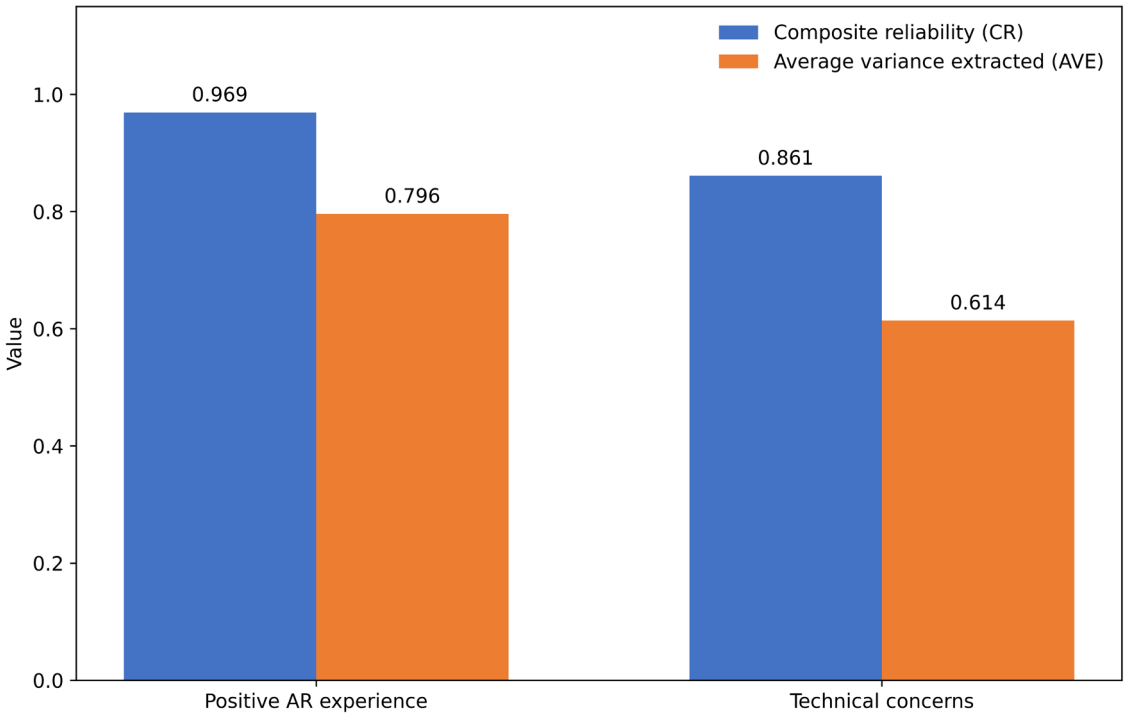


Figure 3. Composite reliability and average variance extracted for the retained constructs
Abbreviation: AR: Augmented reality

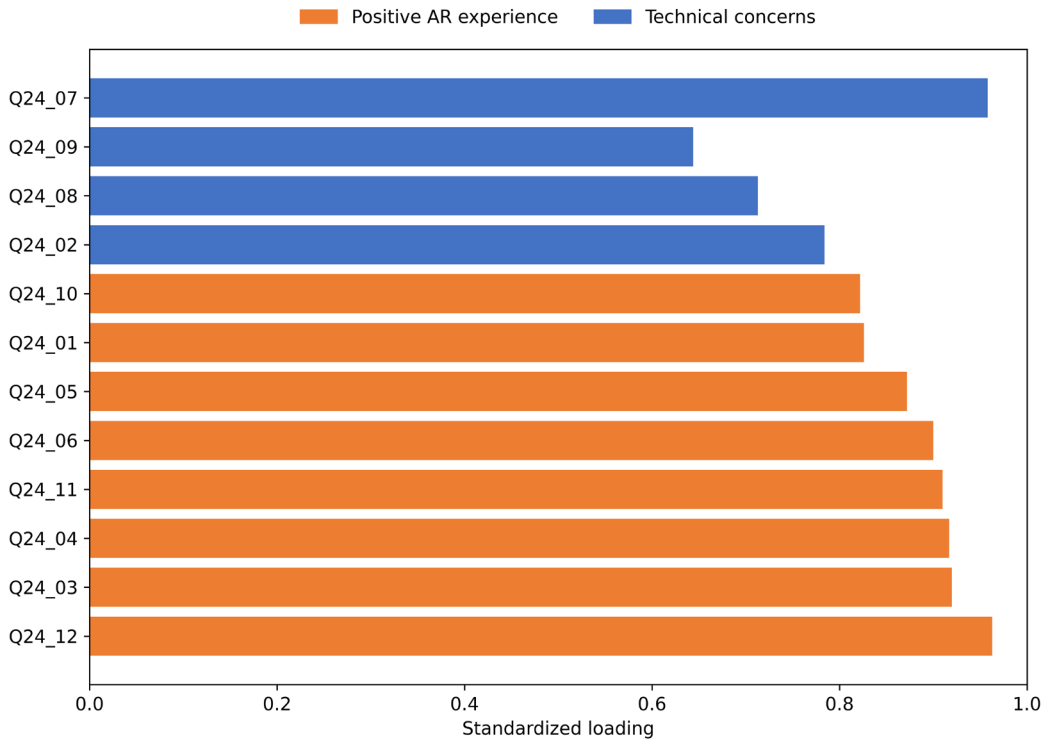


Figure 4. Standardized factor loadings for the retained two-factor model
Abbreviation: AR: Augmented reality

(B = 0.099, standardized β = 0.121, p = 0.032). Variance inflation factors (VIFs) were low (both VIFs = 1.13), indicating that the two-factor specification avoids the previous four-predictor collinearity between information- and interactivity-related composites. Diagnostic screening showed non-normal residuals (Shapiro–Wilk $p < 0.001$), no strong heteroscedasticity signal at the conventional 0.05 threshold in the primary model (approximate Breusch–Pagan p = 0.061), and a small number of observations with Cook's distance above $4/n$ (n = 7; maximum Cook's D = 0.26). Zero-order associations were positive for both predictors (r[positive AR experience, core satisfaction] = 0.788, 95% CI: 0.713–0.844; r[technical concerns, core satisfaction] = 0.370, 95% CI: 0.215–0.508; n = 135). The predictors were also positively correlated (r = 0.333, 95% CI: 0.174–0.476). Therefore, the smaller positive partial coefficient for technical concerns does not reflect a classical sign-reversal suppression pattern; after mutual adjustment, the association is attenuated relative to the zero-order correlation and is not robust under HC3 inference. Under HC3-robust standard errors, positive AR experience remained clearly associated with core satisfaction, whereas technical concerns were no longer statistically distinguishable from zero (HC3 95% CI: −0.005–0.203, p = 0.062).

Table 4. Standardized factor loadings from the confirmatory factor analysis

Indicator	Standardized loading	*p*-value
Q24_01	0.826	<0.001
Q24_03	0.920	<0.001
Q24_04	0.917	<0.001
Q24_05	0.872	<0.001
Q24_06	0.900	<0.001
Q24_10	0.822	<0.001
Q24_11	0.910	<0.001
Q24_12	0.963	<0.001
Q24_02	0.784	<0.001
Q24_07	0.958	<0.001
Q24_08	0.713	<0.001
Q24_09	0.644	<0.001

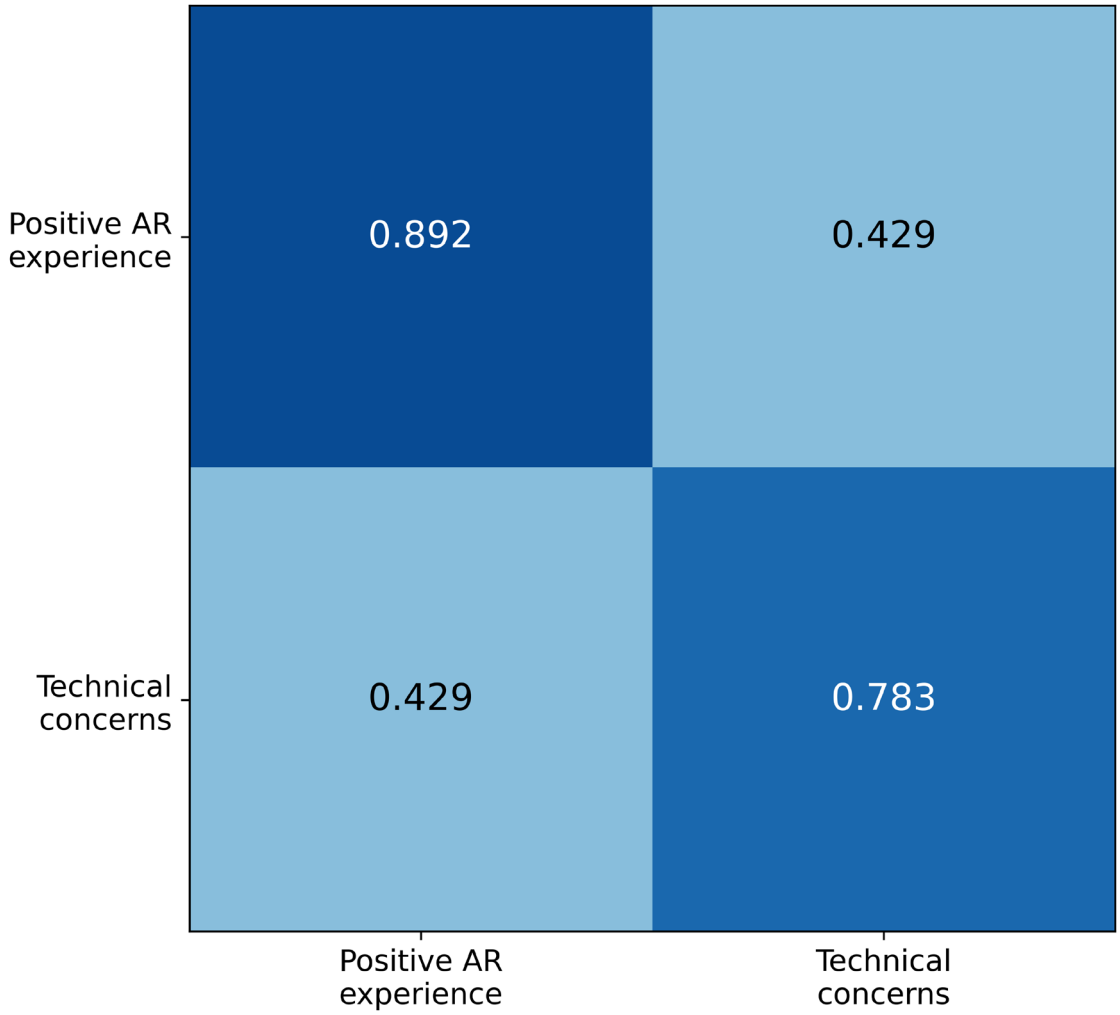


Figure 5. Fornell–Larcker comparison for the retained constructs (diagonal = square root of AVE; off-diagonal = absolute correlation)
AR: Augmented reality; AVE: Average variance extracted

Table 5. Confirmatory factor analysis model-fit indices

Model	CFI	TLI	RMSEA (90% CI)	SRMR
Retained	0.999	0.999	0.049 (0.016– 0.075)	0.042

Abbreviations: CFI: Comparative fit index; RMSEA: Root mean square error of approximation; SRMR: Standardized root mean square residual; TLI: Tucker–Lewis index.

Table 6. Composite reliability and average variance extracted

Construct	CR	AVE
Positive AR Experience	0.969	0.796
Technical Concerns	0.861	0.614

Abbreviations: AR: Augmented reality; AVE: Average variance extracted; CR: Composite reliability.

Covariate-adjusted models yielded the same directional ordering under conventional estimation. After including female gender, an indicator for age 18–24, and prior AR awareness, positive AR experience remained dominant (standardized $\beta = 0.748$, $p < 0.001$), whereas technical concerns retained a smaller positive coefficient (standardized $\beta = 0.135$, $p = 0.021$); demographic covariates were not significant. Adding on-site use frequency did not materially change the focal pattern. Adding average session duration left positive AR experience dominant

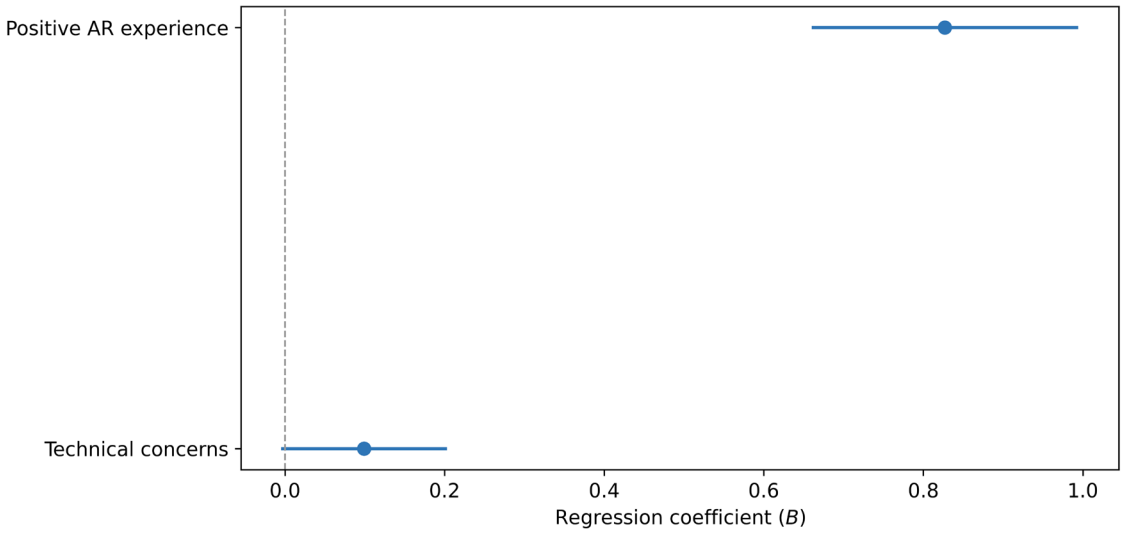


Figure 6. Primary regression coefficients with HC3-robust 95% confidence intervals. Note: HC3 = heteroskedasticity-consistent standard error estimator Abbreviation: AR: Augmented reality

Table 7. Primary regression results with conventional and HC3-robust inference

Predictor	*B*	SE (OLS)	SE (HC3)	Standardized β	95% CI (HC3)	*p*-value (OLS)	*p*-value (HC3)
Intercept	0.350	0.254	0.280	–	−0.20–0.904	0.170	0.213
Positive AR experience	0.827	0.062	0.084	0.747	0.660–0.994	< 0.001	< 0.001
Technical concerns (raw scoring)	0.099	0.046	0.053	0.121	−0.005–0.203	0.032	0.062

Notes: Core satisfaction = mean of Q28_01–Q28_03. Technical concerns were scored in the raw direction (higher values = greater concern endorsement). HC3 = heteroskedasticity-consistent standard error estimator.
Abbreviations: AR: Augmented reality; B: Regression coefficient; SE: Standard error; OLS: Ordinary least squares.

(standardized β = 0.715, $p < 0.001$), and longer reported sessions were also positively associated with core satisfaction (standardized β = 0.136, $p = 0.013$).

A sensitivity model using the full six-item satisfaction mean produced closely similar estimates ($R^2 = 0.663$; positive AR experience standardized β = 0.766, $p < 0.001$; technical concerns standardized β = 0.119, $p = 0.028$). Because several full-scale satisfaction items refer to information, interactivity, and learning, the primary reporting emphasis remains on the core composite; the full-scale model is treated as a robustness check. A further response-pattern sensitivity analysis, following established guidance on careless or low-information response patterns[28], is summarized in Table 8. 35 of the 135 complete Q24–Q28 cases (25.9%) provided identical responses to all 12 Q24 items, and once those fully invariant patterns were excluded, the HC3-robust technical concerns coefficient remained positive but was clearly imprecise ($B = 0.091$, 95% CI: −0.089–0.270, $p = 0.319$). Sensitivity composites that excluded Q24_01 from positive AR experience, or that excluded the forward-looking items Q24_10–Q24_12, left positive AR experience dominant (standardized β ≈ 0.73–0.74) with only modest changes in R^2 (0.618–0.628), indicating that the primary association is not solely an artifact of the single clarity/overall-appraisal item or of behavioral-intention wording. Nevertheless, positive AR experience remains a heterogeneous evaluative composite, and criterion contamination with satisfaction ratings is acknowledged as a limitation.

3.5. Relative contribution of the two dimensions

Comparing standardized coefficients in the primary model, positive AR experience accounts for most of the explained association with core satisfaction, while technical concerns contribute a smaller, secondary association. This ranking is stable under demographic adjustment and under the full satisfaction sensitivity outcome. The result supports a parsimonious account in which on-site satisfaction

Table 8. Response-pattern sensitivity analyses for the technical concerns coefficient

Sensitivity analysis	*n*	*B* for technical concerns	SE (HC3)	95% CI (HC3)	*p*-value (HC3)
Primary model (all complete Q24–Q28 cases)	135	0.099	0.053	−0.005– 0.203	0.062
Exclude respondents with all 12 Q24 responses identical	100	0.091	0.090	−0.089– 0.270	0.319
Restrict to longest within-Q24 run ≤ 10	99	0.091	0.091	−0.089– 0.270	0.319

Note: HC3 = heteroskedasticity-consistent standard error estimator.
Abbreviations: *B*: Regression coefficient; SE: Standard error; OLS: Ordinary least squares.

co-varies chiefly with a broad positive AR experience dimension recovered by factor analysis, rather than with four collinear theoretical facets entered simultaneously.

4. Discussion

The findings provide on-site evidence on how visitors' AR evaluations relate to satisfaction in a Chinese museum setting. Two results structure the interpretation. First, Q24 responses organize into a two-dimensional structure—positive AR experience and technical concerns—rather than four separable validated predictors. Second, in regression models aligned with that structure, positive AR experience shows the dominant positive association with satisfaction. Technical concerns retain a positive coefficient under conventional ordinary least squares estimation, but that coefficient is sensitive to HC3-robust inference and to response-pattern screening. The design is cross-sectional; coefficients are associations, not effects of manipulating AR features.

The prominence of positive AR experience is consistent with museum AR research that emphasizes interpretive value, engagement, and meaningful digital overlay on physical exhibits.[12-15] In the present item set, clarity/usability appraisal, informational insight, engagement, and endorsement are loaded together. For design practice, this coherence implies that visitors in this sample evaluated the AR offer holistically: content that is clear, informative, and engaging tends to travel with overall endorsement of museum AR, rather than as statistically independent levers in a four-variable regression.

Technical concerns formed a second factor whose indicators are all negatively worded, whereas positive AR experience indicators are all positively worded; factor membership is therefore perfectly confounded with item valence in this item set, and the two-factor pattern should not be over-interpreted as fully independent substantive constructs. Under raw scoring, higher concern endorsement was positively associated with satisfaction in the conventional model after adjustment for positive AR experience. However, the coefficient weakened under HC3-robust inference and after exclusion of fully invariant Q24 patterns. A cautious reading is therefore warranted: respondents who progressed through the app-use and evaluation path may have co-reported affirmative experience and residual technical trade-offs, but the data do not support a strong or stable claim that concern endorsement independently predicts higher satisfaction. Mean technical concerns remained near the scale midpoint (M = 3.16, SD = 1.10), so navigation clarity and careful integration with physical exhibits remain practically relevant design issues even if the partial coefficient is inferentially fragile in this sample.

Taken together, the two-factor results separate enabling evaluations from concern-related evaluations, while the regression results show that satisfaction tracks primarily with the enabling composite in this sample.

Single-item ease-of-use content is absorbed within positive AR experience rather than modeled as a standalone predictor. That choice follows the EFA solution and avoids an under-identified ease-of-use latent variable. In a sample dominated by 18–24-year-old visitors with high prior AR awareness, basic operability may function as a baseline condition that co-varies with broader positive appraisal instead of appearing as a separate satisfaction driver.

Methodologically, aligning regression predictors with the two-factor solution resolves the inconsistency that arises when a two-factor measurement model is paired with four composite predictors drawn from the same item pool. The CFA is reported as a same-sample evaluation with exploratory refinement, which matches the data reality of a single field subsample and prevents over-claiming independent scale confirmation. Adding HC3-robust inference and response-pattern sensitivity checks further clarifies which parts of the regression narrative are stable and which remain provisional.

Demographic scope conditions matter for interpretation. The full sample, and especially the regression subsample, is concentrated among young visitors, with high prior AR awareness and a higher female share in analytical samples than in the full intercept sample. Associations should be generalized cautiously to older visitors, first-time AR users, and other museum types.

Returning to the research questions, RQ1 is answered

by the two-dimensional structure of positive AR experience and technical concerns. RQ2 is answered more cautiously: satisfaction co-varies primarily with positive AR experience, whereas the technical concerns coefficient is smaller and inferentially sensitive once robust standard errors and response-pattern screens are applied. RQ3 is answered by the participant flow (420 → 169 → 135) and by the demographic contrasts between the full and analytical samples, which delimit the inferential scope of the multivariate findings.

4.1. Limitations

Several limitations should be noted. First, convenience on-site sampling and path-dependent item display produced analytical subsamples that differ demographically from the full intercept sample. Second, the cross-sectional design cannot identify causal effects of AR design features. Third, EFA and CFA used the same Q24-complete subsample; an independent validation sample was not available. Fourth, the questionnaire was study-developed for this museum AR context: bilingual team review supported basic clarity, but a formal external content-validity panel and a documented cognitive pretest or pilot revision log were not available; psychometric claims are therefore limited to preliminary same-sample structure and reliability evidence. Fifth, residual non-normality and a concentrated subset of invariant Q24 response patterns advise caution in interpreting marginal *p*-values, especially for technical concerns. Sixth, the current archival package does not preserve age-specific guardian-consent records for the under-18 subgroup, so ethics implementation for minors cannot be documented beyond the committee approval statement retained above. Seventh, even with a core satisfaction outcome, survey self-reports can retain shared-method variance.

4.2. Practical implications for Chinese museums

Despite these scope conditions, the study contributes on-site Chinese museum evidence that (i) AR evaluations can be summarized parsimoniously by a positive-experience dimension and a technical/authenticity-concern dimension, and (ii) satisfaction co-varies most strongly with the positive-experience composite when measurement and regression are aligned. For practice, museums can prioritize coherent, informative, and engaging AR content; streamline onboarding and task completion; and embed AR within an edutainment strategy that supports digital cultural heritage interpretation rather than distracting from it. In the present Chinese museum context, this means organizing AR content around clear exhibit narratives, simple scan-and-view interactions, and policy-compatible digital heritage communication, while continuing to monitor navigation clarity and authenticity fit as separable design concerns.

Overall, visitor satisfaction in this AR-enhanced exhibition setting is associated chiefly with positive AR experience. Technical concerns constitute a distinct evaluative dimension, but their raw-score association with satisfaction is smaller and not robust to all inferential checks applied here. The central methodological message is that predictor sets used in explanation should follow the empirically supported measurement structure and should be accompanied by transparent robustness checks when response-quality concerns are plausible. De-identified analytic files (survey_deidentified_N420, analytic_deidentified_key_vars, codebook, item-source map, and de-identification README) are provided in the supplementary package; direct identifiers and open-text responses were removed or transformed as described in the de-identification README.

5. Conclusion

This study investigated associations between AR-related visitor evaluations and satisfaction at Suzhou Museum, China. Using on-site survey data from 420 visitors, with factor analysis on complete Q24 responses (n = 169) and regression on complete Q24–Q28 responses (n = 135), the study estimated a two-dimensional evaluation structure and tested measurement-aligned associations with satisfaction.

Positive AR experience was the dominant positive correlate of visitor satisfaction. Technical concerns formed a separate factor and showed a smaller positive coefficient under conventional estimation, but HC3-robust inference and response-pattern sensitivity checks attenuated those estimates. These results are associative and specific to a predominantly young, AR-aware on-site subsample.

Factor analysis supported a coherent two-dimensional structure rather than four independently validated predictors. Keeping regression aligned with that structure yields a clearer and more defensible account of how on-site AR evaluations relate to satisfaction.

The study contributes field evidence from a Chinese museum context and a more transparent measurement-to-model workflow. For practice, AR implementation should emphasize informative, engaging, and coherent experiences, while continuing to track technical and authenticity-related concerns as a separate evaluative cluster and explicitly evaluating response quality in future survey waves.

Future research should recruit more age-diverse and less AR-familiar visitors, compare multiple museum types,

add independent measurement samples for confirmatory work, and combine surveys with behavioral traces of AR use. Pre-specified coding rules for concern items, formal response-quality screening plans, and preregistered robustness analyses would further strengthen cumulative inference.

Acknowledgments

The authors thank Xi'an Jiaotong-Liverpool University (XJTLU), particularly the Department of Industrial Design, Design School, for institutional support during data collection and manuscript preparation. The authors also thank Suzhou Museum staff and participating visitors.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: Guangyi Wu, Nuno Bernardo
Formal analysis: Guangyi Wu
Investigation: Guangyi Wu
Methodology: Guangyi Wu, Nuno Bernardo
Writing–original draft: Guangyi Wu
Writing–review & editing: Guangyi Wu, Nuno Bernardo

Ethics approval and consent to participate

This study was reviewed and approved by the Ethics Committee of the Design School at Xi'an Jiaotong-Liverpool University (approval no. 00100000897202404151129 56). Participation was voluntary. The questionnaire began with a consent statement explaining that completion of the questionnaire indicated consent to participate.

Consent for publication

Participants were informed that the study findings could be published in academic journals in aggregate form. Completing the questionnaire indicated consent to participate and to publish the findings in this form.

Availability of data

The data are not publicly available due to participant privacy considerations. DDe-identified analytic files (survey_deidentified_N420, analytic_deidentified_key_vars, codebook, item-source map, and de-identification README) are provided in the supplementary package; direct identifiers and open-text responses were removed or transformed as described in the de-identification README.

References

1. Hsieh T-J, Su Y-H, Lin L-S. Augmented Reality Art Museum Mobile Guide for Enhancing User Experience. *IEEE Comput Graph Appl*. 2025;45(1):12–27.

   doi: 10.1109/MCG.2025.3529981

2. International Council of Museums (ICOM). Museum definition. 2022. Accessed on December 22, 2024. https://icom.museum/en/resources/standards-guidelines/museum-definition/.

3. AlGerafi MAM, Zhou Y, Oubibi M, Wijaya TT. Unlocking the Potential: A Comprehensive Evaluation of Augmented Reality and Virtual Reality in Education. *Electronics*. 2023;12(18):3953.

   doi: 10.3390/electronics12183953

4. Galani A. Mixed Reality Museum Visits: Using new technologies to support co-visiting for local and remote visitors. *Museol Rev*. 2003;10:1–15.

5. Cobb PJ, Cobb E, Liang J, Kiyama R, Ng J. Theoretical foundations for archaeological pedagogy with digital 3D, virtual, augmented, and mixed reality technologies. *J Archaeol Educ*. 2024;8(1):1-29.

6. Puspasari S, Ermatita S, Zulkardi S. Constructing Smart Digital Media for Museum Education Post Pandemic Recovery: A Review and Recommendation. 2021 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS). 2021; 238–243.

   doi: 10.1109/ICIMCIS53775.2021.9699345

7. Al Subhi N, Bell D, Lashmar P. Location Based Modelling for Heritage Mobile Applications. UKAIS Conference 2015.

8. Dong, S. Research on the application of digital media technology in museum exhibition design: A case study of the national museum of Singapore. *SHS Web Conf*. 2024;181:04031.

   doi: 10.1051/shsconf/202418104031

9. Milgram P, Kishino F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans Inf*. 1994 ;77(12):1321–1329.

10. Baashar Y, Alkawsi G, Wan Ahmad WN, Alomari MA, Alhussian H, Tiong SK. Towards Wearable Augmented Reality in Healthcare: A Comparative Survey and Analysis of Head-Mounted Displays. *Int J Environ Res Public Health*. 2023;20(5):3940.

    doi: 10.3390/ijerph20053940

11. Singh, S. Designing AR Interfaces for Enhanced User Experience. *J Adv Scholarly Res Allied Educ*. 2024;21(5):96-102.

    doi: 10.29070/v8dpf082

12. Gimeno J, Portalés C, Coma I, Fernández M, Martínez B. Combining traditional and indirect augmented reality for indoor crowded environments. A case study on the Casa Batlló museum. *Comput Graph*. 2017;69:92–103.

doi: 10.1016/j.cag.2017.09.001

13. Pedersen I, Gale N, Mirza-Babaei P, Reid S. More than Meets the Eye: The Benefits of Augmented Reality and Holographic Displays for Digital Cultural Heritage. *J Comput Cult Herit*. 2017;10(2):1-15.

doi: 10.1145/3051480

14. Fenu C, Pittarello F. Svevo tour: The design and the experimentation of an augmented reality application for engaging visitors of a literary museum. *Int J Hum Comput Stud*. 2018;114:20–35.

doi: 10.1016/j.ijhcs.2018.01.009

15. Yoon S, Anderson E, Lin J, Elinich K. How Augmented Reality Enables Conceptual Understanding of Challenging Science Content. *J Educ Technol Soc*. 2017;20(1):156–168.

doi: 10.30191/ETS.201701_20(1).0014

16. Roussou M. Immersive Interactive Virtual Reality and Informal Education. In Proceedings of the i3 Spring Days 2000 Workshop on Interactive Learning Environments for Children, Athens, Greece. 2000.

17. Litvak E, Kuflik T. Enhancing cultural heritage outdoor experience with augmented-reality smart glasses. *Pers Ubiquitous Comput*. 2020;24(6):873–886.

doi: 10.1007/s00779-020-01366-7

18. Tom Dieck MC, Jung TH, Tom Dieck D. Enhancing art gallery visitors' learning experience using wearable augmented reality: Generic learning outcomes perspective. *Curr Issues Tour*. 2018;21(17):2014–2034.

doi: 10.1080/13683500.2016.1224818

19. De Paolis LT, Gatto C, Corchia L, De Luca V. Usability, user experience and mental workload in a mobile Augmented Reality application for digital storytelling in cultural heritage. *Virtual Real*. 2023;27(2):1117–1143.

doi: 10.1007/s10055-022-00712-9

20. Poleszak I, Dzieńkowski M. User experience analysis while visiting selected virtual museums. *J Comput Sci Inst*. 2022;25:384–392.

doi: 10.35784/jcsi.3057

21. Trunfio M, Campana S. A visitors' experience model for mixed reality in the museum. *Curr Issues Tour*. 2020;23(9):1053–1058.

doi: 10.1080/13683500.2019.1586847

22. Jung T, Tom Dieck MC, Lee H, Chung N. Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum. In Information and Communication Technologies in Tourism. Springer, Cham. 2016;621–635.

doi: 10.1007/978-3-319-28231-2_45

23. Loboda O, Nyhan J, Mahony S, Romano D. Towards evaluating the impact of recommender systems on visitor experience in physical museums. In Proceedings of the 2nd Workshop on Mobile Access to Cultural Heritage (MobileCH 2018) co-located with MobileHCI 2018, Barcelona, Spain. 2018. *CEUR Workshop Proc*. 2018;2176.

24. Hair JF, Babin BJ, Anderson RE, Black WC. Multivariate Data Analysis. 8th edition. Cengage Learning; 2019.

25. Kline RB. Principles and Practice of Structural Equation Modeling (4th edition). Guilford Press. 2016.

26. Hu L-T, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1-55.

doi: 10.1080/10705519909540118

27. MacKinnon JG, White H. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *J Econom*. 1985;29(3): 305-325.

doi: 10.1016/0304-4076(85)90158-7

28. Meade AW, Craig SB. Identifying careless responses in survey data. *Psychol Methods*. 2012;17(3):437-455.

doi: 10.1037/a0028085