Computers & Education: X Reality 8 (2026) 100152

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Immersion is not enough: Design quality as the key determinant of perceived effectiveness and usage intention in educational virtual reality

Achmad Samsudin [a,*,1], Muhammad Zahran [a,e], Eki Nugraha [a], Hadi Nasbey [b], Mustafa Sozbilir [c], Nor Farahwahidah Abdul Rahman [d], Takashi Irie [e]

[a] Universitas Pendidikan Indonesia, Bandung, 40154, Indonesia
[b] Universitas Negeri Jakarta, Jakarta, 13220, Indonesia
[c] Atatürk University, Erzurum, 25050, Turkey
[d] Universiti Teknologi Malaysia, Skudai, 81310, Malaysia
[e] Okayama University, Okayama, 700-8530, Japan

ARTICLE INFO

Keywords:
Virtual reality
Science education
Immersive learning
Design quality
PLS-SEM
Technology adoption

ABSTRACT

Immersive technologies such as Virtual Reality (VR) are increasingly used to support conceptual understanding in science education, yet perceived learning benefits depend on the quality of system design rather than immersion alone. This study investigates how four design-quality dimensions—Visual and Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), and Digital Feasibility (DF)—influence learners' Perceived Effectiveness (PE) and Intention to Use (IU) the VARIASI VR simulation, which visualizes particle behavior across phase states of matter. Using an explanatory sequential mixed-methods design, quantitative data from 60 participants were analyzed using PLS-SEM, followed by reflexive thematic analysis of open-ended responses. Results show that VT, AQ, CA, and DF each significantly predict PE, demonstrating that learners' perception of effectiveness depends on the coordinated alignment of conceptual clarity, emotional resonance, and system reliability. However, none of these dimensions directly predicted IU, and the PE → IU relationship was positive but marginal, reflecting learners' need for procedural guidance and comfort to sustain voluntary adoption. Qualitative findings reveal that while VR enhanced conceptual clarity and engagement, continued use depended on familiarity, scaffolding, and physical comfort. The study highlights that effective VR-based science learning experiences emerge when immersive environments are not only well-designed but pedagogically framed to support confidence and sustained use.

1. Introduction

Immersive technologies such as Virtual Reality (VR) and Augmented Reality (AR) have become central to the transformation of science education, providing students with opportunities to visualize and interact with phenomena that are otherwise abstract or invisible. In physics learning, VR has been shown to enhance spatial reasoning, conceptual understanding, and learner motivation by making molecular and macroscopic processes perceptible (Huang & Tseng, 2025; Lui et al., 2023; Zhang et al., 2024). Immersive visualization allows students to directly observe how particles behave during phase transitions, thereby bridging the gap between theoretical representation and experiential understanding (Makransky & Mayer, 2022; Radianti et al., 2020). Such environments can promote curiosity and engagement through embodied learning experiences that traditional instruction struggles to achieve (Kavanagh et al., 2025; Slater & Sanchez-Vives, 2016).

Despite this potential, empirical studies reveal that the impact of immersive learning depends critically on the quality of system design rather than immersion alone (Wenk et al., 2023; Wu et al., 2020). Many existing VR applications adopt static and one-size-fits-all designs that fail to accommodate learners' cognitive diversity or conceptual difficulties, which can limit sustained engagement and long-term learning gains (Crogman et al., 2025; Zekeik et al., 2025). Similarly, inconsistent empirical outcomes—ranging from significant gains in engagement to

* Corresponding author. Department of Physics Education, Indonesia.
E-mail addresses: achmadsamsudin@upi.edu (A. Samsudin), zahranr54@upi.edu (M. Zahran), ekinugraha@upi.edu (E. Nugraha), hadinasbey@unj.ac.id (H. Nasbey), sozbilir@atauni.edu.tr (M. Sozbilir), nfwahidah@utm.my (N.F.A. Rahman), irie@okayama-u.ac.jp (T. Irie).
[1] **Permanent Address**: Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154

https://doi.org/10.1016/j.cexr.2026.100152
Received 14 November 2025; Received in revised form 24 March 2026; Accepted 30 March 2026
Available online 6 April 2026
2949-6780/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

negligible improvements in conceptual understanding—suggest that design quality serves as a pivotal mediating factor between immersion and learning effectiveness (Conrad et al., 2024; Jo & Park, 2023; Makransky & Petersen, 2021). Consequently, current discourse in immersive education is shifting from technological novelty toward evidence-based design, focusing on how particular design dimensions influence users' cognitive, affective, and behavioral responses (Jafarkhani et al., 2024; Zhou & Li, 2024).

This study contributes to that shift by empirically examining how design quality dimensions in a custom-built VR simulation influence learners' Perceived Effectiveness (PE) and Intention to Use (IU). The simulation, developed to visualize particle behavior and phase structures of matter, served not only as a learning medium but also as a research instrument for testing how distinct design attributes shape users' responses.

Four constructs were systematically evaluated—Visual and Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), and Digital Feasibility (DF)—each grounded in theoretical and empirical precedents from human–computer interaction, user experience, and educational design. The following section details the theoretical framework grounding these constructs and the hypotheses regarding their impact on learners' perceptions.

2. Theoretical framework

The effectiveness of immersive learning environments depends less on technological novelty than on how design characteristics support learners' cognitive and affective processes. Research in HCI, UX, and learning sciences consistently shows that design quality—rather than hardware sophistication—determines whether immersive systems enable meaningful learning (Makransky & Petersen, 2021; Wu et al., 2020). To explain these mechanisms, this study integrates Cognitive Load Theory (CLT) (Sweller et al., 2019), Cognitive Theory of Multimedia Learning (CTML) (Mayer, 2024), the Cognitive Affective Model of Immersive Learning (CAMIL) (Makransky & Petersen, 2021) and extended Technology Acceptance Model for virtual reality (Sagnier et al., 2020).

CLT provides the cognitive foundation for curriculum-related design, emphasizing the need to reduce extraneous cognitive load and support schema construction. In immersive contexts, this principle is most directly reflected in Curriculum Alignment (CA), where pedagogically coherent and goal-oriented content minimizes overload and supports learning efficiency. CLT therefore underpins CA and partially informs learners' Perceived Effectiveness (PE) by shaping perceptions of instructional clarity and usefulness.

In contrast, CTML and CAMIL explain how immersive environments support cognitive and affective processing during learning. CTML highlights dual-channel processing and the integration of visual and verbal information, providing a basis for Visual and Technological Design (VT) and Digital Feasibility (DF), as poorly designed visuals or technical disruptions can increase extraneous load. CAMIL further emphasizes presence and agency, linking representational fidelity, aesthetic coherence, and system reliability to engagement and learning outcomes. Accordingly, VT, Aesthetic Quality (AQ), and DF reflect core design affordances that shape learners' immersive experiences and judgments of effectiveness.

To explain learners' adoption of immersive technologies, the study draws on an extended Technology Acceptance Model (TAM), in which Perceived Effectiveness (PE) corresponds to Perceived Usefulness and serves as a key predictor of Intention to Use (IU). Thus, design-quality dimensions (VT, AQ, CA, DF) influence IU both indirectly through Perceived Effectiveness (PE) and, in some cases, directly, integrating cognitive, pedagogical, and affective determinants of technology acceptance. In extended TAM formulations for immersive systems, experiential qualities and practical constraints may also exert direct effects on usage intention, independent of perceived effectiveness, particularly when affective engagement or operational feasibility strongly shapes users' willingness to adopt the technology (Sagnier et al., 2020).

Ultimately, these four theoretical perspectives do not operate in isolation; rather, they form an integrated continuum that captures the complexity of immersive learning. CLT and CTML address the foundational cognitive mechanics required to process information without working-memory overload. CAMIL builds upon this by introducing the affective and embodied dimensions unique to VR, such as presence and aesthetic resonance. Finally, the extended TAM bridges these internal cognitive and affective states to external behavioral adoption. Together, they illustrate that technology adoption in VR relies on a complex interplay: the system must first manage cognitive capacity (CLT, CTML) through aligned and feasible design, while simultaneously fostering the affective engagement (CAMIL) necessary to translate perceived effectiveness into a concrete intention to use (TAM). This structure is summarized in Fig. 1 and provides the theoretical basis for the hypotheses tested in Section 3.

2.1. Visual and technological design (VT)

From an HCI perspective, *visual and technological design* encompasses the interactive and perceptual fidelity of an environment—the degree to which users experience smooth, intuitive, and reliable interaction with system elements. Prior studies identify responsiveness, visual realism, and intuitive navigation as core determinants of immersion and usability (Ahn et al., 2014; Conrad et al., 2024; Slater & Sanchez-Vives, 2016). In educational contexts, these characteristics translate into cognitive affordances that help learners attend to conceptual cues rather than to interface mechanics. Poorly optimized or unintuitive interfaces have been shown to elevate extraneous cognitive load, fragment attention, and reduce persistence (Kavanagh et al., 2025; Makransky et al., 2019). In contrast, high-quality VT supports "cognitive transparency," allowing learners to manipulate virtual phenomena without distraction (Wenk et al., 2023; Whitton, 2014). Accordingly, this framework posits that stronger VT will directly enhance learners' perceived effectiveness of the system. Empirical studies illustrating these relationships are discussed in Section 3.2.

2.2. Aesthetic Quality (AQ)

Aesthetic quality represents the affective and perceptual coherence of a design—the extent to which color, texture, symmetry, and overall visual harmony evoke emotional engagement. In UX theory, aesthetic appeal is not a secondary attribute but a primary determinant of user perception (Hassenzahl, 2004; Sauer & Sonderegger, 2022). Aesthetics influence judgments of usability through the so-called "aesthetic–usability effect," whereby visually pleasing interfaces are perceived as easier to use and more effective (Tractinsky et al., 2000).

Within educational VR, visual beauty and harmony have been linked to emotional flow, motivation, and intrinsic interest (Jafarkhani et al., 2024; Zhou & Li, 2024). Jo and Park (2023) found that coherent graphic styles in educational simulations increased students' sense of enjoyment and attention to scientific detail. Therefore, the present model anticipates that aesthetic coherence enhances Perceived Effectiveness by generating positive affective states that facilitate deeper engagement. Further empirical support for the affective role of aesthetic quality appears in Section 3.3.

2.3. Curriculum Alignment (CA)

Curriculum alignment ensures that the immersive experience connects to instructional objectives and assessment standards. Theories of instructional design stress that technologies become pedagogically meaningful only when integrated coherently with learning goals (Reigeluth & Carr-Chellman, 2009). In VR-based science education,

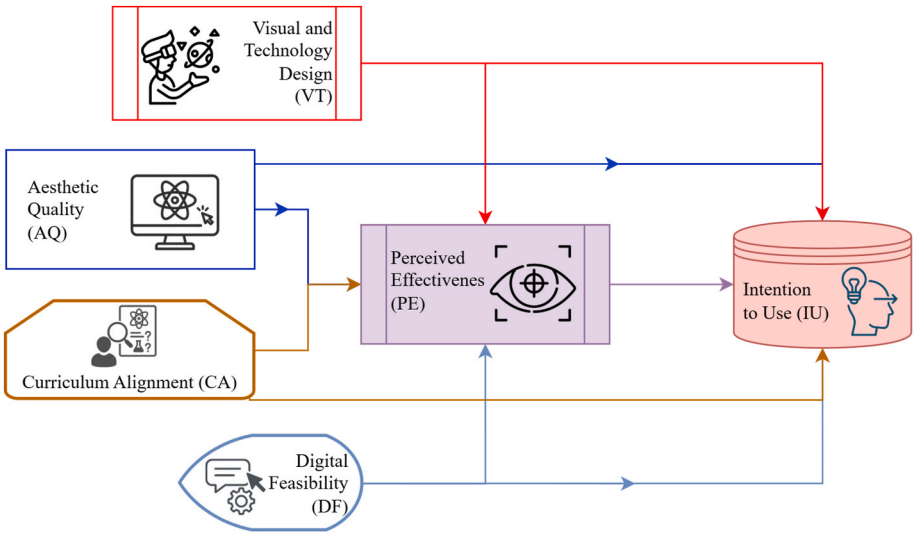

Fig. 1. Hyphotheses summary.

tasks that are tightly mapped to curricular content produce better learning transfer than those emphasizing novelty or entertainment (Adelana et al., 2023; Conrad et al., 2024; Crogman et al., 2025; Zekeik et al., 2025). Misalignment can lead to fragmented conceptual understanding or misconceptions (Makransky & Mayer, 2022). Thus, alignment functions as a *semantic bridge* between immersive engagement and academic purpose. It is hypothesized that higher CA will contribute to learners' PE by reinforcing the perception that the system directly supports curricular mastery. Representative evidence on curricular alignment is reviewed in Section 3.4.

2.4. Digital Feasibility (DF)

Digital feasibility refers to the practical viability of deploying the technology under real classroom conditions—stability, accessibility, compatibility, and ease of integration. Empirical reviews consistently identify feasibility as a decisive determinant of adoption (Huang & Tseng, 2025; Radianti et al., 2020; Zekeik et al., 2025). Lag, crashes, or inconsistent device performance often override positive experiences with immersion, reducing trust and willingness to use (Wenk et al., 2023). A feasible system minimizes friction and ensures that teachers and students can focus on content rather than troubleshooting. Hence, higher DF is expected to improve PE by supporting consistent, reliable interactions that sustain learners' focus on conceptual exploration. Empirical findings on the operational importance of digital feasibility are summarized in Section 3.5.

2.5. Perceived effectiveness (PE) and intention to use (IU)

Perceived Effectiveness (PE) is defined as the learner's cognitive appraisal of how well an immersive simulation supports understanding, engagement, and clarity of concept. It reflects both functional usability and subjective learning value. Prior research indicates that positive perceptions of usefulness and clarity are strong predictors of continued technology use across educational media (Makransky & Petersen, 2021; Wu et al., 2020). In turn, *Intention to Use* (IU) represents the motivational and behavioral inclination to re-engage with or recommend the system. Conceptually, Perceived Effectiveness (PE) represents a *cognitive evaluative judgment* of learning support and system usefulness, whereas Intention to Use (IU) reflects a *motivational–behavioral disposition* toward future adoption. Although closely related, the two constructs are theoretically distinct: learners may recognize an immersive system as effective without necessarily intending to reuse it, particularly when contextual or practical constraints intervene. Accordingly, PE is positioned as a mediating construct that translates learners' cognitive and affective experiences into behavioral intention, while allowing for partial direct effects of design-quality dimensions on IU, consistent with extended TAM formulations for immersive environments. When learners perceive a system as effective, they are more likely to adopt it voluntarily in future learning (Liu et al., 2020; Zhou & Li, 2024). Consequently, PE is positioned as a mediating variable linking design quality to behavioral outcomes. Fig. 1 summarizes the hypothesized relationships:

This framework provides an integrative lens for explaining how multidimensional design quality shapes learners' cognitive evaluations and motivational intentions in immersive science simulations.

Summary of Hypotheses.

H1. Visual and Technological Design (VT) positively influences Perceived Effectiveness (PE).

H2. Aesthetic Quality (AQ) positively influences PE.

H3. Curriculum Alignment (CA) positively influences PE.

H4. Digital Feasibility (DF) positively influences PE.

H5. Perceived Effectiveness (PE) positively influences Intention to Use (IU).

H6. Aesthetic Quality (AQ) positively influences IU.

H7. Curriculum Alignment (CA) positively influences IU.

H8. Digital Feasibility (DF) positively influences IU.

H9. Visual and Technological Design (VT) positively influences Intention to Use (IU).

Having established the conceptual model and hypotheses, the subsequent literature review situates each construct within the broader empirical landscape of immersive learning and instructional design. Section 3 synthesizes prior findings on VR-based science learning, demonstrating how visual-technical clarity, aesthetic coherence, curricular integration, and digital feasibility have been shown to influence learners' cognitive processing and motivational engagement. This synthesis provides the empirical basis for the hypothesized relationships among the design-quality dimensions and Perceived Effectiveness (PE).

Together, the proposed hypotheses form an integrated CLT–CTML–CAMIL–TAM design-quality model that explains how immersive environments support learning not merely through sensory immersion, but through the coordinated reduction of extraneous cognitive load, reinforcement of conceptual coherence, and

enhancement of affective engagement. In this model, each of the four design-quality dimensions is theorized to contribute both to learners' perceived effectiveness of the simulation and to their behavioral intention to engage with it. Finally, Section 4 describes the PLS-SEM procedure used to empirically test this model and evaluate the relative contributions of each design dimension to PE and learners' subsequent Intention to Use (IU) the system.

3. Literature review

Immersive technologies have been widely studied across science education research, yet findings on their effectiveness remain heterogeneous, especially due to variations in instructional design, representational detail, and implementation context. This section reviews empirical evidence related to each of the design-quality dimensions examined in the present study, situating them within the broader landscape of immersive learning research.

3.1. Immersive learning in science education

Immersive technologies have been widely recognized for their capacity to transform science learning by providing embodied, spatially rich environments that make invisible processes visible (Makransky & Mayer, 2022; Zhang et al., 2024), following the conceptual relationships outlined in Section 2, this review synthesizes empirical research for each design dimension. Meta-analyses indicate that VR and AR can significantly improve engagement, motivation, and comprehension—particularly in physics and chemistry, where abstract phenomena are common (Kavanagh et al., 2025; Radianti et al., 2020). However, findings are not uniformly positive.

Some studies report *presence without learning*—where immersive realism increases sensory load and distracts from conceptual focus (Makransky et al., 2019)—while others highlight the risk of novelty effects fading over time (Wu et al., 2020). This divergence suggests that the educational impact of immersive learning depends less on the technology itself and more on the quality of its design, including visual realism, interaction smoothness, aesthetic coherence, curricular relevance, and technical feasibility (Adelana et al., 2023; Conrad et al., 2024; Huang & Tseng, 2025).

Given this variability in learning outcomes, researchers have begun to examine which specific aspects of immersive system design most influence learner experience and cognition. Among these, visual and technological design has emerged as a foundational determinant.

3.2. Visual and technological design (VT)

The visual and technological design of a VR environment—its graphical fidelity, responsiveness, and interactivity—constitutes the foundation of user experience (as defined in Section 2.1). Studies in educational VR consistently associate technical fluency with usability and engagement. Wenk et al. (2023) found that stable frame rates and intuitive navigation were decisive for sustaining flow and minimizing motion sickness. Similarly, Slater and Sanchez-Vives (2016) showed that lag or low responsiveness directly undermines presence. In educational contexts, intuitive visual feedback enables learners to allocate attention to conceptual tasks rather than to interface management (Kavanagh et al., 2025; Whitton, 2014). However, too much graphical detail can paradoxically *increase* cognitive load. Makransky et al. (2019) demonstrated that full-immersion head-mounted displays enhanced presence but reduced transfer learning compared to desktop VR.

This mixed evidence points to an optimal threshold of technological richness, beyond which cognitive overload offsets the benefits of realism. Hence, understanding VT's contribution to Perceived Effectiveness (PE) requires examining not only visual fidelity but also cognitive manageability. While technical fluency shapes how learners interact with a virtual environment, the emotional and perceptual qualities of that environment are equally crucial. These aesthetic elements define how users feel during the experience and are addressed next.

3.3. Aesthetic Quality (AQ)

While VT emphasizes functionality, aesthetic quality (AQ) concerns perceptual harmony, coherence, and emotional appeal (see conceptual description in Section 2.2). The "aesthetic–usability effect" (Hassenzahl, 2004; Sauer & Sonderegger, 2022) suggests that visually pleasing interfaces are perceived as more effective, even when usability is constant. Recent evidence confirms that aesthetic appeal strongly mediates learner engagement and motivation in immersive contexts. Jafarkhani et al. (2024) reported that visual harmony predicted emotional flow and satisfaction in VR learning environments, while Zhou and Li (2024) identified aesthetic coherence as a cross-modal cue that improves perceived clarity and focus. Jo and Park (2023) similarly found that graphic style consistency fosters trust and attentional persistence. Collectively, these findings indicate that aesthetics shape cognitive engagement *indirectly*—by enhancing emotional readiness and perceived usability. However, aesthetic appeal can backfire if incongruent with content; overly stylized or game-like designs may reduce perceived seriousness or cognitive credibility (Wu et al., 2020).

Thus, AQ and VT interact: visual realism must be balanced with aesthetic coherence to achieve optimal PE. Beyond visual and affective design, educational coherence determines whether immersive experiences translate into measurable learning gains. The next construct, curriculum alignment, captures this pedagogical dimension.

3.4. Curriculum Alignment (CA)

Educational technologies achieve meaningful learning only when they integrate coherently with instructional goals, as outlined in Section 2.3 (Reigeluth & Carr-Chellman, 2009). In immersive education, curriculum alignment (CA) ensures that simulations reinforce curricular outcomes rather than functioning as stand-alone experiences. Crogman et al. (2025) and Zekeik et al. (2025) both noted that alignment between VR tasks and learning objectives significantly influences students' conceptual gains and teachers' willingness to adopt such tools. Conversely, poorly aligned simulations can induce misconceptions or be perceived as "edutainment." A comparative review by Zhang et al. (2024) concluded that VR modules designed with explicit curricular mapping yield stronger transfer and assessment outcomes. Therefore, CA functions as a pedagogical anchor, reinforcing the learner's belief that what is experienced virtually has authentic academic value—a factor that directly feeds into perceived effectiveness.

However, even the most pedagogically aligned simulation cannot succeed without reliable technical performance. Thus, digital feasibility represents the operational layer that enables or constrains all other design dimensions.

3.5. Digital Feasibility (DF)

Even the most elegantly designed VR application fails if it cannot operate smoothly within real classroom constraints (see Section 2.4 for theoretical rationale). Digital feasibility (DF) encompasses stability, accessibility, hardware compatibility, and institutional readiness. Zekeik et al. (2025), Huang and Tseng (2025) identify technical friction—connectivity issues, device shortages, or setup complexity—as the foremost barrier to implementation in schools. Wenk et al. (2023) emphasize that user frustration from delays or system crashes rapidly undermines engagement and trust. Moreover, feasibility also includes scalability: whether a system can run across multiple devices or offline. As Makransky and Petersen (2021) argue, reliability is a prerequisite for sustained adoption. Hence, DF contributes both *directly* to perceived effectiveness and *indirectly* to Intention to Use (IU), as learners and teachers value systems that "just work."

Taken together, these dimensions illustrate that immersive learning effectiveness is inherently multidimensional. The next section integrates these strands to explain how design quality translates into learners' perceived effectiveness and intention to use.

3.6. Linking design quality to perception and behavior

Empirical evidence across VR, HCI, and educational psychology converges on a consistent pattern: design quality affects not only performance but also the *perceived usefulness* and *intent to re-engage* with a technology. When users perceive a system as visually intuitive, aesthetically coherent, pedagogically relevant, and technically reliable, they are more likely to evaluate it as effective (PE) and to continue using it (IU). However, most previous studies treat these design aspects separately or qualitatively. Few have tested their combined predictive power using structural modeling approaches in authentic classrooms—particularly in physics learning contexts where microscopic visualization is crucial (Wu et al., 2020). This study addresses that empirical gap by quantitatively modeling the influence of VT, AQ, CA, and DF on PE and IU within an authentic instructional context. The prior findings and research gap motivating this investigation are summarized in Table 1.

3.7. Summary

In sum, prior research underscores the *multifaceted nature of design quality* in immersive education but reveals inconsistent or incomplete evidence regarding how these factors jointly influence learner perception and behavioral intention. By employing Partial Least Squares–Structural Equation Modeling (PLS-SEM) on classroom data from a VR simulation of phase transitions, the present study contributes quantitative validation to this body of work, providing empirical clarity on which design attributes most strongly shape learners' perceptions of effectiveness and their intention to use immersive technologies in science learning. To empirically verify these theoretical linkages, the following section details the methodological procedures, participants, instruments, and analytical techniques employed to test the proposed model using Partial Least Squares–Structural Equation Modeling (PLS-SEM).

4. Methodology

This study employed an explanatory sequential mixed-methods design, selected to examine not only the structural relationships among design-quality dimensions of the VARIASI VR simulation, but also the experiential reasoning that shaped participants' evaluations. In this design, the quantitative phase precedes the qualitative phase, allowing numerical patterns to be interpreted in light of participants' subjective meaning-making (Creswell & Clark, 2017; Venkatesh et al., 2013). The quantitative phase used Partial Least Squares–Structural Equation Modeling (PLS-SEM) to model the predictive relationships among Visual and Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), Digital Feasibility (DF), Perceived Effectiveness (PE), and Intention to Use (IU). PLS-SEM was chosen for its suitability in exploratory research, complex latent constructs, and modest sample sizes (Hair et al., 2019, 2021).

Following the quantitative analysis, participants' open-ended reflections were examined using reflexive thematic analysis (Braun & Clarke, 2021). This qualitative phase did not aim to generate new theory, but to explain why certain design dimensions strongly shaped perceived effectiveness and why perceived effectiveness did not uniformly translate into intention for continued use. Integration occurred during interpretation, where the thematic findings were used to clarify, contextualize, and extend the statistical pathways identified in the structural model (Fetters et al., 2013).

4.1. Research context

Data collection occurred during a university open-house event in which secondary-school students and their teachers visited booths representing different study programs. At the *Physics Education* booth, the research team showcased the VARIASI simulation as part of a science-outreach exhibition. Visitors who voluntarily experienced the simulation through standalone head-mounted displays were invited to complete a brief digital questionnaire concerning their perceptions of the system's design, usefulness, and feasibility.

The exhibition context offered an ecologically valid environment to observe first-time user responses to immersive educational technology. Unlike controlled classroom studies, this setting involved spontaneous participation, allowing researchers to capture natural reactions of both students and teachers encountering educational VR for the first time.

4.2. The VARIASI VR simulation

The Virtual–Augmented Reality Simulation for Phase Structures (VARIASI) is a lightweight immersive environment developed to help learners observe how particle arrangements differ across states of matter. Rather than modeling detailed molecular geometry, the simulation uses colored spheres grouped to represent familiar compounds such as water (H_2O) and nitrogen dioxide (NO_2), as well as a tightly arranged array corresponding to gold (Au) in its solid state. These simplified particle forms were chosen intentionally to focus learners' attention on differences in spacing, motion, and structural organization, which are core to understanding phase changes. Fig. 2 shows the learner's perspective while observing a mixed particle environment.

Interaction in VARIASI is kept intentionally minimal to support cognitive clarity. Learners use basic VR controllers to point, observe, and focus on particle clusters without needing to perform complex manipulations. This approach reduces interface demands and supports learners who are experiencing VR for the first time. Within the gaseous environment, particles are displayed with greater spacing and visible movement, helping learners notice how increased energy leads to freer motion. Fig. 3 illustrates this distributed configuration.

Learners move between environments using teleportation markers

Table 1
Summary of prior findings and research gap.

Design Dimension	Key Empirical Findings	Observed Contradictions/Gaps	Implication for Present Study
Visual & Technological Design (VT)	Smooth interaction and low latency enhance presence and learning (Slater & Sanchez-Vives, 2016; Wenk et al., 2023); over-realism may induce overload (Makransky et al., 2019).	Optimal fidelity threshold unclear; few studies quantify its effect on perceived effectiveness.	Examine VT's direct effect on PE using gathered data.
Aesthetic Quality (AQ)	Aesthetic harmony predicts motivation, satisfaction, and flow (Jafarkhani et al., 2024; Zhou & Li, 2024)	Over-stylization may reduce perceived seriousness; interplay with VT rarely modeled jointly.	Test AQ's independent contribution to PE and interaction with VT.
Curriculum Alignment (CA)	Alignment with learning objectives improves conceptual transfer and teacher acceptance (Crogman et al., 2025; Zekeik et al., 2025)	Many VR tools remain loosely mapped to curricula; effect on PE understudied.	Quantify CA's influence on perceived usefulness and IU.
Digital Feasibility (DF)	Reliability and accessibility predict sustained adoption (Huang & Tseng, 2025; Zekeik et al., 2025)	Technical feasibility often discussed qualitatively, not modeled empirically.	Model DF as a structural predictor of PE and IU.

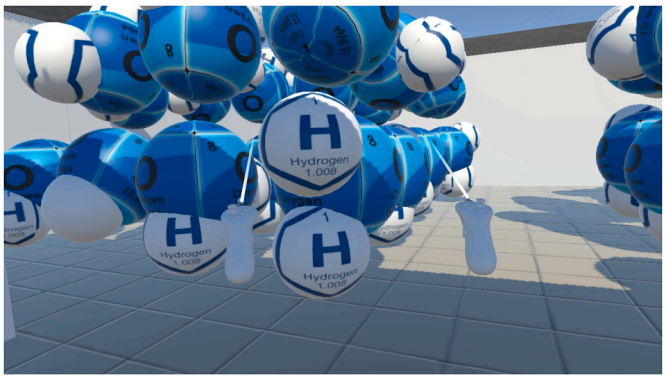

Fig. 2. Learner view of particle groups representing different substances in the immersive environment.

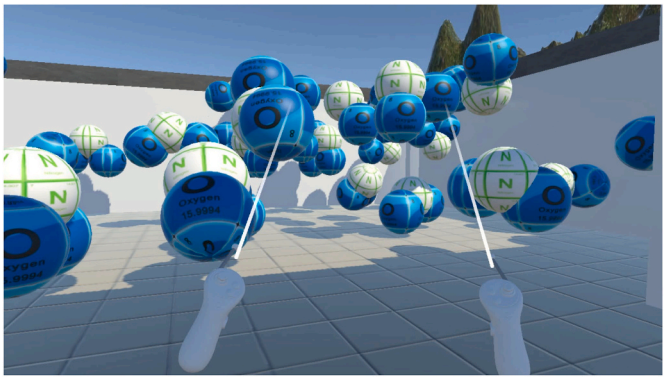

Fig. 3. Particle distribution representing a gas-like state with wider spacing and visible movement.

placed within the scene. Teleportation was chosen rather than continuous locomotion to prevent disorientation and motion sickness, ensuring that visual attention remains on the learning task rather than navigation effort. Fig. 4 shows the teleportation point used to shift between state environments.

The solid-state environment displays closely packed particles arranged in a stable formation, emphasizing the limited motion characteristic of solids. This clear visual contrast allows learners to compare structural rigidity across states without requiring prior familiarity with atomic models or molecular theory. Fig. 5 presents the dense gold particle arrangement used to represent a solid state.

The experience begins at a virtual laboratory entrance designed to frame the simulation as a learning environment rather than a game. This



Fig. 4. Teleportation point for navigating between different phase-state environments within the simulation.

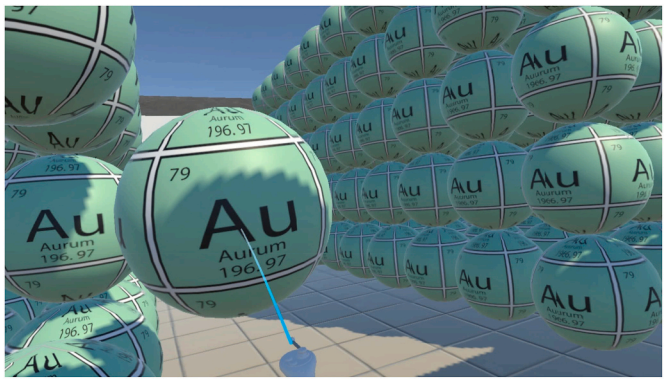

Fig. 5. Dense particle arrangement representing a solid-state structure.

helps establish purpose and educational relevance before the learner enters the particle environments. Fig. 6 shows the entrance scene that serves as the starting point for the simulation.

Together, these visual and interaction choices reflect the design qualities examined in this study. The clarity of movement and interaction demonstrates the role of visual and technological design in maintaining focus on conceptual content. The consistent use of color and spacing contributes to the aesthetic coherence of the environment, supporting comfort and attention. The correspondence between displayed particle behavior and commonly taught explanations of phase change reinforces curriculum alignment. The use of a standalone VR headset ensures that the simulation can be deployed flexibly in both classroom and outreach contexts, reflecting considerations of digital feasibility. In this way, VARIASI functions not only as a tool for learning about particle behavior but also as a research platform for examining how design quality influences learners' perceived effectiveness and intentions to use immersive media in science learning.

4.3. Participants

A total of 60 respondents participated voluntarily, comprising 34 students (M_{age} = 16.4, SD = 5.6; 45% female) and 26 teachers (M_{age} = 34.2, SD = 5.6; 60% female) from various secondary schools. The study employed a nonprobability convenience sampling approach, as participation was open to visitors who interacted with the VR simulation during the exhibition. All participation was voluntary, informed consent was obtained prior to survey completion, and no incentives were provided.

The average duration of the VR experience was approximately 25 min. All participants reported no prior experience with educational VR headsets, ensuring that responses reflected initial adoption and first-



Fig. 6. Virtual laboratory entrance framing the simulation as a structured learning experience.

time user perceptions rather than habituated use. Participants indicated basic familiarity with phase-transition concepts, but none had previously engaged with immersive VR for formal learning purposes.

This mixed sample provided perspectives from both prospective learners and practicing educators. Given the inclusion of both cohorts, a preliminary group comparison of the primary outcome variables was conducted to assess the appropriateness of pooling responses for subsequent analyses. The results of this comparison are reported in Section 5.1.

4.4. Instruments and data-gathering tool

4.4.1. Quantitative measures

Learners' perceptions were measured using a twenty-four-item questionnaire with six constructs: VT, AQ, CA, DF, PE, and IU. Items employed a five-point Likert scale ranging from strong disagreement to strong agreement. The questionnaire was adapted from validated instruments in human–computer interaction, user experience research, and immersive learning studies (Hassenzahl, 2004; Jafarkhani et al., 2024; Makransky & Mayer, 2022; Radianti et al., 2020; Whitton, 2014). The instrument was reviewed by three subject-matter experts (two in physics education, one in educational technology) to ensure content validity, linguistic clarity, and contextual appropriateness.

4.4.2. Qualitative responses

To complement the quantitative findings, the questionnaire included two open-ended prompts designed to elicit participants' reflective evaluations of the simulation experience. These prompts invited respondents to describe (a) which aspects of the VARIASI simulation supported their understanding or engagement, and (b) which aspects posed difficulty or interfered with learning. Specifically, participants were asked:

"What aspects of the VR simulation helped your understanding or engagement?"

"What difficulties or challenges did you experience during the VR simulation?"

The responses to these questions provided narrative accounts that capture learners' cognitive, affective, and practical experiences while interacting with the immersive environment. These qualitative data formed the basis of the explanatory analysis in the second phase of the study, allowing the interpretation of statistical patterns in relation to participants' lived experiences of VR-based learning.

4.4.3. Questionnaire structure

The structure of the questionnaire is summarized in Table 2, showing the number of indicators, an illustrative item, and the theoretical or empirical basis for each construct.

Table 2
Structure of the measurement questionnaire and construct sources.

Construct	No. of Items	Example Indicator	Source Basis
Visual & Technological Design (VT)	5	"The simulation responded smoothly to my actions."	Adapted from Makransky and Mayer (2022); Whitton (2014)
Aesthetic Quality (AQ)	4	"The visual appearance of the simulation was pleasing and coherent."	Adapted from Hassenzahl (2004); Jafarkhani et al. (2024)
Curriculum Alignment (CA)	4	"The simulation content matched the learning objectives discussed in class."	Based on Reigeluth and Carr-Chellman (2009); Zekeik et al. (2025)
Digital Feasibility (DF)	4	"The simulation operated reliably without technical problems."	Adapted from Radianti et al. (2020); Zekeik et al. (2025)
Perceived Effectiveness (PE)	4	"This simulation helped me understand phase changes more clearly."	Derived from Makransky and Mayer (2022)
Intention to Use (IU)	3	"I would like to use this VR system for other science topics."	Adapted from Wu et al. (2020)

4.5. Data collection procedure

Participants were briefly introduced to the purpose of the simulation and the operation of the VR headset before engaging freely with the environment. Interaction time averaged approximately 25 min, during which participants explored the particle-behavior visualization across solid, liquid, and gas phases. Immediately after interaction, participants completed the digital questionnaire. Responses were automatically recorded and exported for analysis.

4.6. Data analysis

4.6.1. Quantitative analysis

Quantitative data were analyzed using SmartPLS 3 to estimate both the measurement and structural models within a partial least squares structural equation modeling (PLS-SEM) framework. PLS-SEM was selected over covariance-based SEM (CB-SEM) due to its higher statistical power in small samples, its robustness to non-normal data, and its suitability for exploratory and prediction-oriented models (Hair et al., 2019). Given the modest sample size (N = 60) and the complexity of the proposed structural model, PLS-SEM was considered an appropriate and reliable analytic approach.

The measurement model was evaluated by examining internal consistency reliability and convergent validity using composite reliability coefficients and average variance extracted (AVE). Discriminant validity was assessed using both the Fornell–Larcker criterion and the heterotrait–monotrait ratio (HTMT) (Fornell & Larcker, 1981; Hair et al., 2019). The structural model was then assessed using bootstrapped path coefficients, along with associated significance levels and effect size indices, to evaluate the hypothesized relationships among constructs.

4.6.2. Qualitative analysis

Open-ended responses were analyzed using reflexive thematic analysis (Braun & Clarke, 2021). The analysis involved iterative reading, identification of meaningful ideas, condensation of recurring patterns, and synthesis into thematic interpretations. The qualitative phase did not attempt to generate grounded theory but served to clarify the experiential basis of the quantitative findings, particularly regarding the nuanced relationship between perceived effectiveness and continued usage intention.

5. Results

To evaluate the proposed relationships among the design-quality dimensions, perceived effectiveness, and behavioral intention, the model was analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM). Fig. 7 displays the resulting structural model, including standardized path coefficients and indicator loadings. The quantitative results are presented first, followed by qualitative findings that provide explanatory depth regarding how participants interpreted and engaged with the VARIASI simulation.

5.1. Preliminary group comparison

Given the mixed composition of the sample (n = 26 teachers, n = 34 students), a preliminary analysis was conducted to determine if there were significant differences in the primary dependent variables—Perceived Effectiveness (PE) and Intention to Use (IU)—between the two

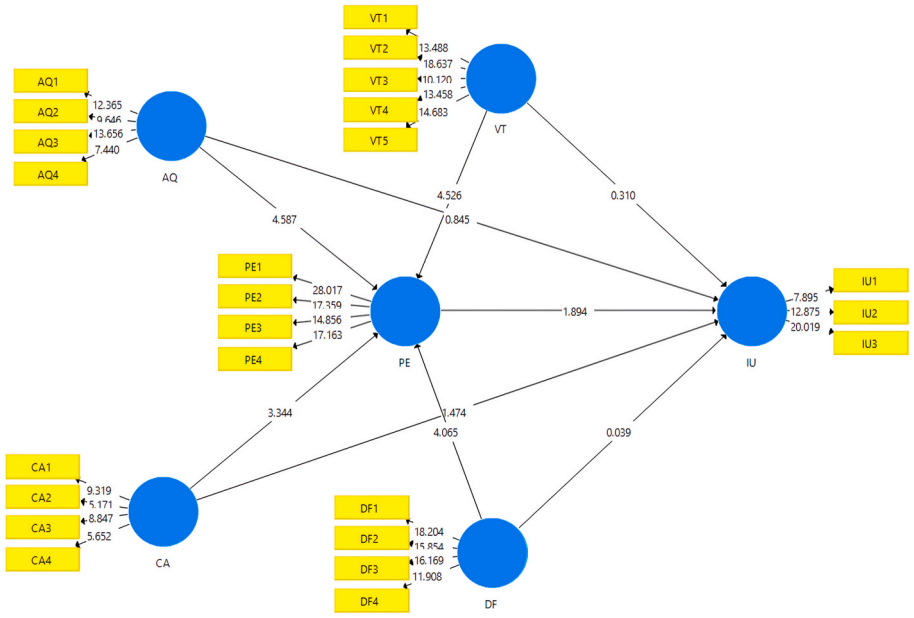

Fig. 7. Structural equation model output from PLS-SEM analysis.

groups. First, a Shapiro-Wilk test was performed to assess the normality of the data distributions. The results indicated that the data for Teachers' Perceived Effectiveness ($W = 0.909$, $p = 0.026$) and Students' Intention to Use ($W = 0.925$, $p = 0.023$) deviated significantly from a normal distribution.

Consequently, a non-parametric Mann-Whitney U test was selected. The results indicated no statistically significant difference between teachers and students for Perceived Effectiveness ($U = 522.5$, $p = 0.229$) or Intention to Use ($U = 411.0$, $p = 0.644$). These findings suggest that both groups perceived the simulation's value and utility in a similar manner, regardless of their role. Therefore, the data were pooled for the subsequent PLS-SEM analysis to examine the structural relationships across the entire cohort.

5.2. Measurement model evaluation

Following confirmation of sample comparability, the measurement model was evaluated to assess the reliability and validity of the latent constructs. Internal consistency, convergent validity, and discriminant validity were examined. As summarized in Table 3, all constructs demonstrated excellent reliability. Cronbach's alpha (α) ranged from 0.776 to 0.920, while composite reliability (CR) values were between 0.870 and 0.939, exceeding the recommended 0.70 threshold (Hair et al., 2019). The Average Variance Extracted (AVE) values ranged from 0.670 to 0.755, confirming convergent validity since all were above the 0.50 criterion (Fornell & Larcker, 1981).

Discriminant validity was assessed using the Fornell–Larcker criterion. As displayed in Table 4, the square root of the AVE for each construct (bolded diagonal values) exceeded its correlations with other constructs, satisfying discriminant validity requirements. Additionally, all HTMT ratios were below 0.90, confirming discriminant separation among latent variables (Henseler et al., 2015). All variance inflation factor (VIF) values ranged from 1.47 to 3.07, well below the conservative threshold of 5, indicating no multicollinearity issues.

Table 3
Construct reliability and validity.

Construct	Cronbach's α	Composite Reliability	AVE
Aesthetic Quality (AQ)	0.879	0.914	0.727
Curriculum Alignment (CA)	0.841	0.890	0.670
Digital Feasibility (DF)	0.888	0.922	0.748
Intention to Use (IU)	0.776	0.870	0.691
Perceived Effectiveness (PE)	0.865	0.907	0.710
Visual and Technological Design (VT)	0.920	0.939	0.755

Table 4
Fornell–Larcker discriminant validity.

Construct	AQ	CA	DF	IU	PE	VT
AQ	**0.852**					
CA	−0.111	**0.819**				
DF	0.005	0.032	**0.865**			
IU	0.237	0.289	0.163	**0.831**		
PE	0.392	0.276	0.409	0.485	**0.843**	
VT	0.167	−0.064	0.043	0.209	0.422	**0.869**

5.3. Structural model evaluation

After confirming the adequacy of the measurement model, the structural model was examined to test the hypothesized relationships among the constructs. Bootstrapping with 5000 resamples was conducted to obtain standardized path coefficients, t-values, and significance levels. Table 5 summarizes the results of the structural model estimation.

The results indicate that all four design dimensions—Visual and

Table 5
Path coefficients and significance.

Path	Original Sample (O)	T-Statistics	p-Values	Interpretation
VT → PE	0.366	4.385	0.000	Significant ***
AQ → PE	0.365	4.730	0.000	Significant ***
CA → PE	0.327	3.289	0.001	Significant **
DF → PE	0.380	4.032	0.000	Significant ***
PE → IU	0.362	1.878	0.060	Marginal *
AQ → IU	0.109	0.835	0.404	Not significant
CA → IU	0.204	1.476	0.140	Not significant
DF → IU	0.006	0.039	0.969	Not significant
VT → IU	0.051	0.304	0.761	Not significant

Note. (*p < 0.10; **p < 0.01; ***p < 0.001).

Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), and Digital Feasibility (DF)—significantly influence Perceived Effectiveness (PE), confirming their centrality to learners' cognitive appraisal of immersive environments. However, none of these design variables had a direct impact on Intention to Use (IU), suggesting that PE functions as a mediating mechanism that links design quality to behavioral engagement.

Following the hypothesis testing results, additional examination was conducted to assess the stability of the PE → IU path estimate. Because this relationship approached the threshold of statistical significance, a bootstrapped sampling distribution of the total effects was inspected to evaluate the robustness of the coefficient across resamples. As displayed in Fig. 8, the resulting distribution is broad and moderately dispersed rather than sharply peaked, indicating that the magnitude of the effect varied substantially across the 5000 bootstrap iterations. This pattern suggests that although the PE → IU path trends positive on average, its influence is not consistently strong across the sample population. The variability observed in the distribution supports the interpretation that Perceived Effectiveness functions as a necessary but insufficient condition for sustained usage intention, and that additional factors external to the simulation itself (e.g., instructional scaffolding, prior familiarity, perceived ease of independent operation) likely moderate the adoption process.

The distribution illustrates the variability of the estimated effect across 5000 bootstrap resamples. The spread and slope indicate a positive yet unstable relationship, consistent with the marginal significance observed in the structural model.

5.3.1. Model explanatory and predictive power

To evaluate the explanatory and predictive power of the model, coefficients of determination (R^2) and cross-validated redundancy (Q^2) values were examined. Table 6 presents the model's predictive relevance for Perceived Effectiveness (PE) and Intention to Use (IU).

The model explained 54.3% of the variance in PE ($R^2 = 0.543$) and 27.2% of the variance in IU ($R^2 = 0.272$), indicating moderate explanatory power (Chin, 1998; Hair et al., 2021). Cross-validated redundancy (Q^2) values were 0.347 for PE and 0.149 for IU, both exceeding zero, which signifies predictive relevance (Geisser, 1975).

5.3.2. Model fit and effect size

The overall model fit was evaluated using the Standardized Root Mean Square Residual (SRMR) and the Normed Fit Index (NFI). The model achieved an SRMR value of 0.087, which is below the recommended 0.10 threshold, indicating an acceptable level of discrepancy between the observed and model-implied correlations. The NFI value of 0.687 falls within the range commonly reported in PLS-SEM studies, where NFI values between 0.60 and 0.90 are considered satisfactory for variance-based modeling (Hair et al., 2019; Henseler et al., 2015, 2016). Together, these indices support that the model demonstrates an acceptable global fit suitable for explanatory and predictive analysis.

In terms of effect size, the f^2 values indicated that Digital Feasibility ($f^2 = 0.316$), Visual and Technological Design ($f^2 = 0.283$), Aesthetic Quality ($f^2 = 0.280$), and Curriculum Alignment ($f^2 = 0.230$) exert medium to large effects on Perceived Effectiveness (PE). This suggests that improvements in any of these design-quality dimensions would meaningfully enhance learners' perceived learning benefits. In contrast, the effect of PE on Intention to Use (IU) was small ($f^2 = 0.082$), aligning with the marginal significance observed in the structural path and further suggesting that while PE contributes to behavioral intention, additional motivational and contextual factors influence learners' willingness to continue using the VR system.

Table 6
Model strength and predictive relevance.

Construct	R^2	Q^2	Interpretation
PE	0.543	0.347	Moderate–High predictive relevance
IU	0.272	0.149	Low–Moderate predictive relevance

5.4. Qualitative explanatory findings

To contextualize the structural model outcomes, participants' open-ended reflections were analyzed using reflexive thematic analysis. The qualitative evidence provides insight into how learners experienced the design-quality dimensions of VARIASI and clarifies why the four design features strongly shaped Perceived Effectiveness (PE) but did not uniformly translate into higher Intention to Use (IU). The analysis yielded three interrelated explanatory themes. The themes are presented below in the order in which they relate to the learning process: initial conceptual clarity during use, sustained engagement and focus, and post-use readiness for continued adoption.

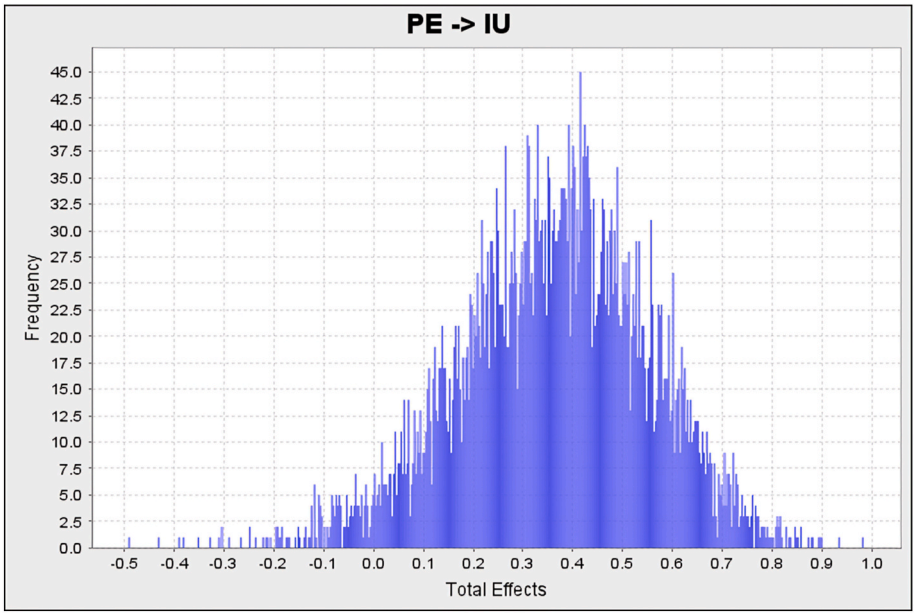

Fig. 8. Bootstrapped total effects distribution for the PE → IU path.

(1) Clarity Through Visual Representation

Participants repeatedly emphasized that visualizing the motion and spacing of particles across solid, liquid, and gas states made the concept "click." Learners described that seeing rather than imagining molecular behavior reduced abstraction and made phase transitions more intuitive:

"*I could finally see how the particles move when the temperature changes. It became clearer than when I only read about it.*"

Such reflections align with CTML's principle of representational alignment, wherein dynamic visualizations reduce intrinsic load and support schema construction. This explains why Visual and Technological Design (VT) and Curriculum Alignment (CA) exhibited strong positive effects on PE: the simulation effectively directed cognitive resources toward conceptual understanding, rather than interface navigation.

(2) Aesthetic Coherence as Cognitive Stabilization

Participants also highlighted how visual harmony and spatial organization created an emotionally and cognitively supportive environment:

"The colors and layout made it feel easy to pay attention. It didn't overwhelm me."

Another commented,

"It felt smooth and relaxing, so I could focus on learning, not adjusting."

These responses suggest that Aesthetic Quality (AQ) served not only motivational purposes but also acted as a regulator of cognitive effort, consistent with CTML findings that coherent and visually balanced environments reduce extraneous cognitive load. This provides qualitative confirmation of the statistically strong AQ → PE path.

(3) Confidence and Guidance as Preconditions for Continued Use

Despite high perceived learning value, many participants expressed uncertainty about using the simulation independently:

"I liked it, but I wasn't always sure what to do next. I needed instructions."

Another shared,

"The headset was a bit heavy, so I couldn't stay too long."

A teacher participant reflected,

"This works well when someone guides the class, but students might struggle if they use it alone."

These reflections indicate that Intention to Use (IU) depends on factors that extend beyond the cognitive benefits of the simulation itself. Learners expressed a need for greater confidence and procedural clarity when navigating the environment, suggesting that the experience was most effective when guidance or scaffolding was present. Several participants also noted the role of physical comfort and familiarity with the device, implying that the headset's weight and the novelty of the interaction required adjustment over time. Taken together, these points help explain the bootstrapped variability observed in the PE → IU path (Fig. 8): even when learners perceived the simulation as helpful for understanding, this did not consistently translate into readiness for continued or autonomous use without structured support.

5.4.1. Interpretive summary

The qualitative findings provide explanatory depth to the quantitative model by revealing the mechanisms behind the relationships among design dimensions, perceived learning benefits, and adoption intentions. Visualization supported conceptual clarity, aesthetic coherence sustained engagement, and digital feasibility reduced frustration, collectively strengthening perceptions of effectiveness. However, continued willingness to use the simulation depended on whether learners felt confident, comfortable, and supported while interacting with the immersive environment. These insights emphasize that effective immersive learning design requires not only representational and technical quality, but also learner-centered scaffolding that ensures comfort, guidance, and accessibility.

Taken together, the quantitative and qualitative findings converge on a consistent interpretation: while the VARIASI simulation effectively supported conceptual understanding through coherent visual, aesthetic, and curricular design, sustained willingness to re-engage with the system depended on the presence of guidance, familiarity, and contextual support. The following discussion elaborates on these implications and situates them within existing immersive learning research.

6. Discussion

This study investigated how four design-quality dimensions of the VARIASI immersive VR simulation—Visual and Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), and Digital Feasibility (DF)—shape learners' Perceived Effectiveness (PE) and their Intention to Use (IU) the system. The combined quantitative and qualitative findings demonstrate that effective learning in VR emerges not from immersion alone, but from the coordinated alignment of cognitive, affective, and pragmatic design factors, in line with Cognitive Load Theory (Sweller et al., 2019) the Cognitive Theory of Multimedia Learning (Mayer, 2024), and the Cognitive Affective Model of Immersive Learning (Makransky & Petersen, 2021). This contributes to a growing body of research arguing that immersive learning environments are most beneficial when they reduce extraneous load (CLT/CTML), enable conceptual focus, and create emotional engagement (CAMIL), which in turn drives technology adoption (TAM) (Makransky & Petersen, 2021; Radianti et al., 2020).

6.1. Interpretation of key findings

The quantitative model showed that VT, AQ, CA, and DF jointly explained over half of the variance in PE, demonstrating that learners' evaluations of effectiveness emerge from multiple interacting design elements rather than from any single dimension. Digital Feasibility exerted the strongest influence, suggesting that technical stability is foundational to cognitive engagement. When VR systems operate smoothly, learners' working-memory resources remain available for conceptual reasoning rather than device management (Makransky & Mayer, 2022; Parong & Mayer, 2021). Conversely, even minor latency can fragment attention and disrupt meaning-making in immersive environments (Saredakis et al., 2020). Thus, DF should be understood not simply as usability, but as a cognitive prerequisite for learning.

Aesthetic Quality also strongly predicted PE. Participants' descriptions of the environment as "comfortable," "calming," and "easy to focus on" reflect findings in HCI and educational psychology showing that visual harmony and emotional resonance can stabilize attention and reduce extraneous cognitive load (Hassenzahl, 2004; Makransky & Lilleholt, 2018; Sauer & Sonderegger, 2022). In immersive science learning, aesthetics act as affective scaffolds, enabling students to maintain concentration long enough to build coherent mental models (Jafarkhani et al., 2024; Parong & Mayer, 2021).

Visual and Technological Design contributed positively by enabling cognitive transparency during interaction. Smooth controller response, intuitive navigation, and visible cause–effect relationships reduce interface-related mental effort (Makransky et al., 2019; Whitton, 2014). This aligns with research showing that VR is most educationally effective when interaction demands do not overshadow conceptual reasoning (Radianti et al., 2020; Saran, 2025).

Curriculum Alignment likewise supported PE by anchoring

experiences to existing conceptual frameworks. Immersive environments yield stronger learning outcomes when integrated with prior knowledge and instructional sequencing (Reigeluth & Carr-Chellman, 2009; Zekeik et al., 2025). VR that is pedagogically unanchored risks producing "wow-effect" engagement without conceptual transfer (Makransky & Petersen, 2021). In VARIASI, alignment with kinetic molecular theory helped avoid that pitfall.

Taken together, these results reinforce that effective VR for science learning requires simultaneous alignment of technical reliability, visual coherence, intuitive interaction, and conceptual relevance—a synergy increasingly recognized in learning sciences research (Bronack et al., 2008; Lindgren & Johnson-Glenberg, 2013).

6.2. Understanding the marginal PE → IU relationship

Although PE positively influenced IU, the effect was marginal and unstable across bootstrap resamples. The qualitative findings clarify that learners valued the conceptual clarity VR provided, but did not consistently feel confident using it independently. This suggests that IU is conditioned by affective readiness, procedural familiarity, and the presence of instructional guidance (Bracq et al., 2019; Hamilton et al., 2021).

Participants' reflections indicated that VR was most effective when supported by teacher facilitation, consistent with research showing that scaffolding is essential to prevent cognitive drifting or surface-level engagement in immersive contexts (Han et al., 2025; Makransky & Petersen, 2021). Physical comfort also played a role; discomfort from headset weight influenced stamina, a well-documented barrier in sustained VR (Saredakis et al., 2020).

Viewed through the integrated lens of our theoretical framework, these findings underline the complexity of adoption-related processes. They confirm that positive cognitive evaluations of usefulness (CLT and CTML driving PE) are necessary but not sufficient for adoption (TAM). Actual adoption heavily depends on the affective and embodied factors central to CAMIL, such as physical comfort, agency, and procedural confidence. Thus, PE is a necessary but not sufficient condition for intention to adopt VR. Adoption requires cognitive trust + procedural confidence + physical comfort, a pathway increasingly recognized in technology-acceptance frameworks for immersive learning that bridge cognitive outcomes with affective readiness (Goulding et al., 2007; Sagnier et al., 2020).

This pattern is consistent with prior immersive learning research, which shows that adoption intentions in educational VR are shaped less by interface qualities per se than by learners' confidence in the learning value and usability of the experience. Accordingly, this suggests that the weak direct paths to IU should not be interpreted as a failure of instructional design, but rather as an indication that physical and ergonomic factors can override positive learning perceptions in immersive systems (Adelana et al., 2023; Conrad et al., 2024).

6.3. Pedagogical implications

These results underscore the need for instructional scaffolding when implementing VR in science learning. While the simulation successfully enhanced conceptual clarity and engagement, sustained use requires structured guidance that orients learners during interaction and helps them transition from initial novelty to autonomous control. Teachers play a central role in linking virtual representations to scientific reasoning, ensuring that immersion supports rather than replaces conceptual discussion. Likewise, gradual familiarization with device controls and consideration of physical comfort can help strengthen long-term usability and student confidence.

6.4. Limitations and future directions

While the model exhibits strong explanatory power, two limitations must be acknowledged. First, the study focused solely on learners' self-reported perceptions, without triangulating with behavioral or performance-based measures. Future research could integrate eye-tracking or in-system analytics to capture more objective indicators of engagement. Second, the sample size was moderate and limited to a specific educational context, suggesting that replication across diverse populations and content domains would strengthen generalizability. Further investigation could explore adaptive mechanisms that dynamically adjust design features based on learners' cognitive profiles, extending the present model toward personalized immersive learning systems.

7. Conclusion

This study examined how four dimensions of design quality—Visual and Technological Design (VT), Aesthetic Quality (AQ), Curriculum Alignment (CA), and Digital Feasibility (DF)—shape learners' Perceived Effectiveness (PE) and Intention to Use (IU) the VARIASI immersive VR simulation for learning about particle behavior and phase structures. The findings demonstrate that all four design dimensions significantly contribute to PE, indicating that learners evaluate immersive educational systems holistically rather than based on a single feature. When visual clarity, emotional comfort, conceptual alignment, and technical stability are simultaneously present, immersive environments more effectively support cognitive focus and conceptual understanding.

However, PE did not uniformly translate into IU. The marginal and variable PE → IU pathway suggests that successful learning experiences do not automatically generate willingness for continued or autonomous use. Qualitative responses reveal that IU is influenced by contextual and experiential conditions such as headset comfort, familiarity with VR controls, and the presence of instructional guidance. Learners valued the clarity and insight provided by VARIASI, but many expressed uncertainty about independent use without teacher facilitation or structured scaffolding. This indicates that adoption of VR for sustained learning depends not only on system design but also on pedagogical integration and user readiness.

Taken together, these results underscore that effective immersive science learning is achieved when VR is designed not only to be visually coherent and technically stable, but also when it is implemented with instructional scaffolding that supports confidence, navigation, and meaning-making. Future research should explore adaptive scaffolding, performance-based learning analytics, and classroom implementation models that strengthen learners' self-efficacy and sustained engagement. Additionally, expanding the sample and incorporating objective measures of conceptual gain would strengthen generalizability and deepen understanding of how immersive environments can best support conceptual change in science education.

CRediT authorship contribution statement

Achmad Samsudin: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Muhammad Zahran:** Writing – review & editing, Validation, Software, Resources, Data curation, Conceptualization. **Eki Nugraha:** Writing – review & editing, Validation, Supervision, Software, Resources, Methodology, Data curation, Conceptualization. **Hadi Nasbey:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Mustafa Sozbilir:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Nor Farahwahidah Abdul Rahman:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Takashi Irie:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Statements on open data and ethics

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to ethical and privacy considerations involving human participants, the data are not publicly available. All procedures involving human participants were conducted in accordance with ethical standards and relevant institutional guidelines.

Ethical statement

This study has been reviewed and approved by the Research Ethics Committee of Universitas Pendidikan Indonesia (Ethics Approval No. 184/UN40.KP/PT.01.01/2025). All procedures were conducted in accordance with relevant laws, institutional guidelines, and ethical standards. Informed consent was obtained from the parents or legal guardians of all participants involved in this study. In addition, assent was obtained from each participant prior to data collection. The privacy and confidentiality of all human subjects were strictly maintained, and all data were securely stored in a protected repository accessible only to the research team.

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. This study was supported by the RKLI Program—Indonesia's Collaborative Teacher Education Research Scheme—under Contract No. 47/UN40.D/PT.01.01/2025. The funding body had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Acknowledgements

This research was funded by the RKLI Program—Indonesia's Collaborative Teacher Education Research Scheme—under Contract No. 47/UN40.D/PT.01.01/2025. The authors also express their sincere appreciation to the Directorate of Research and Community Service of Universitas Pendidikan Indonesia (DPPM UPI) for its support and facilitation throughout the conduct of this study.

References

Adelana, O. P., Ayanwale, M. A., Ishola, A. M., Oladejo, A. I., & Adewuyi, H. O. (2023). Exploring pre-service teachers' intention to use virtual reality: A mixed method approach. *Computers & Education: X Reality, 3*, Article 100045. https://doi.org/10.1016/j.cexr.2023.100045

Ahn, S. J. G., Bailenson, J. N., & Park, D. (2014). Short-and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. *Computers in Human Behavior, 39*, 235–245.

Bracq, M.-S., Michinov, E., Arnaldi, B., Caillaud, B., Gibaud, B., Gouranton, V., & Jannin, P. (2019). Learning procedural skills with a virtual reality simulator: An acceptability study. *Nurse Education Today, 79*, 153–160.

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology, 18*(3), 328–352.

Bronack, S., Sanders, R., Cheney, A., Riedl, R., Tashner, J., & Matzen, N. (2008). Presence pedagogy: Teaching and learning in a 3D virtual immersive world. *International Journal of Teaching and Learning in Higher Education, 20*(1), 59–69.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In *Modern methods for business research* (pp. 295–336). Psychology Press.

Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computers & Education: X Reality, 4*, Article 100053. https://doi.org/10.1016/j.cexr.2024.100053

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research.* Sage publications.

Crogman, H. T., Cano, V. D., Pacheco, E., Sonawane, R. B., & Boroon, R. (2025). Virtual reality, augmented reality, and mixed reality in experiential learning: Transforming educational paradigms. *Education Sciences, 15*(3), 303.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—Principles and practices. *Health Services Research, 48*(6pt2), 2134–2156.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association, 70*(350), 320–328.

Goulding, J., Sexton, M., Zhang, X., Kagioglou, M., Aouad, G. F., & Barrett, P. (2007). Technology adoption: Breaking down barriers using a virtual reality design support tool for hybrid concrete. *Construction Management & Economics, 25*(12), 1239–1250.

Hair, J. F., Astrachan, C. B., Moisescu, O. I., Radomir, L., Sarstedt, M., Vaithilingam, S., & Ringle, C. M. (2021). Executing and interpreting applications of PLS-SEM: Updates for family business researchers. *Journal of Family Business Strategy, 12*(3), Article 100392.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24.

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education, 8*(1), 1–32.

Han, X., Luo, H., Wang, Z., & Zhang, D. (2025). Using virtual reality for teacher education: A systematic review and meta-analysis of literature from 2014 to 2024. *Frontiers in Virtual Reality, 6*, Article 1620905.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction, 19*(4), 319–349.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Huang, T.-C., & Tseng, H.-P. (2025). Extended reality in applied sciences education: A systematic review. *Applied Sciences, 15*(7), 4038.

Jafarkhani, F., Barani, F., Nawaser, K., Rashidi, H., & Gharleghi, B. (2024). Factors affecting the aesthetic experiences in educational games: A qualitative investigation. *Journal of Economy and Technology, 2*, 200–207.

Jo, H., & Park, D.-H. (2023). Affordance, usefulness, enjoyment, and aesthetics in sustaining virtual reality engagement. *Scientific Reports, 13*(1), Article 15097.

Kavanagh, S., Luxton-Reilly, A., Wünsche, B. C., Plimmer, B., & Dunn, S. (2025). Evaluating the use of 360-Degree video in education. *Electronics, 14*(9), 1830.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 445–452.

Liu, R., Wang, L., Lei, J., Wang, Q., & Ren, Y. (2020). Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons. *British Journal of Educational Technology, 51*(6), 2034–2049.

Lui, M., Chong, K.-Y. A., Mullally, M., & McEwen, R. (2023). Facilitated model-based reasoning in immersive virtual reality: Meaning-making and embodied interactions with dynamic processes. *International Journal of Computer-Supported Collaborative Learning, 18*(2), 203–230.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research & Development, 66*(5), 1141–1164.

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. *Educational Psychology Review, 34*(3), 1771–1798.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937–958.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236.

Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review, 36*(1), 8.

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37*(1), 226–241.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778.

Reigeluth, C. M., & Carr-Chellman, A. A. (2009). Domain Th eory for instruction: Mapping attainments to enable learner-centered education. In *Instructional-Design theories and models* (Vol. III, pp. 339–360). Routledge.

Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I., & Valléry, G. (2020). User acceptance of virtual reality: An extended technology acceptance model. *International Journal of Human-Computer Interaction, 36*(11), 993–1007.

Saran, M. (2025). Comparing hand-based and controller-based interactions in virtual reality learning: Effects on presence and interaction performance. *PeerJ Computer Science, 11*, Article e3168.

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A., Rizzo, A., & Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis. *Frontiers in Human Neuroscience, 14*, 96.

Sauer, J., & Sonderegger, A. (2022). Visual aesthetics and user experience: A multiple-session experiment. *International Journal of Human-Computer Studies, 165*, Article 102837.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI, 3*, 74.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is useable. *Interacting with Computers, 13*(2), 127–145.
Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly*, 21–54.
Wenk, N., Penalver-Andres, J., Buetler, K. A., Nef, T., Müri, R. M., & Marchal-Crespo, L. (2023). Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. *Virtual Reality, 27*(1), 307–331.
Whitton, N. (2014). *Digital games and learning: Research and theory*. Routledge.
Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology, 51*(6), 1991–2005.
Zekeik, H., Chahbi, M., Sefian, M. L., & Bakkali, I. (2025). Augmented reality and virtual reality in education: A systematic narrative review on benefits, challenges, and applications. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(9). em2699.
Zhang, Y., Feijoo-Garcia, M. A., Gu, Y., Popescu, V., Benes, B., & Magana, A. J. (2024). Virtual and augmented reality in Science, Technology, Engineering, and Mathematics (STEM) education: An Umbrella Review. *Information, 15*(9), 515.
Zhou, C., & Li, J. (2024). The development of aesthetic experience through virtual and augmented reality. *Scientific Reports, 14*(1), 4290.