Immersion Without Headsets: Expanding the Concept Beyond Virtual Reality Technologies

Miriam Mulders

PII: S2451-9588(26)00093-X

DOI: https://doi.org/10.1016/j.chbr.2026.101019

Reference: CHBR 101019

To appear in: *Computers in Human Behavior Reports*

Received Date: 17 December 2025

Revised Date: 13 March 2026

Accepted Date: 15 March 2026

Please cite this article as: Mulders M., Immersion Without Headsets: Expanding the Concept Beyond Virtual Reality Technologies, *Computers in Human Behavior Reports*, https://doi.org/10.1016/j.chbr.2026.101019.

Immersion Without Headsets: Expanding the Concept Beyond Virtual Reality Technologies

Miriam Mulders [a,*]

* Corresponding author: miriam.mulders@uni-due.de, +49 201 183 5258

[a] University of Duisburg-Essen, Faculty of Educational Sciences, Chair of Educational Technology and Instructional Design, Universitätsstraße 2, 45141 Essen, Germany

Immersion Without Headsets: Expanding the Concept Beyond Virtual Reality Technologies

Abstract

Immersion is commonly conceptualized in virtual reality (VR) research as a technological property, often defined through display fidelity, sensorimotor contingencies, or interactional bandwidth. Such device-centered perspectives capture only one dimension of immersion and do not fully account for the experiential ways in which individuals become absorbed in an environment, narrative, or activity. This article reconceptualizes immersion as a modality-independent experiential state emerging from the interplay of subjective factors, system and content properties, and engagement processes. Immersive experience, it argues, can arise in both mediated and non-mediated contexts, including memorial sites, museum exhibitions, and analogue narratives. From this perspective, VR becomes educationally relevant not because it uniquely produces immersion, but because it can enable immersive experiences when real-world access is limited, impractical, or pedagogically constrained. Interpreting immersion as a modality-independent, relational construct provides a clearer basis for evaluating VR's pedagogical affordances and underscores that immersive learning depends on instructional design and learner preparation rather than technological fidelity alone.

Introduction

Across the fields of virtual reality (VR) research, the concept of immersion has become both central and contested. Traditional accounts describe immersion largely as a technological variable. Slater (2003), for example, conceptualizes immersion as the set of technological features that provide the conditions for a sense of being located within a mediated environment. In his view, the degree of immersion corresponds directly to the technological capabilities of the system: More advanced technology affords a higher level of immersion. This technology-centered understanding has strongly influenced subsequent research on VR, where immersion is frequently operationalized in terms of system fidelity or sensorimotor richness (e.g., Berkman & Akan, 2024; Cummings & Bailenson, 2016). The rapid expansion of VR technologies and educational metaverse applications has further intensified medium-centric interpretations of immersion. In these discourses, immersive learning is often equated with increased technological sophistication, reinforcing assumptions that higher sensory fidelity necessarily

yields deeper experiential engagement (e.g., Buragohain, Meng & Chaudhary, 2025; Mystakidis, 2022).

Building on this conceptualization, VR technologies are frequently categorized as low-, semi-, or high-immersive, based on parameters such as field of view, display resolution, frame rate, or sensorimotor feedback fidelity (Bowman & McMahan, 2007). This taxonomy has undoubtedly been useful for standardizing technological comparisons, but it also risks obscuring the fundamentally experiential nature of what it means to be immersed. When immersion is treated as a fixed attribute of the medium rather than a relational outcome between user, environment, and content, the complexity of immersive learning is flattened into a hardware specification (blinded). However, the term immersion is also frequently used to describe a subjective experience resulting from a technological process (e.g., Witmer & Singer, 1998). High-fidelity spatial modeling, coupled with head- and motion-tracking and spatialized audio, can create conditions in which users perceive the virtual environment with a degree of immediacy that exceeds conventional media. When sensorimotor contingencies are coherent and environmental responses align predictably with users' movements, the mediated space begins to function as an embodied environment rather than a static representation, thereby reinforcing the experiential dimension of immersion (blinded).

This terminological ambiguity has produced persistent conceptual confusion. Even after extensive research efforts, immersion remains conceptually heterogeneous and lacks a universally accepted definition (e.g., Nilsson, Nordahl & Serafin, 2016). Existing conceptual models often employ overlapping terminology or attribute different meanings to the same terms depending on disciplinary or contextual assumptions.

A recent attempt to resolve this definitional inconsistency is provided by Lee's (2025) *unified conceptual model of immersive experience in extended reality*, which distinguishes systematically between *immersive systems* and *immersive experience* and organizes related constructs (e.g., physical presence, social presence, self-presence, and involvement) within a coherent hierarchical structure. This model marks an important step toward conceptual clarity, as it explicitly links experiential states to system properties such as plausibility, interactivity, and interestingness while acknowledging the influence of subjective factors, including prior knowledge and skills. In doing so, Lee offers one of the most comprehensive attempts to consolidate the fragmented landscape of immersion research. Although Lee's unified model provides important conceptual clarity within XR research, its scope remains primarily situated within technologically mediated environments. The cross-modal generalizability of its experiential architecture is not explicitly theorized, leaving open broader structural questions

about immersion beyond VR. Against this backdrop, two conceptual gaps persist in the current literature. *First*, immersion is still predominantly framed as a device-dependent property, leading to a technological determinism that obscures the relational and psychological nature of immersive experience. *Second*, existing models rarely integrate mediated and non-mediated forms of immersion within a single conceptual architecture, resulting in fragmented accounts across VR, narrative media, and physical learning environments. This article addresses these gaps by advancing a modality-independent account of immersion. Modality-independent immersion is defined here as an experiential state emerging from the dynamic interplay of subjective factors, system and content characteristics, and engagement processes, irrespective of whether these determinants are instantiated through digital technologies or physical environments. The primary objective of this article is therefore to reconceptualize immersion as a cross-modal experiential construct by extending Lee's unified model beyond XR contexts. Rather than merely applying Lee's framework to new examples, the article introduces an additional realization layer that makes explicit how the model's determinants can be instantiated across different media configurations. This article therefore adopts a conceptual, theory-building approach. Rather than presenting new empirical data, it develops a theoretical extension of Lee's model to examine whether the determinants identified in the framework operate similarly across mediated and non-mediated contexts. More specifically, the article addresses the following guiding research question: How do the determinants identified in Lee's model manifest across different realization domains, and can immersive experience be understood as structurally invariant yet configurationally variable across media?

In doing so, the article makes three contributions: (1) It advances conceptual clarity by distinguishing immersive system parameters from immersive experience parameters and integrating them within a relational, hierarchical architecture. (2) It reconceptualizes immersion as a modality-independent construct by introducing a realization-domain perspective that accounts for how identical structural determinants manifest across media. (3) It reframes VR not as a privileged source of immersion, but as a configurable realization domain whose pedagogical value depends on how its system properties are aligned with subjective readiness, engagement processes, and the selective amplification of experiential dimensions, particularly in contexts where real-world immersion is constrained or impossible.

The article is structured as follows: The next section reviews relevant literature on immersion in VR research and beyond digital technologies. This is followed by the introduction of Lee's unified conceptual model and its theoretical extension through realization layers. The subsequent section analyzes how immersive experience manifests across different modalities,

before the pedagogical affordances of VR are discussed. The article concludes with a discussion of theoretical implications, limitations, and directions for future research.

Theoretical background

This section establishes the theoretical foundation for the argument by reviewing dominant conceptualizations of immersion in VR research, examining immersive experiences beyond digital technologies, and introducing Lee's (2025) unified conceptual model as the analytical framework guiding the subsequent reconceptualization. The discussion does not constitute a systematic review. Instead, the literature was selected based on its relevance to immersion research and to Lee's model.

Immersion as a Multidimensional Construct in VR Research

The term immersion is typically used in two distinct ways in the literature: as a property of technology and as a characteristic of subjective experience (Lee, 2025). In the former sense, immersion is defined through the objective and quantifiable properties of a VR system that support sensorimotor contingencies and create the potential for presence which refers to the experiential sense of "being in" or "being with" a mediated or represented environment (Slater, 2003; Slater & Wilbur, 1997). Bowman and McMahan (2007) similarly frame immersion as a function of system fidelity, emphasizing the extent to which VR technologies replicate sensory input or afford naturalistic interaction. These system-oriented definitions have been influential, forming the conceptual backbone of numerous empirical studies and experimental manipulations (e.g., Berkman & Akan, 2019; Cummings & Bailenson, 2016). VR systems are often categorized according to their technological characteristics into high-, semi-, or low-immersive configurations, a distinction that typically reflects the use of different technological setups ranging from head-mounted displays to mobile viewers such as cardboard-based solutions, as well as projection-based systems such as CAVEs (Bowman & McMahan, 2007; Cummings & Bailenson, 2016). While such classifications are useful for describing technological configurations, they do not fully capture the experiential conditions under which immersion emerges. Jennett et al. (2008), for example, highlight that immersion encompasses psychological variables such as attention, control, and involvement that are not reducible to hardware features.

In line with these critiques, several studies identify a structural differentiation between system immersion and experiential immersion (Barrett, Pack & Quaid, 2021; Lee, 2025; Makransky, 2021). System immersion refers to the technological parameters of VR, such as head tracking, field of view, sensory input, and representational fidelity, that determine the

system's capacity to occlude the outside world. Experiential immersion, by contrast, reflects the subjective psychological state of being absorbed in a task, characterized by presence, agency, sustained attentional focus, and embodiment, which captures the experience of being situated or represented as an acting self within the environment (Kampling, 2018; Riva, 2008; Southgate et al., 2018). This dual perspective aligns with hierarchical models suggesting that technological features function as precursors enabling experiential immersion (e.g., Makransky & Petersen, 2021). These distinctions underscore that immersion is not reducible to technological fidelity but emerges through its interaction with learner characteristics and situational variables. Together, these perspectives reveal immersion as a multidimensional construct in which technological conditions may initiate immersive potential, but psychological factors determine whether immersive engagement develops and contributes to learning.

Immersive Experiences Beyond Digital Technologies

A growing body of research underscores that immersive states frequently arise in non-digital environments (blinded; Packer & Ballantyne, 2016). Narrative transportation theory likewise demonstrates that individuals can become deeply absorbed in stories, experiencing a vivid sense of entering a narrative world (Green & Brock, 2000). In this context, it becomes evident that the human perceptual system is largely agnostic to the medium through which immersive experience is elicited. Comparable states of absorption may arise from computer-generated environments, curated installations, original artefacts, or interpersonal encounters in physical settings. However, such absorption arises only when individuals actively direct their attention toward the environment and are willing to engage with what it offers. Immersion, whether in virtual or physical settings, is therefore not an automatic property of a medium but the emergent result of an interaction between the medium, the content, and the user (blinded).

This becomes evident in cases where individuals engage deeply with analogue or real-world representations, for instance, when readers of Anne Frank's diary experience intense narrative transportation despite the absence of any technological mediation. Reading Anne Frank's diary can elicit imaginative transportation and empathic identification, as readers mentally simulate the historical situation and its emotional stakes (Derry, 2021; Graver, 2020; Perry, 2018; Schulz, 2019). Filmic adaptations may heighten this experience through dense audiovisual cues. Soundscapes in particular can powerfully shape mood and engagement. However, from a theoretical perspective, cinematic presentation is not inherently more immersive than the written diary. Immersive effectiveness depends on alignment between the adaptation and recipients' expectations, as well as on the quality and plausibility of the filmic realization. Modality alone does not determine immersion. Rather, immersion arises from the interplay of

content, design choices, and the fit between message and medium (blinded). Interactivity adds another pathway to immersive engagement. Immersion can be strengthened when individuals are not merely observers but can act upon and influence the represented world. Even simple affordances, such as navigating a scene or opening a door, create task- and motor-based engagement (Lee, 2025). Digital representations of the Anne Frank hiding place illustrate these principles in practice. Web-based 3D reconstructions and VR applications aim to communicate the cramped and oppressive conditions of the annex and to provoke reflection (Hartmann, 2019; Lichtman, 2021; blinded). Such experiences can render abstract historical facts more tangible, but they also raise ethical concerns. Virtual reconstructions must carefully demarcate documented evidence from speculative reconstruction, and designers must avoid sensationalist or reductive treatments that risk trivialising lived suffering.

From a pedagogical perspective, there is no principled reason to assume that any single medium reliably produces higher immersion or superior learning outcomes. The longstanding Clark-Kozma debate remains instructive: Whereas Clark (1994) argued that media function as mere delivery systems, Kozma (1994) emphasised that media afford distinct pedagogical possibilities that only take effect when aligned with instructional design. Contemporary scholarship synthesizes these views: VR technologies can increase immersion through multisensory and interactive affordances, yet the device alone does not guarantee meaningful learning. Rather, designers must align instructional goals with the specific affordances and limitations of a given medium (blinded; blinded). Without such alignment, immersive features may introduce additional cognitive load, particularly on working memory, that can undermine intended learning outcomes (Makransky, Terkildsen & Mayer, 2019).

In sum, representations of the Anne Frank annex, whether textual, cinematic, curatorial, or virtual, may generate educational value only insofar as they engage an attentive audience and are designed with sensitivity to content, plausibility, ethical responsibility, and pedagogical intent. For both researchers and practitioners, the task is not to privilege one medium over another but to design immersive experiences in which medium, content, and instructional framing are coherently integrated, ethically responsible, and empirically justified.

VR as a Mediator of Experiential Accessibility

Although immersion can occur in many modalities, VR possesses a unique ability to make certain immersive experiences accessible when real-world conditions impose constraints. VR simulations can reconstruct environments that are geographically distant, historically inaccessible, or ethically problematic to visit (Zender et al., 2019). Digitally reconstructed historical environments, such as ancient Rome in VR, allow learners to engage affectively and

cognitively with contexts that no longer exist (Boffi et al., 2023). Similarly, virtual explorations of ecosystems like the Amazon rainforest can offer sensory richness and environmental complexity that might be inaccessible for most learners (blinded). In these examples, VR functions not as the source of immersion but as a mediator that provides the contextual conditions under which immersive experiences can emerge. VR also becomes pedagogically indispensable in domains where immersive real-world learning would be dangerous or prohibitively costly. For example, VR emergency-response simulations allow learners to experience hazardous situations without physical risk, supporting decision-making and situational awareness in ways that would be impossible to replicate safely in the real world (Feng et al., 2018). Likewise, VR-based training for procedural skills such as industrial painting provides high-fidelity practice opportunities without material waste, equipment costs, or safety concerns (blinded).

Lee's Unified Conceptual Model of Immersive Experience in Extended Reality (2025)

Lee (2025) recently offers one of the most comprehensive attempts to resolve the longstanding conceptual ambiguity surrounding immersion (see Figure 1). Rather than treating immersion as a monolithic construct, Lee proposes a multidimensional model of immersive experience, grounded in the distinction between *immersive systems* (i.e., the technological conditions that can enable immersion) and *immersive experience* (i.e., the psychological state that emerges through user engagement). This clarification addresses a major source of inconsistency in the literature, where immersion has alternated between being defined as a technological property and as a subjective experience. Furthermore, the model positions immersive experience at the apex of the hierarchy as the dependent construct, with the immersive system serving as the independent determinant from which immersive experience emerges.

At the core of Lee's model are four experiential dimensions that jointly constitute an immersive experience: physical presence, social presence, self-presence, and involvement. (1) *Physical presence* captures the sense of being physically situated in a virtual environment through sensory simulation (Schubert, Friedmann & Regenbrecht, 2001). (2) *Social presence* describes perceiving virtual or real intelligences as socially and perceptually real (Biocca, 1997). (3) *Self-presence* concerns the feeling of being embodied in a virtual self (Biocca, 1997) and (4) *involvement* reflects deep engagement with a narrative or task (Witmer & Singer, 1998). Crucially, Lee's model holds that none of these dimensions is necessary or sufficient in isolation. Immersive experience can arise from any one of them, but their co-occurrence produces stronger, more complex forms of immersion.

Next, the model identifies three overlapping sub-components: (1) *Sensorimotor engagement* connecting physical presence and self- and social presence, (2) *narrative engagement* linking physical presence and involvement, and (3) *task/motor-related engagement* is situated between involvement and self- and social presence. These types of engagement are assumed to be preconditions of physical, social and self-presence as well as involvement. In this context, sensorimotor engagement refers to the user's perceptual and bodily coupling with the mediated environment. It captures the extent to which visual, auditory, and haptic channels are drawn into interaction with system outputs and how coherently these responses follow the user's movements and actions (O'Regan & Noë, 2001; Slater, 2009). In this sense, sensorimotor engagement constitutes the experiential groundwork upon which physical, social and self-presence can emerge. Task/motor engagement describes the user's cognitive and bodily investment in an activity or task. It involves attending to task goals, understanding the rules or structure of an interaction, and coordinating bodily actions accordingly (Krapp et al., 1992; Witmer & Singer, 1998). In interactive VR scenarios, task/motor engagement is central to the development of social and self-presence. As users become more proficient and fluent in task performance, they tend to experience themselves as competent agents embedded in a meaningful activity, which, in turn, supports deeper involvement. Narrative engagement focuses on the user's alignment with the story, thematic structure, or unfolding events of the content. It may facilitate cognitive absorption in the plot, characters, or scenario and is therefore closely linked to involvement and, indirectly, to physical presence through imaginative transportation (Calleja, 2007; Ryan, 2003). Taken together, these three forms of engagement are conceptualized as lower-level processes that precede immersive experience. Engagement does not itself constitute immersive experience, but it opens the pathway for physical presence, social presence, self-presence, and involvement to arise and intensify. This framing helps explain why high-fidelity VR systems do not automatically produce immersion: Without sufficient sensorimotor, task-related, or narrative engagement, the experiential dimensions that define immersive experience cannot fully develop.

Immersive experience, in Lee's model, depends also on three system and content properties: plausibility, interestingness, and interactivity. *Plausibility* refers to the extent to which a virtual environment, object, or being is perceived as realistic or believable, thereby supporting sensorimotor engagement. Interactivity describes the extent to which a user can actively influence or interact with virtual environments, objects, or beings. This can support task/motor engagement. *Interestingness* captures the degree to which content elicits attention, motivation, and curiosity, thereby fostering narrative engagement.

Lastly, the three system properties are modulated by subjective factors (i.e., internal reference, skills and knowledge, personal preference) which act as confounding variables shaping how a user interprets system cues and engages with content. *Internal reference* refers to the cognitive cues users draw from prior experiences, expectations, or existing knowledge, which influence how plausible a mediated environment appears. *Skills and knowledge* determine users' ability to perform tasks or interact fluently with an immersive system, thereby affecting the degree to which interactivity can support deeper engagement. *Personal preference,* rooted in an individual's enduring interests and dispositions, shapes how interesting or compelling users perceive content to be and thus influences the likelihood of narrative engagement.

By extending Lee's model beyond XR, this article establishes immersion as a modality-independent experiential architecture and specifies how its structural determinants can be realized across different media configurations.

Reconceptualizing Immersion Across Modalities

Although Lee's model is articulated within XR research, its structural architecture is not inherently bound to digital mediation. Building on Lee's framework, this section extends the model beyond XR research by examining how its determinants operate across different realization domains. Figure 1 presents the expanded framework proposed in this article, which introduces realization layers in order to illustrate how the model components operate across different media configurations. The following analysis therefore systematically examines each model component (i.e., subjective factors, system and content properties, engagement preconditions, and core experiential dimensions) to demonstrate their configurational variation across realization layers. The Anne Frank case was selected because it provides a rare multimodal constellation (diary, film, museum, and VR), enabling comparison of immersive

determinants across realization layers.

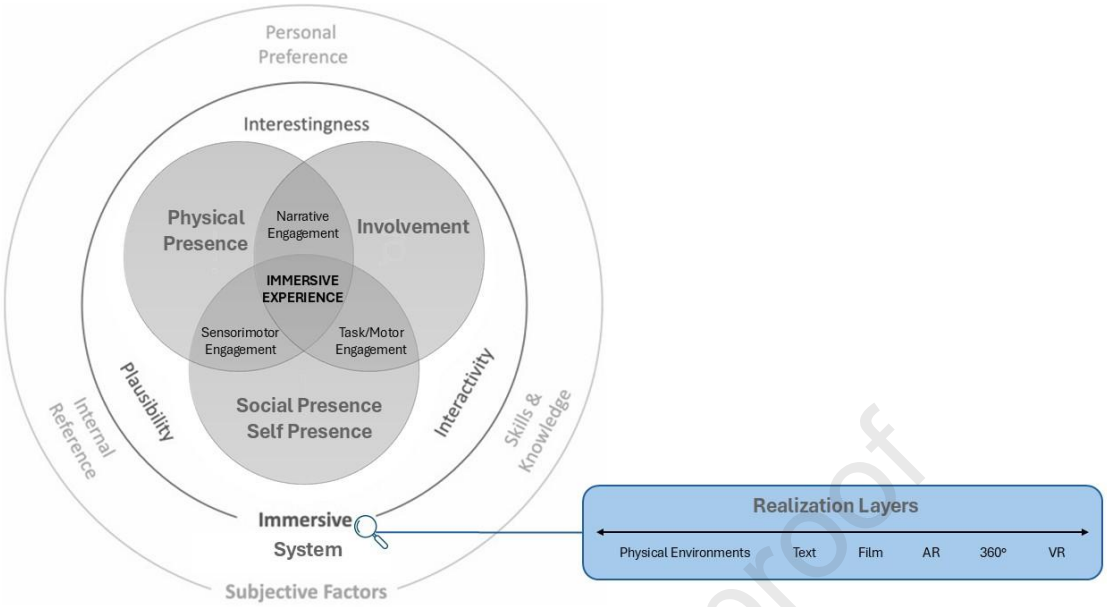

Fig. 1: Expanded model of Lee's (2025) unified conceptual model of immersive experience in extended reality. This version adds a "realization layer" to highlight that immersive experiences can manifest not only in virtual environments but also in physical and analogue settings such as historical sites, artefacts, texts, or interpersonal encounters. The parts of the illustration that originate from the original model are marked in gray, while the newly added components are marked in blue.

Subjective Factors

In Lee's model, subjective factors serve as confounding variables that shape how system and content cues are interpreted and how engagement unfolds. Crucially, these factors can operate identically in VR and non-VR contexts. Internal reference, or the interpretive lens shaped by prior experiences or expectations, can influence plausibility judgments regardless of modality. A learner reading the diary of Anne Frank can draw on historical knowledge, moral frameworks, and prior encounters with Holocaust narratives to make sense of the diary entries. The same internal reference may shape how museum visitors interpret artefacts in the Anne Frank museum in Amsterdam and how VR users evaluate the plausibility of a virtually reconstructed annex. In all cases, immersion can intensify when the representation matches or constructively challenges prior expectations, reinforcing coherence.

Skills and knowledge may play a similar cross-modal role. Readers with stronger historical literacy may experience deeper narrative immersion in the diary. Museum visitors with prior knowledge of Holocaust may navigate the exhibits with greater interpretive fluency and VR users with gaming experience may interact more naturally with a virtual reconstruction of the annex. In each case, domain knowledge and interactional competence can facilitate engagement, just as limited skills can impede immersion irrespective of medium.

Personal preference, finally, can likewise motivation, interest, and openness to engagement. Individuals inclined toward historical narratives may experience greater involvement when reading or visiting the site, while those with low personal interest may require more pedagogical scaffolding, whether in VR or in physical settings, to sustain attention and emotional resonance.

Together, these determinants show that immersion begins prior to exposure, rooted in learners' dispositions and interpretive capacities across contexts.

System- and Content-Level Conditions

Lee's model treats plausibility, interactivity, and interestingness as properties of the immersive system and its content. Although often interpreted technologically, these properties also characterize physical, analogue, and curatorial environments.

Plausibility, defined as the perceived credibility of an experience, is not limited to virtual simulations. In the Anne Frank museum, the cramped physical space itself provides a plausible contextual frame. Curators can enhance plausibility by adhering to historical accuracy in artefacts, room layout, and archival documentation. Similarly, the diary's credibility stems from its authenticity as a historical source. In VR reconstructions, plausibility can be achieved through sensory cues and accurate spatial representation, but the underlying interpretive principle seems to remain the same. Users would need to perceive the representational world as coherent and believable relative to their internal reference.

Interactivity in non-digital contexts can take different but functionally equivalent forms. Turning pages of the diary, navigating rooms in the museum, or physically leaning into displays constitutes a form of task/motor engagement even without digital mediation. These interactions provide agency and influence over the unfolding experience. VR simulations can recreate analogous affordances (e.g., opening doors, pick up a postcard) but the pedagogical function of interactivity remains consistent across modalities.

Interestingness, which can foster narrative engagement, is likewise medium-independent. The diary can captivate through personal voice and emotional authenticity. Museum exhibits can evoke interest through artefacts, testimonies, and spatial dramaturgy. VR can increase situational interest through its multisensory richness and novelty.

Thus, the immersive system, broadly defined, includes not only technological features but also curatorial design, narrative structure, spatial arrangement, and material authenticity in physical environments.

Preconditions for Immersive Experiences

Lee conceptualizes sensorimotor engagement, narrative engagement, and task/motor engagement as preconditions for presence and involvement. In non-digital settings, sensorimotor engagement can be elicited through atmospheric, multisensory signals: the creaking floorboards, the narrow spatial layout that subtly adjusts visitors' posture, or the cadence of a museum guide's voice that anchors attention and shapes emotional tone. VR applications can recreate such embodied attunement through mediated signals, such as spatialized audio that simulates shifting floorboards or the soft echo of footsteps in confined spaces.

Narrative engagement may function similarly in textual, cinematic, curatorial, and virtual modalities. The diary can provide a first-person narrative that invites readers into Anne Frank's world. Films can use audiovisual cues to strengthen emotional resonance. Museums can employ curated narrative pathways through rooms and artefacts. VR can reconstruct the spatial and narrative structure of events, enabling learners to inhabit the physical setting described in the diary. In some applications, this reconstruction may extend further by incorporating Anne Frank's voice or even a virtual avatar resembling her, an affordance that carries significant ethical implications regarding representation, authenticity, and the boundaries of respectful memorialisation.

Task/motor engagement can emerge whenever learners must actively navigate or manipulate elements of an environment. Turning pages or following a museum trail can constitute meaningful task engagement in the real world. VR reconstructions can mimic such embodied coupling through simulated spatial layouts and movement-responsive sensory cues.

Together, these preconditions demonstrate that immersive experience arises whenever sensorimotor, narrative, and task/motor engagements converge, regardless of whether the environment is physical or virtual.

Core Dimensions of Immersive Experience Across Modalities

In Lee's framework, immersive experience is constituted by four core dimensions: physical presence, social presence, self-presence, and involvement. Physical presence in VR settings is typically associated with coherent sensorimotor contingencies and dynamic spatial simulation. However, analogous experiences may arise in non-virtual settings. In the Anne Frank museum, visitors can physically inhabit the cramped spaces of the annex, move through narrow staircases, and orient themselves toward original artefacts. The spatial configuration of the building itself may afford a strong sense of being there. Even reading the diary can evoke a

variant of physical presence through imaginative transportation, as readers are able to mentally construct the spatial layout of the hiding place and experience events as if from within that environment. In all cases, physical presence is not reducible to display technology but arises whenever spatial cues and sensorimotor or imagined engagement cohere into a convincing sense of location.

Social presence equally characterises encounters with other visitors in the museum, with survivors' testimonies presented in audio or video format, and with historical figures mediated through text. When readers of the diary experience Anne Frank as a psychologically vivid interlocutor whose voice they follow over time, they may attribute social reality to a narrated other person despite the absence of any real-time interaction. In VR, social presence can be heightened further by including conversational agents, reconstructed family members, or guides, yet the underlying mechanism remains the same. Learners can experience themselves as being in relation to another whose actions, voice, or story matter for the unfolding experience.

In VR, self-presence is often discussed in terms of avatar embodiment and body ownership (e.g., Kilteni, Groten & Slater, 2012), but similar processes occur in narrative and site-based learning. Readers can experience a form of self-presence when they imaginatively step into Anne Frank's perspective, seeing events through her eyes and aligning their inner monologue with hers. Museum visitors may experience themselves as participants rather than observers when they imagine how it would feel to live in the annex, move along its circulation paths, or hide behind its concealed entrance. Across modalities, self-presence can emerge whenever learners experience themselves as situated agents within the represented situation, whether that situation is physical, narrative, or virtual.

Involvement, finally, denotes deep engagement with a narrative and/or task. It is the dimension that most clearly cuts across media. The diary can invite sustained involvement through its episodic structure. Film adaptations can sustain involvement through pacing, visual framing, and emotional tension. The museum can curate involvement through spatial dramaturgy and interpretive prompts that guide visitors' attention and reflection. VR can foster involvement by coupling narrative progression to user actions, such as moving through rooms, examining objects, or triggering audio excerpts from the diary. Importantly, the depth of involvement seems to be not determined by the medium but by how well narrative, structure, and task demands are aligned with learners' capacities and interests.

Conceptualizing these four dimensions as cross-modal reveals that the psychological architecture of immersive experience remains structurally consistent across VR and non-VR contexts. Physical presence, social presence, self-presence, and involvement can all be

cultivated through carefully designed experiences across media. VR does therefore not introduce fundamentally new ontological categories of immersive experience. However, it can intensify, synchronize, and technologically scaffold these dimensions in ways that may be difficult to replicate in non-VR settings. In particular, the tight coupling of sensorimotor contingencies, spatial simulation, and real-time interactivity can produce heightened forms of embodied attunement and environmental coherence (e.g., Kilteni et al., 2012; Riva, 2008). Research in embodied cognition and virtual embodiment suggests that sensorimotor coupling and virtual body ownership can intensify self-presence and action-based learning, particularly in procedural and spatial domains (e.g., Banakou, Groten & Slater, 2013; Kilteni et al., 2012). In such contexts, VR may amplify experiential dimensions that are only indirectly accessible through text or film. At the same time, VR introduces constraints that analogue and physical modalities do not share. Technological mediation may increase extraneous cognitive load due to high perceptual richness and interactional demands (e.g., Makransky, et al., 2019). Moreover, the heightened realism and embodied simulation of VR environments can raise ethical concerns, particularly when historical trauma or vulnerable identities are represented (e.g., Madary & Metzinger, 2016). Unlike textual or museum-based encounters, immersive enclosure may also reduce opportunities for reflective distancing, which some scholars argue is pedagogically important in emotionally charged contexts (e.g., Markowitz & Bailenson, 2021). Modality independence therefore does not imply experiential equivalence. Rather, different realization domains enable distinct configurations, intensities, and risks of immersive experience within a shared structural architecture.

Immersion as a Modality-Independent Experiential Mode

Synthesizing all components of Lee's framework reveals that immersive experience is not medium-specific. Whether learners read Anne Frank's diary, view film adaptations, visit the museum, or explore a VR reconstruction, immersive experience can arise through different constellations of the same underlying determinants. Importantly, this structural logic is not confined to memorial or historical pedagogy. Comparable realization constellations can be observed in other domains such as emergency-response training (e.g., Sharma, 2020). Here, immersive engagement may be cultivated through written manuals outlining procedural conduct, film-based tutorials demonstrating correct actions, live training exercises with actors simulating crisis situations, or VR-based simulations that recreate hazardous environments. Across these modalities, the same determinants are operative: subjective preparedness shapes plausibility judgments, system and content properties determine perceived realism and relevance, sensorimotor and task/motor engagement vary in intensity depending on the

medium, and different experiential dimensions (e.g., self-presence in procedural enactment or involvement under time pressure) are differentially foregrounded. The realization layer changes, but the structural architecture of immersive experience seems to remain constant.

Taken together, these analyses demonstrate that immersive experience seems to be structurally invariant across realization domains. What changes is not the architecture of immersion but the medium-specific configuration through which its determinants are instantiated. This insight provides the conceptual basis for examining how VR can be pedagogically designed to selectively amplify or constrain specific experiential dimensions.

To clarify how Lee's determinants manifest across realization layers, Table 1 provides a schematic comparison of how the same structural components are instantiated in physical sites, text-based narratives, film and VR (along the selected Anne Frank case).

	VR Simulation	**Physical Site** (e.g., Museum)	**Text-Based Narrative**	**Film**
Subjective Factors (internal reference, skills, preference)	prior VR familiarity, domain knowledge, openness to immersive technology	historical knowledge, affective readiness, spatial familiarity	reading competence, narrative literacy	media literacy, genre expectations, emotional disposition
Plausibility	spatial fidelity, sensorimotor coherence, audiovisual realism	material authenticity, archival integrity, spatial credibility	perceived historical authenticity, narrative credibility	cinematic realism, coherence of adaptation
Interactivity	real-time navigation, object manipulation, avatar embodiment	physical navigation, spatial orientation, guided interaction	page turning, interpretive imagination	head and eye movements
Interestingness	multisensory novelty, situational immersion	artefacts, testimonies, curated dramaturgy	first-person voice, emotional immediacy	audiovisual dramatization, pacing, sound design

Sensorimotor Engagement	head tracking, spatial audio, embodied movement	bodily presence in space, acoustic atmosphere	imagined spatial simulation	visual and auditory immersion
Narrative Engagement	spatialized narrative progression	curated narrative pathways	episodic diary structure	scripted plot progression
Task/Motor Engagement	object interaction, procedural actions	navigating rooms, following guided prompts	cognitive interpretation	head and eye movements
Physical Presence	simulated spatial enclosure	actual spatial enclosure	imaginative spatial presence	visual scene immersion
Social Presence	virtual agents, reconstructed characters	other visitors, testimonies	narrated voice as interlocutor	characters on screen
Self-Presence	avatar embodiment, perspective alignment	reflective identification	perspective-taking through narrative	emotional identification
Involvement	action-driven narrative coupling	spatial dramaturgy	sustained narrative absorption	emotional pacing

Table 1. Cross-Modal Realization of Lee's (2025) Immersive Experience Determinants

Pedagogical Affordances and the Conditions for Immersive Learning in VR

Building on the cross-modal invariance established above, VR can be conceptualized as a configurable realization domain in which the determinants of immersive experience can be deliberately orchestrated. A pedagogically grounded account of VR must therefore ask not whether a simulation is immersive in a technological sense, but how its design supports the psychological processes that underlie meaningful presence and involvement, especially in situations where real-world immersion is unavailable.

The outermost component of Lee's model points to subjective factors as the starting point for design. In VR, these factors cannot be fully controlled, but they can be actively supported. Internal reference, skills and knowledge, and personal preferences shape how learners recognize plausibility, manage interaction, and engage with content. Instructional design can strengthen these subjective resources before the headset is even put on. In the context of the

Anne Frank example, historical briefings, introductory narratives, or exposure to primary sources can align expectations and provide interpretive frames that render the subsequent VR experience more coherent. Similarly, pre-training on basic interaction mechanics in a neutral VR environment can prevent learners from expending scarce cognitive resources on navigating menus or mastering locomotion once they are inside a demanding, emotionally charged scenario. In memorial contexts, affective framing also becomes essential. Signalling the seriousness of the experience, delineating its purpose, and articulating ethical boundaries all prepare learners to enter the simulation with an appropriate stance. VR design must therefore be embedded in an instructional sequence that prepares subjective readiness rather than presuming it.

The second model component, concerning system and content properties, identifies plausibility, interactivity, and interestingness as central levers. In VR, these levers are often treated as purely technological, with design efforts focused on pushing fidelity and interactivity to their maximum. From a pedagogical perspective, however, the question is not how much plausibility and interactivity can be achieved, but which forms and degrees are functional for the learning objectives. Plausibility should be pursued to the extent that it supports sensorimotor and narrative coherence without obscuring the constructed nature of the simulation. In a VR memorial, this means privileging historically warranted spatial layouts, artefacts, and soundscapes over spectacular but anachronistic dramatization. Interactivity likewise needs to be meaningful rather than abundant. Interactions that are tightly coupled to the conceptual or ethical focus of the scenario (e.g., opening the hidden door to the annex) can deepen task engagement, whereas gratuitous manipulation of objects risks shifting attention towards the interface itself. Interestingness, finally, should be anchored in the intrinsic power of the content rather than in technological novelty. In contexts where direct access to authentic sites or artefacts is impossible, situational interest may initially be triggered by the unusual experience of being inside a virtual environment, but sustained involvement depends on the quality of the narrative, the relevance of the tasks, and the perceived significance of the subject matter. VR affords fine-grained control over these parameters. Pedagogical design determines whether this control is used to support or distract from learning.

At the level of experiential preconditions, VR is particularly well placed to orchestrate sensorimotor, narrative, and task/motor engagement in a coordinated way. Sensorimotor engagement can be fostered through carefully tuned spatial audio, lighting, and environmental responsiveness that invite embodied attunement without overwhelming the senses. In a virtual reconstruction of the annex, the creaking of floorboards when the user shifts weight, the

constrained field of view in narrow staircases, or the muffled ambience of the hidden rooms can scaffold a sense of bodily situatedness that prepares the ground for physical presence. Narrative engagement can be strengthened by aligning spatial progression with the unfolding of a story. Moving from room to room can correspond to temporal progress in the historical events, with diary excerpts, photographs, or testimonies triggered at key spatial thresholds. This spatial-narrative coupling helps learners organize the experience into a meaningful whole rather than a sequence of loosely related scenes. Task/motor engagement should then be designed to reinforce these narrative and sensorimotor structures rather than compete with them. Interactions such as opening a cupboard to reveal confiscated belongings can make the narrative tangible, whereas complex game-like mechanics with scores, timers, or unrelated challenges may fragment attention and compromise immersion. VR's pedagogical affordance lies in its ability to synchronize these three forms of engagement. Effective design ensures that embodying, understanding, and acting in the environment all point in the same conceptual and ethical direction.

The inner core of the model (i.e., physical presence, social presence, self-presence, and involvement) can thus be seen as the outcome of a series of design decisions rather than as properties of the device. For educational purposes, maximizing all four dimensions equally is neither feasible nor pedagogically warranted. Instead, their relative weighting should be aligned with the instructional intention. In a historically relevant VR scenario, designers might prioritize physical presence and involvement, ensuring that learners feel located within the environment and absorbed in its narrative, while treating social and self-presence with particular ethical care. Introducing highly realistic avatars of historical victims or perpetrators, or simulating the learner's perspective as that of a specific historical person, may increase social and self-presence but also raises questions about appropriation, identification, and the representation of suffering. In skill-based , for example emergency-response, training, by contrast, self-presence and task-related involvement may be foregrounded. Learners need to experience themselves as competent agents within the simulated task, with physical presence serving as a backdrop for procedural rehearsal. Across these different scenarios, the pedagogical affordance of VR lies in the possibility to configure which dimensions of immersive experience are cultivated, to what degree, and for which learners.

Although the focus of this section has been on VR, the underlying logic is not confined to head-mounted displays. Museums, textual narratives, and cinematic representations can likewise be designed to support subjective readiness, to shape system and content properties, and to orchestrate experiential preconditions in ways that foster presence and involvement.

However, VR becomes particularly valuable when comparable real-world experiences are inaccessible, because the site no longer exists, access is restricted, or the situation would be dangerous or ethically unacceptable to stage. In such cases, Lee's framework offers more than a diagnostic tool for describing immersion. It provides a design heuristic for educational VR usage. By asking, how subjective factors will be prepared, which system and content properties should be emphasized, how sensorimotor, narrative and task engagement will be coordinated, and which experiential dimensions are to be foregrounded, designers can move beyond the assumption that immersion is an automatic by-product of advanced technology. Instead, VR is positioned as a pedagogical medium whose affordances must be intentionally aligned with instructional aims in order to yield immersive experiences that are not only compelling, but also ethically responsible and educationally meaningful.

If Lee's model is extended across realization layers, a clear empirical implication follows: immersive experience should exhibit structural continuity alongside configurational variation. In other words, if the underlying experiential architecture remains constant while its material instantiation differs (e.g., VR simulation, museum site, analogue narrative), systematic cross-modal patterns should be observable. The framework advanced in this article therefore yields a set of empirically testable propositions:

(1) Structural Invariance Hypothesis: Across different realization layers, immersive experience will be predicted by the same core experiential dimensions (i.e., physical presence, social presence, self-presence, and involvement) even if their relative intensities vary.

(2) Dimensional Weighting Hypothesis: Distinct realization layers will systematically privilege particular experiential dimensions (e.g., VR emphasizing physical and self-presence while text-based narratives emphasizing involvement, and physical sites emphasizing physical and social presence.), resulting in differentiated immersion profiles rather than uniform immersive states.

(3) Amplification Hypothesis: VR-based realizations will not introduce fundamentally new experiential dimensions but may produce stronger synchrony between sensorimotor engagement and narrative engagement compared to non-digital realizations.

(4) Ethical Boundary Condition Hypothesis: In historically sensitive simulations, increases in technologically mediated self- and social-presence (e.g., avatar embodiment of historical figures) will differentially influence reflective depth and

moral evaluation compared to lower-embodiment realizations, thereby moderating the educational effects of immersion.

Together, these propositions enable systematic cross-modal research capable of determining whether immersive technologies primarily amplify shared experiential determinants or produce qualitatively distinct experiential configurations under specific boundary conditions.

Discussion

Theoretical Contribution and Implications

As a conceptual contribution, the argument developed throughout this article can be summarized in three central propositions.

First, the article offers conceptual clarification by disentangling immersive system parameters from immersive experience parameters and situating both within a relational architecture. By distinguishing subjective factors, system properties, engagement processes, and experiential dimensions, the article aims to reduce definitional ambiguity that has long characterized immersion research.

Second, the article extends Lee's model through the introduction of realization layers, demonstrating that immersive experience can be instantiated not only in VR but equally in physical sites, analogue media, and narrative formats. This extension reframes immersion beyond hardware-dependent interpretations by conceptualizing it as a modality-independent experiential architecture. While the underlying experiential structure remains invariant, different media configurations instantiate and weight its dimensions in distinct ways.

Third, the article provides a pedagogically grounded design heuristic for VR. Rather than assuming that immersion follows from VR`s technological fidelity, the article specifies how subjective preparation, system design, and engagement alignment should be orchestrated to cultivate meaningful presence and involvement, particularly in contexts where real-world immersion is inaccessible, unsafe, or ethically constrained.

The three contributions outlined above are inherently interdisciplinary in scope. They speak to media psychology by clarifying the structure of immersive experience, to human–computer interaction by conceptualizing how technological systems shape experiential engagement, to media and communication studies by situating immersion across different representational formats, and to educational technology and museum pedagogy by specifying the instructional and ethical conditions under which immersive experiences can support learning.

The article further generates integrative design implications that apply across realization domains. Designing immersive learning, whether in VR, physical environments, narrative

media, or hybrid formats, requires coordinated alignment across four structural layers: 1. Subjective readiness must be cultivated through the activation of prior knowledge, the alignment of expectations, and, where relevant, ethical framing. 2. System and content properties such as plausibility, interactivity, and interestingness need to be calibrated in relation to clearly defined instructional objectives rather than maximized indiscriminately. 3. Sensorimotor, narrative, and task/motor engagement should be orchestrated so that they reinforce one another instead of fragmenting attention. 4. In addition, specific experiential dimensions ought to be selectively foregrounded in accordance with pedagogical intention, recognizing that not all dimensions require equal emphasis in every instructional context. Importantly, these layers seems to be interdependent. Amplification at one level, such as intensified embodiment or heightened narrative dramatization, may generate trade-offs at another, including increased cognitive load, reduced interpretive openness, or diminished ethical sensitivity. Effective immersive design is therefore not a matter of maximizing immersion per se, but of structurally configuring its determinants in a goal-consistent and context-sensitive manner.

Furthermore, the notion of realization layers also carries implications for accessibility and inclusivity. If immersive experience is not bound to a single technological medium but can emerge across different realization domains, educators can intentionally select modalities that accommodate diverse learner needs and contextual constraints. For instance, learners with visual impairments may engage more strongly through narrative or auditory immersion than through visually dominant VR environments, while geographically distant learners may access otherwise unreachable historical sites through virtual reconstructions. In this sense, the cross-modal architecture of immersion expands the range of pedagogical entry points through which learners can encounter immersive experiences.

Limitations

Several limitations should be acknowledged. *First*, the argument is conceptual and illustrative rather than empirical. The framework proposed in this article therefore represents a theoretical synthesis rather than an empirically validated model. Although testable propositions have been formulated, empirical investigations are required to examine structural invariance, dimensional weighting, and amplification effects across realization domains using controlled comparative designs. In particular, the present article does not include direct experimental comparisons between VR systems and other realization domains, which would be necessary to empirically evaluate the proposed relationships.

Second, the proposed realization-layer extension simplifies complex sociocultural and contextual influences that may shape immersive experience beyond the structural determinants outlined here. Institutional framing and learner diversity may interact with immersion in ways that exceed the current model's scope. Moreover, the concept of immersion itself is defined differently across disciplines such as media psychology, human-computer interaction, and educational technology research. The present framework adopts Lee's model as a reference structure, but alternative conceptualizations may emphasize additional experiential dimensions.

Third, the article does not resolve ongoing debates in embodied cognition and VR research regarding potential qualitative shifts in phenomenological experience. While the present account interprets VR effects primarily as reconfigurations or amplifications of shared experiential dimensions, it remains possible that certain forms of technologically mediated embodiment introduce emergent phenomenological qualities under specific conditions.

Future research should therefore pursue multi-method approaches combining experimental manipulation, psychometric measurement, and phenomenological inquiry to investigate whether immersive technologies operate through structural intensification, dimensional redistribution, or genuinely novel experiential patterns. Empirical validation of the proposed framework across different realization domains therefore represents a central avenue for future research. In particular, comparative studies across realization domains would allow researchers to test whether immersive experience is best understood as invariant in architecture but variable in configuration.

Conclusion

Taken together, the present analysis conceptualizes immersion as a modality-independent yet configurationally variable experiential architecture. Rather than equating immersion with technological sophistication, it proposes structural alignment between subjective factors, system properties, engagement processes, and experiential dimensions as the decisive criterion for educational value. In doing so, the article provides a theoretically grounded framework for evaluating and designing immersive learning environments across realization layers.

References

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, *110*(31), 12846-12851. https://doi.org/10.1073/pnas.130677911

Barrett, A. J., Pack, A., & Quaid, E. D. (2021). Understanding learners' acceptance of high-immersion virtual reality systems: Insights from confirmatory and exploratory PLS-SEM

analyses. *Computers & Education*, *169*, 104214. https://doi.org/10.1016/j.compedu.2021.104214

Berkman, M. I., & Akan, E. (2024). Presence and immersion in virtual reality. In N. Lee (Ed.), *Encyclopedia of computer graphics and games* (pp. 1461–1470). Springer International Publishing. https://doi.org/10.1007/978-3-031-23161-2_162

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of computer-mediated communication*, *3*(2), JCMC324. https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

Boffi, P., Clerici, M., Gallace, A., & Lanzi, P. L. (2023). An educational experience in ancient Rome to evaluate the impact of virtual reality on human learning processes. *Computers & Education: X Reality*, *2*, 100014. https://doi.org/10.1016/j.cexr.2023.100014

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough?. *Computer, 40*(7), 36-43. https://doi.org/10.1109/MC.2007.257

Buragohain, D., Meng, Y., & Chaudhary, S. (2025). Metaverse impact and trends in sustainable integration of immersive technologies in education. *Computer Applications in Engineering Education*, *33*(3), e70024. https://doi.org/10.1002/cae.70024

Clark, R. E. (1994). Media will never influence learning. *Educational Technology: Research & Development 42*(2), 21–9. https://doi.org/10.1007/BF02299088

Calleja, G. (2007). Revising immersion: A conceptual model for the analysis of digital game involvement. *3rd digital games research association international conference: "Situated play", DiGRA 2007,* 83-90.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740

Derry, K. (2021). Where Is Anne Frank? *Journal of Religion and Film 25*(2). https://doi.org/10.32873/uno.dc.jrf.25.02.016.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, *127*, 252-266. https://doi.org/10.1016/j.compedu.2018.09.002

Graver, L. (2020). *An Obsession with Anne Frank: Meyer Levin and the Diary*. University of California Press. https://doi.org/10.1525/9780520313231

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of personality and social psychology*, *79*(5), 701.

Hartmann, R. (2019). Virtualities in the new tourism landscape: The case of the Anne Frank House Virtual Tour and of the visualizations of the Berlin Wall in the Cold War context. In M. Gravari-Barbas, N. Graburn & J. F. Staszak (Eds.), *Tourism Fictions, Simulacra and Virtualities* (pp. 211-223). London: Routledge.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, *66*(9), 641-661. https://doi.org/10.1016/j.ijhcs.2008.04.004

Kampling, H. (2018). The role of immersive virtual reality in individual learning. *Hawaii International Conference on System Sciences.* https://doi.org/10. 24251/HICSS.2018.173

blinded

blinded

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373-387. https://doi.org/10.1162/PRES_a_00124

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology: Research & Development, 42,* 7–19. https://doi.org/10.1007/BF02299087

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning and development. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role ofinterest in learning and development* (pp. 3–25). Hillsdale, NJ: Erlbaum.

Lee, H. (2025). A conceptual model of immersive experience in extended reality. *Computers in Human Behavior Reports*, 100663. https://doi.org/10.1016/j.chbr.2025.100663

Lichtman, S. A. (2021). Designing Absence at the Anne Frank House Museum, Amsterdam, and the Secret Annex Online: Exhibition design, virtual reality, and historic preservation. In G. Lees Maffai & R. House (Eds.), *Design and Heritage. The Construction and Identity of Belonging.* London: Routledge.

Madary, M., & Metzinger, T. K. (2016). Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Frontiers in Robotics and AI*, *3*, 180932. https://doi.org/10.3389/frobt.2016.00003

Makransky, G. (2021). The immersion principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 296–303). New York: Cambridge University Press.

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. https://doi.org/10.1007/s10648-020-09586-2

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. https://doi.org/10.1016/j.learninstruc.2017.12.007

Markowitz, D. M., & Bailenson, J. (2021). Virtual reality and emotion: A 5-Year systematic review of empirical research (2015-2019). PsyArXiv. https://doi.org/ 10.31234/osf.io/tpsm

blinded

blinded

blinded

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, *2*(1), 486-497. https://doi.org/10.3390/encyclopedia2010031

Nilsson, N., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human technology*, *12*(2), 108-134. https://doi.org/10.17011/ht/urn.201611174652

O'Regan, J. K., & Noë, A. (2001). A otor account of vision and visual consciousness. *Behavioral and brain sciences*, *24*(5), 939-973. https://doi.org/10.1017/S0140525X01000115

Packer, J., & Ballantyne, R. (2016). Conceptualizing the visitor experience: A review of literature and development of a multifaceted model. *Visitor Studies*, *19*(2), 128-143. https://doi.org/10.1080/10645578.2016.1144023

Perry, G. W. (2019). *The Legacy of Anne Frank*. Grub Street Publishers.

Riva, G. (2008). From virtual to real body: virtual reality as embodied technology. *J. Cyber Ther. Rehabil. 1,* 7–22.

Ryan, M. (2003). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media.* London: Johns Hopkins University Press.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, *10*(3), 266-281. https://doi.org/10.1162/105474601300343603

Schulz, F. (2019). Figures of Memory: Das Tagebuch der Anne Frank zwischen Text und Bild, Wort und Symbol. *Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung,* 51–62. https://doi.org/10.21248/gkjf-jb.37

Sharma, S. (2020). Improving emergency response training and decision making using a collaborative virtual reality environment for building evacuation. In *International Conference on Human-Computer Interaction* (pp. 213-224). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-59990-4_17

Slater, M. (2003). A note on presence terminology. *Presence connect*, *3*(3), 1-5.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549-3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Southgate, E., Buchanan, R., Cividino, C., Saxby, S., Eather, G., Smith, S. P., Bergin, C., Kilham, J., Summerville, D. & Scevak, J. (2018). Highly immersive virtual reality. *Scan: The Journal for Educators*, *37*, 130-143. https://doi.org/10.3316/informit.243743601845101

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225-240. https://doi.org/10.1162/105474698565686

Zender, R., Knoth, A. H., Fischer, M. H., & Lucke, U. (2019). Potentials of virtual reality as an instrument for research and education. *i-com*, *18*(1), 3-15. https://doi.org/10.1515/icom-2018-0042

Highlights

- Reconceptualizes immersion as a modality-independent construct.
- Extends Lee's (2025) model across realization layers.
- Reframes VR as pedagogically conditional, not inherently immersive.
- Derives cross-modal, empirically testable hypotheses.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Credit authorship contribution statement

Miriam Mulders: Conceptualization, Writing original draft, review & editing.

Data availability

No empirical data were generated or analyzed for the research presented in this article.

Declaration of interests

☒ 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.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: