https://doi.org/10.65166/ss92y378

Journal of Frontier Science and Futures
Volume I, Issue 1, July 2026
https://journal.ijhba.com/index.php/jfsf
https://slsipress.com

Presence, Cognitive Load, and Learning in Extended Reality: A Thematic Analytical Review

Ramon George O. Atento, PhD[1]; Leah F. Quinto, PhD[2]; Andrea Gwyneth B. Atento[3]; Victor Jolo Maranan[4]; Charmaine O. Castañeda, MS Bio[5]

[1]First Asia Institute of Technology and the Humanities
[2]De La Salle Medical and Health Sciences Institute
[3]SLSI Press
[4]CEU-Manila Graduate School
[5]Adamson University-QC; FEU-Manila

Abstract

Extended reality (XR) is increasingly used in education and training, yet claims that greater immersion necessarily improves learning remain theoretically and empirically unsettled. This thematic analytical review examines how presence, cognitive load, and learning outcomes are conceptualized, measured, and related across virtual, augmented, and mixed reality research. English-language literature published from 1990 to 2026 was identified through Scopus, Web of Science, IEEE Xplore, PsycINFO, and the ACM Digital Library, supplemented by backward and forward citation tracking. The literature was synthesized across five themes: presence conceptualization and measurement; cognitive load in XR; empirical relationships among presence, load, and learning; the immersion assumption; and moderating effects of task, learner, and instructional design characteristics. The evidence does not support a uniform relationship between technological immersion, presence, and learning. Presence appears most educationally useful when it is task-congruent, particularly in spatial, procedural, and experiential learning, and when usability, guidance, cueing, and segmentation limit extraneous processing. Conversely, complex interfaces, irrelevant realism, unguided exploration, and poor alignment between XR affordances and learning objectives may increase cognitive burden and weaken learning outcomes. Measurement inconsistency and the frequent conflation of immersion with presence continue to limit cumulative inference. The review proposes a contingent framework in which XR effectiveness depends on task–affordance alignment, learner characteristics, and instructional design rather than immersion alone. These findings support more selective, evidence-based XR adoption and more rigorous, multidimensional evaluation of presence, cognitive load, retention, and transfer.

Keywords: *extended reality; presence; cognitive load theory; immersive learning; instructional design; learning outcomes*

Article History: Received: March 21, 2026 Revised April 26, 2026 Accepted: June 16, 2026

1. Introduction

Extended reality (XR) has become an increasingly prominent component of educational technology, professional training, and simulation-based learning. The term encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), each of which combines physical and digitally generated environments in different ways. XR systems allow learners to interact with three-dimensional representations, rehearse procedures in controlled settings, explore environments that may otherwise be inaccessible, and visualize spatial or dynamic processes that are difficult to communicate through conventional instructional media. These affordances have encouraged the adoption of XR in fields such as medicine, engineering, science education, vocational training, and workplace learning (Crogman et al., 2025; Hamilton et al., 2020; Tusher et al., 2024).

Much of the educational interest in XR is associated with its capacity to create a sense of presence. Presence refers to the subjective experience of being located within or psychologically connected to a mediated environment, even when the learner is aware that the environment is technologically generated (Grassini & Laumann, 2020; Slater, 1999). Presence is related to, but conceptually distinct from, immersion. Immersion concerns the objective technological properties of a system, including display fidelity, field of view, sensory richness, tracking accuracy, and interactivity. Presence, by contrast, describes the psychological response produced through the interaction of system characteristics, task demands, individual differences, and contextual factors (Johnson-Glenberg, 2018; Kelly et al., 2022).

This distinction is important because educational discussions frequently treat immersion and presence as interchangeable. Highly immersive systems are often assumed to produce stronger presence, and stronger presence is subsequently assumed to improve engagement, motivation, retention, and learning. This reasoning has contributed to what may be described as the **immersion assumption**: the proposition that increasingly realistic or technologically immersive environments are inherently more effective for learning. The assumption appears not only in promotional discourse but also in research designs that compare head-mounted VR systems with desktop, video, or text-based conditions without adequately controlling for instructional design, novelty, exposure time, or content quality (Lawson et al., 2024; Mayer & Bailenson, 2024; Mulders, 2023).

The empirical evidence, however, does not support a uniformly positive relationship among immersion, presence, and learning. Some studies have reported that immersive XR improves spatial understanding, procedural performance, engagement, or knowledge acquisition, particularly when the learning objective requires interaction with three-dimensional, dynamic, or experiential information (Johnson et al., 2022; Makransky & Mayer, 2022; Pedram et al., 2024). Other studies have found no meaningful learning advantage over less immersive alternatives, even when participants report substantially higher presence. In some cases, immersive conditions have produced weaker learning outcomes, suggesting that technological richness and subjective involvement do not automatically translate into effective cognitive processing (Johnsdorf et al., 2023; Ochs & Sonderegger, 2022; Poupard et al., 2024).

Cognitive load theory offers a useful framework for understanding these inconsistent findings. The theory maintains that learning is constrained by the limited capacity of working memory and that instructional environments must manage the mental demands placed on learners. Cognitive load is commonly differentiated into intrinsic load, which arises from the complexity and interactivity of the material; extraneous load, which results from inefficient presentation or unnecessary processing demands; and germane processing, which concerns the mental effort directed toward meaningful learning and schema construction (Krieglstein et al., 2023; Skulmowski & Xu, 2021).

XR may influence each of these forms of cognitive processing. When an immersive representation corresponds closely with the learning task, it may reduce the need for mental transformation and allow learners to perceive spatial, procedural, or dynamic relations more directly. For example, an interactive anatomical model may make structural relationships easier to understand than a static two-dimensional image. Similarly, a simulated procedure may allow learners to connect conceptual knowledge with embodied action. Under such conditions, XR may support the construction of task-relevant mental models and direct attention toward meaningful learning processes.

The same features can nevertheless become cognitively counterproductive when they are poorly aligned with the instructional objective. Highly detailed environments may contain irrelevant visual information; complex controllers may require attention that would otherwise be allocated to the learning task; navigational freedom may lead learners away from essential content; and excessive interactivity may increase physical, attentional, or procedural demands. Cybersickness, unfamiliar interfaces, and technological interruptions may further consume cognitive resources. Consequently, XR can increase extraneous load even when it produces high levels of presence (Breves & Stein, 2022; Weech et al., 2019).

The relationship between presence and cognitive load is therefore unlikely to be linear. Presence may support learning when it directs attention toward task-relevant information and facilitates the cognitive operations required by the learning objective. It may hinder learning when it is produced through features that attract attention without contributing to understanding. This distinction suggests that the educational value of presence depends on whether it is **task-congruent**. Task-congruent presence arises when the learner's sense of being situated within the environment directly supports perception, action, interpretation, or practice. Task-incongruent presence arises when realism, interactivity, or sensory stimulation increases involvement but does not contribute to the target knowledge or skill.

Several moderating factors further complicate this relationship. Task characteristics are particularly important. Spatial, procedural, and experiential tasks may benefit from the representational affordances of XR, whereas abstract, symbolic, or text-intensive learning may be more efficiently supported through less immersive media (Baceviciute et al., 2021; Yu et al., 2025). Learner characteristics also matter. Prior knowledge, spatial ability, familiarity with digital interfaces, gaming experience, age, and susceptibility to motion sickness may influence both presence and cognitive load (Brucker et al., 2024; Han et al., 2023; Thorp et al., 2024). Novices may benefit from concrete visualization but may also be more vulnerable to interface complexity and incidental detail. More experienced learners may manage the technological environment more efficiently but may derive less benefit from representations that duplicate knowledge structures they already possess.

Instructional design may be the most immediately controllable moderator. Evidence increasingly indicates that guidance, cueing, segmentation, scaffolding, and opportunities for reflection improve the effectiveness of XR learning environments (Albus & Seufert, 2022; Mulders et al., 2020; Weerasinghe et al., 2022). Unguided exploration may generate interest and presence while producing limited learning because learners do not necessarily attend to, interpret, or organize the information required by the instructional objective. Effective XR design must therefore do more than reproduce an environment realistically. It must structure attention, sequence complexity, simplify interaction, and connect immersive activity with explicit learning goals.

Despite the expanding literature, the field remains conceptually and methodologically fragmented. Presence is variously treated as a unitary state or as a multidimensional construct involving spatial presence, involvement, realism, and related experiences. Cognitive load is frequently measured through global workload instruments that do not distinguish among theoretically different sources of load. Many studies use technological immersion as a proxy for presence without measuring subjective experience directly. Others compare media conditions that differ not only in immersion but also in content, interactivity, novelty, and instructional structure. These variations limit cross-study comparison and make it difficult to determine whether learning outcomes are attributable to presence, system characteristics, cognitive load, instructional design, or combinations of these factors (Grassini & Laumann, 2020; Lawson et al., 2024; Mayer & Bailenson, 2024).

Current research also remains disproportionately concentrated on immediate post-intervention outcomes. Retention, transfer, repeated use, and the decline of novelty effects are less frequently examined. Moreover, VR has received substantially more attention than AR and MR, despite differences in how these technologies integrate digital information with the physical environment. The field therefore requires an integrative account that moves beyond the question of whether XR improves learning and instead examines the conditions under which presence and immersive affordances support, burden, or redirect cognitive processing.

The present thematic analytical review addresses this need by synthesizing conceptual, empirical, methodological, and critical literature concerning presence, cognitive load, and learning in XR environments. It treats XR as an instructional medium whose effectiveness depends on task–affordance alignment, learner characteristics, technological usability, and pedagogical design. It also critically examines the immersion assumption and develops a contingent perspective in which presence is neither inherently beneficial nor inherently detrimental. Its value depends on the functions it performs within a particular learning environment.

Objectives of the Study

This review aimed to critically examine the relationships among presence, cognitive load, and learning in extended reality systems. Specifically, it sought to:

1. analyze how presence and cognitive load have been conceptualized, operationalized, and measured in XR learning research;
2. synthesize empirical findings concerning the relationships among immersion, presence, cognitive load, and learning outcomes;
3. examine the extent to which the immersion assumption is supported or contradicted by available evidence;
4. identify task, learner, technological, and instructional factors that moderate the effectiveness of XR learning environments;

5. determine the major conceptual, methodological, and empirical limitations affecting the cumulative interpretation of XR research; and
6. propose a contingent framework explaining the conditions under which presence may support or impede learning through its interaction with cognitive load.

2. Literature Review

2.1 Presence and Immersion in Extended Reality

Presence is one of the principal psychological constructs used to explain how individuals experience extended reality environments. It is commonly defined as the subjective sense of being situated within a mediated environment despite awareness that the experience is technologically generated. Foundational approaches distinguish presence from immersion: immersion refers to the objective technological characteristics of a system, whereas presence refers to the user's subjective psychological response to those characteristics (Slater, 1999; Schuemie et al., 2001).

This distinction has important implications for XR learning research. Systems with wide fields of view, high-resolution displays, responsive tracking, spatial audio, and natural interaction may increase the likelihood of presence, but they do not guarantee it. The experience of presence is also shaped by usability, task engagement, prior experience, individual differences, and the coherence of the simulated environment (Kelly et al., 2022; Maneuvrier et al., 2020). Technological immersion should therefore be treated as an enabling condition rather than a direct equivalent of presence.

Presence has also been conceptualized as multidimensional. Common dimensions include spatial presence, involvement, and experienced realism. Spatial presence concerns the perceived sense of being physically located within the environment; involvement concerns the degree of attention devoted to the mediated experience; and realism refers to the perceived plausibility or authenticity of the environment (Berkman & Çatak, 2021; Grassini & Laumann, 2020). These dimensions may not contribute equally to learning. Spatial presence may assist tasks requiring navigation, mental rotation, or embodied action, whereas realism may introduce visually detailed but instructionally irrelevant information.

Measurement practices reflect this conceptual diversity. Self-report instruments such as the Igroup Presence Questionnaire, the Presence Questionnaire, and the Slater–Usoh–Steed scale remain widely used because they are accessible and comparatively easy to administer (Grassini & Laumann, 2020). However, retrospective self-reports may be vulnerable to memory distortion, demand characteristics, and post-experience rationalization. They also provide limited information about fluctuations in presence during different stages of a task.

Behavioral and physiological measures have therefore been proposed as complementary approaches. Head movement, reaction time, interruptions in task performance, heart-rate variability, electrodermal activity, and other physiological indicators may provide continuous or near-real-time information (Chandio et al., 2023; Khenak et al., 2018, 2020). Nevertheless, such indicators are not specific to presence. Physiological arousal may reflect emotional response, cognitive effort, novelty, anxiety, or cybersickness rather than presence alone (Kerrebroeck et al., 2021; Martinez et al., 2020).

A recurring problem is the use of system type as a proxy for presence. Comparisons between head-mounted displays and desktop environments often assume that the former automatically represents a high-presence condition. Such designs obscure whether observed effects arise from subjective presence, technological characteristics, interaction demands, visual fidelity, or media novelty (Johnson-Glenberg, 2018; Lawson et al., 2024). Presence must therefore be measured directly and interpreted separately from immersion.

2.2 Cognitive Load in XR Learning Environments

Cognitive load theory explains learning in relation to the limited processing capacity of working memory. It distinguishes among intrinsic load, extraneous load, and germane processing. Intrinsic load is associated with the inherent complexity and interactivity of the learning material. Extraneous load results from instructional features that require mental effort without contributing directly to learning. Germane processing refers to cognitive activity directed toward organizing, integrating, and constructing usable knowledge structures (Krieglstein et al., 2023; Skulmowski & Xu, 2021).

XR can alter these forms of cognitive demand in different ways. Its representational affordances may reduce the need for learners to mentally translate two-dimensional or symbolic information into spatial or procedural models. Three-dimensional anatomical representations, simulations of machinery, and embodied procedural training may allow learners to perceive relationships more directly. In such cases, XR may make complex content more manageable by aligning the presentation format with the perceptual and motor requirements of the task (Kenneally et al., 2026; Zou et al., 2025).

At the same time, immersive environments can impose substantial extraneous load. Learners may need to manage unfamiliar controllers, navigation systems, menus, visual cues, or physical movements while simultaneously processing instructional content. Rich environmental detail may compete for attention, and high interactivity may require decision-making unrelated to the learning objective. Cybersickness and visual discomfort can further disrupt attention and working-memory resources (Breves & Stein, 2022; Weech et al., 2019).

The instructional value of XR therefore depends partly on whether its cognitive demands are relevant to the task. An interaction that supports procedural rehearsal may contribute to meaningful processing, while the same interaction may become extraneous if the objective is the acquisition of factual or symbolic knowledge. The distinction cannot be determined solely by the technological feature itself; it depends on how that feature functions within the instructional activity.

Cognitive load is most frequently measured using post-task self-report instruments, including the NASA Task Load Index and single-item mental-effort scales. Although these instruments provide practical measures of perceived workload, they often combine mental, physical, temporal, emotional, and effort-related demands into a single score (Brachten et al., 2020; Vidal-Balea et al., 2024). Such aggregation is problematic for studies attempting to explain whether XR influences intrinsic, extraneous, or germane processing.

More differentiated instruments have been developed to distinguish sources of cognitive load, including multidimensional scales aligned with cognitive load theory (Andersen & Makransky, 2021; Krieglstein et al., 2023). These measures offer stronger theoretical correspondence but remain less consistently applied in XR studies. Physiological and behavioral measures, such as pupil dilation, electroencephalography, reaction time, and interaction traces, may provide additional temporal resolution, although their interpretation is complicated by emotional arousal, motor activity, visual stimulation, and individual variability (Ahmed et al., 2025; Kosel et al., 2024).

The literature consequently lacks a standardized measurement strategy. Studies frequently employ instruments that assess different constructs under the common label of cognitive load, limiting cross-study comparison. Greater methodological precision is required if XR research is to identify the mechanisms through which immersive features facilitate or interfere with learning.

2.3 Presence, Cognitive Load, and Learning Outcomes

Studies examining the relationship among immersion, presence, cognitive load, and learning have produced inconsistent findings. Positive effects are most frequently reported in tasks involving spatial reasoning, procedural knowledge, visualization, and experiential learning. In these settings, immersive systems may provide representations or forms of interaction that correspond directly with the structure of the content (Conrad et al., 2024; Johnson et al., 2022; Pedram et al., 2024).

Virtual field trips, medical simulations, assembly training, and three-dimensional scientific representations illustrate contexts in which XR may support learning by making relationships perceptually accessible or allowing learners to rehearse actions in realistic but controlled environments (Makransky & Mayer, 2022; Pietschmann et al., 2023). Presence may contribute to these benefits by sustaining attention, increasing situational involvement, or supporting the construction of coherent spatial and procedural models.

However, higher presence does not consistently produce better learning. Several studies have reported that immersive conditions generate stronger presence or engagement without corresponding improvements in test performance (Johnsdorf et al., 2023; Ochs & Sonderegger, 2022). Some comparisons have found that desktop or less immersive formats produce equivalent or superior outcomes, particularly when the learning content is factual, abstract, symbolic, or language-based (Baceviciute et al., 2021; Srivastava et al., 2019; Yu et al., 2025).

These findings challenge the immersion assumption—the belief that increasingly immersive technologies inherently improve learning. Greater immersion may intensify the experience without improving the cognitive processes required for knowledge acquisition. Engagement, satisfaction, and presence are meaningful educational variables, but they cannot be treated as substitutes for retention, transfer, procedural accuracy, or conceptual understanding (Mayer & Bailenson, 2024; Schwarze et al., 2019).

The relationship may depend on cognitive load. Immersive features can enhance learning when they reduce the need for mental transformation or direct attention toward relevant information. Conversely, the same features can weaken performance when they introduce interface demands, irrelevant detail, or navigational complexity (Breves & Stein, 2022; Poupard et al., 2024). Cognitive load may therefore help explain why presence is beneficial in some conditions and ineffective or detrimental in others.

Nevertheless, current evidence does not justify the categorical claim that cognitive load always mediates the relationship between presence and learning. Direct mediation studies remain comparatively limited, and reported findings vary according to task, instrument, participant characteristics, and analytical model. Some research suggests that immersive features increase extraneous load and subsequently reduce performance, whereas other studies indicate that presence may improve motivation or involvement without substantially changing perceived load (Breves & Stein, 2022; Thorp et al., 2024).

Another limitation is the prevalence of media-comparison designs. Comparisons between XR and conventional instruction often vary several factors simultaneously, including novelty, interactivity, display modality, content format, instructional duration, and learner control. As a result, observed differences cannot be attributed confidently to presence or immersion alone (Hamilton et al., 2020; Lawson et al., 2024). More informative studies would manipulate theoretically relevant design features while directly measuring presence, differentiated cognitive load, and learning outcomes.

2.4 Moderating Effects of Task, Learner, and Instructional Design Characteristics

The mixed findings in XR research have prompted greater attention to moderating factors. Task characteristics, learner attributes, technological usability, and instructional design appear to determine whether immersive environments support or impede learning.

Task–affordance alignment is one of the most consistent principles emerging from the literature. XR is particularly suited to content requiring spatial representation, embodied interaction, procedural rehearsal, or experiential access. Tasks involving anatomy, equipment operation, navigation, physical assembly, or simulated environments may benefit because XR represents information in a format closely aligned with the required cognitive and motor operations (Johnson et al., 2022; Tusher et al., 2024). By contrast, abstract, symbolic, or text-intensive tasks may not benefit from immersion and may be processed more efficiently through simpler instructional media (Baceviciute et al., 2021; Yu et al., 2025).

Prior knowledge can alter these effects. Novices may benefit from concrete visualizations because they lack established schemas for interpreting abstract content. However, they may also be more vulnerable to interface complexity, irrelevant detail, and navigational demands. Experts may manage immersive environments more efficiently because they can identify relevant information and disregard distractions, although they may derive less benefit from representations that merely duplicate knowledge they already possess (Han et al., 2023; Poupard et al., 2024).

Spatial ability and technological familiarity are also relevant. Learners with stronger visuospatial skills may navigate and interpret three-dimensional environments more successfully, whereas learners with lower spatial ability may experience additional cognitive demands (Brucker et al., 2024; Lin & Suh, 2021). Prior gaming or XR experience may similarly reduce the cognitive cost of system use, although familiarity should not be confused with instructional mastery.

Technological usability affects both presence and cognitive load. Natural, responsive, and reliable interactions may support presence without diverting substantial cognitive resources from the learning objective. Complex, delayed, or unnatural controls can disrupt presence and increase extraneous load (Bailey et al., 2018; Chandio et al., 2024). The educational quality of XR therefore depends not only on fidelity but also on whether learners can use the system with minimal cognitive effort.

Instructional design is a central moderator. Unguided exploration may generate curiosity and involvement but often produces weak learning outcomes because learners do not necessarily attend to the information required by the objective (Ferguson & Oostendorp, 2020; Mulders et al., 2020). Guidance, cueing, signaling, segmentation, and scaffolding can direct attention, regulate complexity, and reduce task-irrelevant processing (Albus & Seufert, 2022; Weerasinghe et al., 2022).

Adaptive guidance may be particularly valuable because learners differ in prior knowledge, spatial ability, and technological familiarity. Novices may require simplified interactions, explicit instructions, and frequent cues, while advanced learners may benefit from greater autonomy and complexity. However, excessive guidance may reduce exploration or create redundancy for experienced users. Instructional support should therefore be calibrated to learner needs and task demands rather than applied uniformly.

The literature suggests that these moderators interact. Task type influences which technological features are useful; learner characteristics affect the cognitive cost of those features; and instructional design determines whether attention is directed toward relevant processing. Yet most studies examine only one or two moderators at a time. Integrated studies capable of testing interactions among task, learner, technology, and pedagogy remain limited.

2.5 Synthesis and Research Gaps

The literature collectively indicates that presence is neither an inherently beneficial nor inherently detrimental condition for learning. Its educational value depends on whether the immersive experience supports the cognitive operations required by the learning task. Presence may facilitate learning when it is task-congruent, when interaction is usable, and when instructional design directs attention toward relevant information. It may impede learning when it is generated through excessive realism, complex controls, unguided exploration, or other features that increase extraneous processing (Breves & Stein, 2022; Makransky & Petersen, 2021; Poupard et al., 2024).

Several gaps remain. First, the distinction between immersion and presence is not consistently maintained, and many studies infer psychological presence from hardware conditions. Second, presence and cognitive load are measured using heterogeneous instruments, limiting cumulative comparison. Third, differentiated forms of cognitive load are seldom assessed alongside multidimensional presence. Fourth, mediation and interaction models remain underdeveloped, particularly those examining task type, prior knowledge, spatial ability, usability, and guidance simultaneously. Fifth, most studies emphasize immediate learning outcomes rather than retention, transfer, and repeated use. Finally, the evidence base remains concentrated on VR, while AR, MR, workplace learning, and social or affective outcomes receive less systematic attention.

The present review responds to these gaps by integrating conceptual, empirical, methodological, and critical literature into a contingent framework. Rather than treating immersion as a universal educational advantage, it examines how task–affordance alignment, cognitive load, learner characteristics, and instructional design jointly determine the effectiveness of XR learning environments.

3. Methodology

3.1 Research Design

This study employed a **thematic analytical review** to examine the conceptual, methodological, and empirical literature concerning presence, cognitive load, and learning in extended reality environments. This design was selected because the literature does not consist solely of comparable empirical studies suitable for quantitative aggregation. It also includes theoretical models, measurement research, experimental comparisons, systematic reviews, critical analyses, and discussions of the broader discourse surrounding immersive educational technologies.

Unlike a meta-analysis, the review did not calculate pooled effect sizes. Instead, it organized and critically interpreted the literature to identify recurring patterns, conceptual disagreements, methodological limitations, and conditions under which XR may support or impede learning. The analytical emphasis was therefore placed on the consistency of definitions, alignment between constructs and measures, direction of reported findings, and the contextual factors that may account for divergent results.

The review was structured around five principal analytical themes:

1. conceptualization and measurement of presence;
2. cognitive load in XR environments;
3. relationships among presence, cognitive load, and learning outcomes;
4. the immersion assumption in scholarly and broader technological discourse; and
5. moderating effects of task characteristics, learner attributes, technological usability, and instructional design.

These themes provided the initial framework for organizing the literature while allowing additional subthemes and relationships to emerge during synthesis.

3.2 Information Sources and Search Strategy

The literature search covered English-language publications issued between 1990 and 2026. The starting period was selected to capture foundational scholarship on presence and virtual environments, while the endpoint allowed the review to include recent developments in virtual, augmented, mixed, and extended reality research.

The principal databases were:

- Scopus;
- Web of Science;
- IEEE Xplore;
- PsycINFO; and
- ACM Digital Library.

Search terms were organized into four conceptual clusters:

1. **XR technologies:** “virtual reality,” “augmented reality,” “mixed reality,” “extended reality,” and “immersive technology”;
2. **presence:** “presence,” “spatial presence,” “telepresence,” and “sense of presence”;
3. **cognitive load:** “cognitive load,” “cognitive load theory,” “mental workload,” “intrinsic load,” “extraneous load,” and “germane load”; and
4. **learning:** “learning,” “learning outcomes,” “education,” “training,” and “skill acquisition.”

Boolean combinations required terms representing XR technology and learning together with presence, cognitive load, or both. The database search was supplemented by backward reference checking and forward citation tracking of foundational and frequently cited studies, including major works on presence, cognitive load theory, and immersive learning.

The search strategy was designed to achieve broad conceptual coverage rather than to limit retrieval to a single discipline. This was necessary because relevant studies are distributed across educational psychology, human–computer interaction, cognitive science, simulation, medical education, engineering education, and computer-supported learning.

3.3 Eligibility Criteria

Sources were included when they met one or more of the following criteria:

1. they examined virtual, augmented, mixed, or extended reality in an educational, instructional, training, or skill-development context;
2. they conceptualized or measured presence, immersion, cognitive load, mental workload, or closely related constructs;
3. they reported learning, performance, retention, transfer, engagement, or task-related outcomes relevant to the presence–load–learning relationship;
4. they developed or evaluated theoretical models relevant to immersive learning;
5. they reviewed or synthesized XR learning research; or
6. they provided substantive critical analysis of claims concerning immersion and educational effectiveness.

Eligible source types included peer-reviewed journal articles, conference proceedings, scholarly books or chapters, dissertations or institutional research outputs where directly relevant, and selected preprints when they addressed recent developments not yet represented adequately in the peer-reviewed literature.

Sources were excluded when they:

1. focused exclusively on software development, hardware specifications, rendering performance, or engineering implementation without addressing learning or cognitive outcomes;
2. mentioned presence or cognitive load only incidentally;
3. concerned entertainment or gaming without a clear learning, training, or cognitive-performance component;
4. were not available in English;
5. lacked sufficient conceptual or methodological information for evaluation; or
6. repeated findings already represented by a more complete or formally published version of the same work.

Selected non-scholarly or critical materials were considered only for examining the immersion assumption and broader technological discourse. These materials were not treated as equivalent in evidentiary weight to peer-reviewed empirical research.

3.4 Screening and Source Selection

The manuscript reports that the initial search produced **1,247 records after duplicate removal**. Titles and abstracts were screened to determine whether the sources addressed XR learning together with presence, cognitive load, or related cognitive and instructional outcomes.

Following preliminary screening, **236 records underwent full-text assessment**. Full texts were reviewed to determine whether the source contributed relevant conceptual definitions, measurement approaches, empirical findings, theoretical interpretations, moderating variables, or methodological critiques.

A total of **158 sources were reported as included in the final synthesis**. Sources were retained according to their direct relevance to one or more of the five analytical themes rather than solely according to study design.

Because the present review was analytical rather than meta-analytic, studies were not excluded simply because they used small samples, non-randomized designs, or heterogeneous measures. Instead, these methodological characteristics were considered when assigning evidentiary weight and interpreting the strength and generalizability of their conclusions.

3.5 Data Extraction and Thematic Organization

A structured extraction framework was used to organize information from the included sources. The following elements were examined where available:

1. definition and conceptualization of presence;
2. distinction between presence and technological immersion;
3. cognitive load framework or workload construct employed;
4. instruments or indicators used to measure presence and cognitive load;
5. XR modality, including VR, AR, MR, or broader XR;
6. educational or training context;
7. participant characteristics;
8. learning task and outcome measures;
9. direction and statistical significance of reported relationships;
10. moderating or mediating variables;
11. instructional design characteristics;
12. technological usability or interaction features;
13. methodological limitations; and
14. implications for theory, research, and practice.

The extracted material was initially grouped according to the five predefined themes. Within each theme, the analysis identified recurring propositions, areas of convergence, contradictory findings, and possible methodological explanations for disagreement.

Cross-theme synthesis was subsequently conducted to examine how task characteristics, learner differences, technological properties, and instructional design jointly shaped the relationship among presence, cognitive load, and learning. This process supported the development of the contingent framework presented in the Results and Discussion section.

3.6 Evaluation of Evidentiary Strength

The sources were evaluated according to three broad criteria:

Theoretical coherence

The review examined whether constructs were clearly defined and whether the measures used were consistent with those definitions. Particular attention was given to studies that distinguished technological immersion from subjective presence and differentiated among forms of cognitive load.

Methodological rigor

Studies were considered in relation to their research design, sample characteristics, comparison conditions, measurement validity, control of potential confounding variables, and appropriateness of statistical or qualitative analysis.

Consistency and explanatory value

Greater evidentiary weight was assigned to findings replicated across multiple studies, contexts, or review-level sources. Isolated findings, studies with serious confounding, and claims based primarily on engagement or satisfaction rather than learning outcomes were interpreted cautiously.

The review distinguished among:

- findings supported consistently across the literature;
- plausible interpretations supported by partial or mixed evidence; and
- speculative claims requiring further empirical testing.

This distinction was particularly important when evaluating assertions that greater technological immersion necessarily produces stronger presence or superior learning.

3.7 Analytical Procedure

The analytical process proceeded in three stages.

First, **within-theme analysis** was conducted to identify dominant definitions, measures, findings, and methodological concerns within each thematic area.

Second, **cross-theme comparison** examined how the themes intersected. For example, studies reporting high presence but weak learning were assessed in relation to cognitive load, task characteristics, instructional structure, and interface demands.

Third, **integrative interpretation** was used to formulate a contingent explanation of XR learning effectiveness. This interpretation proposed that presence is educationally valuable primarily when it is aligned with the learning task, supported by usable technology, and structured through appropriate instructional design. Conversely, immersive features may impair learning when they impose task-irrelevant processing or exceed learners' cognitive and technological capacities.

The resulting framework is theoretical and interpretive rather than statistically validated. It is intended to organize the available evidence and generate propositions for future empirical testing.

3.8 Methodological Limitations

Several limitations apply to the review. First, the study did not conduct a quantitative meta-analysis and therefore cannot estimate the pooled magnitude of the relationships among presence, cognitive load, and learning. Second, restriction to English-language sources may have excluded relevant research published in other languages. Third, XR technologies have evolved rapidly, and results obtained using older hardware may not generalize fully to current systems. Fourth, heterogeneity in tasks, measures, devices, learning outcomes, and participant populations limits direct comparison across studies. Fifth, the inclusion of selected preprints, dissertations, and non-scholarly materials introduces variation in source quality, although these sources were interpreted according to their evidentiary status. Sixth, the thematic organization involved analytical judgment, and some sources addressed more than one theme. Finally, the proposed contingent framework was derived from literature synthesis and has not yet been tested as an integrated empirical model. Its propositions should therefore be treated as theoretically grounded directions for future research rather than definitive causal conclusions.

3.9 Ethical Considerations

The study involved the analysis of publicly available published and scholarly materials and did not include direct human participation, intervention, or collection of identifiable personal information. Institutional ethics approval and individual informed consent were therefore not applicable.

4. Results and Discussion

The thematic synthesis identified five interrelated findings concerning presence, cognitive load, and learning in extended reality environments. These findings show that the educational effects of XR are not determined by technological immersion alone. Instead, outcomes depend on how presence is produced, how cognitive demands are distributed, whether XR affordances correspond with the learning task, and how instructional and learner-related factors shape the experience. Table 1 summarizes the principal thematic findings.

Table 1. *Summary of Thematic Findings on Presence, Cognitive Load, and Learning in XR*

Theme	Principal Finding	Main Limitation in the Literature	Implication
Presence and immersion	Immersion increases the potential for presence but does not guarantee it	Presence is often inferred from hardware rather than measured directly	Immersion and presence should be treated as distinct constructs
Cognitive load	XR may reduce some processing demands while increasing task-irrelevant demands	Global workload measures often fail to distinguish types of cognitive load	Cognitive load should be measured multidimensionally
Learning outcomes	XR benefits are strongest for spatial, procedural, and experiential tasks	Media-comparison studies often contain uncontrolled differences	Effectiveness depends on task–affordance alignment
Immersion assumption	Higher immersion does not consistently produce better learning	Engagement and presence are sometimes treated as proxies for learning	XR claims should be based on learning outcomes, retention, and transfer
Moderating factors	Task, learner, technology, and instructional design jointly shape outcomes	Moderators are frequently studied in isolation	XR effectiveness should be understood through a contingent framework

4.1 Presence Is Multidimensional and Distinct from Technological Immersion

The first major finding concerns the conceptual distinction between presence and immersion. Across the literature, immersion is primarily associated with the technological properties of an XR system, whereas presence refers to the learner's subjective experience of being situated within or psychologically connected to a mediated environment. Although greater technological immersion can increase the likelihood of presence, the relationship is neither automatic nor uniform (Johnson-Glenberg, 2018; Kelly et al., 2022).

Display resolution, field of view, motion tracking, sensory fidelity, and interactivity can strengthen the conditions under which presence develops. However, presence also depends on system usability, coherence of the environment, interaction naturalness, task relevance, user expectations, and individual susceptibility to immersive experiences. A technologically advanced system may therefore generate weak presence when interaction is unreliable or when learners are preoccupied with operating the interface. Conversely, a less technologically intensive environment may generate meaningful presence if the experience is coherent, responsive, and closely aligned with the user's task.

The synthesis also indicates that presence should not be treated as a single, undifferentiated state. Studies commonly distinguish among spatial presence, involvement, and experienced realism. Spatial presence concerns the sense of being physically located within the mediated environment. Involvement refers to attentional and psychological absorption. Experienced realism concerns whether the environment appears plausible or authentic (Berkman & Çatak, 2021; Grassini & Laumann, 2020).

These dimensions may affect learning differently. Spatial presence may assist navigation, procedural rehearsal, and spatial reasoning because it allows the learner to experience relations among objects, locations, and actions more directly. Involvement may support sustained attention, but it can also produce concentration on visually compelling or emotionally salient elements that are unrelated to the learning objective. Realism may improve contextual understanding in some tasks, yet excessive realism can introduce irrelevant details that compete for limited cognitive resources.

This differentiation explains why higher overall presence scores do not consistently predict superior learning. A learner may experience strong involvement and realism while attending to aspects of the environment that do not support the target knowledge or skill. Presence should therefore be evaluated functionally: its educational significance depends on which dimensions are activated and whether those dimensions contribute to task-relevant cognitive processing.

Measurement practices remain a substantial limitation. Self-report questionnaires such as the Igroup Presence Questionnaire and other presence scales are widely used, but they capture retrospective judgments rather than continuous experience. Physiological and behavioral measures offer greater temporal sensitivity but are not specific to presence. Heart-rate variation, electrodermal activity, reaction time, and head movement may also reflect arousal, effort, anxiety, novelty, or cybersickness (Chandio et al., 2023; Kerrebroeck et al., 2021).

The most defensible conclusion is therefore that presence is a multidimensional psychological response influenced by, but not reducible to, technological immersion. Research designs that classify head-mounted displays as “high presence” and desktop displays as “low presence” without directly measuring subjective experience provide only limited evidence about presence itself.

4.2 Cognitive Load Functions as a Conditional Mechanism

The second finding is that cognitive load provides a plausible mechanism for explaining why XR supports learning in some conditions and impedes it in others. However, this mechanism is conditional rather than universal.

XR can alter cognitive demand by changing how learners perceive, manipulate, and interpret information. For spatial or procedural tasks, three-dimensional representation may reduce the need to mentally convert symbolic or two-dimensional information into spatial models. Embodied interaction can also externalize certain cognitive operations by allowing learners to manipulate objects, perform procedures, or observe dynamic relationships directly (Kenneally et al., 2026; Zou et al., 2025).

For example, understanding an anatomical structure through an interactive three-dimensional model may be cognitively more manageable than reconstructing the same structure from several static diagrams. Similarly, practicing a procedure in a simulated environment may allow learners to coordinate conceptual knowledge with action. In these cases, XR may reduce unnecessary mental transformation and facilitate task-relevant schema construction.

Nevertheless, the same environment can generate additional cognitive demands. Learners may need to operate controllers, manage navigation, interpret system feedback, monitor physical movement, and respond to unfamiliar interface conventions while processing the instructional content. Highly detailed environments may also contain objects, sounds, animations, or spatial features that attract attention without contributing to the learning objective. These demands increase extraneous processing and may offset the representational benefits of XR (Breves & Stein, 2022; Skulmowski & Xu, 2021).

The findings therefore do not support the proposition that immersive environments simply reduce or increase cognitive load. Rather, XR redistributes cognitive demands. Some demands may be reduced because information is represented more directly, while others may increase because of interaction, navigation, sensory stimulation, or technological complexity.

This redistribution is especially important when distinguishing among intrinsic, extraneous, and germane processing. A realistic environment may support meaningful processing in a navigation task but function as irrelevant detail in an abstract conceptual lesson. Similarly, interactivity may be central to procedural learning but unnecessary for factual recall. The cognitive classification of an XR feature therefore depends on the instructional function it performs.

Measurement limitations complicate this interpretation. Many XR studies rely on global workload instruments such as the NASA Task Load Index, which combines mental demand, physical demand, temporal pressure, effort, and frustration. Although useful for estimating overall workload, such instruments cannot determine whether cognitive demand is caused by the inherent complexity of the material, poor interface design, physical discomfort, or productive engagement with the content (Brachten et al., 2020; Vidal-Balea et al., 2024).

More differentiated measures are required. Instruments that distinguish among sources of cognitive load are more theoretically informative, particularly when used alongside behavioral or physiological measures. Even then, interpretation must remain cautious because physiological arousal or mental effort does not independently reveal whether processing is productive or extraneous.

The review consequently supports a qualified conclusion: cognitive load may mediate the relationship between presence and learning under certain conditions, particularly when immersive features either facilitate task-relevant processing or impose unnecessary cognitive demands. However, the evidence does not establish a universal mediation pathway. The relationship varies according to task, learner, system, and instructional design.

4.3 Learning Benefits Depend on Task–Affordance Alignment

The third major finding is that XR learning outcomes depend strongly on the alignment between the affordances of the technology and the cognitive requirements of the task.

Positive effects are most consistently observed in tasks involving spatial understanding, procedural rehearsal, visualization, embodied interaction, or experiential learning. These tasks benefit from the ability of XR to represent three-dimensional relationships, simulate dynamic processes, and permit action within controlled environments (Conrad et al., 2024; Johnson et al., 2022; Pedram et al., 2024).

Medical and health education provides several examples. XR can support anatomical visualization, procedural practice, and simulated clinical decision-making without exposing patients or learners to unnecessary risk. Engineering and vocational training can similarly benefit from virtual assembly, equipment operation, and guided rehearsal. Virtual field trips may enable learners to examine environments that are geographically distant, hazardous, or otherwise inaccessible (Makransky & Mayer, 2022; Pietschmann et al., 2023).

In these contexts, XR does more than increase engagement. It provides representational and experiential capabilities that correspond directly with the learning objective. The learner can inspect spatial relations, perform actions, receive feedback, and construct mental models through interaction.

The evidence is less favorable when the task is primarily abstract, symbolic, language-based, or dependent on concentrated reading and conceptual analysis. For such tasks, immersive environments may provide limited representational advantage and may introduce additional interface or attentional demands. Studies have consequently reported null or negative effects for some forms of factual, declarative, or abstract learning, even when immersive conditions generate higher presence or enjoyment (Johnsdorf et al., 2023; Ochs & Sonderegger, 2022; Yu et al., 2025).

This pattern supports a task–affordance matching principle. XR is most effective when its distinctive features are necessary or materially useful for representing the content or performing the target task. It is less likely to provide educational value when immersive features are added to content that can be communicated more efficiently through text, diagrams, video, or desktop interaction.

The findings also reveal that XR should not be evaluated as a single instructional treatment. Different XR systems provide different combinations of visual fidelity, interactivity, embodiment, learner control, and environmental complexity. Similarly, learning tasks differ in the extent to which they require spatial, procedural, or experiential representation. General claims about whether XR "works" are therefore analytically weak unless the specific affordances and task demands are identified.

Media-comparison studies frequently fail to isolate these factors. Comparisons between head-mounted VR and conventional instruction may differ not only in immersion but also in instructional time, learner control, novelty, content sequencing, and interaction. Consequently, superior performance in one condition cannot always be attributed to immersion or presence, while equivalent performance cannot necessarily be interpreted as evidence that XR lacks value (Hamilton et al., 2020; Lawson et al., 2024).

More informative studies should compare instructional designs that are equivalent in content and duration while varying theoretically relevant XR features. They should also assess whether the technology produces unique gains in retention, transfer, procedural accuracy, or performance under realistic conditions.

4.4 The Immersion Assumption Exceeds the Available Evidence

The fourth finding concerns the persistent immersion assumption: the belief that greater technological immersion naturally produces stronger engagement, deeper learning, and improved performance.

The literature does not support this proposition as a general rule. Higher immersion may increase presence, enjoyment, novelty, or subjective engagement, but these outcomes are distinct from learning. Learners can report that an XR experience is compelling or realistic without demonstrating stronger retention, transfer, or conceptual understanding (Mayer & Bailenson, 2024; Schwarze et al., 2019).

This distinction is particularly important because engagement is often treated as an indirect indicator of educational effectiveness. Engagement may be beneficial when it sustains attention toward relevant content, but it may also reflect interest in the medium, visual novelty, or exploratory freedom. Without corresponding evidence of learning, engagement should not be interpreted as proof of instructional success.

The immersion assumption is reinforced by promotional and technological discourse. XR systems are frequently described as transformative, revolutionary, or inherently experiential. These descriptions may influence procurement decisions, policy priorities, and research agendas before sufficient evidence has established whether a particular application improves learning relative to less costly or less technically demanding alternatives.

The assumption is also reproduced indirectly in research. Studies sometimes begin with the expectation that immersive conditions should outperform conventional ones and interpret null findings primarily as design failures rather than as evidence that immersion may not be necessary for the task. Other studies use presence, satisfaction, or intention to use as principal outcomes while making broader claims about learning effectiveness.

The synthesis indicates that XR should be treated as a medium rather than an instructional method. Its educational value depends on the pedagogical structure embedded within it. High-fidelity simulation without guidance may be less effective than a simpler but carefully scaffolded learning activity. Conversely, a well-designed XR environment can provide advantages when its affordances are closely matched to the task.

The immersion assumption should therefore be replaced by an evidence-based adoption principle: technological sophistication is educationally justified only when it supports a demonstrable instructional function. Evaluation should include not only user experience but also knowledge retention, procedural competence, transfer to new contexts, durability of learning, and cost relative to alternative approaches.

4.5 Moderating Factors and the Contingent Framework

The fifth finding is that XR effectiveness is shaped by interacting moderators at four levels: task, learner, technology, and instruction.

Task-level factors

Task characteristics determine whether XR provides meaningful representational advantages. Spatial, procedural, dynamic, and experiential tasks are generally better aligned with XR than abstract, symbolic, or text-intensive tasks. Task complexity also matters. Highly complex content may benefit from segmentation or visualization, but it may overwhelm learners when combined with extensive navigation or interaction.

Learner-level factors

Prior knowledge, spatial ability, technological familiarity, gaming experience, age, and susceptibility to cybersickness influence both presence and cognitive load. Novices may benefit from concrete representation, but they may also require more guidance and simplified interaction. Experts may navigate complex environments more efficiently but may derive less benefit from representations that duplicate knowledge already encoded in established schemas (Brucker et al., 2024; Han et al., 2023).

Spatial ability is particularly relevant because many XR environments require learners to interpret perspective, orientation, distance, and three-dimensional relationships. Learners with weaker visuospatial ability may need additional cues, guided perspectives, or reduced navigational demands.

Technology-level factors

Usability, interaction naturalness, system latency, tracking accuracy, visual clarity, and physical comfort affect both presence and cognitive demand. A highly immersive system with poor controls may generate lower educational value than a simpler system with intuitive interaction. Technological quality must therefore be assessed according to functional usability rather than fidelity alone (Bailey et al., 2018; Chandio et al., 2024).

Instruction-level factors

Guidance, cueing, signaling, segmentation, scaffolding, feedback, and opportunities for reflection are central to effective XR learning. Unstructured exploration may support discovery in limited circumstances, but it frequently allows attention to shift toward incidental features. Guided exploration is more likely to ensure that immersive activity contributes to the intended learning outcome (Albus & Seufert, 2022; Mulders et al., 2020; Weerasinghe et al., 2022).

These factors form the basis of a contingent framework. The framework proposes that presence supports learning when four conditions are substantially satisfied:

1. XR affordances correspond with the cognitive demands of the task;
2. the interface can be used without excessive cognitive or physical burden;
3. instructional design directs attention toward relevant content and manages complexity; and
4. the level of guidance and interaction is appropriate to learner characteristics.

When these conditions are present, XR may reduce unnecessary mental transformation, support embodied or spatial understanding, and promote meaningful processing. When they are absent, immersion may generate distraction, interface burden, irrelevant realism, or unguided activity.

The framework does not imply that all four conditions must be maximized. Excessive guidance may restrict productive exploration, while excessive realism may increase unnecessary detail. The goal is not maximum immersion, presence, or interactivity, but an appropriate configuration of features for the learner and task.

4.6 Discussion

The findings support a functional contingency interpretation of XR learning. Presence is not a direct or universally positive cause of learning. Its effects depend on the cognitive functions it supports and the processing demands it imposes.

This interpretation resolves several apparent contradictions in the literature. Studies reporting positive effects of immersive XR generally involve tasks in which spatial, procedural, or experiential representation is central. Under these conditions, the sense of being situated within the environment can support perception, action, and mental-model construction. Studies reporting null or negative effects often involve tasks that do not require immersive representation or systems that impose substantial interface, navigation, or attentional demands.

The key distinction is therefore between task-congruent and task-incongruent presence. Task-congruent presence supports the operations necessary for learning. Task-incongruent presence increases involvement or realism without advancing the objective and may therefore contribute to extraneous processing.

The findings also extend the application of cognitive load theory in XR. Cognitive demands in immersive environments cannot be understood solely as properties of the content. They emerge from the interaction among content complexity, representational format, interface demands, learner characteristics, and instructional structure. XR may reduce some demands while generating others. The relevant analytical question is not whether XR increases or decreases load overall, but which demands are reduced, which are introduced, and whether the resulting processing contributes to learning.

A further implication concerns the treatment of presence as multidimensional. Spatial presence, involvement, and realism should not be assumed to have identical effects. Spatial presence may be beneficial for navigation or procedural tasks, while realism may become counterproductive when it adds irrelevant detail. Involvement may sustain attention, but only when attention remains aligned with the instructional objective. Future models should therefore distinguish among dimensions of presence and examine their separate and combined relationships with cognitive processing.

The findings also expose the limitations of current measurement practices. Global presence and workload scores may conceal the mechanisms that matter most. Studies should combine multidimensional self-report measures with behavioral evidence, task performance, learning outcomes, and, where appropriate, physiological indicators. No single indicator should be treated as definitive.

For research, the contingent framework suggests moving beyond simple comparisons between XR and conventional media. Studies should manipulate theoretically meaningful features such as guidance, realism, interactivity, learner control, and representational alignment. Mediation and moderation models should test whether specific forms of cognitive demand explain the relationship between presence and learning under defined conditions.

For practice, XR adoption should begin with the instructional problem rather than the technology. Educators and developers should determine whether spatial, procedural, dynamic, or experiential representation is necessary. They should then design the least cognitively burdensome interaction capable of supporting the objective. Greater immersion is justified only when it provides an instructional advantage that cannot be achieved as efficiently through simpler media.

Overall, the thematic evidence does not support technological determinism. XR can create meaningful learning opportunities, but its effectiveness depends on design discipline. Presence becomes educationally valuable when it serves cognition rather than spectacle, and immersion becomes defensible when it performs a necessary instructional function rather than merely increasing technological novelty.

5. Conclusions, Recommendations, and Implications

5.1 Conclusions

This thematic analytical review examined how presence, cognitive load, and learning interact in extended reality environments. The reviewed evidence indicates that the educational effectiveness of XR cannot be explained by technological immersion alone. Higher display fidelity, broader sensory stimulation, or increased interactivity may enhance the possibility of presence, but these features do not automatically produce stronger learning outcomes.

Presence and immersion should therefore be treated as related but distinct constructs. Immersion refers primarily to the objective capabilities of an XR system, whereas presence concerns the learner's subjective experience of being situated within or psychologically connected to the mediated environment. The frequent conflation of these constructs has weakened the interpretation of XR research because learning effects attributed to presence may instead arise from differences in interface demands, content presentation, novelty, usability, or instructional design.

The literature further indicates that presence is multidimensional. Spatial presence, involvement, and perceived realism may affect learning differently depending on the objective and context. Spatial presence may support navigation, procedural rehearsal, and three-dimensional understanding, while involvement may sustain attention. However, realism and sensory richness may also introduce irrelevant information or distraction when they do not contribute directly to the learning task.

Cognitive load offers a plausible but conditional explanation for the inconsistent effects of XR. Immersive environments may reduce the need for mental transformation when they present spatial, procedural, or dynamic information in directly perceptible forms. At the same time, XR may impose additional demands through complex controls, navigation, environmental detail, physical movement, cybersickness, and unfamiliar interface conventions. XR therefore does not simply increase or decrease cognitive load; it redistributes cognitive demands across content processing, system operation, and environmental interaction.

The review does not establish that cognitive load universally mediates the relationship between presence and learning. Rather, the evidence suggests that cognitive load may function as a mediating mechanism under particular conditions. Presence is more likely to support learning when immersive features facilitate task-relevant processing and less likely to do so when they impose extraneous demands.

The strongest pattern across the literature concerns task–affordance alignment. XR appears most beneficial for spatial, procedural, experiential, and dynamic learning tasks in which three-dimensional representation, embodied action, or simulated practice provides a meaningful instructional advantage. Its benefits are less consistent for abstract, symbolic, textual, or factual tasks that may be communicated more efficiently through less immersive media.

Learner characteristics also shape outcomes. Prior knowledge, spatial ability, technological familiarity, and susceptibility to cybersickness influence the cognitive cost and educational value of XR experiences. Novices may benefit from concrete visualization but may require greater guidance and simplified interaction. More experienced learners may navigate immersive systems more efficiently but may derive less benefit from highly explicit representations.

Instructional design is a decisive factor. Guidance, cueing, segmentation, scaffolding, feedback, and opportunities for reflection can reduce task-irrelevant processing and direct attention toward the intended learning objectives. In contrast, unstructured exploration may increase presence and interest without necessarily improving retention, transfer, or conceptual understanding.

The review therefore rejects the immersion assumption that more technologically immersive systems are inherently better for learning. XR should be understood as a medium whose value depends on how its affordances are matched with task requirements, learner characteristics, usability constraints, and pedagogical design.

The resulting contingent framework proposes that presence is educationally valuable when it is task-congruent, cognitively manageable, instructionally guided, and appropriate to the learner. Under these conditions, XR may support meaningful knowledge construction and skill development. When these conditions are absent, immersion may increase distraction, interface burden, or cognitive overload without generating corresponding learning benefits.

5.2 Recommendations

Recommendations for Researchers

Future studies should distinguish explicitly between immersion and presence. Technological conditions should not be labeled as high- or low-presence environments solely on the basis of hardware. Presence should be measured directly using validated instruments and, where possible, treated as a multidimensional construct.

Researchers should adopt cognitive load measures that distinguish among sources of processing demand. Global workload scores may be useful for estimating overall effort, but they are insufficient for determining whether load arises from content complexity, interface difficulty, physical demand, frustration, or productive learning activity. Multidimensional self-report instruments should be complemented, where appropriate, by behavioral, performance-based, and physiological indicators.

Research designs should move beyond broad media comparisons. Rather than comparing XR with conventional instruction as undifferentiated treatments, future studies should manipulate specific features such as realism, interactivity, guidance, learner control, feedback, and representational format. This would permit stronger conclusions about which components of XR contribute to learning and which impose unnecessary demands.

Future research should test mediation and moderation models directly. Particular attention should be given to whether differentiated forms of cognitive load explain the relationship between dimensions of presence and learning under specific task conditions. Interaction effects involving task type, prior knowledge, spatial ability, usability, and instructional guidance also require adequately powered investigation.

Longitudinal research is needed to assess retention, transfer, repeated exposure, and the decline of novelty effects. Immediate post-tests may overestimate educational value when learner interest is driven primarily by technological novelty. Studies should examine whether gains persist and whether skills acquired in XR transfer to real-world or unfamiliar contexts.

Research should also expand beyond virtual reality. Augmented and mixed reality may produce different relationships among presence, cognitive load, and learning because they integrate digital information with the physical environment rather than replacing it entirely. Workplace training, informal learning, collaborative learning, affective outcomes, and social learning should receive greater attention.

Future reviews should provide reproducible search procedures, complete Boolean strategies, transparent screening processes, documented quality appraisal, and clear treatment of preprints and non-peer-reviewed sources. These practices are necessary to improve methodological transparency and cumulative knowledge.

Recommendations for Educators and Instructional Designers

The decision to use XR should begin with an analysis of the learning objective. XR should be selected when spatial, procedural, dynamic, or experiential representation offers a clear advantage over simpler media. It should not be adopted solely because it is innovative or highly engaging.

Instructional designers should prioritize task–affordance alignment. Each immersive feature should serve a defined cognitive or instructional function. Realistic detail, interaction, environmental freedom, and multimedia elements should be retained only when they contribute to understanding, practice, or transfer.

Usability should be treated as an instructional requirement. Controls, navigation, feedback, and system responses should be intuitive, reliable, and consistent. Learners should receive sufficient orientation before beginning the instructional task so that cognitive resources are not consumed by learning how to operate the system.

Guidance should be embedded within the XR environment. Cueing, signaling, sequencing, segmentation, feedback, and reflection prompts should direct attention toward relevant information. Unguided exploration should be used selectively and only when discovery or autonomous navigation is itself part of the learning objective.

Instructional support should be differentiated according to learner characteristics. Novices may require simplified environments, more explicit cues, and greater procedural support. Experienced learners may benefit from reduced guidance, more complex scenarios, or greater autonomy.

XR learning should be evaluated using educational outcomes rather than user satisfaction alone. Engagement, enjoyment, and presence are useful indicators, but they should be assessed alongside knowledge acquisition, procedural accuracy, retention, transfer, and real-world performance.

Recommendations for XR Developers

Developers should design for cognitive efficiency rather than maximum technological complexity. High fidelity, extensive interactivity, and realistic detail should not be treated as inherently desirable. The relevant question is whether each feature improves the learner's ability to perceive, understand, practice, or apply the target knowledge or skill.

Systems should support adaptive guidance, configurable complexity, and multiple interaction modes. Such flexibility would allow XR applications to accommodate learners with different levels of prior knowledge, spatial ability, physical capability, and technological familiarity.

Applications should include analytics capable of recording navigation patterns, interaction errors, time on task, cue use, and performance progression. These data can help researchers and instructors distinguish productive engagement from confusion, distraction, or inefficient exploration.

Accessibility and cybersickness mitigation should be incorporated from the beginning of development. Alternative input systems, adjustable movement options, seated modes, readable visual displays, and reduced-motion settings can broaden participation and reduce cognitive or physical burden.

Recommendations for Institutions and Policymakers

Educational institutions should avoid procurement decisions based primarily on technological novelty or promotional claims. XR investments should be supported by a clearly defined instructional need, evidence of usability, and evaluation against less costly alternatives.

Pilot testing should precede large-scale adoption. Institutions should examine learning effectiveness, faculty readiness, technical support requirements, accessibility, maintenance costs, and learner acceptance before committing substantial resources.

Policies and funding frameworks should require evidence beyond engagement and satisfaction. Vendors and project proponents should demonstrate how an XR application supports specific learning outcomes and how those outcomes will be evaluated over time.

Institutions should also establish standards for data privacy, learner safety, accessibility, content accuracy, and ethical use. XR systems may collect detailed behavioral, movement, gaze, and performance data, making governance and informed use essential.

5.3 Implications of the Study

Theoretical Implications

The review supports a contingent rather than deterministic theory of XR learning. Presence should not be positioned as a universally positive antecedent of learning. Its effects depend on whether it supports the cognitive operations required by the task and whether it does so without imposing excessive extraneous demands.

The findings also suggest that cognitive load theory should be applied dynamically in immersive environments. Cognitive demands arise not only from the complexity of the content but also from representational format, embodied interaction, interface use, and environmental navigation. The same XR feature may support productive processing in one task and create extraneous load in another.

Presence research would benefit from greater dimensional specificity. Spatial presence, involvement, and realism may have different relationships with cognitive load and learning. Future models should therefore specify which dimension is expected to affect which learning process and under what conditions.

The proposed functional contingency perspective integrates these insights. It treats task congruence as the central criterion for determining whether presence contributes to or interferes with learning.

Methodological Implications

The study highlights the limitations of relying on single post-task questionnaires. Presence and cognitive load fluctuate throughout an XR experience, and retrospective global scores may conceal important changes occurring during specific interactions or instructional stages.

A multimethod approach is therefore necessary. Self-reports should be combined with learning performance, behavioral traces, interaction logs, and, where appropriate, physiological indicators. Greater measurement precision would permit researchers to distinguish presence from arousal, workload from productive effort, and engagement from learning.

The review also demonstrates the need for stronger control in media-comparison studies. Without equivalent content, instructional time, guidance, and assessment, differences between XR and conventional formats cannot confidently be attributed to immersion or presence.

Practical and Educational Implications

The primary practical implication is that XR should be designed from the instructional objective backward. The technology should serve the learning task rather than determine it.

This principle has direct consequences for curriculum design, faculty training, and technology adoption. Educators require support not only in operating XR equipment but also in determining when immersive representation is pedagogically justified, how cognitive load can be managed, and how learning outcomes should be assessed.

The findings further imply that simpler technologies may sometimes be preferable. When content is primarily symbolic, textual, or abstract, conventional media may provide greater clarity, lower cognitive burden, and better cost efficiency. Selecting a less immersive medium is not a technological failure when it better serves the educational objective.

Institutional and Policy Implications

Institutions should adopt evidence-based criteria for evaluating XR initiatives. These criteria should include instructional necessity, usability, accessibility, learning outcomes, transfer, sustainability, and cost-effectiveness.

Policy guidance should also counter the tendency to equate technological sophistication with educational quality. Funding and procurement systems should reward demonstrated instructional value rather than maximum immersion or visual realism.

The review additionally underscores the need for responsible communication. Researchers, developers, and institutions should avoid presenting preliminary or context-specific findings as proof that XR universally improves learning. More accurate communication can reduce technological overclaiming and support realistic expectations among educators, learners, policymakers, and investors.

Scholarly Implication

The broader scholarly contribution of the review lies in reframing the central question. The issue is not simply whether XR works. The more productive questions are:

- For which learning tasks is XR appropriate?
- Which dimensions of presence support those tasks?
- What cognitive demands are reduced or introduced?
- Which learners require additional support?
- What design conditions produce retention and transfer?
- When is a simpler medium more effective?

This reframing provides a more rigorous foundation for the study and design of learning in extended reality.

References

Ahmed, N., Wu, P., Huang, K., Jung, S., Rheem, H., Tan, G., Imani, M., & Islam, R. (2025). Human task performance and associated internal states in extended reality: A systematic review of cognitive, psychophysiological, and physiological dimensions. *Frontiers in Virtual Reality, 6*, Article 1589256. https://doi.org/10.3389/frvir.2025.1589256

Albus, P., & Seufert, T. (2022). Signaling in 360° desktop virtual reality influences learning outcome and cognitive load. *Frontiers in Education, 7*, Article 916105. https://doi.org/10.3389/feduc.2022.916105

Andersen, M. S., & Makransky, G. (2021). The validation and further development of the Multidimensional Cognitive Load Scale for physical and online lectures (MCLS-POL). *Frontiers in Psychology, 12*, Article 642084. https://doi.org/10.3389/fpsyg.2021.642084

Baceviciute, S., Lucas, G., Terkildsen, T., & Makransky, G. (2021). Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study. *Journal of Computer Assisted Learning, 38*(1), https://doi.org/10.1111/jcal.12595

Bailey, S.K.T., Johnson, C.I., Sims, V.K. (2019). Using Natural Gesture Interactions Leads to Higher Usability and Presence in a Computer Lesson. In: Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T., Fujita, Y. (eds) Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). IEA 2018. Advances in Intelligent Systems and Computing, vol 826. Springer, Cham. https://doi.org/10.1007/978-3-319-96065-4_70

Berkman, M. İ., & Çatak, G. (2021). Igroup Presence Questionnaire: Psychometrically revised English version. *Muğla Journal of Science and Technology, 7*. https://doi.org/10.22531/muglajsci.882271

Brachten, F., Brünker, F., Frick, N., Roß, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: An experimental study. *Information Systems and E-Business Management, 18*(2). https://doi.org/10.1007/s10257-020-00471-7

Breves, P., & Stein, J. (2022). Cognitive load in immersive media settings: The role of spatial presence and cybersickness. *Virtual Reality, 27*(2). https://doi.org/10.1007/s10055-022-00697-5

Brucker, B., Pardi, G., Uehlin, F., Moosmann, L., Lachmair, M., Halfmann, M., & Gerjets, P. (2024). How learners' visuospatial ability and different ways of changing the perspective influence learning about movements in desktop and immersive virtual reality environments. *Educational Psychology Review, 36*(3). https://doi.org/10.1007/s10648-024-09895-w

Chandio, Y., Bashir, N., Interrante, V., & Anwar, F. M. (2023). Investigating the correlation between presence and reaction time in mixed reality. *IEEE Transactions on Visualization and Computer Graphics, 30*(9). https://doi.org/10.1109/TVCG.2023.3319563

Chandio, Y., Interrante, V., & Anwar, F. M. (2024). Tap into reality: Understanding the impact of interactions on presence and reaction time in mixed reality. *arXiv*. https://doi.org/10.48550/arXiv.2411.05272

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

Crogman, H. T., Cano, V. D., Pacheco, E., Sonawane, R., & Boroon, R. (2025). Virtual reality, augmented reality, and mixed reality in experiential learning: Transforming educational paradigms. *Education Sciences, 15*(3), Article 303. https://doi.org/10.3390/educsci15030303

Ferguson, C., & van Oostendorp, H. (2020). Lost in learning: Hypertext navigational efficiency measures are valid for predicting learning in virtual reality educational games. *Frontiers in Psychology, 11*, Article 578154. https://doi.org/10.3389/fpsyg.2020.578154

Grassini, S., & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology, 11*, Article 349. https://doi.org/10.3389/fpsyg.2020.00349

Hamilton, D. E., McKechnie, J., Edgerton, E., & Wilson, C. (2020). 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). https://doi.org/10.1007/s40692-020-00169-2

Han, J., Liu, G., & Zheng, Q. (2023). Prior knowledge as a moderator between signaling and learning performance in immersive virtual reality laboratories. *Frontiers in Psychology, 14*, Article 1118174. https://doi.org/10.3389/fpsyg.2023.1118174

Johnsdorf, M., Pham, K. A., Schmidt, T., Truong, V. L., Wohnig, A., Kisker, J., Gruber, T., & Schöne, B. (2023). New is not always better: Virtual reality does not necessarily enhance mnemonic processing. *Frontiers in Psychology, 14*, Article 1089725. https://doi.org/10.3389/fpsyg.2023.1089725

Johnson, C. I., Bailey, S. K. T., Schroeder, B. L., & Marraffino, M. D. (2022). Procedural learning in virtual reality: The role of immersion, interactivity, and spatial ability. *Technology, Mind, and Behavior, 3*(4). https://doi.org/10.1037/tmb0000087

Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI, 5*, Article 81. https://doi.org/10.3389/frobt.2018.00081

Kelly, N. J., Hallam, J., & Bignell, S. (2022). Using interpretative phenomenological analysis to gain a qualitative understanding of presence in virtual reality. *Virtual Reality, 27*(2). https://doi.org/10.1007/s10055-022-00719-2

Kenneally, C. D., Bentley, B., & Willison, J. (2026). Cognitive load, extended reality, and visuospatial abilities in physical science education: A systematic review. *Frontiers in Psychology, 17*, Article 1767304. https://doi.org/10.3389/fpsyg.2026.1767304

Kerrebroeck, B. V., Caruso, G., & Maes, P. (2021). A methodological framework for assessing social presence in music interactions in virtual reality. *Frontiers in Psychology, 12*, Article 663725. https://doi.org/10.3389/fpsyg.2021.663725

Khenak, N., Vézien, J.-M., Théry, D., & Bourdot, P. (2018). Spatial presence in real and remote immersive environments and the effect of multisensory stimulation. *Presence: Virtual and Augmented Reality, 27*(3). https://doi.org/10.1162/pres_a_00332

Khenak, N., Vézien, J.-M., & Bourdot, P. (2020). Spatial presence, performance, and behavior between real, remote, and virtual immersive environments. *IEEE Transactions on Visualization and Computer Graphics, 26*(12). https://doi.org/10.1109/TVCG.2020.3023574

Kosel, C., Michel, S., Seidel, T., & Foerster, M. (2024). Exploring the dynamic interplay of cognitive load and emotional arousal by using multimodal measurements: Correlation of pupil diameter and emotional arousal in emotionally engaging tasks. *arXiv*. https://doi.org/10.48550/arXiv.2403.00366

Krieglstein, F., Beege, M., Rey, G. D., Sanchez-Stockhammer, C., & Schneider, S. (2023). Development and validation of a theory-based questionnaire to measure different types of cognitive load. *Educational Psychology Review, 35*(1). https://doi.org/10.1007/s10648-023-09738-0

Lawson, A. P., Martella, A. M., LaBonte, K., Delgado, C. Y., Zhao, F., Gluck, J. A., Munns, M. E., Leroy, A., & Mayer, R. E. (2024). Confounded or controlled? A systematic review of media comparison studies involving immersive virtual reality for STEM education. *Educational Psychology Review, 36*(3). https://doi.org/10.1007/s10648-024-09908-8

Lin, Y., & Suh, A. (2021). The role of spatial ability in learning with virtual reality: A literature review. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2021.011

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). https://doi.org/10.1007/s10648-022-09675-4

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

Maneuvrier, A., Decker, L. M., Ceyte, H., Fleury, P., & Renaud, P. (2020). Presence promotes performance on a virtual spatial cognition task: Impact of human factors on virtual reality assessment. *Frontiers in Virtual Reality, 1*, Article 571713. https://doi.org/10.3389/frvir.2020.571713

Martinez, M., Arbelaez Garces, G., Dupont, L., Hily, A., Camargo, M., Jacob, C., & Dinet, J. (2020, April). Physiological assessment of user experience supported by immersive environments: First input from a literature review. In ConVRgence (VRIC) Virtual Reality International Conference (pp. 89–108). Laval Virtual. https://hal.science/hal-02551341

Mayer, R. E., & Bailenson, J. N. (2024). Advances in research on learning in immersive virtual reality. *Technology, Mind, and Behavior, 5*(4). https://doi.org/10.1037/tmb0000146

Mulders, M. (2023). Confounding in educational research: An overview of research approaches investigating virtual and augmented reality. *Digital Psychology, 4*. https://doi.org/10.24989/dp.v4i1s.2227

Mulders, M., Buchner, J., & Kerres, M. (2020). A framework for the use of immersive virtual reality in learning environments. *International Journal of Emerging Technologies in Learning, 15*(24). https://doi.org/10.3991/ijet.v15i24.16615

Ochs, C., & Sonderegger, A. (2022). The interplay between presence and learning. *Frontiers in Virtual Reality, 3*, Article 742509. https://doi.org/10.3389/frvir.2022.742509

Pedram, S., Kennedy, G., & Sanzone, S. (2024). Assessing the validity of VR as a training tool for medical students. *Virtual Reality, 28*(1). https://doi.org/10.1007/s10055-023-00912-x

Pietschmann, L., Zuercher, P.-D., Bubík, E., Chen, Z., Pfister, H., & Bohné, T. (2023). Quantifying the impact of XR visual guidance on user performance using a large-scale virtual assembly experiment. *arXiv*. https://doi.org/10.48550/arXiv.2308.03390

Poupard, M., Larrue, F., Sauzeon, H., & Tricot, A. (2024). A systematic review of immersive technologies for education: Learning performance, cognitive load, and intrinsic motivation. *British Journal of Educational Technology, 56*(1). https://doi.org/10.1111/bjet.13503

Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. P. G. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behavior, 4*(2). https://doi.org/10.1089/109493101300117884

Schwarze, A., Kampling, H., Heger, O., & Niehaves, B. (2019). Is virtual reality the future of learning? A critical reflection. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2019.214

Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review, 34*(1). https://doi.org/10.1007/s10648-021-09624-7

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 8*(5). https://doi.org/10.1162/105474699566477

Srivastava, P., Rimzhim, A., Vijay, P., Singh, S., & Chandra, S. (2019). Desktop VR is better than non-ambulatory HMD VR for spatial learning. *Frontiers in Robotics and AI, 6*, Article 50. https://doi.org/10.3389/frobt.2019.00050

Thorp, S., Rimol, L. M., Lervik, S., Evensmoen, H. R., & Grassini, S. (2024). Comparative analysis of spatial ability in immersive and non-immersive virtual reality: The role of sense of presence, simulation sickness, and cognitive load. *Frontiers in Virtual Reality, 5*, Article 1343872. https://doi.org/10.3389/frvir.2024.1343872

Tusher, H. M., Mallam, S., & Nazir, S. (2024). A systematic review of virtual reality features for skill training. *Technology, Knowledge and Learning, 29*(2). https://doi.org/10.1007/s10758-023-09713-2

Vidal-Balea, A., Fraga-Lamas, P., & Fernández-Caramés, T. M. (2024). Advancing NASA-TLX: Automatic user interaction analysis for workload evaluation in XR scenarios. In *2024 IEEE Conference on Games*. https://doi.org/10.1109/GEM61861.2024.10585425

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology, 10*, Article 158. https://doi.org/10.3389/fpsyg.2019.00158

Weerasinghe, M., Quigley, A., Pucihar, K. Č., Toniolo, A., Miguel, A., & Kljun, M. (2022). Arigatō: Effects of adaptive guidance on engagement and performance in augmented reality learning environments. *IEEE Transactions on Visualization and Computer Graphics, 28*(11). https://doi.org/10.1109/TVCG.2022.3203088

Yu, N., Shi, W., Dong, W., & Kang, R. (2025). The impact of virtual reality immersion on learning outcomes: A comparative study of declarative and procedural knowledge acquisition. *Behavioral Sciences, 15*(10), Article 1322. https://doi.org/10.3390/bs15101322

Zou, L., Zhang, Z., Mavilidi, M. F., Chen, Y., Herold, F., Ouwehand, K., & Paas, F. (2025). The synergy of embodied cognition and cognitive load theory for optimized learning. *Nature Human Behaviour, 9*(5). https://doi.org/10.1038/s41562-025-02152-2