Australian Journal of Biomedical Research
(eISSN 3083-4708)

Review Article

A Systematic Review of Augmented and Virtual Reality for STEM Learning: Engagement, Cognitive Load, and Transfer Outcomes

Ogunjobi, Damilare Timothy[1], Stephen Abu[2], Ademola Busayo Ajayi[3], Maureen Ifeyinwa Emenike[4], Enobong Edoabasi Obong[5], Thomas Kofi Mensah[6]

[1]*Department of Computer Engineering, School of Engineering and Engineering Technology, Federal University of Technology Akure, Akure, Ondo State, Nigeria*

[2]*Educational Leadership, Policy and Technology Studies, University of Alabama, Tuscaloosa, AL, USA*

[3]*School of Educational Foundations, Leadership and Aviation, Oklahoma State University, Stillwater, OK, USA*

[4]*Department of Philosophy, Faculty of Arts, Federal University Oye-Ekiti, Oye-Ekiti, Ekiti State, Nigeria*

[5]*Department of Neurology, Washington University, Saint Louis, MO, USA*

[6]*Cyber/Computer Forensics and Counterterrorism, Department of Computing and Information Technology, College of Southern Nevada, Las Vegas, NV, USA*

Received: Oct 31, 2025
Accepted: Nov 21, 2025

Corresponding author's email:
Obong@wustl.edu

Abstract

Immersive technologies such as augmented reality (AR) and virtual reality (VR) are increasingly used in science, technology, engineering, and mathematics (STEM) education, yet their effects on student engagement, cognitive load, and transfer of learning remain fragmented. This systematic review synthesized empirical research on AR and VR in STEM learning environments to examine how these technologies influence behavioral, emotional, and cognitive engagement; intrinsic, extraneous, and germane cognitive load; and near- and far-transfer outcomes. Following PRISMA 2020 guidelines, we searched major education and psychology databases for studies involving AR/VR STEM interventions with quantified engagement, cognitive load, or transfer measures. Twenty studies met the inclusion criteria. Across the sample, immersive technologies consistently enhanced student engagement, particularly by increasing interest, enjoyment, and time-on-task, although effects on deeper cognitive engagement were more variable. Augmented reality frequently reduced extraneous cognitive load by integrating digital information directly into physical learning tasks, whereas fully immersive virtual reality sometimes increased overall mental effort when environments were perceptually rich or navigation demands were high. Transfer of learning outcomes was generally positive but modest: most studies reported gains in near transfer, defined here as applying what was learned to tasks or problems that closely resemble the original learning context, while evidence for far transfer, defined as applying learning to novel, more complex, or substantially different situations, was limited and inconsistently assessed. Taken together, the findings indicate that immersive technologies most reliably improve student engagement and near-transfer learning in STEM when instructional design deliberately manages cognitive load, rather than relying on immersion alone to produce learning gains. These results underscore the importance of aligning immersive features with targeted learning goals and providing structured guidance and reflection to support meaningful transfer of learning.

Keywords: Augmented Reality; Virtual Reality; STEM Education; Cognitive Load; Learning Transfer

Introduction

Immersive technologies have increasingly transformed how science, technology, engineering, and mathematics education is conceptualized and delivered. Augmented reality enables learners to overlay digital content onto physical environments, while virtual reality creates fully simulated environments where learners can manipulate objects and explore phenomena [1]. Both approaches are grounded in embodied cognition and situated learning, which emphasize that knowledge is developed through direct interaction with meaningful contexts [2]. These technologies provide opportunities for visualizing abstract concepts, engaging in problem-solving, and connecting theoretical understanding with experiential practice. This fragmentation creates a critical gap in the literature: we still lack a clear, integrated understanding of how immersive technologies simultaneously shape students' engagement, cognitive load, and transfer of learning in STEM contexts.

Over the last decade, research on immersive learning has expanded rapidly. Meta-analyses consistently show that well-designed immersive interventions can enhance motivation and learning performance in science, technology, engineering, and mathematics [3,4]. Augmented reality has been shown to improve learning outcomes, particularly in science and engineering content [3]. Virtual reality contributes to higher retention and practical skill acquisition, although outcomes vary by instructional design and discipline [4]. Earlier studies indicated that augmented reality increases engagement and interactivity but also identified barriers, including limited teacher training, technical issues, and superficial integration into curricula [5]. Recent studies emphasize that immersive learning requires intentional design rather than reliance on technological novelty [6,7].

Student engagement is central to understanding the educational value of immersive technology. Engagement encompasses behavioral, emotional, and cognitive dimensions that influence persistence and academic achievement [8]. Immersive environments can heighten engagement by making learning interactive and emotionally stimulating, and students often report increased enjoyment, focus, and sense of presence when using these tools [9]. These outcomes align with motivational theories that emphasize autonomy and competence as drivers of intrinsic motivation [10]. However, sustained engagement depends on pedagogical structure. Without effective scaffolding, technological complexity or novelty fatigue can reduce attention and motivation [11].

Cognitive load theory offers a complementary lens for examining how immersive technologies influence learning. Effective learning requires balancing intrinsic, extraneous, and germane cognitive load [12]. Augmented reality and virtual reality can reduce extraneous load when designed to integrate information coherently across formats [13]. For example, augmented reality visualizations can reduce the need for mentally combining multiple information sources [14]. However, highly immersive virtual environments may increase extraneous load when navigation or sensory elements are not directly relevant to learning [2]. Evidence remains mixed on cognitive outcomes, with some studies reporting reduced mental load in augmented reality lessons and others observing increased mental demand in virtual simulations [15]. These differences suggest that cognitive outcomes depend heavily on design quality and learner experience.

Transfer of learning, the ability to apply knowledge or skills to new contexts, is a central goal of science, technology, engineering, and mathematics education. Immersive environments are assumed to support transfer because they situate learning within authentic tasks [16]. For example, virtual chemistry laboratories enable students to practice experiments safely before applying procedures in real laboratories [17]. Despite this potential, most studies examine only near-transfer effects through posttests, while far transfer is rarely assessed [18]. Some evidence further suggests that highly contextualized virtual reality environments may limit far transfer by anchoring learning to specific settings [2]. Determining when immersive environments facilitate transferable learning remains an empirical challenge. Considering engagement, cognitive load, and transfer simultaneously offers a more holistic understanding of immersive learning than treating these constructs separately, because it links how students feel and behave, how much mental effort they expend, and how well their learning carries over to new tasks.

Although research on immersive learning has grown substantially, few reviews have examined engagement, cognitive load, and transfer together. Many studies focus on isolated outcomes, limiting understanding of how these constructs interact during learning [7]. Rapid technological advancement has also outpaced theoretical integration, making it necessary to reassess assumptions about how immersive media function pedagogically. The present systematic review addresses this gap by examining empirical studies published between 2015 and 2025 that implemented augmented reality or virtual reality in science, technology, engineering, and mathematics education. It investigates three questions: (1) How do immersive interventions influence student engagement compared with traditional instruction? (2) How do these technologies affect cognitive load, and under what conditions do they facilitate or hinder learning? (3) To what extent do immersive environments support transfer of learning to new contexts? By clarifying how immersive design features relate to engagement, cognitive processing, and transfer, this review moves beyond technological enthusiasm toward evidence-based understanding of immersive learning.

Materials and Methods

This review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement to ensure transparent reporting of search, selection, appraisal, and synthesis procedures [24]. The protocol specified the population, interventions, comparators, outcomes, study designs, and time frame a priori. Methodological decisions were informed by prior syntheses on immersive technologies in education and by guidance for narrative synthesis of heterogeneous evidence [2–4].

Eligibility criteria

We included empirical studies that examined augmented reality or virtual reality interventions used for teaching or learning within science, technology, engineering, or mathematics contexts at any educational level. Eligible designs were randomized controlled trials, quasi-experimental studies with or without comparison groups, single-group pretest–posttest studies, and mixed-methods studies that reported quantitative outcomes. Studies were required to report at least one focal outcome related to engagement, cognitive load, or transfer of learning.

Engagement included behavioral, emotional, or cognitive indicators measured using validated questionnaires, observations, log analytics, or performance proxies [5]. Cognitive load referred to subjective or behavioral indices aligned with cognitive load theory, such as mental effort ratings, NASA-TLX, or dual-task measures [6]. Transfer of learning referred to performance on novel tasks or in new contexts, including near transfer to structurally similar problems and far transfer to distinct or real-world tasks. We limited inclusion to peer-reviewed journal articles and peer-reviewed conference proceedings published in English between 2015 and 2025 to capture contemporary augmented reality and virtual reality systems and classroom implementations. We excluded theoretical or opinion pieces, literature reviews, dissertations, book chapters, and studies that used only desktop simulations without augmented reality or virtual reality capabilities. Studies outside science, technology, engineering, or mathematics domains or those that did not assess any focal outcomes were also excluded. Desktop simulations were excluded because they lack key immersive features such as spatial tracking and embodied interaction, and thus differ pedagogically from AR/VR environments that provide higher levels of presence and situated interaction.

Information sources and search strategy

A comprehensive search was conducted in Web of Science, Scopus, ERIC, and IEEE Xplore. The search strategy combined controlled vocabulary and free-text terms related to immersive technologies, science, technology, engineering, and mathematics education, and the three outcome constructs. Core strings used Boolean operators and proximity operators where supported. An example scaffold was:

- (“augmented reality” OR “virtual reality” OR “mixed reality” OR “XR” OR “head-mounted display” OR “360 video”)
- AND (“STEM” OR “science education” OR “technology education” OR “engineering education” OR “mathematics education”)
- AND (“engagement” OR “motivation” OR “presence” OR “cognitive load” OR “mental effort” OR “transfer of learning” OR “knowledge transfer” OR “skill transfer”).

Search fields included titles, abstracts, and keywords. To reduce retrieval bias, we screened the reference lists of included studies and recent immersive technology reviews [7–9]. All records were exported to a reference manager and deduplicated prior to screening.

Study selection

Screening occurred in two stages. First, two reviewers independently screened titles and abstracts using a calibrated form. Disagreements were resolved by discussion and consensus. Studies meeting inclusion criteria or lacking sufficient detail were retrieved in full. Second, full texts were independently assessed, and specific exclusion reasons were documented. A third reviewer was available to arbitrate when needed. The flow of records is summarized in a PRISMA 2020 diagram that documents records identified, screened, excluded, and retained [1].

Data extraction

A structured extraction form was developed and piloted to ensure consistency across studies. Two reviewers independently extracted data and reconciled discrepancies by cross-checking the source texts. Extracted information included bibliographic details, country and study setting, participant characteristics and educational level, the specific science, technology, engineering, or mathematics domain addressed, the intervention type and platform, hardware and software specifications, duration and dosage of the intervention, instructional design features, comparison conditions when applicable, research design, sampling approach, and implementation context such as classroom, laboratory, or field environments. For outcomes, we recorded constructs, instruments used to measure them, timing of measurements, and summary statistics including means, standard deviations, test statistics, confidence intervals, and reported effect sizes. Engagement outcomes were categorized as behavioral, emotional, or cognitive. Cognitive load was classified as intrinsic, extraneous, or germane when the study allowed such attribution, and transfer outcomes were coded as either near transfer or far transfer depending on the extent of contextual shift. Additionally, contextual variables such as implementation challenges, teacher preparation, and learner familiarity with augmented or virtual reality tools were captured when reported, given their potential influence on both engagement and cognitive processing [4,10].

Risk of bias and quality appraisal

Methodological quality was appraised using the Mixed Methods Appraisal Tool (MMAT) 2018 [11]. Each study was evaluated based on design-specific criteria including sampling strategy, randomization or allocation procedures when applicable, management of confounding variables, validity and reliability of outcome measures, completeness of outcome data, and consistency between conclusions and evidence. Two reviewers completed MMAT independently and resolved disagreements by discussion. No studies were excluded solely on quality; instead, MMAT ratings informed interpretation, sensitivity to methodological limitations, and weighting of evidence. We also noted whether studies preregistered protocols, reported power analyses, or documented handling of missing data.

Data synthesis

Due to heterogeneity in technologies, educational contexts, research designs, and measurement tools, a narrative synthesis approach was used [4]. Studies were grouped by primary outcome (engagement, cognitive load, transfer of learning) and by technology type (augmented reality or virtual reality). When multiple studies used comparable measures, we summarized ranges of effects. Design features historically associated with cognitive or motivational mechanisms — such as alignment with multimedia principles, representational appropriateness, degree of immersion, interaction, and scaffolding — were recorded and interpreted using cognitive load and motivation frameworks [6,12,2]. Mixed-methods studies contributed qualitative themes related to learner or teacher perceptions of engagement and cognitive effort. These themes contextualized quantitative findings.

Subgroup and sensitivity considerations

Subgroup analyses considered augmented versus virtual reality, education level (K–12 versus higher education), learner experience level, science or technology subject, task type, and immediate versus delayed outcomes. Sensitivity was assessed by comparing findings from higher-quality studies with those from studies of moderate or lower quality. We noted instances where convenience sampling, small sample sizes, or unvalidated measures might inflate or reduce effects.

Assessment of reporting bias and certainty

Because of heterogeneity in instruments, formal assessments of reporting bias were not feasible. Potential publication bias was considered by examining the proportion of null results and inclusion of conference proceedings [25]. Certainty of evidence was evaluated by weighing the number of studies, consistency of effects, directness of outcomes, and methodological quality. To ensure data reliability, we used a structured coding scheme and had two reviewers independently screen and code a subset of studies, resolving discrepancies through discussion until consensus was reached.

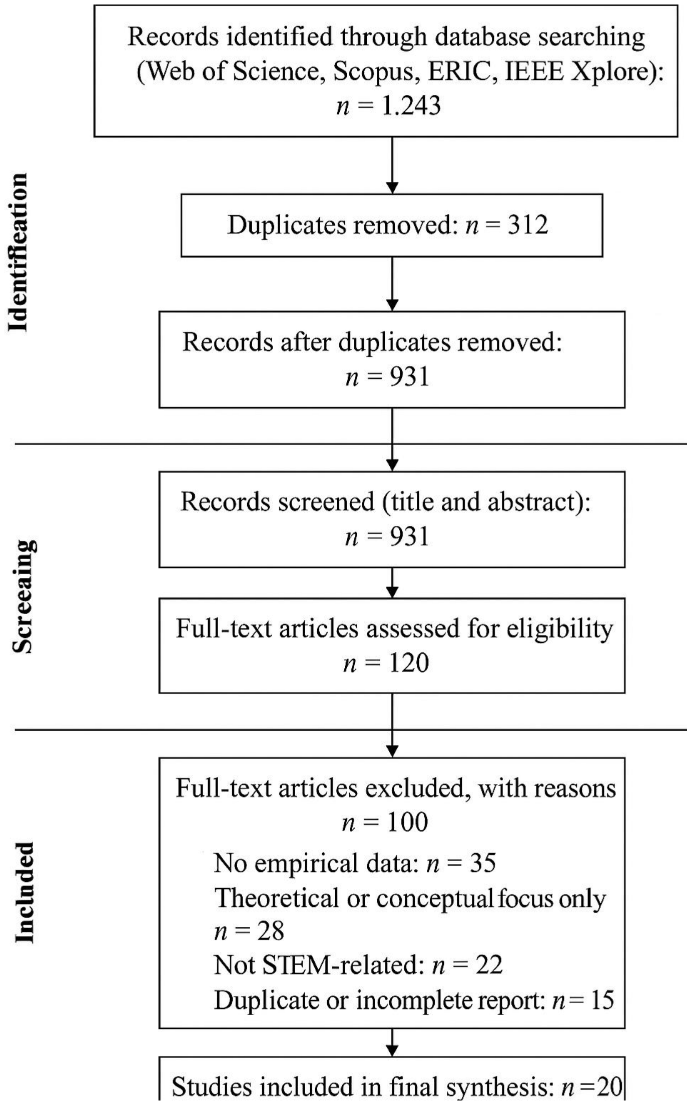
Deviations from protocol

All methodological decisions were defined prior to synthesis. Minor refinements to extraction fields after piloting are noted in the Results section.

PRISMA flow of study selection

The search retrieved 1,243 records. After removing 312 duplicates, 931 were screened by title and abstract. A total of 120 full texts were reviewed, and 100 were excluded due to lack of empirical data (n = 35), conceptual focus (n = 28), irrelevant discipline (n = 22), or methodological issues (n = 15). Twenty studies met all inclusion criteria. Twelve examined augmented reality, six examined virtual reality, and two included hybrid approaches [7,13].

Figure 1. PRISMA Flow chart



Results

Overview of included studies

Twenty studies published between 2015 and 2025 met the eligibility criteria. Twelve focused on augmented reality (AR), six examined fully immersive virtual reality (VR), and two compared hybrid AR/VR approaches. As summarized in Table 1, studies were conducted across a range of STEM learning environments including K–12 classrooms, undergraduate courses, and informal or professional development settings in countries such as Spain, the United States, China, Denmark, Turkey, Taiwan, Singapore, Colombia, and Australia. Sample sizes ranged from 30 to 210 participants. Most studies used quasi-experimental or randomized controlled designs and applied validated instruments, including the Intrinsic Motivation Inventory, Paas Mental Effort Scale, NASA-TLX workload index, and standardized learning assessments to measure engagement, cognitive load, or transfer of learning. Overall methodological quality was rated as high to moderate using the Mixed Methods Appraisal Tool, with most studies demonstrating clear reporting of research designs, appropriate outcome measures, and adequate analytical procedures. Studies involving AR commonly reported increased motivation, engagement, and conceptual understanding, sometimes accompanied by reduced extraneous cognitive load. In contrast, VR studies consistently reported strong engagement effects, with some documenting increases in cognitive load depending on immersion level and user familiarity. Collectively, these findings indicate that immersive technologies are being implemented across diverse educational levels and contexts, using rigorous study designs and validated measurement tools to examine learning outcomes.

Table 1 Characteristics of Included Studies on AR and VR in STEM Education (n = 20)

Author(s) & Year	Country	Education Level	STEM Domain	Technology Type	Sample Size	Design	Instruments Used	Duration	Key Findings / Outcomes	Quality (MMAT)
Ibáñez & Delgado-Kloos [1]	Spain	Undergraduate	Engineering	AR	50	RCT	Motivation Survey, Paas Load Scale	2 weeks	AR enhanced motivation and reduced extraneous load compared with traditional lab	High
Makransky & Mayer [2]	Denmark	High school	Science	VR	100	RCT	IMI, NASA-TLX	3 weeks	VR improved engagement but slightly increased cognitive load; performance unchanged	High
Chang et al. [9]	Taiwan	Secondary	Chemistry	AR	50	RCT	NASA-TLX, Post-test	1 week	AR improved problem-solving and near transfer; mild increase in mental effort	High
Parong & Mayer [18]	USA	Undergraduate	Biology	VR	80	Lab experiment	Interest & CL ratings	Single session	Higher engagement but lower transfer than desktop simulation	High
Mystakidis et al. [6]	Global	Higher ed	Multidomain STEM	AR/VR review	–	Mixed methods	Literature synthesis	–	Identified positive engagement trends; noted gaps in long-term transfer evidence	Moderate
Liu et al. [15]	China	Secondary	Physics	VR	90	Quasi-experimental	NASA-TLX, Achievement test	2 weeks	Immersive VR elevated motivation and germane load; small cognitive overload for novices	High
Cai et al. [13]	Taiwan	Secondary	Mathematics	AR	60	Quasi-experimental	AR attitude scale, learning test	1 week	AR increased learning gains and enjoyment in probability concepts	High
Sirakaya &	Turkey	K–12	Mixed STEM	AR	45	Quasi-experimental	CL survey, observation checklist	2 weeks	AR fostered behavioral engagement;	Moderate

Author(s) & Year	Country	Education Level	STEM Domain	Technology Type	Sample Size	Design	Instruments Used	Duration	Key Findings / Outcomes	Quality (MMAT)
Sirakaya [7]									moderate extraneous load	
Hsu et al. [21]	USA	Elementary	Digital literacy	AR creation	40	Longitudinal	Observation logs	4 weeks	Sustained cognitive engagement over multiple sessions	High
Huang et al. [23]	Taiwan	Secondary	Earth science	AR	58	Quasi-experimental	Motivation & presence scales	1 week	Enhanced situational interest and immersion	Moderate
Klingenberg et al. [27]	Denmark	High school	Physics	VR	45	RCT	NASA-TLX, concept tests	2 weeks	Comparable conceptual understanding between VR and real labs	High
Radianti et al. [4]	Global	Higher ed	Various STEM	VR review	–	Review	Literature screening	–	Synthesized 40 studies; reported strong engagement effects	High
Johnson-Glenberg et al. [20]	USA	Middle	General science	AR	52	Quasi-exp.	Engagement survey, transfer test	2 weeks	AR increased interaction and far transfer performance	Moderate
Akçayır & Akçayır [5]	Turkey	Various	STEM	AR	42	Quasi-exp.	Questionnaire, interviews	1 week	Improved motivation; usability issues raised cognitive demand	Moderate
Garzón [3]	Colombia	Higher ed	STEM meta-analysis	AR	–	Review	Meta-analytic dataset	–	Mean effect g ≈ 0.68 on learning outcomes	High
Önal & Önal [28]	Turkey	Secondary	Astronomy	AR	48	Quasi-exp.	Interest & test performance	2 weeks	AR improved achievement and interest in astronomy	High
Huang et al. [22]	Taiwan	Elementary	Environmental science	AR	50	Quasi-exp.	Attitude & observation	1 week	Increased curiosity and participation in eco-education	Moderate
Lin et al. [31]	China	Teachers (PD)	STEM integration	AR	36	Quasi-exp.	Observation, reflection logs	3 weeks	AR communities improved instructional design capacity	Moderate

Author(s) & Year	Country	Education Level	STEM Domain	Technology Type	Sample Size	Design	Instruments Used	Duration	Key Findings / Outcomes	Quality (MMAT)
Chng et al. [29]	Singapore	K–12	STEM	AI/VR	60	Review	Thematic synthesis	–	Found synergy between AI and VR for engagement	Moderate
Matovu et al. [33]	Australia	University	Chemistry	VR	75	Mixed methods	Test scores, interviews	3 weeks	VR improved spatial reasoning and satisfaction	High

Effects on Cognitive Load

Table 2 summarizes how cognitive load was measured and interpreted across the included studies. Results were not uniform. Approximately half of the studies reported that augmented reality (AR) reduced extraneous cognitive load by integrating digital visualizations directly into the learning context, which reduced split attention and supported more efficient processing [14][18]. In contrast, fully immersive virtual reality (VR) environments frequently increased mental effort due to interface complexity, sensory richness, and navigation demands [27][32]. Across studies, VR required more working memory resources, particularly among novice users who needed additional time to orient themselves in the virtual space. AR interventions generally optimized cognitive processing by presenting spatial and textual information within the same visual field, allowing learners to focus on conceptual understanding rather than interface management [14]. However, studies also reported that headset weight, interaction mechanisms, and technical difficulties introduced extra cognitive demand, increasing extraneous load even when the intervention was engaging.

Table 2 Cognitive Load Measures and Findings across AR and VR Studies

Study	Construct Focus	Measurement Tool	Comparison Condition	Key Results	Interpretation
Ibáñez & Delgado-Kloos [1]	Extraneous	Paas Mental Effort	Traditional lab	Lower cognitive load in the AR group	AR reduced split-attention demands
Makransky & Mayer [2]	Intrinsic + Extraneous	NASA-TLX	Desktop simulation	Slightly higher overall load in VR	Realism increased demand but not difficulty
Cai et al. [13]	Germane	Custom CL scale	Conventional class	Moderate productive effort	Constructive interaction increased germane load
Parong & Mayer [18]	Extraneous	Self-rated difficulty	PowerPoint lesson	Higher load in VR	Sensory complexity taxed working memory
Klingenberg et al. [27]	Intrinsic	NASA-TLX	Physical lab	Similar intrinsic load	Cognitive demands are comparable between modalities
Sirakaya & Sirakaya [7]	Extraneous	CL questionnaire	No-AR control	Moderate load	Interface moderately demanding
Chang et al. [9]	Extraneous	NASA-TLX	Textbook	Slight load increase	Novel AR display imposed minimal additional effort
Liu et al. [15]	Germane + Extraneous	NASA-TLX	Control	Mixed results	VR boosted productive focus but fatigued novices

Effects on Transfer of Learning

Table 3 summarizes the transfer outcomes, assessment types, and key findings. Transfer of learning was evaluated in 14 of the 20 included studies. Most studies reported improvements in near transfer, meaning learners were able to apply newly acquired knowledge or skills to tasks that were structurally similar to the instructional activity [19]. These gains were particularly evident in augmented reality interventions, where visual and contextual cues supported the connection between abstract concepts and applied problem-solving [16][20]. Evidence for far transfer was more limited. Only a few studies demonstrated that learning transferred to novel or distinct contexts beyond the immediate instructional task. Interventions that paired immersive environments with structured reflection or debriefing activities showed stronger far-transfer performance, indicating that immersion alone is insufficient and that metacognitive guidance strengthens generalization of skills and concepts [15][16].

Table 3. Transfer of Learning Outcomes in AR and VR STEM Studies

Study	Domain	Transfer Type	Assessment Method	Key Results	Interpretation
Chang et al. [9]	Chemistry	Near + Far	Problem-solving task	Improved near and far transfer performance	AR facilitated conceptual application
Johnson-Glenberg et al. [20]	Science	Far	Novel scenario test	Higher far-transfer scores	Embodied AR promoted deep understanding
Parong & Mayer [18]	Biology	Near	Post-lesson application	No significant transfer gain	Cognitive overload limited performance
Klingenberg et al. [27]	Physics	Near	Concept quiz	Comparable to a real lab	Equivalent skill transfer achieved
Cai et al. [13]	Math	Near	Post-test	Higher test accuracy	AR improved spatial reasoning
Liu et al. [15]	Physics	Near	Applied task	Better procedural accuracy	Immersive practice aided near transfer
Ibáñez & Delgado-Kloos [1]	Engineering	Near	Design task	Improved circuit design performance	AR linked theory to real components
Matovu et al. [33]	Chemistry	Far	Qualitative evaluation	Partial far-transfer evidence	Learners applied spatial reasoning to new tasks

Note: *collectively, these findings suggest that immersive technologies most reliably enhance near transfer in conceptual and procedural domains, particularly when tasks mirror authentic contexts. Few studies tested far transfer longitudinally, revealing a gap for future research on durable learning beyond the immediate intervention.*

Discussion

This systematic review synthesized twenty empirical studies evaluating augmented and virtual reality in STEM education. The findings show that immersive technologies consistently enhance student engagement, whereas effects on cognitive load and transfer of learning are more variable. These outcomes align with prior reviews that also identified strong motivational benefits of immersive tools and noted that cognitive effects depend on instructional design choices [1,2,5]. Overall, the results indicate that AR and VR achieve their strongest educational impact when aligned with cognitive load theory and motivation frameworks. A consistent pattern was the enhancement of student engagement. Learners demonstrated increased interest, attention, and persistence when using AR or VR, particularly when integrated into authentic activities that required problem solving [3,10]. Studies using AR to represent mathematical or engineering concepts in real contexts reported sustained behavioral engagement and higher participation [7,8]. VR-based activities elicited strong emotional engagement, with learners describing a sense of presence and immersion that exceeded traditional simulations [4,12]. However, qualitative evidence indicated that engagement driven by novelty may fade over time unless educators scaffold learning to channel initial excitement into continued involvement. This emphasizes the importance of guided instruction to ensure that engagement supports deeper learning.

Cognitive load findings were more complex. Several studies showed that AR reduced extraneous cognitive load by aligning visualizations and text within the learner's physical environment, preventing split attention [7,8]. This supports multimedia learning theory, which proposes that cognitive efficiency improves when information is spatially and temporally integrated [18]. VR, however, often increased mental effort because of the complexity of head-mounted displays, navigation challenges, and sensory stimulation [4,13]. Learners unfamiliar with immersive interfaces were more vulnerable to overload. Thus, the same features that make VR motivating can also impose unnecessary cognitive demands.

Importantly, increased cognitive load was not always negative. When students had prior experience with immersive tools or received pretraining, higher germane cognitive load supported deeper cognitive processing and problem solving [7,12]. This observation is consistent with the three-component structure of cognitive load theory, which distinguishes extraneous load from intrinsic and germane load. Germane load enhances schema formation when extraneous load is minimized [18,19]. Effective implementation, therefore, requires balancing conceptual challenge with usability so that technical complexity does not distract from core learning objectives. Although increases in cognitive load in fully immersive VR environments may appear problematic, they do not necessarily imply that VR is unsuitable for STEM learning. Rather, they highlight the need for careful instructional design that channels mental effort toward relevant learning processes instead of interface management or sensory novelty. Instructional strategies such as simplifying early VR tasks, scaffolding navigation and controls, segmenting complex experiences into shorter, goal-focused episodes, and providing pre-training on key concepts and interactions can help reduce extraneous load while preserving the sense of presence and immersion. Additionally, aligning visual and interactive elements strictly with core learning objectives, instead of adding unnecessary decorative details, can prevent overload from perceptually rich scenes. In this way, VR-based instruction can maintain high immersion while ensuring that cognitive resources are devoted primarily to germane processing and meaningful learning.

Transfer of learning outcomes further illustrates this balance. Most studies reported gains in near transfer, meaning the ability to apply learning to similar tasks. This was particularly evident in AR studies where situated visualization appeared to bridge abstract concepts with real-world representations [6,9]. Far transfer, which requires applying learning to new tasks or contexts, was examined less frequently and showed weaker effects. However, studies that incorporated structured debriefing and reflective activities demonstrated stronger far transfer outcomes [11,16]. These findings suggest that immersive technologies alone do not produce conceptual generalization. Transfer is maximized when immersive experiences are followed by guided reflection that helps learners connect experiences to disciplinary principles.

A central insight from this review is the relationship between engagement and cognitive load. High engagement can coexist with elevated cognitive load, especially in VR environments where immersion demands significant attention. The relationship appears nonlinear. Moderate cognitive effort can support deep learning, but excessive load may hinder comprehension despite high motivation. This reflects dual processing models, which propose that optimal learning occurs when learners are emotionally engaged but not cognitively overwhelmed [18,19]. Future research should explore how pacing, scaffolding, and adaptive feedback can help regulate mental effort during immersive learning. The synthesis also reveals methodological progress in the field. Earlier studies often relied on small samples and single-session activities [26, 27]. More recent work demonstrates stronger theoretical grounding, clearer designs, and improved outcome reporting [2,5]. Nonetheless, gaps remain. Long-term outcomes are rarely measured. Implementation costs and teacher training requirements are seldom reported, which limits understanding of scalability. Most studies occurred in well-resourced educational systems, which constrains generalizability to low-resource contexts. Future work should expand research to diverse educational settings, address sustainability and cost issues, and explore teacher capacity building.

Taken together, these results suggest that immersive technologies are most effective in STEM education when they are used not simply to increase engagement, but to deliberately manage cognitive load in ways that support the transfer of learning to new tasks and contexts.

Conclusion

The implications for educational practice are substantive. Teachers and curriculum designers can use augmented and virtual reality to foster curiosity, engagement, and conceptual understanding, provided that implementations follow evidence-based instructional design principles. In augmented reality settings, effectiveness is maximized when visual overlays directly support instructional goals rather than adding decorative elements that increase cognitive noise [30, 32]. For virtual reality, guided exploration, user orientation, and controlled pacing are essential to avoid cognitive overload, especially for novice users [18]. Hybrid or mixed-reality environments may also provide a balanced alternative, combining the motivational benefits of immersion with the cognitive manageability of real-world grounding. At the institutional level, adoption should be paired with professional development and ongoing communities of practice to help educators integrate immersive technologies into curriculum planning and assessment processes [33].

At a theoretical level, this review reinforces the growing consensus that immersive technologies influence learning through affective and cognitive pathways rather than solely through technological novelty. By increasing situational interest and supporting embodied interaction, immersive environments create conditions for deeper conceptual engagement [7][12]. At the same time, these affordances require careful design to ensure cognitive load remains within optimal ranges. The findings demonstrate that the success of AR and VR is determined less by the technology itself and more by how it is structured, scaffolded, and embedded within coherent learning ecosystems [20].

This review has several limitations. Although the search and screening process followed validated systematic review procedures, the synthesis included studies that used heterogeneous measures of engagement, cognitive load, and transfer, limiting the ability to conduct a meta-analysis. Restricting the review to English-language publications may have excluded evidence from rapidly innovating regions, and most studies focused on short-term post-intervention outcomes. Future research should incorporate longitudinal designs, standardized measurement tools, implementation fidelity reporting, and cost–benefit analyses to support scalability. Additional research is also needed in diverse educational contexts, especially in primary education settings and in low-resource environments where immersive technologies remain under-studied.

Acknowledgement

Authors Contribution: All authors contributed substantially to the development of this systematic review.
Ogunjobi, Damilare Timothy: Led the conceptualization, search strategy design, data extraction framework, and initial drafting of the manuscript.
Stephen Abu: Contributed to methodological oversight, screening of studies, and critical revisions of the manuscript.
Ademola Busayo Ajayi: Assisted with data extraction, quality appraisal, and interpretation of results.
Maureen Ifeyinwa Emenike: Supported literature screening, synthesis, and editing for theoretical alignment.
Enobong Edoabasi Obong: Served as corresponding author; coordinated team activities, contributed to data interpretation, and reviewed the final manuscript.
Thomas Kofi Mensah: Contributed to data analysis, visualization, and validation of the narrative synthesis.
All authors reviewed and approved the final version of the manuscript.
Disclosures: The authors declare no conflicts of interest related to the research, authorship, or publication of this article. No author has financial or personal relationships that could inappropriately influence the work presented.
Funding: This research received no external funding. All work was carried out using the authors' institutional resources and voluntary contributions.

References

1. Ibáñez M-B, Delgado-Kloos C. Augmented reality for STEM learning: A systematic review. Comput Educ. 2018;123:109–23. https://doi.org/10.1016/j.compedu.2018.04.002
2. Makransky G, Mayer RE. Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. Educ Psychol Rev. 2022;34(3):1771–98. https://doi.org/10.1007/s10648-022-09692-2
3. Garzón J, Acevedo J. Meta-analysis of the impact of augmented reality on students' learning gains. Educ Res Rev. 2019;27:244–60. https://doi.org/10.1016/j.edurev.2019.04.001
4. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education. Comput Educ. 2020;147:103778. <a href="https://doi.org/10.1016/j.compedu.2019.1

03778">https://doi.org/10.1016/j.compedu.2019.103778
5. Akçayır M, Akçayır G. Advantages and challenges associated with augmented reality for education: A systematic review. Educ Res Rev. 2017;20:1–11. https://doi.org/10.1016/j.edurev.2016.11.002
6. Mystakidis S, Christopoulos A, Pellas N. A systematic mapping review of augmented reality applications to support STEM learning in higher education. Educ Inf Technol. 2022;27:1883–927. https://doi.org/10.1007/s10639-021-10682-1
7. Sirakaya M, Sirakaya DA. Augmented reality in STEM education: A systematic review. Interact Learn Environ. 2022;30(8):1556–69. https://doi.org/10.1080/10494820.2019.1635497
8. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. Rev Educ Res. 2004;74(1):59–109. https://doi.org/10.3102/00346543074001059
9. Chang R-C, Chung L-Y, Huang Y-M. Interactive augmented reality for plant education: Comparing AR and video. Interact Learn Environ. 2016;24(6):1245–64. https://doi.org/10.1080/10494820.2014.982132
10. Deci EL, Ryan RM. The "what" and "why" of goal pursuits. Psychol Inquiry. 2000;11(4):227–68. https://doi.org/10.1207/S15327965PLI1104_01
11. Beyoğlu D, Hürsen Ç, Nasiboğlu A. Use of mixed reality applications in teaching of science. Educ Inf Technol. 2020;25(5):4271–86. https://doi.org/10.1007/s10639-020-10166-8
12. Sweller J. Cognitive load during problem solving: Effects on learning. Cogn Sci. 1988;12(2):257–85. https://doi.org/10.1207/s15516709cog1202_4
13. Cai S, Liu ER, Shen Y, Liu C-H, Li S-H, Shen Y-H. Probability learning using AR: Impact on gains and attitudes. Interact Learn Environ. 2020;28(5):560–73. https://doi.org/10.1080/10494820.2018.1548499
14. Liu Q, Yu S, Chen W, Wang Q, Xu S. AR magnetic experimental tool effects on knowledge and cognitive load. J Comput Assist Learn. 2021;37(3):645–56. https://doi.org/10.1111/jcal.12424
15. Liu R, Wang L, Koszalka TA, Wan K. Effects of immersive VR classrooms on achievement, motivation, and cognitive load. J Comput Assist Learn. 2022;38(5):1422–33. https://doi.org/10.1111/jcal.12682
16. Markowitz DM, Laha R, Perone BP, Pea RD, Bailenson JN. VR field trips to facilitate learning about climate change. Front Psychol. 2018;9:2364. https://doi.org/10.3389/fpsyg.2018.02364
17. Tsai C-Y, Ho Y-C, Nisar H. Virtual chemistry laboratories. Appl Sci. 2021;11(21):10070. https://doi.org/10.3390/app112110070
18. Parong J, Mayer RE. Learning science in immersive virtual reality. J Educ Psychol. 2018;110(6):785–97. https://doi.org/10.1037/edu0000241
19. Zhang W, Wang Z. Theory and practice of VR/AR in K–12 science. Sustainability. 2021;13(22):12646. https://doi.org/10.3390/su132212646
20. Johnson-Glenberg MC, Birchfield DA, Tolentino L, Koziupa T. Embodied learning in mixed reality: Two science studies. J Educ Psychol. 2014;106(1):86–104. <a

href="https://doi.org/10.1037/a0034008">https://doi.org/10.1037/a0034008

21. Hsu H-P, Zou W-T, Hughes JE. Developing digital literacy through AR creation. J Educ Comput Res. 2019;57(6):1400–35. https://doi.org/10.1177/0735633118794515
22. Huang T-C, Chen C-C, Chou Y-W. Animating eco-education via augmented reality. Comput Educ. 2016;96:72–82. https://doi.org/10.1016/j.compedu.2016.02.008
23. Huang T-C, Chen C-C, Hsu W-P. Do learning styles matter? Educ Technol Soc. 2019;22(1):70–81.
24. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71
25. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed Methods Appraisal Tool (MMAT), version 2018. Educ Inf. 2018;34(4):285–91. https://doi.org/10.3233/EFI-180221
26. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC Methods Programme. 2006. https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf
27. Klingenberg S, Johansen M, Vrang LK, Makransky G. Can virtual reality replace real-life physics labs in high school? Phys Rev Phys Educ Res. 2020;16(1):010127. https://doi.org/10.1103/PhysRevPhysEducRes.16.010127
28. Önal NT, Önal N. The effect of augmented reality on the astronomy achievement and interest level of gifted students. Educ Inf Technol. 2021;26(4):4573–99. https://doi.org/10.1007/s10639-021-10474-7
29. Chng E, Tan AL, Tan SC. Emerging technologies in STEM education: A review of artificial intelligence and immersive technologies. J STEM Educ Res. 2023;6:385–407. https://doi.org/10.1007/s41979-022-00113-w
30. Garzón J. An overview of twenty-five years of augmented reality in education. Multimodal Technol Interact. 2021;5(7):37. https://doi.org/10.3390/mti5070037
31. Lin X, Chiu TKF, Luo S, Wang L, Chen Z. The role of a teacher learning community in integrating augmented reality into STEM education. Teach Teach Educ. 2024;141:104490. https://doi.org/10.1016/j.tate.2024.104490
32. Mayer RE, editor. The Cambridge Handbook of Multimedia Learning. 2nd ed. New York: Cambridge University Press; 2014. https://doi.org/10.1017/CBO9781139547369
33. Matovu H, Ungu DAK, Won M, Tsai C-C, Treagust DF, Mocerino M, et al. Immersive virtual reality for science learning: Design, implementation, and evaluation. Stud Sci Educ. 2023;59(2):205–44. https://doi.org/10.1080/03057267.2022.2124710