Journal of Digital Media and Artificial Intelligence
Vol. 3 No. 2 (2026) PP 99- 118.

Human-AI Symbiotic Learning Spaces: An Integrative Review of Applications, Future Innovations, and Responsible Governance

Samah Zakareya Ahmad*[1]

[1]Department of English, College of Science and Humanities-Jubail, Imam Abdulrahman bin Faisal University, Dammam, KSA.

Received 2 April 2026
Revised 26 April. 2026
Accepted 6 May. 2026

Published 1 May. 2026

Cited as:
S. Zakareya Ahmad, Human-AI Symbiotic Learning Spaces: An Integrative Review of Applications, Future Innovations, and Responsible Governance, J. Dig Media. AI. Vol. 3 No. 2 (2026) PP 99- 118.
DOI: 10.18576/jdmai /030202

Abstract: This integrative review synthesizes evidence on how artificial intelligence is reshaping education. Drawing on 185 empirical and conceptual studies published between 2011 and 2025, it combines systematic search with critical, theory-informed synthesis. Established applications—intelligent tutoring systems, adaptive platforms, automated writing evaluation, learning analytics dashboards, and simulation-based training—yield small-to-moderate gains, especially when tightly embedded in curriculum and pedagogy. It addresses fragmentation between effectiveness studies, speculative futures, and ethical analyses, lacking a unifying framework linking learning sciences theory with AI design and governance. Following Torraco's methodology, the review involved multi-database search, dual-screening, quality appraisal, and thematic synthesis integrating quantitative effects with qualitative themes. Findings confirm these tools as core infrastructure, with emerging generative and multimodal AI, retrieval-augmented architectures, cross-course agents, and institutional digital twins shifting toward collaborative companions and analytics centers—while raising concerns over surveillance, inferential privacy, cognitive offloading, and bias. The review proposes the Human-AI Symbiotic Learning Space, a multidimensional framework organized around Agency, Alignment, Adaptivity, and Accountability to guide design, implementation, and governance. Intended for researchers, practitioners, and policymakers, it provides actionable guidance for AI systems enhancing learning while protecting equity and autonomy. It advocates scripted human–AI collaboration over educator replacement, coupling gains with protections for autonomy, equity, and oversight, plus testable hypotheses for future research.

Keywords: Artificial intelligence in education; Human–AI symbiosis; Generative AI and multimodal AI; Learning analytics dashboards; Simulation-based training; Intelligent tutoring systems

1. Introduction

The early twenty-first century has witnessed rapid technological change, with Artificial Intelligence (AI) becoming a transformative force across sectors (Russell & Norvig, 2021). While industries have swiftly adopted AI, education is still negotiating issues of equity, effectiveness, and pedagogy (Chen et al., 2020). Traditional teaching models face challenges in balancing scalability and personalization (Anderson & Dron, 2011). AI addresses this through adaptive technologies such as intelligent tutoring systems, learning

*Corresponding author E-mail: szamohammed@iau.edu.sa

analytics dashboards, and automated assessments (Mousavinasab et al., 2021; Luckin et al., 2022). These tools enhance personalized learning and redefine teacher roles, despite ongoing concerns about privacy and bias.

Despite the growing volume of empirical work on these technologies, there remains a critical need for comprehensive analyses that systematically map the current landscape while projecting future trajectories. Existing reviews often emphasize either historical effectiveness or hypothetical future scenarios, and few synthesize these perspectives in ways that directly inform policy and practice. This article addresses this gap by offering a holistic analysis of AI in Education (AIED) (2011–2025), integrating insights from learning sciences and AI research to articulate forward-looking pathways for the responsible use of AI.

To navigate this emerging frontier, the following research questions guide the current investigation:

1. What are the primary, established applications of AI currently being utilized across different domains of education?
2. What are the most significant prospective applications and emerging technologies expected to reshape institutional efficiency in the near future?
3. What are the major ethical considerations and policy implications associated with the responsible integration of these current and prospective AI applications?
4. The primary conceptual contribution of this research is the proposal of the Human–AI Symbiotic Learning Space, a multidimensional framework designed to evaluate how AI can augment human potential through the lenses of Agency, Alignment, Adaptivity, and Accountability.

1.1 Definition and Conceptual Foundations of AIED

AI involves designing systems that perform human-like cognitive tasks such as reasoning and language understanding (Russell & Norvig, 2021). In education, it supports teaching, learning, and administration through systems that perceive, reason, and learn (Ferster, 2014). Drawing on machine learning, natural language processing, and learning analytics (Lee & Kwon, 2024), AI advances beyond static technologies toward adaptive systems that model learners and make real-time, context-sensitive instructional decisions (Nguyen, 2023).

AIED draws on theories portraying learning as active, constructive, and socially situated. Constructivist, self-regulatory, socio-cultural, connectivist, and cognitive frameworks provide a multidimensional lens for human-AI interplay, anchoring AI's potential. Constructivism posits learners build knowledge through interaction within the Zone of Proximal Development (ZPD) (Vygotsky, 1978); Intelligent Tutoring Systems (ITS) and Generative AI (GenAI) apply this via adaptive scaffolding, prompting, and feedback to deepen understanding (VanLehn, 2011; Holmes et al., 2019). Self-Regulated Learning (SRL) emphasizes planning, monitoring, and reflection (Zimmerman, 2000); AI tools like Learning Analytics Dashboards (LADs) and adaptive models support metacognition and load regulation (Verbert et al., 2014; Wang & Lajoie, 2023). Socio-cultural theory views learning as mediated by cultural tools and collaboration (Vygotsky, 1978); AI enhances co-construction through dialogue management and group structuring (Rosé et al., 2008). Connectivism frames learning as navigating networks (Siemens, 2005); AI curates connections via recommenders and adaptive pathways (Brusilovsky, 2001). Cognitive Load Theory (CLT) stresses managing cognitive load for schema building (Ouwehand et al., 2025), though excessive AI support risks cognitive offloading and superficial learning (Kelly & Risko, 2019).

1.2 Current Applications of AIED

AI technologies are reshaping traditional paradigms of teaching and learning. From K–12 settings to higher education, vocational training, and beyond, these innovations demonstrate both measurable learning improvements and emerging challenges related to transparency, equity, and pedagogical alignment.

Classroom-facing AI applications span intelligent tutoring systems (ITS), adaptive learning platforms, automated writing evaluation (AWE), learning analytics dashboards (LADs), AI-supported assessment, simulation-based training, and generative AI companions. Across these tools, reported effects on achievement and skill performance are generally small to moderate (0.4–0.6 SD for ITS under strong conditions; medium effects for AWE in process-oriented pedagogy), with the most consistent gains accruing to lower-performing learners when AI is tightly integrated into curriculum and accompanied by active teacher orchestration. Detailed evidence, effect ranges, and risk profiles are presented in Section 3.1 and Table 1.

Micro-credential ecosystems increasingly rely on recommender systems to create adaptive learning pathways informed by prior achievements and career objectives (Gašević et al., 2015; Joksimović et al., 2018). While these systems enable efficient progression, they often emphasize completion metrics rather than individualized learning goals, raising concerns about transparency and learner autonomy (Drachsler & Kalz, 2016; Tsai et al., 2020). Similarly, language-learning platforms leverage spaced repetition and reinforcement learning to personalize tasks and offer real-time feedback on pronunciation and grammar, achieving outcomes comparable to traditional instruction for foundational skills (Loewen et al., 2019; Vesselinov & Grego, 2012; Verspoor et al., 2021). Despite the benefits of gamification and dashboards that encourage self-regulated learning (Godwin-Jones, 2017), micro-task formats may foster superficial engagement or "gaming the system," limiting transfer to authentic communication (Sockett & Toffoli, 2012; Loewen et al., 2019).

Within workplace and professional learning, AI operates as just-in-time performance support, offering adaptive guidance and skill reinforcement. In fields such as healthcare and aviation, these systems adjust to learners' cognitive and emotional states to improve procedural fluency and retention (Graesser et al., 2012; Sottilare et al., 2018). AI-powered knowledge management tools similarly recommend resources aligned with emerging skill demands, promoting self-directed upskilling (Majumder & Dey, 2022; Milligan & Littlejohn, 2014). Nevertheless, their effectiveness often hinges on organizational culture and user trust in AI-generated content (Castañeda & Selwyn, 2018).

1.3 Prospective Applications and Emerging Technologies in AIED

Recent research highlights a decisive shift from narrow automation toward more sophisticated forms of human–AI collaboration, intelligent learning ecosystems, and integrated analytics infrastructures that are expected to transform both learning and institutional operations (Chiu & Rospigliosi, 2025; Adil, 2025; Ge, 2025). As educational systems evolve, attention increasingly turns to how generative and multimodal AI, retrieval-augmented generation frameworks, digital twins, and AI-driven curriculum design can converge within a cohesive pedagogical and governance model (Holmes et al., 2022; United States Department of Education [USDOE], 2023; Xiaoyu et al., 2025). These developments represent a pivotal moment in reimagining how AI supports teaching, learning, and assessment through human-centered, transparent, and ethically grounded approaches.

GenAI has become a cornerstone of educational technology, offering tools for idea generation, feedback, problem-solving, and multimodal content creation. When GenAI use is scaffolded, learners experience notable gains in productivity and creativity (Alfarwan, 2025; Li, 2025; Song et al., 2025). The evolution of these systems reflects a shift from isolated chatbots toward integrated "learning companions" embedded within institutional platforms, providing on-demand support (Holmes et al., 2022; Bozkurt et al., 2024). Multimodal Large Language Models (MLLMs), combining text, vision, and speech, further extend these capabilities by enabling dialogic tutoring, visual explanations, and accessibility-sensitive interaction (Beale, 2025; Lee et al., 2024). These tools align with human-centered, co-creative learning models in which AI augments teacher and peer collaboration (Holmes et al., 2022; Luckin, 2018). When deployed as simulation or dialogue partners, GenAI fosters metacognition, argumentation, and critical thinking (Borge et al., 2024; Helal et al., 2025; Singh et al., 2025). However, successful implementation requires transparent policies, integrity safeguards, and AI literacy (Luckin, 2018; Bozkurt et al., 2024).

In K–12 STEM and mathematics education, multimodal AI plays a critical role in addressing the complexity of diagram-rich content, although many models still struggle with authentic diagram understanding (Caffagni et al., 2024; Zhang et al., 2024). Emerging education-oriented MLLMs such as UniEDU aim to deliver unified, efficient processing for image-based question types (Chu et al., 2025), while data mining approaches leverage these models to extract knowledge components and improve transparency (Moon et al., 2025). Collectively, these advancements signal the emergence of integrated, grounded, and evaluable AI systems designed for authentic educational use.

Retrieval-Augmented Generation (RAG) models further enhance institutional control over information quality by minimizing hallucinations and ensuring policy alignment (Swacha & Gracel, 2025). Improvements in retrieval and reranking methods significantly reduce errors (Zhou et al., 2024; Xu et al., 2025), and practical deployments show that RAG supports smaller, cost-efficient models while maintaining content provenance and auditable bias control (Ayala & Bechard, 2024). AI roadmaps anticipate a transition from isolated intelligent tutors to cross-course tutoring and coaching agents linked by integrated learner models (Holmes et al., 2022; Adil, 2025). These agents support inquiry and project-based learning by mediating peer feedback, collaborating dynamically with learners, and adapting instructional roles (Chiu & Rospigliosi, 2025; Ge, 2025). Hybrid constellations of pedagogical and administrative agents will coordinate routine academic and operational support, allowing educators to focus on nuanced judgment and socio-emotional guidance (Luckin, 2018; Ge, 2025).

Institutional transformation is also being powered by LADs, digital twins, and integrated "nerve centres" that synthesize data for real-time decision-making. Educational digital twins draw from sensor and analytics data to produce multidimensional campus views that enable predictive maintenance, policy simulation, and learning optimization (Jones et al., 2020; Xie et al., 2024). These infrastructures merge information from multiple systems into dynamic dashboards supporting early intervention, resource planning, and strategic governance (USDOE, 2023; Zawacki-Richter et al., 2019). Coupled with process mining and automation, they enhance administrative efficiency and enable more flexible credentialing pathways (Adil, 2025).

AI is also increasingly central to curriculum, assessment, and program design, supporting the co-development of educational materials aligned with standards and responsive to learner data (Ge, 2025). Generative approaches enable dynamic rubric generation, formative feedback, and problem variation (Chiu & Rospigliosi, 2025; Xiaoyu et al., 2025). Looking ahead, competency-based, adaptive assessment ecosystems will integrate digital credentialing frameworks to reflect evolving labour market demands (Adil, 2025; USDOE, 2023).

At the core of these transformations lies human–AI collaboration, emphasizing pedagogical innovation that enhances diagnostic and design capacities while fostering learner agency and co-creation (Chiu & Rospigliosi, 2025; Ge, 2025). Future institutional strategies will intertwine "learning with," "learning about," and "preparing for" AI, scaffolding these dimensions through professional development, infrastructure, and governance (Holmes et al., 2022; USDOE, 2023). The near future thus points to interconnected socio-technical systems reshaping educational roles, learning experiences, and organizational practices (Adil, 2025; Zawacki-Richter et al., 2019).

1.4 Ethical, legal, and socio-technical challenges of AIED

AIED presents interrelated ethical, legal, and socio-technical challenges involving data use, algorithmic decisions, and institutional governance (Dignum, 2021). Key concerns include the "democratisation of surveillance" through inferential profiling, cognitive offloading undermining metacognitive development, persistent algorithmic bias disadvantaging marginalised learners, and implementation inequities shaped by uneven infrastructure and professional capacity (Gerlich, 2025; Risko & Gilbert, 2016; Boateng & Boateng, 2025; Wieczorek et al., 2025). These challenges are elaborated in Section 3.3, which presents the empirical evidence base and proposed safeguards.

2. Methodology

This study used an integrative review to connect fragmented efficacy research, speculative AI applications, and ethical-policy analyses across K–12, higher, vocational, and lifelong learning. Guided by Torraco's framework, it systematically mapped established and emerging AI uses and examined their ethical, legal, and socio-technical implications.

2.1 Design and standards

An integrative review was conducted following Torraco's (2016) four phases: (a) conceptual formulation, (b) systematic search, (c) data extraction and quality appraisal, and (d) critical, theory-informed synthesis.

2.2 Conceptual formulation

A priori concepts drawn from the learning sciences and AI governance scholarship—such as adaptivity, self-regulated learning, surveillance and inferential privacy, human–AI collaboration, and institutional governance—were used to frame the review questions and extraction fields. These sensitizing constructs were organized under the Human–AI Symbiotic Learning Space dimensions (Agency, Alignment, Adaptivity, Accountability), providing a scaffold for later mapping of empirical findings to mechanisms and design–governance propositions in the Results and Discussion.

2.3 Systematic search

Searches were conducted across Scopus, Web of Science, ERIC, PsycINFO, and selected publisher portals for the period 2011–2025, using keyword families covering: (a) AI application types (*intelligent tutoring system*, *adaptive learning*, *automated writing evaluation*, *learning analytics*, *educational simulation*, *generative AI*, *large language model*, *retrieval-augmented generation*, *digital twin*); (b) educational context (*K–12*, *higher education*, *vocational training*, *lifelong learning*, *MOOC*); (c) outcomes (*achievement*, *self-regulated learning*, *writing quality*, *engagement*, *equity*, *bias*); and (d) governance and ethics (*algorithmic bias*, *inferential privacy*, *surveillance*, *responsible AI*). Boolean combinations used AND/OR logic (e.g., "artificial intelligence AND education") across Scopus, Web of Science, ERIC, PsycINFO, and selected publisher portals. The search targeted developments from rule-based to generative and multimodal systems, including intelligent tutoring, adaptive learning, analytics dashboards, simulations, and policy frameworks.

Brief backward and forward citation chasing expanded coverage, resulting in 3,361 unique records for initial screening.

2.4 Screening and inclusion

Two reviewers independently screened 3,361 records, excluding 2,784 during title–abstract review and 392 after full-text assessment, resulting in 185 studies meeting the inclusion criteria across educational levels, study designs, and AIED relevance. Inclusion criteria required that studies: (a) reported on an AI application in a formal, non-formal, or informal educational context; (b) were published between 2011 and 2025 in English; (c) employed an empirical, conceptual, or systematic design; and (d) addressed at least one of the following: learning outcomes, AI design features, ethical or governance considerations. Studies were excluded if they: (a) focused exclusively on technical AI performance without educational application; (b) were conference abstracts without full-text data; (c) were opinion pieces or editorials lacking substantive evidence; or (d) examined non-educational AI contexts. Disagreements between the two reviewers during screening (κ = .81, indicating strong agreement) were resolved through discussion until consensus was reached. Screening decisions followed predefined criteria, with disagreements resolved through discussion, and the full selection process is detailed in the PRISMA-style flow diagram (Figure).

2.5 Data extraction

For each study, data were extracted on context (sector, subject, setting), AI type and maturity (e.g., ITS, LADs, AWE, simulations, generative/MLLM tools, RAG, digital twins), research design, sample characteristics, measures, and key outcomes (achievement, writing quality, engagement, SRL, equity/fairness/privacy indicators, teacher and learner agency). Extraction templates were piloted and refined to ensure that variables necessary to address RQ1–RQ3 and to map findings to Agency, Alignment, Adaptivity, and Accountability were captured consistently across quantitative and qualitative evidence.

2.6 Quality appraisal

Study quality was assessed using RoB 2 for RCTs, ROBINS-I for non-randomized and quasi-experiments, and the CASP Qualitative Checklist (CASP, 2018) for qualitative studies. High-quality evidence informed claims about efficacy and mechanisms, while studies with serious or critical bias were used only for contextual or ethical insights. Qualitative studies with strong CASP ratings enriched understanding of implementation, agency, and governance rather than supporting causal claims.

2.7 Thematic synthesis

Following Thomas and Harden's (2008) three-step approach, qualitative data underwent line-by-line coding, development of descriptive themes, and generation of analytical themes. Descriptive themes captured patterns across AI applications, while analytical themes linked them to learning sciences mechanisms and the four Human–AI Symbiotic Learning Space dimensions.

2.8 Evidence integration

Quantitative effect estimates (e.g., achievement and writing gains, engagement metrics) and qualitative analytical themes were then integrated in a configurative synthesis to answer RQ1–RQ3 and to shape the Human–AI Symbiotic Learning Space propositions. This involved aligning patterns in effect sizes and implementation findings with theoretical constructs and governance concerns to characterize Human–AI symbiosis across established applications, emerging infrastructures, and socio-technical risks, as elaborated in the Results and Discussion sections. The Figure presents a PRISMA-style flow diagram summarizing the number of records identified, screened, excluded, and retained at each stage of the review process.

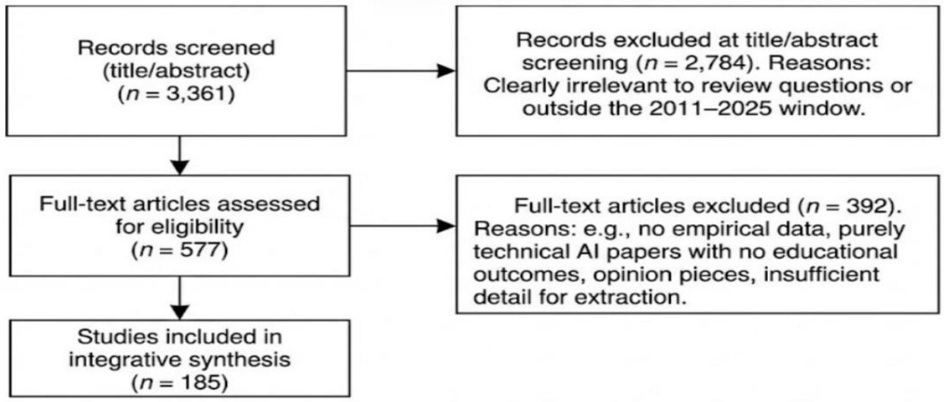

Figure. PRISMA-style flow of study selection.

2.9 Synthesis of effect sizes and diverse evidence

Because this integrative review spans heterogeneous study designs, a narrative meta-analytic approach was adopted rather than statistical pooling. Effect sizes (Cohen's d or equivalent standardised mean differences) were extracted where reported and organised by application type, educational level, and implementation condition to enable cross-study comparison (see Table 1). Where multiple studies addressed the same application, effect size ranges and central tendencies are reported alongside moderator conditions (e.g., curricular integration, teacher orchestration) to contextualise variation. Qualitative findings—including implementation barriers, governance challenges, and equity concerns—were synthesised thematically following Thomas and Harden's (2008) three-step approach, then integrated with quantitative estimates in a configurative synthesis to generate the Human–AI Symbiotic Learning Space propositions. This mixed-evidence approach is consistent with Torraco's (2016) integrative review methodology, which explicitly accommodates diverse epistemologies and study designs.

3. Results

The synthesis of the identified literature provides a multifaceted perspective on the current efficacy, structural evolution, and socio-technical implications of AIED. The findings are structured to respond directly to the research questions, offering a detailed analysis of established applications, emerging generative architectures, and the complex ethical-legal frameworks currently under development.

3.1 Current classroom-facing AI applications (RQ1)

Across applications, classroom-facing AI tools generally yield small-to-moderate improvements in achievement, writing quality, and skill performance, with ITS and adaptive platforms often in the 0.4–0.6 SD range under strong implementation conditions. These gains are most consistently realized when systems are tightly integrated with curriculum, accompanied by active teacher orchestration, and used within robust pedagogical designs such as mastery learning, process-oriented writing instruction, and scaffolded practice. Key risks across tools include reduced learner agency through opaque personalization and surveillance-oriented analytics, as well as over-reliance on AI feedback that narrows learning to what is easily measurable or encourages superficial engagement (see Table 1 for detailed effect ranges, contexts, and risk profiles).

However, the evidence base is not uniformly positive, and several important tensions emerge. First, ITS effects are moderated substantially by comparison condition: while gains versus no-treatment controls are consistent, comparisons against high-quality non-intelligent software or expert human tutoring show meaningfully smaller advantages (Kulik & Fletcher, 2016; Létourneau et al., 2025), raising questions about whether AI per se drives gains or whether any structured, personalised feedback would suffice. Second, AWE benefits appear contingent on pedagogical framing: studies embedding AWE in process-oriented writing instruction report medium positive effects, whereas implementations focused on surface-level accuracy show limited or null impacts on rhetorical quality (Fleckenstein et al., 2023; Liu, 2024). Third, LADs have produced inconsistent learning outcomes despite widespread implementation, with some studies reporting improved intervention targeting and others finding no significant effects on achievement, suggesting that dashboard availability does not automatically translate into effective teacher or student action (Paolucci et al., 2024; Liu et al., 2025). Fourth, GenAI evidence, while promising for productivity and higher-order thinking in scaffolded conditions, remains predominantly short-term and methodologically heterogeneous, limiting claims about sustained learning transfer (Borge et al., 2024; Song et al., 2025). These tensions collectively indicate that AI effectiveness is less a property of the technology and more a function of design quality, instructional integration, and institutional context.

Table 1. Evidence map of major AI application types in education, including learning outcomes, contexts, evidence strength, and associated risks.

AI application type	Core educational functions	Main learning outcomes reported	Contexts most represented	Evidence strength & key risks
ITS	Adaptive sequencing, hints, step-by-step guidance, mastery learning support	Consistent achievement gains in math/STEM; strongest benefits for lower-performing learners (Ma et al., 2014; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011)	Mainly K–12 STEM and mathematics; some higher education	Strong evidence. Meta-analytic effects often moderate; benefits may taper vs. high-quality non-intelligent software (Kulik & Fletcher, 2016; Létourneau et al., 2025). Risks: learner dependence, reduced agency, and opacity of learner models
Adaptive learning platforms	Personalized pacing, content selection, performance-based recommendations	Improved completion and exam performance in gateway STEM courses (Chen et al., 2020; Rivera Muñoz et al., 2022; Ryoo, 2021); personalization benefits lower-performing learners (Judijanto, 2025; Nkambou et al., 2010)	K–12 and higher education; large-enrollment courses	Moderate to strong evidence. Success depends on pedagogical transparency and teacher orchestration (Peña-Ayala, 2018). Risks: autonomy reduction, “black-box” personalization, and equity concerns
Automated Writing Evaluation (AWE) / AI feedback	Automated feedback on accuracy, organization, revision, coherence; iterative drafting support	Medium positive effects on writing quality and writing self-efficacy when integrated into process-oriented pedagogy (Fleckenstein et al., 2023; Liu, 2024; Sari & Han, 2024)	Higher education; EFL/ESL writing contexts	Moderate evidence. Effects vary with feedback integration and teacher mediation. Risks: narrowing writing to surface-level features; bias against linguistic variation; over-reliance
Learning analytics dashboards (LADs) and early-warning systems	Real-time monitoring, predictive risk alerts, progress visualization	Supports intervention timing and monitoring, but raises interpretation and equity challenges (Paolucci et al., 2024; Sohail et al., 2022)	Higher education and LMS-based systems; increasing interest institution-wide	Moderate evidence. Widely implemented but learning impact is inconsistent across settings. Risks: inferential privacy, surveillance pedagogies, and algorithmic profiling

AI application type	Core educational functions	Main learning outcomes reported	Contexts most represented	Evidence strength & key risks
AI-supported assessment and formative feedback at scale	Automated feedback generation, rubric support, scalable formative assessment	Increased efficiency and formative feedback capacity; challenges remain in transparency and alignment with instructor criteria (Mousa, 2025; Trajkovski & Hayes, 2025; Zakareya et al., 2024)	Higher education	Moderate evidence. Human-in-the-loop models increasingly preferred (Figueras et al., 2025). Risks: fairness, contestability, and instructor–AI misalignment
AI-enhanced simulation and virtual patients	Adaptive simulation scenarios, deliberate practice, performance monitoring	Strengthens clinical reasoning and psychomotor skills (Elendu et al., 2024; Motola et al., 2013); blended models preferred for empathy preservation (Harder et al., 2025; Rodger et al., 2025)	Medical, health professions, and professional education	Moderate evidence. Simulation is strong; AI enhancement is growing. Risks: over-optimization of measurable performance, safety implications of scoring
GenAI tutors / conversational learning companions	Dialogue-based tutoring, explanation generation, idea generation, writing support, inquiry scaffolding	Gains in productivity and creativity when scaffolded (Alfarwan, 2025; Li, 2025; Song et al., 2025); supports metacognition and critical thinking in dialogic roles (Borge et al., 2024; Helal et al., 2025; Singh et al., 2025)	Higher education; emerging K–12 pilots	Emerging evidence. Rapid adoption but limited longitudinal and rigorous comparative research. Risks: hallucinations, cognitive offloading, academic integrity issues (Bozkurt et al., 2024; Luckin, 2018)
RAG-enhanced educational chatbots (retrieval-augmented generation)	Policy-aligned answering, grounded institutional knowledge, reduced hallucinations	Improves reliability and institutional control through provenance-based retrieval (Swacha & Gracel, 2025); reduced errors via retrieval/reranking advances (Zhou et al., 2024; Xu et al., 2025)	Higher education support systems; institutional deployments	Emerging evidence. Promising for governance and trust. Risks: data governance, bias in curated knowledge bases, and unequal access to institutional resources (Ayala & Bechard, 2024)
Multimodal AI (MLLMs) for STEM learning	Diagram interpretation, visual explanations, multimodal tutoring	Supports diagram-rich STEM learning but struggles with authentic diagram understanding (Caffagni et al., 2024; Zhang et al., 2024); education-oriented models are emerging (Chu et al., 2025)	K–12 STEM and mathematics; early higher education studies	Emerging evidence. High potential but not yet mature. Risks: over-trust in visual reasoning outputs; transparency and reliability challenges (Moon et al., 2025)
Institution-level AI infrastructures (digital twins, “nerve centres”)	Real-time decision dashboards, policy simulation, predictive maintenance, strategic governance	Enables predictive maintenance, policy simulation, and learning optimization (Jones et al., 2020; Xie et al., 2024); supports intervention and planning via integrated dashboards (USDOE, 2023; Zawacki-Richter et al., 2019)	Higher education; institution-wide systems	Emerging evidence. Mostly conceptual and technical reports, limited learning-outcome evaluation. Risks: inferential privacy, governance concentration, “democratization of surveillance,” and institutional bias (Adil, 2025)

Note. Evidence strength reflects the maturity and consistency of the reviewed empirical literature. “Strong” indicates sustained evidence bases supported by meta-analytic findings; “Moderate” indicates consistent but context-dependent results; and “Emerging” indicates rapidly developing areas where evaluations remain limited, short-term, or methodologically heterogeneous.

3.2 Institutional infrastructures and emerging AI technologies (RQ2)

GenAI and multimodal AI are evolving into long term learning companions that remember prior interactions, monitor progress, and collaborate with instructors. They support drafting, coding, and complex problem solving, and when designed around Socratic prompting and critique, they can enhance creativity and higher order thinking rather than simply delivering answers. Key risks include hallucinations, over reliance, and context specific evidence, which make claims about large scale transformation necessarily provisional.

To strengthen institutional control and academic integrity, many emerging systems combine these models with retrieval mechanisms that ground responses in vetted curricular resources, policies, and open educational materials. This approach helps reduce hallucinations, supports use of smaller and more affordable models, and preserves provenance so that cross course tutoring and advising agents remain aligned with institutional standards. Such agents can offer coherent support across subjects while still keeping instructors "in the loop" for judgment and assessment.

Institutionally, AI is increasingly embedded in digital twins and integrated analytics "nerve centres" that aggregate data from learning platforms, sensors, and administrative systems. Early implementations focus on operational benefits such as predictive maintenance and energy management, but their educational promise lies in enabling leaders to simulate policy scenarios—such as timetable, admission, or support changes—before implementation. Used responsibly, these infrastructures can inform decisions about retention, equity, and resource allocation while raising important questions about governance, privacy, and stakeholder participation.

3.3 Ethical, legal, and socio-technical risks and safeguards (RQ3)

A central ethical concern is the "democratization of surveillance," where AI infers sensitive traits—such as learning disabilities, mental health, or socio-economic status—from routine behavioral data (e.g., pacing or feedback patterns). Such inference occurs silently and often falls outside current privacy laws, risking profiling and inequitable resource allocation. Scholars advocate "Inference Transparency Audits" and explicit opt in protocols for analytics extending beyond direct instructional uses.

Research identifies a "cognitive paradox" in which reliance on AI produces "cognitive debt"—reduced engagement that weakens durable learning and metacognition. Cognitive offloading eases cognitive load but undermines retrieval and monitoring processes vital for schema formation. Frequent AI use correlates negatively with critical thinking, particularly among learners aged 17–25. To counteract this, "effort-aware design" and "scaffolded AI debrief protocols" prompt students to justify and reflect on their reasoning rather than accept AI outputs unquestioningly.

Algorithmic bias persists across assessment and prediction systems, with biased data and opaque models disadvantaging marginalized learners. The "transparency gap" between AI capabilities and users' understanding diminishes teachers' agency and accountability. Responsible AI frameworks call for embedding equity and social justice values into AI governance rather than treating fairness as a technical adjustment.

Implementation challenges stem from uneven infrastructure, high computational costs, and limited professional training. Institutional approaches vary from centralized compliance regimes to faculty driven models, with evidence favouring the latter for balancing innovation and autonomy. AI literacy for educators and students—covering ethics, factual limits of LLMs, and algorithmic bias—is essential for responsible integration. Effective adoption requires viewing AI as part of an evolving socio technical ecosystem reshaping institutional roles and practices.

3.4 Theoretical intersections and mechanism analysis

The review findings align strongly with the conceptual foundations of learning sciences, emphasizing the interplay between technological design and pedagogical theory.

The effectiveness of ITS and GenAI tutors rests on their capacity for dynamic scaffolding—recruiting interest, modelling tasks, and providing adaptive support similar to human tutoring. However, scaffolding must be intentionally “faded” to prevent dependency. LADs dashboards reinforce SRL by visualizing progress and supporting metacognitive monitoring and goal regulation.

AI’s relationship with CLT is dual: while adaptive sequencing reduces extraneous load, excessive automation can encourage unproductive cognitive offloading. Research underscores that germane load—effort devoted to schema building—emerges when AI prompts learners to self-explain reasoning. Developmentally aligned exposure is essential, calibrating AI assistance to learners’ expertise and growth level.

AI applications align with socio cultural and connectivist theories, functioning as partners and “active nodes” in distributed learning networks. These agentic systems scaffold collaborative discourse and help students navigate complex information flows, reflecting the connectivist view of learning as building and managing interlinked knowledge nodes. Table 2. maps AI application types to key pedagogical mechanisms and outcomes.

Table 2. AI application types mapped to key pedagogical mechanisms and outcomes

AI application	Dominant mechanisms	Typical outcomes (illustrative)
Intelligent Tutoring Systems (ITS)	Scaffolded practice in ZPD, stepwise feedback, mastery learning	Moderate gains in maths and science achievement; strongest benefits for lower-performing learners
Adaptive learning platforms	Data-driven pacing and content selection, formative assessment loops	Improved course completion and exam performance in large-enrolment courses
Automated Writing Evaluation (AWE)	Iterative feedback, focused revision cycles, metalinguistic awareness	Medium positive effects on writing quality and self-efficacy when integrated into process-oriented pedagogy
Learning analytics dashboards	SRL monitoring and reflection, teacher decision support	Mixed effects on learning; improved targeting of support where used formatively
AI-enhanced simulations	Deliberate practice, immediate feedback, situated learning	Gains in clinical reasoning and psychomotor skills comparable to high-quality simulation training
GenAI tutors	Dialogic scaffolding, self-explanation prompts, co-creation	Short-term gains in productivity, creativity, and higher-order thinking when scaffolded and governed carefully

The Human–AI Symbiotic Learning Space is distinguished from prior frameworks in three ways. First, unlike technology acceptance models (e.g., TAM) or AI readiness frameworks that focus on adoption attitudes, the Symbiotic Learning Space centres on the redistribution of cognitive and institutional work between humans and AI as a design principle. Second, unlike single-dimension ethical frameworks (e.g., fairness toolkits or explainability checklists), it integrates learning sciences mechanisms (scaffolding, SRL, CLT) with governance concerns under four interdependent dimensions rather than treating pedagogy and ethics as separate concerns. Third, unlike prior integrative reviews that propose descriptive typologies of AI tools, the framework generates testable design–governance propositions (see 4.3) linking dimension-specific features to measurable outcomes, enabling empirical evaluation.

4. Discussion

4.1 Overview of main contributions

Across sectors, the reviewed evidence confirms that established AI applications such as ITS, adaptive platforms, simulation-based training, and AWE now function as core infrastructure rather than fanciful add ons, with effect sizes in the small to moderate range and the greatest benefits for lower performing and underserved learners when tools are tightly integrated into curriculum and pedagogy. At the same time, the field is pivoting toward generative and multimodal architectures, agentic learning companions, and institution level analytics "nerve centres," which extend AI's reach from micro level tutoring to meso and macro level orchestration of learning and operations. The proposed Human–AI Symbiotic Learning Space offers a unifying lens on this evolution by foregrounding four interdependent design–governance dimensions—Agency, Alignment, Adaptivity, and Accountability—that cut across current and emerging applications.

4.2 Positioning within prior AIED and learning sciences research

The findings extend a long trajectory in AIED from rule based ITS and adaptive hypermedia toward data driven and generative systems, corroborating meta analytic evidence that ITS can approach one to one human tutoring when learner models are precise and scaffolding is sustained over time. However, the review also shows that once comparison conditions involve strong non intelligent software or high quality teaching, the marginal advantage of AI is smaller, emphasizing that design quality and instructional integration matter more than AI per se. By systematically mapping empirical outcomes to constructivism, SRL, socio cultural theory, connectivism, and CLT, the review answers critiques that AI research has often been theoretically thin, showing instead that many high impact systems operationalize well established mechanisms such as scaffolding, SRL monitoring–control cycles, dialogic collaboration, and load management—albeit sometimes implicitly.

4.3 Human–AI symbiosis: Unpacking Agency, Alignment, Adaptivity, Accountability

Framing the future of AI in education as a Human–AI Symbiotic Learning Space allows a move beyond "AI as tool" or "AI as replacement" narratives toward a focus on how socio technical designs redistribute cognitive and institutional work. The Agency dimension highlights that the same analytics and generative capabilities that enable personalization can also undermine learner and teacher autonomy through over scaffolding, opaque recommendations, and convenience driven cognitive offloading; the reviewed evidence suggests that designs which make goals, data, and decision logics visible and editable to learners support SRL more effectively than answer giving bots. Alignment emphasizes that improvements in achievement or efficiency are insufficient if systems are not explicitly aligned with robust learning theories and equity goals:

tools that simply optimize time on task without engaging metacognition, collaboration, or deep processing risk producing superficial gains that do not transfer. Alignment can be assessed through depth-of-learning and near/far transfer outcomes, collaborative interaction quality ratings, and reductions in achievement gaps between demographic subgroups across implementation cycles.

Adaptivity emerges as a multi layer construct extending beyond item by item difficulty adjustment to include adaptive modalities, feedback timing, role shifts between human and AI partners, and developmentally calibrated exposure to assistance; this is particularly salient in the context of multimodal LLMs and STEM tutoring, where visual reasoning and conceptual articulation must be balanced with productive struggle. Adaptivity indicators include learning curve trajectories across AI-assisted sessions, the ratio of student-initiated to AI-initiated interactions, and fading metrics tracking decreasing assistance levels as learner competence grows Finally, Accountability addresses the democratization of surveillance, inferential privacy, algorithmic bias, and the transparency gap, arguing that institutional adoption must pair technical safeguards (e.g., privacy preserving architectures, retrieval augmented generation, bias audits) with participatory governance structures that give students and educators meaningful voice in the design and oversight of AI systems. Accountability indicators include AI recommendation contestation rates and their resolution outcomes, student- and teacher-perceived fairness ratings, and documented governance actions (e.g., bias audit findings and resulting policy changes). Table 3 shows design questions and indicators of the human-AI symbiotic learning space.

Table 3. Human-AI Symbiotic Learning Space: design questions and indicators

Dimension	Guiding question	Example design features	Example indicators
Agency	Who can see, shape, and contest goals, data, and AI-supported actions?	Learner-editable goals and profiles; co-configured dashboard views; opt-in/opt-out options for advanced analytics	Changes in SRL measures; perceived autonomy; frequency of learner-initiated adjustments
Alignment	With which learning theories, curricular goals, and equity commitments is the system explicitly aligned?	Scaffolding and fading aligned with ZPD; prompts for self-explanation; collaborative tasks; bias-aware content curation	Depth-of-learning and transfer outcomes; collaborative quality; reductions in achievement gaps
Adaptivity	What adapts, when, and for whom, and how is over-automation avoided?	Adaptive sequencing with effort-aware prompts; modality shifts (text/visual/simulation); graduated assistance and fading	Learning curves; time-on-task profiles; balance between independence and support
Accountability	Who is answerable for AI-supported decisions and on what basis?	Audit trails of AI recommendations; explainable models; inference transparency audits; participatory governance processes	Contestation rates and outcomes; perceived fairness; documented policy and governance actions

To illustrate how these dimensions operate in practice, consider three concrete use cases. (1) Agency in a LAD deployment: A university implements an early-warning dashboard but allows students to toggle the visibility of their own risk scores, set personal learning goals within the system, and request human advisor review of any algorithmic flag. Agency indicators—frequency of student-initiated goal adjustments and contestation rates—are tracked termly. (2) Adaptivity in a GenAI writing companion: A secondary school deploys a GenAI writing tool configured with effort-aware prompts that require students to explain their revision choices before the AI provides the next feedback cycle, and the level of AI assistance is graduated based on demonstrated self-explanation quality. Adaptivity indicators—balance between AI-initiated and student-initiated interactions, and time-on-task profiles—are monitored across drafting cycles. (3) Accountability in an AI-supported assessment system: A higher education institution implements AI-assisted grading with a mandatory audit trail: every AI-generated grade is accompanied by an explainability summary, reviewed by the course instructor, and subject to a structured student contestation pathway. Accountability indicators—contestation rates, resolution outcomes, and student-perceived fairness—are reviewed by an ethics committee annually.

4.4 Implications for Practice, Leadership, and Design

From an educational innovation perspective, the findings underscore that the impact of AI depends less on the sophistication of algorithms and more on how technologies are embedded within teaching, learning, and institutional systems. The review highlights that AI can support pedagogical transformation when it is used to strengthen feedback, scaffolding, and self-regulated learning rather than to automate learning tasks or displace teacher judgment.

4.4.1 Implications for Classroom Practice and Teacher Learning

The synthesis offers teachers five design principles for AI-supported instruction.

1. Pair GenAI with self-explanation prompts to counter cognitive offloading.
2. Use LADs as SRL mirrors via co-interpretation, not grading.
3. Embed AI in mastery learning or cognitive apprenticeship.
4. Script human–AI roles so teachers focus on diagnosis and mentoring.
5. Implement AI debrief routines comparing AI outputs to authoritative sources for critical evaluation.

4.4.2 Implications for Institutional Leaders and Policymakers

For institutional leaders and policymakers, a pragmatic checklist emerging from the Human–AI Symbiotic Learning Space includes:

1. Agency: Establish policies that guarantee students and staff visibility into, and meaningful control over, how their data are used in AI systems.
2. Alignment: Require that major AI deployments be explicitly mapped to institutional pedagogical models and equity goals, not only to efficiency objectives.
3. Adaptivity: Prioritize systems that support graduated assistance and fading, and monitor for over-automation that erodes learners' self-regulation and critical thinking.
4. Accountability: Implement multi-layered governance (technical audits, ethics committees, participatory forums) and clear procedures for contesting AI-supported decisions.
5. Develop coherent AI roadmaps that integrate infrastructure, professional learning, and AI literacy for students, rather than relying on ad hoc tool adoption.

4.4.3 Implications for Edtech Designers and Researchers

For edtech designers and researchers, the framework highlights the importance of: (a) encoding learning-theoretic principles such as scaffolding and SRL cycles into AI tools; (b) supporting explainability features that make learner models and recommendations intelligible; and (c) co-designing systems with educators and students from diverse contexts to surface equity and governance issues early in development.

4.5 Limitations

Several limitations of this review warrant acknowledgement. First, the reliance on English-language publications introduces a language bias that may underrepresent AIED research conducted in non-Anglophone contexts, limiting the generalisability of conclusions to global educational settings. Second, despite the use of multi-database searching and citation chasing, the rapid pace of publication in GenAI education research means that some studies published in late 2024 and early 2025 may not be fully captured, particularly those in conference proceedings. Third, although quality appraisal was conducted using established tools (RoB 2, ROBINS-I, CASP), the inherent heterogeneity of study designs, outcome measures, and AI application types precludes statistical pooling of effect sizes; the reported effect ranges should therefore be interpreted as indicative rather than precise estimates. Fourth, the Human–AI Symbiotic Learning Space framework, while grounded in the reviewed evidence, has not yet been prospectively tested as a design or evaluation instrument; its propositions remain theoretical pending longitudinal and design-based validation. Fifth, the review focuses on the direct educational uses of AI rather than the broader institutional and labour-market implications of AI adoption, which represents a complementary but distinct area of inquiry.

5. Conclusion

AI in education is shifting from isolated tools to socio-technical systems that redistribute cognitive and institutional work across learners, teachers, and institutions, making human–AI symbiosis a central design challenge rather than a distant prospect. The Human–AI Symbiotic Learning Space proposed in this review offers a practical, theory-grounded framework for balancing achievement and efficiency gains with learner agency, equity, and ethical governance.

Building on the evidence synthesised here, the following testable research directions are proposed for future empirical investigation:

(1) Agency and SRL: Designs that make AI goals, data, and decision logic visible and editable to learners will produce stronger gains in self-regulated learning (SRL) measures and perceived autonomy than equivalent AI tools operating opaquely, particularly in adaptive learning and LAD contexts.

(2) Alignment and transfer: AI-supported interventions explicitly aligned with scaffolding-and-fading principles and collaborative learning designs will produce greater near and far transfer of learning than AI implementations optimising for time-on-task or completion metrics alone.

(3) Adaptivity and cognitive outcomes: Effort-aware AI designs that require learner self-explanation before providing assistance will reduce cognitive offloading and produce stronger metacognitive regulation and critical thinking outcomes among 17–25-year-old learners compared with answer-giving AI tools.

(4) Accountability and equity: Institutions implementing participatory AI governance structures—including inference transparency audits, bias reviews, and student contestation pathways—will demonstrate smaller

AI-related achievement gaps across demographic subgroups than institutions relying on compliance-only governance.

(5) Longitudinal impact: Multi-year studies tracking learners who transition from heavily AI-assisted to independently managed learning contexts are needed to establish whether scaffolded human–AI collaboration supports durable self-regulation or produces dependency effects that erode performance once AI support is withdrawn.

These hypotheses are intended as an actionable research agenda for longitudinal, design-based, and comparative studies in diverse educational settings, and they represent the primary directions in which the Human–AI Symbiotic Learning Space framework can be empirically refined and validated.

References

[1]. Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. In Frontiers in Education, 10, 1647573. https://doi.org/10.3389/feduc.2025.1647573

[2]. Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. International Review of Research in Open and Distributed Learning, 12(3), 80–97. https://doi.org/10.19173/irrodl.v12i3.890

[3]. Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y., Nerantzi, C., Moore, S., ... & Asino, T. (Eds.). (2024). The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. Open Praxis, 16(4), 487-513. https://doi.org/10.55982/openpraxis.16.4.777

[4]. Ayala, O., & Bechard, P. (2024). Reducing hallucination in structured outputs via Retrieval-Augmented Generation. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 6: Industry Track) (pp. 228-238). https://doi.org/10.18653/v1/2024.naacl-industry.19

[5]. Beale, R. (2025). Dialogic Pedagogy for Large Language Models: Aligning Conversational AI with Proven Theories of Learning. arXiv preprint arXiv:2506.19484. https://doi.org/10.48550/arXiv.2506.19484

[6]. Borge, M., Smith, B. K., & Aldemir, T. (2024). Using generative ai as a simulation to support higher-order thinking. International Journal of Computer-Supported Collaborative Learning, 19(4), 479-532.

[7]. Brusilovsky, P. (2001). Adaptive hypermedia. User Modeling and User-Adapted Interaction, 11(1–2), 87–110.

[8]. Caffagni, D., Cocchi, F., Barsellotti, L., Moratelli, N., Sarto, S., Baraldi, L., Cornia, M., & Cucchiara, R. (2024). The revolution of multimodal large language models: A survey. In Findings of the Association for Computational Linguistics: ACL 2024 (pp. 13590–13618). https://doi.org/10.48550/arXiv.2402.12451

[9]. Castañeda, L., & Selwyn, N. (2018). More than tools? Making sense of the ongoing digitalization of higher education. International Journal of Educational Technology in Higher Education, 15, Article 22. https://doi.org/10.1186/s41239-018-0109-y

[10]. CASP. (2018). Critical Appraisal Skills Programme. CASP qualitative checklist, https://casp-uk.net/casp-tools-checklists/qualitative-studies-checklist/

[11]. Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

[12]. Chiu, T., & Rospigliosi, P. (2025). Encouraging human-AI collaboration in interactive learning environments. Interactive Learning Environments, 33(2), 921-924. https://doi.org/10.1080/10494820.2025.2471199

[13]. Chu, Z., Xie, J., Wang, S., Wang, Z., & Wen, Q. (2025). UniEDU: Toward unified and efficient large multimodal models for educational tasks. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track. Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.emnlp-industry.68

[14]. Dignum, V. (2021). The role and challenges of education for responsible AI. London Review of Education, 19(1), 1–11. https://doi.org/10.14324/LRE.19.1.01

[15]. Drachsler, H., & Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges. Journal of Computer Assisted Learning, 32(3), 281–290. https://doi.org/10.1111/jcal.12135

[16]. Elendu, C., Amaechi, D., Okatta, A., Amaechi, E., Elendu, T., Ezeh, C., & Elendu, I. (2024). The impact of simulation-based training in medical education: A review. Medicine, 103(27), e38813.

[17]. Figueras, C., Farazouli, A., Cerratto Pargman, T., McGrath, C., & Rossitto, C. (2025). Promises and breakages of automated grading systems: A qualitative study in computer science education. Education Inquiry, 1–22. https://doi.org/10.1080/20004508.2025.2464996

[18]. Ferster, B. (2014). Teaching machines: Learning from the intersection of education and technology. Johns Hopkins University Press.

[19]. Fleckenstein, J., Liebenow, L., & Meyer, J. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. Frontiers in Artificial Intelligence, 6, 1162454. https://doi.org/10.3389/frai.2023.1162454

[20]. Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

[21]. Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

[22]. Godwin-Jones, R. (2017). Smartphones and language learning. Language Learning & Technology, 21(2), 3–17. http://llt.msu.edu/issues/june2017/emerging.pdf

[23]. Graesser, A., Conley, M., & Olney, A. (2012). Intelligent tutoring systems. In K. Harris, S. Graham, T. Urdan, A. Bus, S. Major, & H. Swanson (Eds.), APA educational psychology handbook, Vol. 3. Application to learning and teaching (pp. 451–473). American Psychological Association. https://doi.org/10.1037/13275-018

[24]. Helal, M., Elgendy, I., Albashrawi, M., Dwivedi, Y., Al-Ahmadi, M., & Jeon, I. (2025). The impact of generative AI on critical thinking skills: a systematic review, conceptual framework and future research directions. Information Discovery and Delivery.

[25]. Harder, N., McNelis, A., & Garcia, M. (2025). Comparing artificial intelligence enhanced virtual reality with standardized patient simulation in undergraduate nursing education. Clinical Simulation in Nursing, 105, 101780. https://doi.org/10.1016/j.ecns.2025.101780

[26]. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

[27]. Holmes, W., Bialik, M., & Fadel, C. (2022). Artificial intelligence in education: Promises and implications for teaching and learning (2nd ed.). Center for Curriculum Redesign.

[28]. Majumder, S., & Dey, N. (2022). AI-empowered knowledge management. Springer.

[29]. Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the digital twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology, 29, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

[30]. Joksimović, S., Dowell, N., Poquet, O., Mills, C., Gašević, D., Dawson, S., & Graesser, A. (2018). Exploring development of social capital in a cMOOC through language and discourse. The Internet and Higher Education, 36, 54–64. https://doi.org/10.1016/j.iheduc.2017.09.004

[31]. Judijanto, L. (2025). A Bibliometric Analysis of Adaptive Learning in K-12 Education. The Eastasouth Journal of Learning and Educations, 3(01), 75–86. https://doi.org/10.58812/esle.v3i01.496

[32]. Kelly, M., & Risko, E. (2019). Offloading memory: Serial position effects. Psychonomic Bulletin & Review, 26, 1347–1353.

[33]. Kulik, J., & Fletcher, J. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. Review of Educational Research, 86(1), 42–78. https://doi.org/10.3102/0034654315581420

[34]. Lee, U., Jeon, M., Lee, Y., Byun, G., Son, Y., Shin, J., ... & Kim, H. (2024). LLaVA-docent: Instruction tuning with

multimodal large language model to support art appreciation education. Computers and Education: Artificial Intelligence, 7, 100297. https://doi.org/10.1016/j.caeai.2024.100297

[35]. Lee, S., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. Computers and Education: Artificial Intelligence, 6, 100211. https://doi.org/10.1016/j.caeai.2024.100211

[36]. Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K–12 education. npj Science of Learning, 10(1), 29. https://doi.org/10.1038/s41539-025-00320-7

[37]. Li, S. (2025). Generative AI and second language writing. Digital Studies in Language and Literature, 2(1),122-152. https://doi.org/10.1515/dsll-2025-0007

[38]. Liu, W. (2024). A Systematic review of automated writing evaluation feedback: Validity, effects and students' engagement. Language Teaching Research Quarterly, 45, 86-105. https://doi.org/10.32038/ltrq.2024.45.05

[39]. Liu, Y., Wang, W., & Xu, E. (2025). The effectiveness of learning analytics-based interventions in enhancing students' learning effect: A meta-analysis of empirical studies. SAGE Open, 15(2), https://doi.org/10.1177/21582440251336707

[40]. Loewen, S., Crowther, D., Isbell, D., Kim, K., Maloney, J., Miller, Z., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. ReCALL, 31(3), 293–311. https://doi.org/10.1017/S0958344019000065

[41]. Luckin, R. (2018). Machine learning and human intelligence: The future of education for the 21st century. UCL IOE Press.

[42]. Luckin, R., Cukurova, M., Kent, C., & Du Boulay, B. (2022). Empowering educators to be AI-ready. Computers and Education: Artificial Intelligence, 3, 100076.

[43]. Ma, W., Adesope, O., Nesbit, J., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. http://dx.doi.org/10.1037/a0037123

[44]. Milligan, C., & Littlejohn, A. (2014). Supporting professional learning in a massive open online course. International Review of Research in Open and Distributed Learning, 15(5), 197–213. https://doi.org/10.19173/irrodl.v15i5.1855

[45]. Moon, H., Davis, R., Neshaei, S., & Dillenbourg, P. (2025). Using large multimodal models to extract knowledge components for knowledge tracing from multimedia question information. In Proceedings of the 18th International Conference on Educational Data Mining. https://doi.org/10.48550/arXiv.2409.20167

[46]. Motola, I., Devine, L., Chung, H., Sullivan, J., & Issenberg, S. (2013). Simulation in healthcare education: A best evidence practical guide. AMEE Guide No. 82. Medical Teacher, 35(10), e1511–e1530. https://www.tandfonline.com/doi/full/10.3109/0142159X.2013.818632

[47]. Mousa, M. (2025). AI-Supported Formative Assessments: Enhancing Student-Centered Learning and Teacher Perceptions. Journal of Pedagogy and Education Science, 4(02), 127-141. https://doi.org/10.56741/jpes.v4i02.857

[48]. Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. Interactive Learning Environments, 29(1), 142–163.

[49]. Nguyen, N. (2023). Exploring the role of AI in education. London Journal of Social Sciences, 6, 84-95. https://doi.org/10.31039/ljss.2023.6.108

[50]. Nkambou, R., Bourdeau, J., & Mizoguchi, R. (Eds.). (2010). Advances in intelligent tutoring systems. Springer.

[51]. Adil, J. (2025). AI in education: A systematic literature review of emerging trends, benefits, and challenges. In Seminars in Medical Writing and Education, 4, 795. https://doi.org/10.56294/mw2025795

[52]. Ouwehand, K., Lespiau, F., Tricot, A., & Paas, F. (2025). Cognitive load theory: Emerging trends and innovations. Education Sciences, 15(4), 458. https://doi.org/10.3390/educsci15040458

[53]. Paolucci, C., Ifenthaler, D., Gašević, D., Araya, R., Rummel, N., & Ferguson, R. (2024). A review of learning

analytics opportunities and challenges for K–12 education. Heliyon, 10(4), e25767.

[54]. *Peña-Ayala, A. (Ed.). (2018). Learning analytics: Fundaments, applications, and trends. Springer.*

[55]. *Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: Scoping review. JMIR Medical Education, 9, e48785. https://doi.org/10.2196/48785*

[56]. *Risko, E., & Gilbert, S. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002*

[57]. *Rivera Muñoz, J., Moscoso Ojeda, F., Aparicio Jurado, D., Puga Peña, P., Martel Carranza, C., Quispe Berríos, H., Ugarte Molina, S., Maldonado Farfan, A., Arias-Gonzáles, J., & Vasquez-Pauca, M. (2022). Systematic review of adaptive learning technology for learning in higher education. Eurasian Journal of Educational Research, 98, 221–233.*

[58]. *Rodger, D., Mann, S., Earp, B., Savulescu, J., Bobier, C., & Blackshaw, B. P. (2025). Generative AI in healthcare education: How AI literacy gaps could compromise learning and patient safety. Nurse Education in Practice, 87, 104461. https://doi.org/10.1016/j.nepr.2025.104461*

[59]. *Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analysing collaborative learning processes automatically. International Journal of Computer Supported Collaborative Learning, 3(3), 237–271. https://doi.org/10.1007/s11412-007-9034-0*

[60]. *Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.*

[61]. *Ryoo, J. (2021). Innovative learning environments in STEM higher education. Springer. https://library.oapen.org/handle/20.500.12657/47325*

[62]. *Sari, E., & Han, T. (2024). The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. Journal of Computer Assisted Learning, 40(5), 2065-2080.*

[63]. *Siemens, G. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1).*

[64]. *Singh, A., Guan, Z., & Rieh, S. (2025). Enhancing critical thinking in generative AI search with metacognitive prompts. Proceedings of the Association for Information Science and Technology, 62(1), 672-684.*

[65]. *Sohail, S., Alvi, A., & Khanum, A. (2022). Interpretable and Adaptable Early Warning Learning Analytics Model. Computers, Materials & Continua, 71(2). https://doi.org/10.32604/cmc.2022.023560*

[66]. *Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: A dynamic systems view of the informal learning of English in virtual online communities. ReCALL, 24(2), 138–151. https://doi.org/10.1017/S0958344012000031*

[67]. *Song, D., Zhang, P., Zhu, Y., Qi, S., Yang, Y., Gong, L., & Zhou, L. (2025). Effects of generative artificial intelligence on higher-order thinking skills and artificial intelligence literacy in nursing undergraduates: A quasi-experimental study. Nurse Education in Practice, 88, 104549. https://doi.org/10.1016/j.nepr.2025.104549*

[68]. *Sottilare, R., Graesser, A., Hu, X., & Goldberg, B. (Eds.). (2018). Design recommendations for intelligent tutoring systems: Volume 6—Team tutoring. U.S. Army Research Laboratory.*

[69]. *Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. Journal of Educational Psychology, 106(4), 901–918. http://dx.doi.org/10.1037/a0032447*

[70]. *Swacha, J., & Gracel, M. (2025). Retrieval-augmented generation (RAG) chatbots for education: A survey of applications. Applied Sciences, 15(8), 4234. https://doi.org/10.3390/app15084234*

[71]. *Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8, 45. https://doi.org/10.1186/1471-2288-8-45*

[72]. *Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. Human Resource Development Review, 15(4), 404–428. https://doi.org/10.1177/1534484316671606*

[73]. *Trajkovski, G., & Hayes, H. (2025). AI Assisted formative assessment and feedback. In AI assisted assessment in education: Transforming assessment and measuring learning (pp. 283–312). Springer. https://doi.org/10.1007/978-3-*

031-88252-4_7

[74]. Tsai, Y., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2020). Complexity leadership in learning analytics: Drivers, challenges and opportunities. British Journal of Educational Technology, 51(6), 2783–2799. https://doi.org/10.1111/bjet.12846

[75]. United States Department of Education, Office of Educational Technology. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations. U.S. Department of Education. https://files.eric.ed.gov/fulltext/ED631097.pdf

[76]. VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

[77]. Verbert, K., Govaerts, S., Duval, E., Santos, J., Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. Personal and Ubiquitous Computing, 18(6), 1499–1514.

[78]. Verspoor, M., de Bot, K., & Lowie, W. (2021). A dynamic approach to second language development: Methods and techniques. John Benjamins.

[79]. Vesselinov, R., & Grego, J. (2012). Duolingo effectiveness study. City University of New York.

[80]. Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

[81]. Wang, T., & Lajoie, S. (2023). How does cognitive load interact with self-regulated learning? A dynamic and integrative model. Educational Psychology Review, 35, Article 69.

[82]. Wieczorek, M., Hosseini, M., & Gordijn, B. (2025). Unpacking the ethics of using AI in primary and secondary education: A systematic literature review. AI and Ethics, 1-19. https://doi.org/10.1007/s43681-025-00770-0

[83]. Xiaoyu, L., Dutta, S., & Khine, M. (2025). Generative artificial intelligence in pedagogical practices: A systematic review of empirical studies (2022–2024). Cogent Education, 12(1), 2485499. https://doi.org/10.1080/2331186X.2025.2485499

[84]. Xie, Y., Zhan, N., Zhu, Q., Zhan, J., Guo, Z., Qiao, C., Zhu, J., & Xu, B. (2024). Multimodal data visualization method for digital twin campus construction. International Journal of Digital Earth, 17(1), 2431624. https://doi.org/10.1080/17538947.2024.2431624

[85]. Xu, S., Yan, Z., Dai, C., & Wu, F. (2025). MEGA RAG: A retrieval augmented generation framework with multi evidence guided answer refinement for mitigating hallucinations of LLMs in public health. Frontiers in Public Health, 13, 1635381. https://doi.org/10.3389/fpubh.2025.1635381

[86]. Ge, Y. (2025). Integrating Human-AI Collaboration in Education: A New Approach to Curriculum Design. Educational Innovation Research, 3(1), 35-39. https://doi.org/10.18063/eir.v3i1.614

[87]. Zakareya, S., Alsaleem, N., Alnaghmaish, A., Alnaim, N., & Alojail, F. (2024). Evaluating the discrimination index of AI-generated vs. human-generated multiple-choice questions: Action research. In ICERI2024 Proceedings (pp. 221-226). IATED. https://doi.org/10.21125/iceri.2024.0137

[88]. Boateng, O., & Boateng, B. (2025). Algorithmic bias in educational systems: Examining the impact of AI-driven decision making in modern education. World Journal of Advanced Research and Reviews, 25(1), 2012-2017. https://doi.org/10.30574/wjarr.2025.25.1.0253

[89]. Zawacki-Richter, O., Marín, V., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? International Journal of Educational Technology in Higher Education, 16, 39. https://doi.org/10.1186/s41239-019-0171-0

[90]. Zhang, M., Jiang, D., Zhang, Y., Lin, H., Guo, Z., Qiu, P., ... Li, H. (2024). MathVerse: Does your multi modal LLM truly see the diagrams in visual math problems? ECCV 2024 (LNCS 15066), 169–186. https://doi.org/10.1007/978-3-031-73242-3_10

[91]. Zhou, Y., Liu, Y., Li, X., Jin, J., Qian, H., Liu, Z., ... & Yu, P. S. (2024). Trustworthiness in retrieval-augmented generation systems: A survey. arXiv preprint arXiv:2409.10102. https://doi.org/10.48550/arXiv.2409.10102

[92]. Zimmerman, B. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M.

Zeidner (Eds.), Handbook of self-regulation (pp. 13–39). Academic Press.