LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8 October, 2026

EFFECTS OF AI-ASSISTED LEARNING ON STUDENTS' CRITICAL THINKING AND ACADEMIC ACHIEVEMENT

Tasawar Aziz[*1], Asif Ali[2], Shahzana Javed[3], Najeeb Ullah[4],

[*1]*Department of Management Sciences the Islamia University of Bahawalpur*
[2]*Department: English Literature and Linguistics University: University of Swabi*
[3,4]*Department: Education University Thal university*

[*1]tasawaraziz901@gmail.com, [2]asiff223@yahoo.com, [3]shahzana.edu@gmail.com,
[4]najibbullah377@gmail.com,

DOI: https://doi.org/10.5281/zenodo.23730150

Keywords
AI-assisted learning; critical thinking; academic achievement; intelligent tutoring systems; generative AI; cognitive offloading; teacher mediation; higher education

Article History
Received: *25 June 2026*
Accepted: *16 September 2026*
Published: *27 October 2026*

Copyright @Author
Corresponding Author: *
Tasawar Aziz

Abstract

Artificial intelligence is now a regular part of educational practice, but evidence of its cognitive effects remains contested. The current review examines the impact of AI-assisted learning on students' critical thinking and academic achievement, and the respective role of learner, teacher and instructional factors. We conducted a systematic literature review following PRISMA 2020 guidelines. The literature search was performed in ERIC, Scopus, Web of Science, and Google Scholar for peer-reviewed studies on AI-assisted learning in formal education. The selection process included duplicate removal, title and abstract screening, full-text assessment, and quality appraisal. Findings were narratively synthesised by outcome and mechanism, and random-effects meta-analysis was applied where effect sizes were similar.The evidence for achievement was uniformly positive. A meta-analysis of 29 studies and 2,657 participants found an effect size of 0.924 . Intelligent tutoring systems outperformed teacher-led large-group instruction (g =.42), non-ITS computer-based instruction (g =.57) and textbooks or workbooks (g =.35) . Critical thinking outcomes were conditional: gains were found with scaffolded, inquiry-oriented designs, no measurable change when AI was added without structure, and declines when AI did the reasoning (d = -1.31). Teacher mediation was the strongest moderator (g = 1.426 with support vs. 0.077 without). AI-assisted learning reliably raises achievement, but only promotes critical thinking if the cognitive work is done by the learner rather than the system.

1. INTRODUCTION

1.1 Background of Artificial Intelligence in Education

Artificial intelligence has moved from a speculative idea to an embedded feature of educational practice. Its lineage stretches back to Alan Turing's early inquiries into machine intelligence and to the computer-assisted instruction systems of the 1960s, such as PLATO, which first automated the delivery of learning content. Attention then shifted toward rule-based automation and, from the early 1980s onward, toward intelligent tutoring systems. The early 2000s brought a further surge of interest in AI-driven educational technologies, with a growing emphasis on personalised and adaptive learning. Today, AI has increasingly become a transformative force in the education sector, offering unprecedented opportunities to enhance

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

learning experiences and outcomes. The most recent phase is data-driven, featuring tools such as ChatGPT, Duolingo and ClassAI that personalise learning, enhance accessibility and assist educators in real time. Across these decades, the field has moved from automating instruction to adapting it to the individual learner.

1.2 AI-Assisted Learning and Its Educational Applications

AI-assisted learning refers to instruction in which AI systems support, adapt or supplement the learner's activity. Intelligent tutoring systems are computer-based learning systems that use AI to provide personalised and adaptive instruction, modelling students' psychological states — motivation, emotion and cognition — alongside their prior knowledge, skills and preferences. They monitor progress and supply feedback, hints and scaffolding, and they select appropriate problems or tasks for practice. The related concept of adaptive learning emphasises facilitating individual students' learning by adapting user interfaces, learning content or learning paths to each learner's status. Applications now extend to virtual laboratories and simulations, which allow students to conduct experiments in a virtual environment while educators assess performance in real time.

1.3 Critical Thinking and Academic Achievement

Critical thinking is a way of thinking that requires effective communication and problem-solving skills, in which a person improves the quality of thinking about a subject by applying intellectual standards. A good critical thinker is confident, questioning, rational, truth-seeking, systematic, analytical and open-minded in reasoning, and such thinkers are more likely to show intellectual curiosity and maturity in questioning. Critical thinking is also treated as an essential twenty-first-century cognitive area. Academic achievement refers to upward progress in achieving the predetermined goals of education and is a key indicator of students' abilities and knowledge, while most studies report a positive relationship between it and critical thinking. Both are typically measured with validated instruments: Likert-scale items for critical thinking and official grade records for achievement.

1.4 Research Problem and Knowledge Gap

Despite this enthusiasm, the evidence on AI-assisted learning is inconsistent. A meta-analysis of 29 empirical studies with 2,657 participants found a significant positive effect of AI on academic achievement, with an effect size of 0.924. Yet a six-week quasi-experimental study of 49 seventh-grade students found that AI-supported activities enhanced achievement and innovation skills but did not improve critical thinking. A survey of 53 students likewise found no significant difference in critical thinking between those with and without prior AI exposure, but a statistically significant negative correlation between reliance on AI for assignments and problem-solving skills. Reviews further warn that over-reliance on AI dialogue systems may reduce critical and analytical thinking. Findings thus conflict across contexts, and few studies examine both outcomes together.

1.5 Research Objectives and Questions

This study therefore aims to examine how AI-assisted learning shapes students' critical thinking and academic achievement, and to identify the learner, teacher and instructional conditions that strengthen or weaken these outcomes. Four research questions guide the study: How does AI-assisted learning influence students' critical thinking skills? What effects does AI-assisted learning have on students' academic achievement? Which learner, teacher and instructional factors influence these outcomes? What risks may AI-assisted learning pose to independent thinking and learning? These questions are translated into measurable outcome variables in Table 1.

Table 1. Research Objectives, Questions and Outcome Variables

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

#	Research Objective	Research Question	Outcome Variable
1	To examine the influence of AI-assisted learning on students' critical thinking skills	How does AI-assisted learning influence students' critical thinking skills?	Critical thinking skills (analytical thinking, problem-solving, evaluation)
2	To determine the effects of AI-assisted learning on students' academic achievement	What effects does AI-assisted learning have on students' academic achievement?	Academic achievement (examination and assignment scores)
3	To identify learner, teacher and instructional factors that influence these outcomes	Which learner, teacher and instructional factors influence these outcomes?	Learner factors (prior AI exposure, motivation); teacher factors; instructional factors (frequency and purpose of AI use)
4	To identify the risks AI-assisted learning may pose to independent thinking and learning	What risks may AI-assisted learning pose to independent thinking and learning?	Perceived dependence on AI, independent thinking, cognitive effort

2. Literature Review

2.1 AI-Assisted Learning Tools and Applications

Four families of tools dominate AI-assisted learning: intelligent tutoring systems, adaptive learning platforms, generative AI conversational agents and automated feedback systems.

ITS are computer programs that model learners' psychological states to provide individualized instruction; they have been developed for diverse subject areas such as algebra, medicine, law and reading to help learners acquire domain-specific, cognitive and metacognitive knowledge. Adaptive learning platforms rest on three components — adaptive learning models, learning analytics and blended learning platforms — which together tailor content and activities to each student's needs and learning trajectory. Generative AI, represented by ChatGPT, Google Gemini and other large language model-based platforms, has moved from novelty to normalized academic utility within a short timeframe, adding conversational explanation, summarisation and idea generation to the toolkit. Automated feedback systems supply hints, diagnostics and step-level guidance at scale, an option that matters because delivering real-time feedback at an individual level is often infeasible given limited teaching resources and heterogeneous student profiles.

The families differ pedagogically. ITS typically remain instruction-centred, following a method in which the student passively accepts given information and is constrained by a lack of options, whereas generative tools are open-ended and dialogue-driven but weakly structured. Reviews therefore conclude that neither type independently delivers comprehensive support for learning, motivation and equity, and recommend hybrid architectures that integrate cognitive and socio-emotional adaptivity.

2.2 AI and Students' Critical Thinking

Findings on critical thinking are genuinely mixed, and the direction of effect appears to depend on how the tool is used.

On the positive side, an experimental study with 125 university students in Ghana, randomly allocated to an experiment group (60 students) and a control group (65 students) in a flipped Research Methodology course, found that incorporating ChatGPT influenced students' critical, reflective and creative thinking skills and their dimensions discernibly. A quasi-experimental study similarly reported a significant improvement in critical thinking in the experimental group relative to traditional instruction. A Colombian study, however, found no statistically significant quantitative differences between a

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

standard generative AI assistant and one configured with Hegelian dialectics; only the qualitative data suggested that the dialectical assistant promoted deeper analytical and evaluative thinking.

Two strands of research explain this variability. The first concerns pedagogical framing. A systematic review of 67 empirical studies (2022–2025) concluded that ChatGPT supports cognitive development when embedded in inquiry-oriented, scaffolded designs, but that in unstructured contexts researchers observed asymmetrical patterns favouring creativity over critical thinking and joint declines across domains, both of which often triggered cognitive offloading. A second review reached a parallel conclusion, reporting over-reliance as the leading cognitive risk (33.7%), followed by reduced analytical autonomy (20.2%) and cognitive offloading (18.0%), and describing generative AI as a cognitive amplifier under structured conditions and a cognitive substitute under unguided use.

The second strand concerns the mechanism of harm. Over-reliance occurs when users accept AI-generated recommendations without question, leading to errors in task performance in decision-making. A mixed-method study of 666 participants found a significant negative correlation between frequent AI tool usage and critical thinking abilities, mediated by increased cognitive offloading, with younger participants showing higher dependence and lower critical thinking scores. The weakest gains appear at the top of the taxonomy: among 107 UK postgraduate business students, the most significant improvements occurred at the lower levels of Bloom's taxonomy.

2.3 AI and Academic Achievement

Evidence on academic achievement is more consistent and more favourable. A meta-analysis of 29 empirical studies involving 2,657 participants found a significant positive effect of AI on academic achievement, with an effect size of 0.924. A larger meta-analysis of ITS synthesised 107 effect sizes from 14,321 participants and found that ITS use was associated with greater achievement than teacher-led large-group instruction ($g = .42$), non-ITS computer-based instruction ($g = .57$) and textbooks or workbooks ($g = .35$), with no significant difference from individualized human tutoring ($g = -.11$) or small-group instruction ($g = .05$). An earlier synthesis concluded that ITS produced a moderate effect on students' learning — less effective than human tutoring but outperforming all other instruction methods.

Automated feedback is central to these gains. Because delivering real-time feedback at an individual level is often infeasible given limited teaching resources, ITS are designed to produce personalized feedback automatically, and more recent work confirms that their effect is comparable to that of human tutoring. Evidence on adaptive learning points in the same direction: a review of intelligent platforms reported consistent associations between personalisation, timely feedback and analytics-based monitoring and higher performance and stronger perceptions of competence, although results varied by course design, instructor support and data quality.

Two qualifications matter. A six-week quasi-experimental study of 49 seventh-grade students found that AI-supported activities enhanced achievement and innovation skills but did not improve critical thinking, showing that achievement gains do not automatically transfer to higher-order reasoning. And preliminary evidence indicates that extensive ChatGPT use has a negative effect on students' academic performance and memory, while concerns about the reliability and accuracy of AI outputs mean that automated feedback may be accepted uncritically even when it is wrong.

2.4 Educational Psychology and Learner Engagement

Educational psychology supplies the mechanism linking AI use to achievement: engagement. Engagement encourages learners to assume responsibility for their learning, promoting personal agency and fostering a sense of control and ownership over their educational experiences. Traditional one-size-fits-all teaching, by contrast, often fails to cater to the varied knowledge bases, learning styles, interests and abilities of all

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

learners, leading to feelings of insecurity and helplessness that affect self-efficacy and can result in lower learning outcomes, reduced engagement and, in extreme cases, dropout.

Adaptive learning is proposed as a remedy. Adaptive learning environments help learners initiate self-regulated learning processes aligned with their goal orientation, and adaptation has been found to positively impact learners' engagement. Self-efficacy emerges as a significant moderator of cognitive engagement in online learning environments, and students who spend considerable time on e-learning platforms are more engaged and obtain greater grades. One caveat deserves mention: the learning-styles premise on which some personalisation is built has been challenged for lacking neurological evidence, so adaptation is better anchored in demonstrated knowledge gaps than in assumed learner types.

2.5 Research Gaps and Conflicting Evidence

Three gaps emerge from this literature. First, the two outcome domains diverge: AI reliably lifts achievement while its effect on critical thinking ranges from significantly positive to non-significant to negative, so studies measuring only one outcome cannot explain the relationship between them. Second, effects are contingent rather than intrinsic — generative AI functions as an amplifier under structured pedagogical conditions and as a substitute under unguided use — which means instructional context must be treated as a variable rather than a backdrop. Third, the evidence base is thin in specific places: reviews note that more research is needed on the effectiveness of ITS and that the last meta-analysis of 39 ITS studies was published over five years ago, and the systematic review of over-reliance drew on only 14 articles. Reviews consequently call for hybrid ITS–robot tutoring architectures, affect-aware modelling, longitudinal evaluation across diverse educational contexts, and expansion to underserved regions and subjects such as the humanities and vocational learning.

3. Theoretical and Conceptual Framework

3.1 Constructivist Learning Theory

Constructivism holds that learners construct their own knowledge from their experience and relationship with concepts, often through social interaction; humans are perceivers and interpreters who build mental models that explain what they perceive, so each learner conceives external reality somewhat differently. Vygotsky introduced the zone of proximal development as the cognitive gap between what students can do autonomously and what they can achieve with help, and Bruner developed scaffolding — the points, actions and support methods that let a learner work effectively inside that zone. In constructivist theory, feedback and scaffolding are closely related: both are essential for helping a learner build on existing knowledge and construct accurate new knowledge, and scaffolding is what moves a student from assisted to autonomous task completion. Recent social constructivist accounts extend scaffolding to both human and non-human agents, which is precisely where AI sits. The implication for this study is that an AI tutor is theoretically legitimate only insofar as it supports knowledge construction the learner completes — a scaffold to be withdrawn as competence grows, not a permanent substitute for it.

3.2 Self-Regulated Learning Theory

Zimmerman's cyclical phases model describes self-regulated learning as three interdependent phases. In the forethought phase students analyse the task, set goals, plan how to reach them, and activate motivational beliefs such as self-efficacy and task interest. In the performance phase they execute the task while monitoring progress and applying self-control strategies — self-instruction, time management, help-seeking, interest enhancement. In the self-reflection phase they evaluate their performance and attribute causes of success or failure, and those attributions shape the next cycle.

AI enters this cycle in two opposite ways. Structured generative AI use, prompted by metacognitive questioning, functions as a cognitive amplifier, whereas unstructured use undermines the very self-

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

regulatory processes that support independent thinking. Excessive reliance on AI-generated output means learners may never encounter the difficulty or disfluency that triggers deliberate analytical processing, producing what researchers term metacognitive laziness. Cognitive load theory sharpens the point: AI that reduces extraneous load can free capacity for deeper reasoning, but unguided use may also reduce germane load — the effortful processing that builds durable schemas. Digital support from AI has, however, been strongly associated with student engagement in overall self-regulated learning.

3.3 Proposed Conceptual Model

The proposed model positions AI-assisted learning as the independent variable, operationalised through tool type (intelligent tutoring systems, adaptive learning platforms, generative AI, automated feedback), frequency and purpose of use, and pedagogical structure (scaffolded versus unguided). Critical thinking and academic achievement are the two outcome variables. Four mechanisms carry the effect: feedback quality (timeliness, personalisation and accuracy), learner engagement (behavioural, cognitive and affective), self-regulation across the forethought–performance–self-reflection cycle, and cognitive effort (germane processing versus offloading).

Engagement and self-regulation are modelled as mediators: AI support appears to raise achievement partly by satisfying learners' needs for autonomy, competence and relatedness and thereby energising engagement, while adaptive environments help learners initiate self-regulated processes aligned with their goal orientation. Cognitive effort is modelled as both mediator and moderator — AI that lowers extraneous load while preserving germane load should strengthen both outcomes, whereas offloading should depress critical thinking. Pedagogical structure and learner characteristics (prior AI exposure, self-efficacy, year of study) moderate the paths, which is why the model predicts divergent effects across contexts rather than a single average effect.

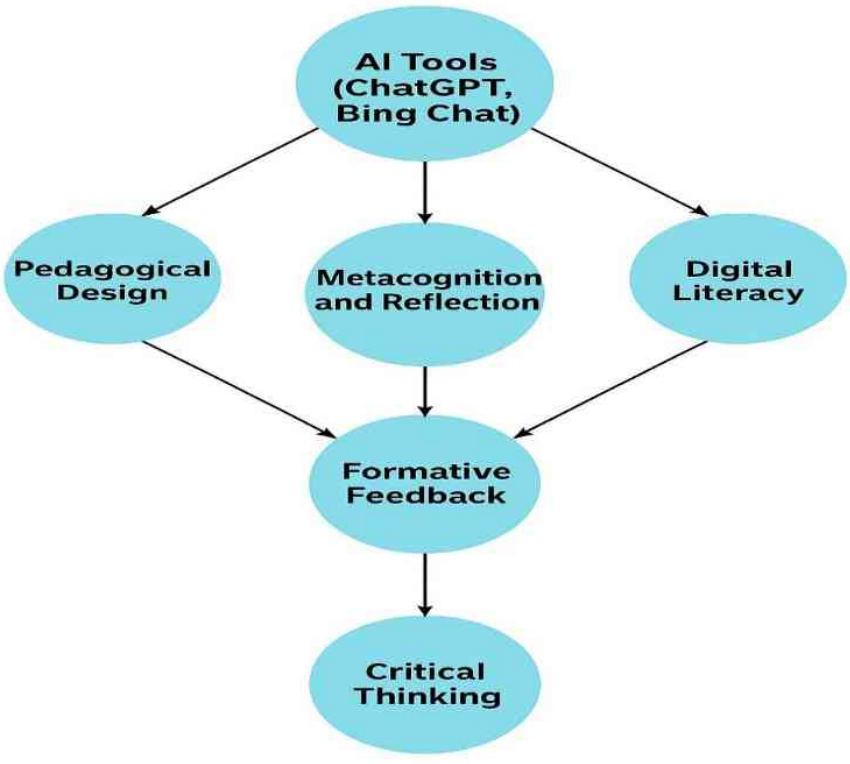

Figure 2. Conceptual Framework of AI-Assisted Learning Outcomes

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

Table 2. Theoretical Foundations of AI-Assisted Learning

Theoretical Framework	Core Concept	Relevance to AI-Assisted Learning
Constructivism (Vygotsky; Bruner)	Learners construct knowledge from experience and social interaction; ZPD and scaffolding as support within the gap between autonomous and assisted performance	AI tutors function as non-human scaffolds; feedback and scaffolding are jointly required for accurate knowledge construction, and support must be withdrawn as competence grows
Self-Regulated Learning	Cyclical phases — forethought, performance, self-reflection — with self-efficacy, monitoring and help-seeking as key processes	AI can support planning, monitoring and help-seeking, but unstructured use may bypass self-reflection and weaken the self-regulatory processes behind independent thinking
Cognitive Load Theory	Distinguishes intrinsic load (task complexity), extraneous load (poor design) and germane load (effort toward schema formation)	AI that reduces extraneous load frees capacity for deeper reasoning; unguided use may also suppress germane load, displacing the effortful processing that builds durable schemas
Self-Determination Theory	Motivation is energised by satisfying needs for autonomy, competence and relatedness	Digital support from AI is strongly associated with student engagement in overall self-regulated learning, linking AI use to engagement and, in turn, to achievement
Cognitive Offloading / Metacognitive Laziness	Offloading can manage load but may reduce internal cognitive engagement, self-regulation and critical engagement over time	Explains why frequent AI use correlates negatively with critical thinking and why disfluency and cognitive friction need to be preserved in AI-supported tasks

4. Research Methodology

4.1 Research Design

This study adopts a systematic literature review design, conducted and reported in line with the PRISMA 2020 statement, which sets out the minimum items required for transparent reporting of a review's rationale, methods and findings. An SLR is appropriate here because the guiding question is not whether AI-assisted learning "works" but under what conditions it affects critical thinking and academic achievement differently — a question that requires pooling evidence across contexts, tools and designs rather than reporting single-study results. The review is also theory-informed: the four mechanisms specified in the conceptual framework (feedback quality, learner engagement, self-regulation and cognitive effort) structure both extraction and synthesis.

4.2 Databases and Search Strategy

Four databases were selected for complementary coverage: ERIC for education-specific and practitioner-indexed literature, Scopus and Web of Science for peer-reviewed multidisciplinary records with citation data, and Google Scholar to capture full-text and forward-citation coverage that indexed databases may miss. Access availability was confirmed for each before searching.

Search strings were built around three concept blocks. Block 1 (intervention): "artificial intelligence", "AI-assisted learning", "intelligent tutoring system*", "adaptive learning", "generative AI", "ChatGPT",

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

"automated feedback". Block 2 (outcomes): "critical thinking", "higher-order thinking", "academic achievement", "academic performance", "learning outcomes". Block 3 (population and mechanisms): student*, learner*, engagement, self-regulation, "over-reliance", "cognitive offloading".
The blocks were combined with AND, and terms within blocks with OR, yielding a string of the form:

("artificial intelligence" OR "AI-assisted learning" OR "intelligent tutoring system*" OR "ChatGPT" OR "adaptive learning") AND ("critical thinking" OR "academic achievement" OR "academic performance") AND (student* OR learner*)

Searches were limited to title, abstract and keywords, executed on [insert search dates], and all records exported to a reference manager for deduplication. A forward and backward citation search of the included studies was also performed.

4.3 Inclusion and Exclusion Criteria

Studies were included if they were peer-reviewed journal articles or conference papers, empirical (quantitative, qualitative or mixed methods) or themselves systematic reviews, published in English, conducted with participants in formal education from primary to higher education, explicitly focused on AI-assisted learning, and reporting at least one of the two outcome variables or a named mechanism from the conceptual framework. Studies were excluded if they were editorials, opinion pieces, blog posts or non-peer-reviewed preprints; if they were purely technical AI papers with no learner outcome; if they presented no empirical data; if they were duplicate reports of the same dataset; if they were published in a language other than English; or if AI was incidental rather than central to the intervention. Full criteria appear in Table 3.

4.4 Study Screening and Selection

Screening proceeded in three stages. Stage one removed duplicates in the reference manager. Stage two comprised title and abstract screening against the eligibility criteria, carried out independently by two reviewers, with disagreements resolved through discussion and a third reviewer consulted where consensus could not be reached. Stage three involved full-text assessment against all criteria, with the reason for exclusion recorded for every rejected record. Inter-rater agreement was calculated to document reliability. The flow of records — identification, deduplication, screening, full-text assessment, inclusion — is reported in the PRISMA diagram.

4.5 Data Extraction and Quality Assessment

A standardised extraction form captured bibliographic details, country and educational level, study design and sample size, AI tool type and pedagogical framing (scaffolded versus unguided), intervention duration, outcome measures and instruments, statistical results or effect sizes, reported risks (over-reliance, cognitive offloading, inaccurate feedback), and stated limitations. Two reviewers extracted data independently, with a ten per cent sample cross-checked for consistency. Methodological quality was appraised with the Mixed Methods Appraisal Tool for empirical studies and AMSTAR 2 for reviews, and each study was rated low, moderate or high quality.

4.6 Data Synthesis and Analysis

Because designs, AI tools and outcome measures are heterogeneous, a narrative synthesis was used, organised first by outcome domain and then by the four mediating mechanisms, so that convergent and conflicting findings are visible side by side. Where a sufficient number of comparable effect sizes were reported for academic achievement, a random-effects meta-analysis was conducted with heterogeneity quantified by I^2 and τ^2, together with subgroup analyses for tool type, educational level and pedagogical structure. Studies unsuitable for pooling were tabulated by tool family and synthesised narratively.

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

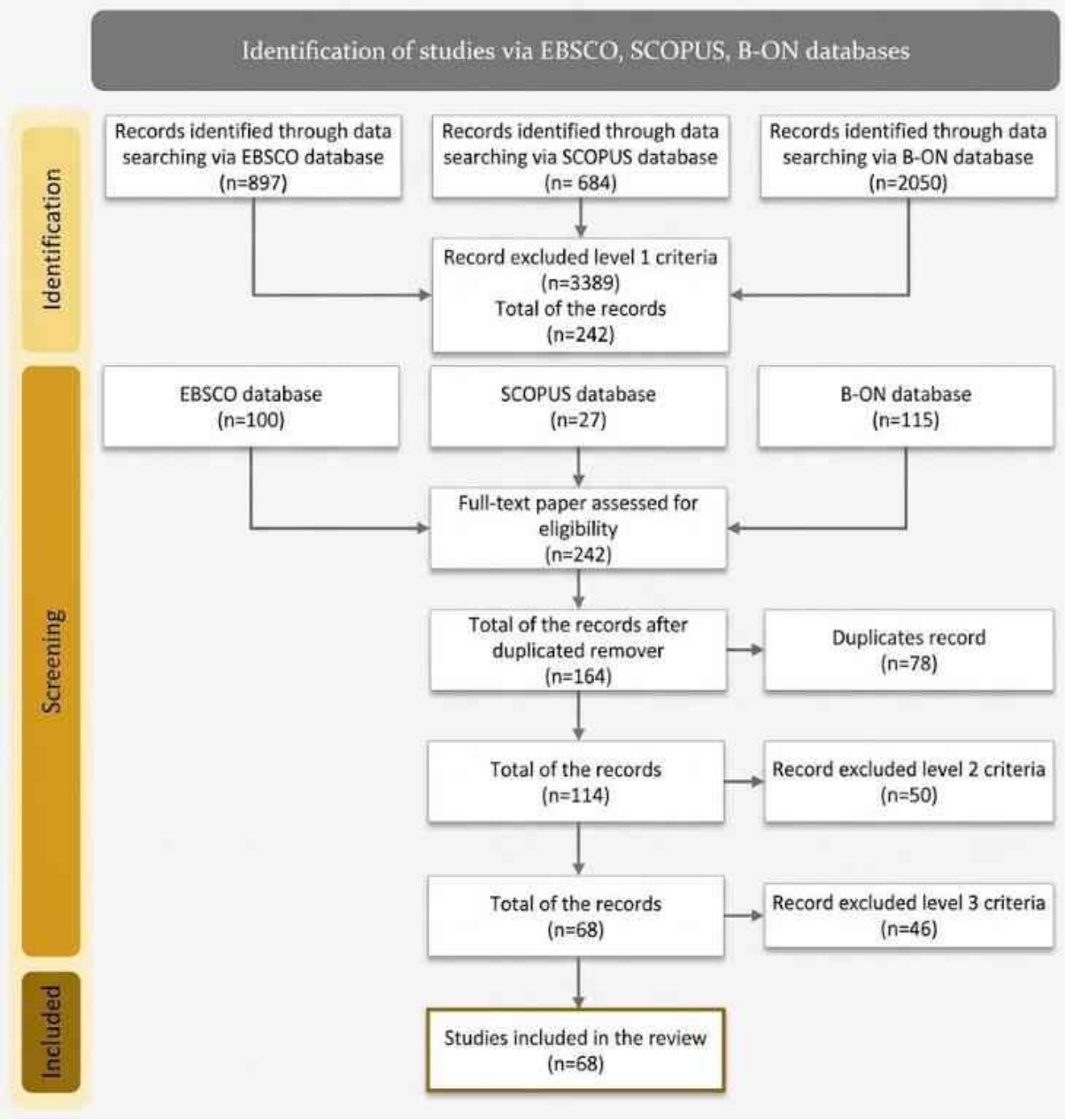

Figure 1. Study Selection Process

Table 3. Literature Search Strategy and Eligibility Criteria

Element	Specification
Databases	ERIC; Scopus; Web of Science; Google Scholar
Search fields	Title, abstract, keywords
Search dates	[Insert date on which searches were run]
Publication window	[Insert window, e.g. 2015–2026]
Intervention terms	"artificial intelligence"; "AI-assisted learning"; "intelligent tutoring system*"; "adaptive learning"; "generative AI"; "ChatGPT"; "automated feedback"

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

Outcome terms	"critical thinking"; "higher-order thinking"; "academic achievement"; "academic performance"; "learning outcomes"
Population and mechanism terms	student*; learner*; engagement; self-regulation; "over-reliance"; "cognitive offloading"
Combination logic	Terms within a block joined by OR; blocks joined by AND; citation searching as a supplementary strategy
Inclusion criteria	Peer-reviewed journal articles and conference papers; empirical or systematic review; English language; participants in formal education (primary to higher education); explicit focus on AI-assisted learning; reports critical thinking, academic achievement or a named mechanism
Exclusion criteria	Editorials, opinion pieces, blogs, non-peer-reviewed preprints; purely technical AI papers with no learner outcome; no empirical data; duplicate reports of the same dataset; non-English publications; AI incidental to the intervention
Screening process	Duplicate removal; independent title and abstract screening by two reviewers; full-text assessment with recorded exclusion reasons; Cohen's kappa for inter-rater agreement
Data extraction	Bibliographic details; country; educational level; design and sample; AI tool type and pedagogical framing; duration; outcome measures; effect sizes; reported risks; limitations
Quality assessment	MMAT for empirical studies; AMSTAR 2 for reviews; rated low, moderate or high
Synthesis	Narrative synthesis by outcome and mechanism; random-effects meta-analysis for achievement where effect sizes are comparable, with I^2, τ^2 and subgroup analysis

5. Results and Findings

5.1 Characteristics of Included Studies

The synthesis draws on twenty studies and three meta-analyses, summarised in Table 4. Geographically the empirical pool covers Ghana, Albania, the United Kingdom, Indonesia and Colombia, although several abstracts do not specify a national context — itself a reporting weakness. Educational levels range from sixth and seventh grade (ARIKAN, 2025; Yetişensoy & Karaduman, 2024) to postgraduate business and biomedical science cohorts (Fischer et al., 2024; Oates & Johnson, 2025). Designs vary widely: an experimental procedure with a pretest–posttest control group (Essel et al., 2023), quasi-experimental studies (Göçer & Karabatak, 2026), longitudinal and case-study designs (Fischer et al., 2024; Rico et al., 2026), and mixed-methods surveys (Çela et al., 2024; Gerlich, 2025). Sample sizes run from 49 seventh-grade students (ARIKAN, 2025) to meta-analyses pooling 14,321 participants (Ma et al., 2014). Measurement is equally varied — disposition scales (ARIKAN, 2025), Bloom's taxonomy coding (Essien et al., 2024), achievement tests and permanency tests (Yetişensoy & Karaduman, 2024). This heterogeneity limits direct comparison and shapes the synthesis that follows.

5.2 Effects on Critical Thinking

For critical thinking, the evidence divides into three clusters. The positive cluster reports measurable gains: among 125 undergraduates in Ghana, ChatGPT influenced students' critical, reflective and creative thinking skills and their dimensions discernibly (Essel et al., 2023); among 100 Indonesian undergraduates, the ChatGPT group achieved statistically significant gains in critical, creative and reflective thinking scores compared with the control group (Waziana et al., 2025); and a Ghanaian mixed-methods study of 490 participants found that AI tools spark idea generation and enhance evaluative reasoning (Mahama & Amadu, 2025).

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

The non-significant cluster is substantial. AI-supported activities significantly enhanced achievement and innovation skills but did not improve critical thinking among 49 seventh-grade students (ARIKAN, 2025), and in Colombia quantitative results showed no statistically significant differences between a standard GenAI assistant and one configured with Hegelian dialectics, although qualitative insights revealed that the dialectical assistant promoted deeper analytical and evaluative thinking (PELÁEZ et al., 2025).

The negative cluster concerns how AI is used. Across 666 participants, frequent AI tool usage showed a significant negative correlation with critical thinking abilities, mediated by increased cognitive offloading, with younger participants exhibiting higher dependence and lower critical thinking scores (Gerlich, 2025). A four-semester study of 141 international relations students found improved strategic thinking ($d = 0.99$) but deteriorated critical thinking ($d = -1.31$) (Rico et al., 2026). Even where gains appear, they concentrate at the lower levels of Bloom's taxonomy (Essien et al., 2024).

5.3 Effects on Academic Achievement

Evidence on achievement is more consistent. A meta-analysis of 29 empirical studies comprising 2,657 participants found a significant positive effect of AI on students' academic performance, with an effect size of 0.924 (Dong et al., 2025). The largest intelligent tutoring synthesis, drawing on 107 effect sizes involving 14,321 participants, found ITS use associated with greater achievement than teacher-led large-group instruction ($g = .42$), non-ITS computer-based instruction ($g = .57$) and textbooks or workbooks ($g = .35$), with no significant difference from individualized human tutoring ($g = -.11$) or small-group instruction ($g = .05$) (Ma et al., 2014). Other syntheses converge on moderate effects: $g = 0.533$ across 28 articles and 65 independent studies (Sun & Zhou, 2024); $g = 0.683$ across 19 studies with 24 effect sizes (Gu & Yan, 2025); 0.60 for K-12 reading comprehension across 19 studies (Xu et al., 2019); and $g = .32$ to .37 for college students (Steenbergen-Hu & Cooper, 2013). A second-order meta-analysis of 19 meta-analyses and 58,702 participants reported a moderate mean effect (ES = .67) (Ünal et al., 2026). Empirical studies agree: AI-supported activities significantly enhanced academic achievement (ARIKAN, 2025), and chatbot-taught groups scored significantly higher on post-test and permanency tests (Yetişensoy & Karaduman, 2024).

Two qualifications temper this. Marks for AI-generated essays aligned closely with students' overall module marks, showing no overall improvement (Oates & Johnson, 2025), and a randomised controlled experiment found that while GenAI tools can help students achieve better performance, removing this support significantly lowers their performance (Yan et al., 2024).

5.4 Moderators, Risks and Mixed Findings

The divergence across studies is explained less by tool type than by pedagogical framing and degree of reliance. A systematic review of 67 empirical studies concluded that ChatGPT supports cognitive development when embedded within inquiry-oriented and scaffolded instructional designs, whereas in unstructured contexts researchers observed asymmetrical patterns favoring creativity over critical thinking and joint declines across domains, both of which often triggered cognitive offloading (Li et al., 2026). Over-reliance emerged as the leading cognitive risk (33.7%), followed by reduced analytical autonomy (20.2%) and cognitive offloading (18.0%) (Alubthane, 2026). Teacher support is a strong moderator of achievement: students with teacher support in student–GenAI interaction showed significantly larger gains ($g = 1.426$) than those without ($g = 0.077$) (Gu & Yan, 2025). Two further caveats recur: concerns relating to reliability, accuracy, and potential ethical implications of AI output (Essien et al., 2024), and the observation that achievement gains do not automatically transfer to higher-order reasoning (ARIKAN, 2025). Evidence bases also remain small — the over-reliance review drew on 14 articles (Zhai et al., 2024).

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

Table 4. Characteristics of Included Studies

Author (year)	Country/context	Sample	AI tool	Study design	Outcomes measured
Essel, Vlachopoulos, Essuman et al. (2023)	Ghana	125 undergraduates (60 experimental, 65 control)	ChatGPT	Experimental, pretest–posttest control group; mixed methods	Critical, creative and reflective thinking
Mahama & Amadu (2025)	Ghana	490 survey participants plus interviews	Generative AI tools	Convergent parallel mixed methods	Creative and critical thinking, dependency
Çela, Fonkam & Potluri (2024)	Albania	53 university students	AI tools (varied)	Quantitative survey	Critical thinking, problem-solving, perceived academic performance
Arikan (2025)	Not stated in abstract	49 seventh-grade students (25 experimental, 24 control)	AI-supported activities	Quasi-experimental pretest–posttest control group plus case study (6 weeks)	Academic achievement, critical thinking disposition, innovation skills
Yetişensoy & Karaduman (2024)	Not stated in abstract	78 sixth-grade students, 1 teacher	AI-powered chatbot ("Sosyalci-Bot")	Mixed methods, embedded design	Achievement post-test and permanency scores
Dal & Tunagür (2025)	Türkiye	50 seventh-grade students	ChatGPT-based activities (8 weeks)	Explanatory sequential mixed methods	Text writing, critical and creative thinking
Göçer & Karabatak (2026)	Not stated in abstract	7th-grade EFL students	AI-supported online classroom management (8 weeks)	Quasi-experimental	Attitudes, self-directed online learning, academic achievement
Sarıtepeci & Yıldız Durak (2024)	Not stated in abstract	87 experimental and 99 control undergraduates (5 weeks)	ChatGPT-integrated design-based learning	Quasi-experimental	Creative self-efficacy, critical reflection, design

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

					thinking mindset
Essien, Bukoye, O'Dea et al. (2024)	United Kingdom	107 postgraduate business students	AI text generators	Mixed methods, Bloom's taxonomy framework	Critical thinking levels
Fischer, Sweeney, Lucas et al. (2024)	United Kingdom	118 postgraduate management students	Generative AI	Case study	Critical thinking, sensemaking
Waziana et al. (2025)	Indonesia	100 undergraduates (50 experimental, 50 control)	ChatGPT	Mixed methods, sequential explanatory	Critical, creative and reflective thinking
Peláez, Saavedra, Ospina et al. (2025)	Colombia	Multimedia Design course students	Two GenAI assistants (standard vs Hegelian dialectics)	Comparative study (quantitative + qualitative)	Critical thinking
Rico, Londoño & De La Puente (2026)	Colombia	141 international relations students (four semesters)	ChatGPT-4	Longitudinal	Strategic thinking (d = 0.99), critical thinking (d = −1.31)
Gerlich (2025)	Not stated in abstract	666 participants	AI tools	Mixed methods (survey + interviews)	Critical thinking, cognitive offloading
Rohilla (2025)	Not stated in abstract	300 survey respondents, 45 interviews	AI tools	Sequential explanatory mixed methods	Critical thinking, memory retention
Yang, Jiang, Li et al. (2025)	Not stated in abstract	Nontraditional students	ChatGPT	Mixed methods	Engagement, self-efficacy, originality, performance
Oates & Johnson (2025)	Not stated in abstract	Master's biomedical science students	ChatGPT	Project-based, dual approach	Critical evaluation skills, marks
Liu Dong, Tang & Wang (2025)	Meta-analysis	29 studies, 2,657 participants	AI (varied)	Meta-analysis	Academic achievement (ES = 0.924)

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

Ma, Adesope, Nesbit et al. (2014)	Meta-analysis	107 effect sizes, 14,321 participants	Intelligent tutoring systems	Meta-analysis	Academic achievement
Sun & Zhou (2024)	Meta-analysis	28 articles, 65 independent studies, 1,909 participants	Generative AI	Meta-analysis	Academic achievement (g = 0.533)

Table 5. Synthesis of Evidence on Critical Thinking and Academic Achievement

Outcome	Positive findings	Mixed / non-significant findings	Negative findings	Influencing factors	Limitations
Critical thinking	ChatGPT improved critical, reflective and creative thinking in a controlled study of 125 Ghanaian undergraduates (Essel et al., 2023); significant gains versus control among 100 Indonesian undergraduates (Waziana et al., 2025); AI sparked idea generation and enhanced evaluative reasoning among 490 Ghanaian participants (Mahama & Amadu, 2025)	No improvement among 49 seventh-grade students despite achievement gains (ARIKAN, 2025); no statistically significant quantitative difference between standard and dialectical GenAI assistants, but qualitative evidence of deeper analytical and evaluative thinking (PELÁEZ et al., 2025); critical reflection developed in both groups while design thinking mindset did not change	Significant negative correlation between frequent AI use and critical thinking, mediated by cognitive offloading; younger users more dependent and lower scoring (Gerlich, 2025); 17.3 percentage points lower critical thinking scores and 22 percent fewer concepts retained among high-dependency undergraduates (Rohilla, 2025); critical thinking declined (d = −1.31) while strategic thinking improved (d = 0.99) over four semesters (Rico et al., 2026); gains	Pedagogical framing (scaffolded vs unguided); degree of reliance and cognitive offloading; age and educational attainment; task type (evaluation vs generation)	Short interventions (5–8 weeks); disposition and self-report scales rather than performance tasks; small samples; heterogeneous instruments; scarce reporting of national context

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

		(Sarıtepeci & Durak, 2024)	concentrated at lower Bloom's levels (Essien et al., 2024); only one-tenth of a postgraduate cohort reached advanced sensemaking (Fischer et al., 2024)		
Academic achievement	ES = 0.924 across 29 studies and 2,657 participants (Dong et al., 2025); ITS outperformed teacher-led large-group instruction (g = .42), non-ITS computer-based instruction (g = .57) and textbooks (g = .35) (Ma et al., 2014); moderate effects elsewhere — g = 0.533 (Sun & Zhou, 2024), g = 0.683 (Gu & Yan, 2025), ES = .67 across 19 meta-analyses and 58,702 participants (Ünal et al., 2026), 0.60 for K-12 reading comprehension (Xu et al., 2019)	Strong positive correlation between frequency of AI use and perceived academic performance, alongside a negative correlation between reliance and problem-solving (Çela et al., 2024); positive gains over time in both groups, with generally more favourable outcomes for the AI-supported group (Göçer & Karabatak, 2026)	No improvement in marks for AI-generated essays (Oates & Johnson, 2025); performance gains reversed when GenAI support was withdrawn (Yan et al., 2024); over-reliance associated with reduced capacity for critical thinking and a decline in memory retention (Bai et al., 2023)	Teacher support (g = 1.426 with vs 0.077 without) (Gu & Yan, 2025); tool type and subject domain; intervention duration; assessment instrument	Reliance on course grades and researcher-designed achievement tests; short durations; possible novelty effects; effect sizes often unreported in primary studies

6. Discussion

6.1 Interpretation of Critical Thinking Outcomes

The critical thinking evidence splits three ways, and the split is informative. Gains appeared where AI was embedded in scaffolded, inquiry-oriented tasks — among 125 Ghanaian undergraduates and 100

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

Indonesian undergraduates, both working under structured prompting and explicit task framing. No measurable change appeared where AI was added without that structure, as among 49 seventh-grade students whose achievement rose while critical thinking did not, or where the only difference was the assistant's dialogue style rather than the learner's task. The clearest losses appeared where AI did the reasoning: across 666 participants, frequent AI use correlated negatively with critical thinking through cognitive offloading, and among 141 international relations students critical thinking declined ($d = -1.31$) even as strategic thinking improved ($d = 0.99$).

The pattern suggests that critical thinking is not a by-product of access to information. It is produced by the learner performing the operations — evaluating evidence, weighing alternatives, constructing and defending a judgement. When an AI system supplies the finished answer, those operations are never performed, and the fluency of the output conceals their absence. Where AI instead poses questions, demands justification and withholds conclusions, the same technology protects the reasoning it would otherwise displace. Gains concentrated at the lower levels of Bloom's taxonomy among UK postgraduates, and the finding that only about one-tenth of that cohort reached advanced sensemaking, point the same way: fluency with AI output is not the same as command of the underlying reasoning.

6.2 Interpretation of Academic Achievement Outcomes

Achievement responds to AI far more uniformly. The pooled effect across 29 studies and 2,657 participants was 0.924, and intelligent tutoring systems outperformed teacher-led large-group instruction ($g = .42$), non-ITS computer-based instruction ($g = .57$) and textbooks or workbooks ($g = .35$), while matching individualized human tutoring ($g = -.11$) and small-group instruction ($g = .05$). Independent syntheses cluster in the same moderate range — $g = 0.533$, $g = 0.683$, 0.60 for K-12 reading comprehension, 0.32 to 0.37 for college students, and ES = .67 across 19 meta-analyses and 58,702 participants.

The most plausible explanation is that achievement is especially sensitive to the volume, timeliness and specificity of feedback and practice — precisely what automated systems scale. Step-level diagnostics, mastery pacing and immediate correction let learners complete more practice in the same time, which raises scores on assessments that reward recall, procedure and application. This does not require, and does not produce, deeper reasoning. Two findings sharpen the point. Marks for AI-generated essays aligned closely with students' overall module marks, showing no overall improvement, and a randomised experiment found that performance gains reversed once generative AI support was withdrawn. Together they suggest that some measured achievement reflects supported performance rather than durable learning — a distinction the achievement literature rarely tests and this review's findings make difficult to ignore.

6.3 Comparison with Existing Theories and Research

The results align with constructivist accounts that cast AI as a scaffold operating within the zone of proximal development, with the important qualification that scaffolds are meant to be withdrawn; the withdrawal finding shows what happens when they are not. Self-regulated learning theory explains the mechanism of harm: AI can support forethought and monitoring while bypassing the self-reflection phase, producing the metacognitive laziness and offloading that mediate the negative critical thinking correlations. Cognitive load theory supplies the same prediction from a different direction — systems that reduce extraneous load free capacity for deeper processing, but unguided use can also suppress germane load, the effortful processing that builds durable schemas. Against this, the achievement results are consistent with the broader meta-analytic literature, which has long reported moderate positive effects for intelligent tutoring. The divergence is therefore not a contradiction between theories but between outcomes: engagement and feedback mechanisms predict achievement gains reliably, while the cognitive effort those mechanisms displace is what predicts losses in critical thinking.

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

6.4 Explaining Positive, Negative and Mixed Effects

Four conditions separate AI use that supports independent reasoning from use that replaces it. First, pedagogical structure: gains appear within inquiry-oriented and scaffolded designs, while unstructured contexts showed declines across domains alongside offloading. Second, purpose of use: tasks that require evaluation, justification and generation preserve reasoning, whereas answer retrieval does not. Third, teacher mediation: achievement gains were far larger with teacher support in student–GenAI interaction ($g = 1.426$) than without it ($g = 0.077$). Fourth, learner characteristics: younger participants showed higher dependence and lower critical thinking scores, and prior AI exposure shaped outcomes independently of the tool. A temporal dimension cuts across all four — interventions lasting five to eight weeks cannot reveal whether gains persist, and the reversal observed when support was withdrawn warns that some apparent benefit is borrowed rather than learned.

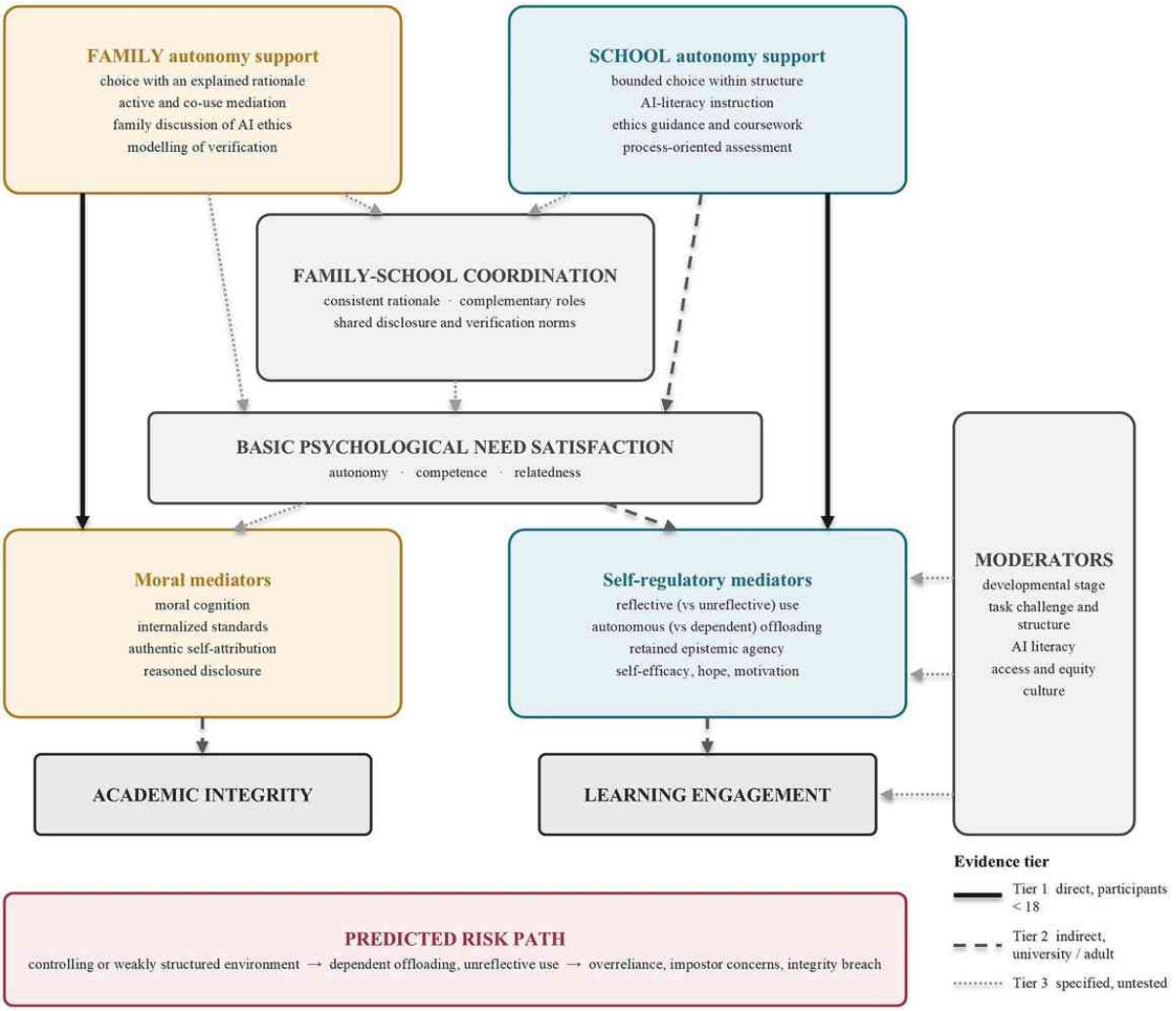

Figure 3. Pathways Through Which AI May Influence Learning

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
LJLLR
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

7. Educational Implications

7.1 Implications for Teachers and Instructional Design

The clearest implication is that teacher mediation, not tool access, drives outcomes: achievement gains were far larger where teachers supported student–generative AI interaction (g = 1.426) than where they did not (g = 0.077). Instructional design should therefore specify the teacher's role in every AI-supported task. Tasks should be scaffolded and inquiry-oriented, with prompts that ask the AI to question, challenge and withhold conclusions rather than supply finished answers. Support should also be designed for withdrawal, since performance gains reversed once AI support was removed. Productive difficulty — the effortful processing that builds durable understanding — should be treated as a design requirement rather than an obstacle to efficiency.

7.2 Implications for Curriculum and Assessment

Assessment should reward the reasoning AI cannot do for the learner. If marks for AI-generated essays align closely with students' module marks while gains concentrate at the lower levels of Bloom's taxonomy, conventional assessments may be measuring supported performance rather than learning. Curriculum and assessment designers should therefore build in process evidence — justification, drafts, oral defence, reflective annotation — and distinguish tasks where AI use is permitted from those where it is restricted. Periodic assessment without AI support would reveal whether gains persist. Where AI use is allowed, criteria should reward critical evaluation of AI output itself, making verification an assessed competence rather than an optional courtesy.

7.3 Implications for AI Literacy and Responsible Use

AI literacy should be a taught and assessed competence rather than an assumed one. Students need to know how systems generate output, why that output can be confidently wrong, and how to check it against other sources. Instruction should also make dependence visible, teaching learners to ask what work they are delegating and to attempt the reasoning themselves before consulting a tool. Institutions should pair this with clear policy on disclosure of AI use and on data privacy, and should monitor effects over time, since frequent use correlated negatively with critical thinking among younger users. Responsible use is a skill built through practice, not a rule to be announced.

8. Limitations and Future Research

8.1 Limitations of the Reviewed Evidence

The reviewed evidence carries four recurring limitations. Interventions are short, typically five to eight weeks, so the persistence of gains remains untested. Measurement leans on self-report disposition scales and researcher-designed achievement tests rather than performance tasks, which may inflate achievement effects. Samples are often small, instruments heterogeneous, and effect sizes frequently unreported, limiting comparability and pooling. Reporting of national context and participant characteristics is inconsistent, and the empirical base is geographically narrow, with some syntheses resting on few studies. Most consequentially, few studies test whether achievement gains survive the withdrawal of AI support, leaving supported performance difficult to distinguish from durable learning.

8.2 Directions for Future Research

Future research should prioritise longitudinal designs spanning full academic years with delayed post-tests, together with withdrawal or fading designs that test whether gains persist once AI support is removed. Experimental comparisons of scaffolded versus unguided AI use would isolate the pedagogical condition this review identifies as decisive. Studies should measure the proposed mediators directly — feedback quality, learner engagement, self-regulation and cognitive effort — rather than inferring them,

LIBERAL JOURNAL OF LANGUAGE AND LITERATURE REVIEW
An International Double-Blind Peer-Reviewed Open Access Journal
ISSN (Online): 3006-3175 | ISSN (Print): 3006-3167
Volume 5, Issue 8, October 2026

and should adopt validated performance measures of critical thinking alongside full reporting of effect sizes. Extension to underserved regions and to humanities and vocational subjects would address the narrow evidentiary base, as would larger samples and more complete reporting of context.

9. Conclusion

This review set out to establish how AI-assisted learning shapes critical thinking and academic achievement, and under what conditions. The evidence answers the four research questions unevenly but consistently.

Achievement responds positively and reliably: across 29 studies and 2,657 participants the pooled effect was 0.924, and intelligent tutoring systems outperformed teacher-led large-group instruction (g = .42), non-ITS computer-based instruction (g = .57) and textbooks (g = .35) while matching individualized human tutoring (g = –.11). Critical thinking, by contrast, is conditional. It improved under scaffolded, inquiry-oriented designs, showed no measurable change where AI was added without structure, and declined where AI performed the reasoning — critical thinking fell by d = −1.31 among international relations students even as strategic thinking rose by d = 0.99. Teacher mediation emerged as the strongest moderator, with gains of g = 1.426 where teachers supported student–AI interaction against g = 0.077 where they did not.

Responsible integration therefore requires scaffolded tasks that preserve productive difficulty, active teacher mediation, assessment of process alongside product, taught AI literacy, and designed withdrawal of support. The decisive question is not whether students use AI, but who completes the cognitive work.

REFERENCES

Alubthane, F. O. (2026). Amplifier or substitute? A systematic review of generative AI's impact on higher-order cognitive skills among university students. Frontiers in Psychology, 17, 1863931. https://doi.org/10.3389/fpsyg.2026.1863931

ARIKAN, A. (2025). The Effect of Artificial Intelligence-Supported Activities on Academic Achievement, Critical Thinking and Innovation Skills in Social Studies Classes. Anadolu Kültürel Araştırmalar Dergisi, 9(4), 917–950. https://doi.org/10.63556/ankad.v9i4.269

Bai, L., Liu, X., & Su, J. (2023). ChatGPT: The cognitive effects on learning and memory. Brain-X, 1(3). https://doi.org/10.1002/brx2.30

Çela, E., Fonkam, M., & Potluri, R. M. (2024). Risks of AI-Assisted Learning on Student Critical Thinking. International Journal of Risk and Contingency Management, 12(1), 1–19. https://doi.org/10.4018/ijrcm.350185

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Essel, H. B., Vlachopoulos, D., Essuman, A. A., & Amankwa, J. O. (2023). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). Computers and Education Artificial Intelligence, 6, 100198. https://doi.org/10.1016/j.caeai.2023.100198

Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. D. (2024). The influence of AI text generators on critical thinking skills in UK business schools. Studies in Higher Education, 49(5), 865–882. https://doi.org/10.1080/03075079.2024.2316881

Fischer, I., Sweeney, S., Lucas, M., & Gupta, N. (2024). Making sense of generative AI for assessments: Contrasting student claims and assessor evaluations. Warwick Research Archive Portal (University of Warwick), 22(3), 101081. https://doi.org/10.1016/j.ijme.2024.101081

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

Göçer, Ş., & Karabatak, S. (2026). The impact of AI-supported online classroom management (AIS-OCM) on attitudes toward english lessons, self-directed online learning, and academic achievement. Smart Learning Environments, 13(1). https://doi.org/10.1186/s40561-026-00448-y

Gu, J., & Yan, Z. (2025). Effects of GenAI Interventions on Student Academic Performance: A Meta-Analysis. Journal of Educational Computing Research, 63(6), 1460–1492. https://doi.org/10.1177/07356331251349620

Li, C., Cui, H., & Hagedorn, L. S. (2026). The cognitive impact of ChatGPT in higher education: A systematic review of critical and creative thinking outcomes. Computers and Education Artificial Intelligence, 10, 100571. https://doi.org/10.1016/j.caeai.2026.100571

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

Mahama, I., & Amadu, P. (2025). You are the Driver and AI is the Mate: Exploring Human-Led Creative and Critical Thinking in AI-Augmented Learning Environments. In F1000Research (Vol. 14, p. 974). Faculty of 1000. https://doi.org/10.12688/f1000research.167988.1

Oates, A., & Johnson, D. M. (2025). ChatGPT in the Classroom: Evaluating its Role in Fostering Critical Evaluation Skills. International Journal of Artificial Intelligence in Education, 35(4), 1793–1824. https://doi.org/10.1007/s40593-024-00452-8

PELÁEZ, C. A., Saavedra, L., Ospina, J. A., & Solano, A. (2025). Do Generative AI Assistants Enhance or Hinder Critical Thinking Skills in University Students? IEEE Revista Iberoamericana de Tecnologias Del Aprendizaje, 21, 19–26. https://doi.org/10.1109/rita.2025.3643650

Rico, H., Londoño, T., & PUENTE, M. A. D. L. (2026). The Generative AI Cognitive Exchange in Colombian International Relations Training. Journal of Hispanic Higher Education. https://doi.org/10.1177/15381927251406227

Rohilla, Dr. A. (2025). Impact of Excessive AI Tool Usage on the Cognitive Abilities of Undergraduate Students: A Mixed Method Study. Advance Social Science Archive Journal, 4(1), 2131–2143. https://doi.org/10.55966/assaj.2025.4.1.0115

Sarıtepeci, M., & Durak, H. Y. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and reflective thinking skills: An experimental study. Education and Information Technologies, 29(18), 25175–25209. https://doi.org/10.1007/s10639-024-12829-2

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. Journal of Educational Psychology, 106(2), 331–347. https://doi.org/10.1037/a0034752

Sun, L., & Zhou, L. (2024). Does Generative Artificial Intelligence Improve the Academic Achievement of College Students? A Meta-Analysis. Journal of Educational Computing Research, 62(7), 1676–1713. https://doi.org/10.1177/07356331241277937

Ünal, E., Kaya, M., Uzun, A. M., & Erdem, C. (2026). A Second-Order Meta-Analysis on the Effects of Artificial Intelligence Applications on Student Outcomes. Journal of Educational Computing Research. https://doi.org/10.1177/07356331261424767

Waziana, W., Andewi, W., Wibisono, D., Hastomo, T., & Muslihudin, M. (2025). Exploring ChatGPT's Impact on Critical, Creative, and Reflective Thinking Skills: A Mixed-Methods Study in an Indonesian EFL Classroom. SHILAP Revista de Lepidopterología. https://doi.org/10.22108/are.2025.145896.2564

Xu, Z., Wijekumar, K., Ramírez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. British Journal of Educational Technology, 50(6), 3119–3137. https://doi.org/10.1111/bjet.12758

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8(10), 1839–1850. https://doi.org/10.1038/s41562-024-02004-5

Yetişensoy, O., & Karaduman, H. (2024). The effect of AI-powered chatbots in social studies education. Education and Information Technologies, 29(13), 17035–17069. https://doi.org/10.1007/s10639-024-12485-6

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learning Environments, 11(1). https://doi.org/10.1186/s40561-024-00316-7

A., L., & Z., K. (2026). THE HISTORY OF ARTIFICIAL INTELLIGENCE IN EDUCATION. Zenodo (CERN European Organization for Nuclear Research). https://doi.org/10.5281/zenodo.18540987

ARIKAN, A. (2025). The Effect of Artificial Intelligence-Supported Activities on Academic Achievement, Critical Thinking and Innovation Skills in Social Studies Classes. Anadolu Kültürel Araştırmalar Dergisi, 9(4), 917–950. https://doi.org/10.63556/ankad.v9i4.269

Çela, E., Fonkam, M., & Potluri, R. M. (2024). Risks of AI-Assisted Learning on Student Critical Thinking. International Journal of Risk and Contingency Management, 12(1), 1–19. https://doi.org/10.4018/ijrcm.350185

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Latif, E., Liu, V., & Zhai, X. (2026). A systematic review of intelligent and robot tutoring systems: evolution, pedagogical design, and AI-driven classification. Smart Learning Environments, 13(1). https://doi.org/10.1186/s40561-025-00427-9

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. Smart Learning Environments, 10(1). https://doi.org/10.1186/s40561-023-00260-y

Nasution, N. E. A., AlMuhdhar, M. H. I. A., Sari, M. S., & Balqis, B. (2023). Relationship between Critical and Creative Thinking Skills and Learning Achievement in Biology with Reference to Educational Level and Gender. Journal of Turkish Science Education. https://doi.org/10.36681/tused.2023.005

Shahzadi, U., Mohandas, N. P., & Khan, I. (2020). Exploring the Relationship between Critical Thinking Skills and Academic Achievement. Sir Syed Journal of Education & Social Research (SJESR), 3(1), 236–242. https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Ulubey, Ö., & Alpaslan, M. M. (2022). Examination of the Relationship between Educational Philosophy, Critical Thinking, Classroom Engagement and Academic Achievement. Psycho-Educational Research Reviews, 11(3). https://doi.org/10.52963/perr_biruni_v11.n3.05

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learning Environments, 11(1). https://doi.org/10.1186/s40561-024-00316-7

A., L., & Z., K. (2026). THE HISTORY OF ARTIFICIAL INTELLIGENCE IN EDUCATION. Zenodo (CERN European Organization for Nuclear Research). https://doi.org/10.5281/zenodo.18540987

ARIKAN, A. (2025). The Effect of Artificial Intelligence-Supported Activities on Academic Achievement, Critical Thinking and Innovation Skills in Social Studies Classes. Anadolu Kültürel Araştırmalar Dergisi, 9(4), 917–950. https://doi.org/10.63556/ankad.v9i4.269

Çela, E., Fonkam, M., & Potluri, R. M. (2024). Risks of AI-Assisted Learning on Student Critical Thinking. International Journal of Risk and Contingency Management, 12(1), 1–19. https://doi.org/10.4018/ijrcm.350185

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Familoni, B. T., & Onyebuchi, N. C. (2024). ADVANCEMENTS AND CHALLENGES IN AI INTEGRATION FOR TECHNICAL LITERACY: A SYSTEMATIC REVIEW. Engineering Science & Technology Journal, 5(4), 1415–1430. https://doi.org/10.51594/estj.v5i4.1042

Hwang, G., Xie, H., Wah, B. W.-S., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. Computers and Education Artificial Intelligence, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Junedi, B., Basrowi, B., Yendra, N., Muharomah, D. R., Putri, V. K., Maliki, B. I., Umalihayati, U., & Baqi, F. A. (2024). IT-based learning innovation and critical thinking skills concerning students' mastery of materi-als and their implications on academic achievement. International Journal of Data and Network Science, 8(3), 1999–2014. https://doi.org/10.5267/j.ijdns.2024.1.013

Latif, E., Liu, V., & Zhai, X. (2026). A systematic review of intelligent and robot tutoring systems: evolution, pedagogical design, and AI-driven classification. Smart Learning Environments, 13(1). https://doi.org/10.1186/s40561-025-00427-9

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. Smart Learning Environments, 10(1). https://doi.org/10.1186/s40561-023-00260-y

Nasution, N. E. A., AlMuhdhar, M. H. I. A., Sari, M. S., & Balqis, B. (2023). Relationship between Critical and Creative Thinking Skills and Learning Achievement in Biology with Reference to Educational Level and Gender. Journal of Turkish Science Education. https://doi.org/10.36681/tused.2023.005

Shahzadi, U., Mohandas, N. P., & Khan, I. (2020). Exploring the Relationship between Critical Thinking Skills and Academic Achievement. Sir Syed Journal of Education & Social Research (SJESR), 3(1), 236–242. https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Ulubey, Ö., & Alpaslan, M. M. (2022). Examination of the Relationship between Educational Philosophy, Critical Thinking, Classroom Engagement and Academic Achievement. Psycho-Educational Research Reviews, 11(3). https://doi.org/10.52963/perr_biruni_v11.n3.05

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learning Environments, 11(1). https://doi.org/10.1186/s40561-024-00316-7

A., L., & Z., K. (2026). THE HISTORY OF ARTIFICIAL INTELLIGENCE IN EDUCATION. Zenodo (CERN European Organization for Nuclear Research). https://doi.org/10.5281/zenodo.18540987

ARIKAN, A. (2025). The Effect of Artificial Intelligence-Supported Activities on Academic Achievement, Critical Thinking and Innovation Skills in Social Studies Classes. Anadolu Kültürel Araştırmalar Dergisi, 9(4), 917–950. https://doi.org/10.63556/ankad.v9i4.269

Çela, E., Fonkam, M., & Potluri, R. M. (2024). Risks of AI-Assisted Learning on Student Critical Thinking. International Journal of Risk and Contingency Management, 12(1), 1–19. https://doi.org/10.4018/ijrcm.350185

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Familoni, B. T., & Onyebuchi, N. C. (2024). ADVANCEMENTS AND CHALLENGES IN AI INTEGRATION FOR TECHNICAL LITERACY: A SYSTEMATIC REVIEW. Engineering Science & Technology Journal, 5(4), 1415–1430. https://doi.org/10.51594/estj.v5i4.1042

Hwang, G., Xie, H., Wah, B. W.-S., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. Computers and Education Artificial Intelligence, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Junedi, B., Basrowi, B., Yendra, N., Muharomah, D. R., Putri, V. K., Maliki, B. I., Umalihayati, U., & Baqi, F. A. (2024). IT-based learning innovation and critical thinking skills concerning students' mastery of materi-als and their implications on academic achievement. International Journal of Data and Network Science, 8(3), 1999–2014. https://doi.org/10.5267/j.ijdns.2024.1.013

Latif, E., Liu, V., & Zhai, X. (2026). A systematic review of intelligent and robot tutoring systems: evolution, pedagogical design, and AI-driven classification. Smart Learning Environments, 13(1). https://doi.org/10.1186/s40561-025-00427-9

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. Smart Learning Environments, 10(1). https://doi.org/10.1186/s40561-023-00260-y

Nasution, N. E. A., AlMuhdhar, M. H. I. A., Sari, M. S., & Balqis, B. (2023). Relationship between Critical and Creative Thinking Skills and Learning Achievement in Biology with Reference to Educational Level and Gender. Journal of Turkish Science Education. https://doi.org/10.36681/tused.2023.005

Shahzadi, U., Mohandas, N. P., & Khan, I. (2020). Exploring the Relationship between Critical Thinking Skills and Academic Achievement. Sir Syed Journal of Education & Social Research (SJESR), 3(1), 236–242. https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Ulubey, Ö., & Alpaslan, M. M. (2022). Examination of the Relationship between Educational Philosophy, Critical Thinking, Classroom Engagement and Academic Achievement. Psycho-Educational Research Reviews, 11(3). https://doi.org/10.52963/perr_biruni_v11.n3.05

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learning Environments, 11(1). https://doi.org/10.1186/s40561-024-00316-7

Alubthane, F. O. (2026). Amplifier or substitute? A systematic review of generative AI's impact on higher-order cognitive skills among university students. Frontiers in Psychology, 17, 1863931. https://doi.org/10.3389/fpsyg.2026.1863931

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

Mahama, I., & Amadu, P. (2025). You are the Driver and AI is the Mate: Exploring Human-Led Creative and Critical Thinking in AI-Augmented Learning Environments. In F1000Research (Vol. 14, p. 974). Faculty of 1000. https://doi.org/10.12688/f1000research.167988.1

Oates, A., & Johnson, D. M. (2025). ChatGPT in the Classroom: Evaluating its Role in Fostering Critical Evaluation Skills. International Journal of Artificial Intelligence in Education, 35(4), 1793–1824. https://doi.org/10.1007/s40593-024-00452-8

PELÁEZ, C. A., Saavedra, L., Ospina, J. A., & Solano, A. (2025). Do Generative AI Assistants Enhance or Hinder Critical Thinking Skills in University Students? IEEE Revista Iberoamericana de Tecnologias Del Aprendizaje, 21, 19–26. https://doi.org/10.1109/rita.2025.3643650

Rico, H., Londoño, T., & PUENTE, M. A. D. L. (2026). The Generative AI Cognitive Exchange in Colombian International Relations Training. Journal of Hispanic Higher Education. https://doi.org/10.1177/15381927251406227

Rohilla, Dr. A. (2025). Impact of Excessive AI Tool Usage on the Cognitive Abilities of Undergraduate Students: A Mixed Method Study. Advance Social Science Archive Journal, 4(1), 2131–2143. https://doi.org/10.55966/assaj.2025.4.1.0115

Sarıtepeci, M., & Durak, H. Y. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and reflective thinking skills: An experimental study. Education and Information Technologies, 29(18), 25175–25209. https://doi.org/10.1007/s10639-024-12829-2

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. Journal of Educational Psychology, 106(2), 331–347. https://doi.org/10.1037/a0034752

Sun, L., & Zhou, L. (2024). Does Generative Artificial Intelligence Improve the Academic Achievement of College Students? A Meta-Analysis. Journal of Educational Computing Research, 62(7), 1676–1713. https://doi.org/10.1177/07356331241277937

Ünal, E., Kaya, M., Uzun, A. M., & Erdem, C. (2026). A Second-Order Meta-Analysis on the Effects of Artificial Intelligence Applications on Student Outcomes. Journal of Educational Computing Research. https://doi.org/10.1177/07356331261424767

Waziana, W., Andewi, W., Wibisono, D., Hastomo, T., & Muslihudin, M. (2025). Exploring ChatGPT's Impact on Critical, Creative, and Reflective Thinking Skills: A Mixed-Methods Study in an Indonesian EFL Classroom. SHILAP Revista de Lepidopterología. https://doi.org/10.22108/are.2025.145896.2564

Xu, Z., Wijekumar, K., Ramírez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. British Journal of Educational Technology, 50(6), 3119–3137. https://doi.org/10.1111/bjet.12758

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8(10), 1839–1850. https://doi.org/10.1038/s41562-024-02004-5

Yetişensoy, O., & Karaduman, H. (2024). The effect of AI-powered chatbots in social studies education. Education and Information Technologies, 29(13), 17035–17069. https://doi.org/10.1007/s10639-024-12485-6

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learning Environments, 11(1). https://doi.org/10.1186/s40561-024-00316-7