

RESEARCH

Free and Open Access

Exploring the effects of generative AI (GenAI) on learning outcomes: A scoping review

Ngo Cong-Lem ^{1*} and Hongming Fan²

*Correspondence:
lemnc@dlu.edu.vn
Faculty of Foreign Languages,
Dalat University, 01 Phu Dong
Thien Vuong St., Lam Vien – Da
Lat, Lam Dong, Vietnam
Full list of author information is
available at the end of the article

Abstract

Generative artificial intelligence (GenAI) is being rapidly incorporated into educational settings; however, empirical evidence regarding its effects on learning remains fragmented and conceptually uneven. This review synthesises evidence from 31 empirical studies, identified through a structured search of Scopus and Web of Science, to examine how GenAI tools shape learning outcomes across educational levels and disciplinary contexts. Thematic analysis of the included studies identified four outcome domains through which GenAI's educational impact has been examined: achievements, skills, dispositions, and learning processes. Among these, achievement outcomes were the most frequently measured (71.0%), followed by dispositions and skills (58.1% each), whereas learning processes were comparatively under-examined (32.3%). Although less frequently examined, learners' engagement patterns offered important insight into how GenAI supported or constrained learning across different contexts. Specifically, GenAI appeared to be educationally productive when learners engaged in cycles of explanation, feedback uptake, monitoring, and iterative refinement, but less productive when AI outputs were used to bypass cognitive effort, defer judgement, or produce superficially improved work without deeper understanding. Based on these findings, the review proposes a cognitive–affective–behavioural model that conceptualises GenAI-supported learning as unfolding through productive scaffolding and risk-oriented offloading pathways. These findings underscore the need for AI literacy development, reflective learning tasks, process-oriented assessment, and opportunities for social interaction to ensure that GenAI supports meaningful learning rather than superficial task completion.

Keywords: generative AI, ChatGPT, intervention, learner, achievement, review



© The Author(s). 2026 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

The rapid evolution of artificial intelligence (AI) has begun to transform educational practices, with the emergence of generative AI (GenAI) tools such as ChatGPT introducing new forms of language-based, interactive support for learners. These systems can generate explanations, examples, feedback, and content on demand, enabling applications such as real-time tutoring, writing support, and task scaffolding across disciplines (e.g., Alsofyani & Barzanji, 2024; Celik et al., 2025; Cong-Lem, 2026; Shaikh et al., 2023). In recent classroom-based research, GenAI tools have been integrated into learning activities ranging from academic writing and feedback-supported revision to technical and problem-solving tasks in computing-related fields (e.g., Bašić et al., 2023; Sun et al., 2024). Because GenAI can provide immediate guidance and alternative representations of knowledge, it is often positioned as an instructional aid that may support learners' understanding and performance, particularly when embedded within structured learning activities rather than used as a stand-alone solution (Kosar et al., 2024; Sun et al., 2024). At the same time, emerging evidence indicates that the educational impact of GenAI is not uniformly positive: several studies report limited or null effects on performance relative to comparison conditions, suggesting that learning benefits may depend on how tools are used and what forms of student engagement are required (Bašić et al., 2023; Kosar et al., 2024; Naamati-Schneider & Alt, 2024).

Despite the growing integration of GenAI tools in educational contexts, empirical findings on their influence on student learning outcomes remain fragmented and context-dependent. Studies vary substantially in terms of disciplines, student populations, learning tasks, intervention structures, and outcome measures, making it difficult to derive a coherent picture of GenAI's educational value without synthesis (Cong-Lem et al., 2025; Einarsson et al., 2024). Importantly, patterns observed in recent interventions suggest that outcomes may be shaped by factors such as task design, learner proficiency, and instructional guidance. For example, behavioural changes in engagement or learning processes do not always translate into measurable differences in performance outcomes, indicating complex mechanisms of impact (Sun et al., 2024; Wu & Ho, 2025). Similarly, some GenAI-supported designs show improvements in specific domains (e.g., evaluative skills), while producing limited gains in others, raising questions about what forms of learning are most responsive to GenAI mediation (Naamati-Schneider & Alt, 2024). These mixed findings highlight the need for broader evidence mapping that captures both reported benefits and boundary conditions.

To address these gaps, this scoping review explores the impact of generative AI tools, such as ChatGPT, on student learning outcomes, focusing on academic performance, learning experiences, and skill development. It aims to map the breadth of evidence on how these technologies are associated with learning-related outcomes and to identify contextual and pedagogical factors that may shape their effects. The central research question guiding this review is: *What is the impact of generative AI tools on students' learning outcomes?* By synthesising empirical evidence across diverse educational settings and applications, this review seeks to clarify the learning domains most frequently examined, the extent to which GenAI effects are consistent, and the contextual factors associated with stronger or weaker outcomes.

Given the transformative possibilities of GenAI, this review offers an evidence-informed foundation for educators, policymakers, and developers seeking to incorporate these tools responsibly. By highlighting both reported learning benefits and documented limitations, including cases where GenAI produced no added performance advantage, the review supports more realistic expectations about GenAI's role in learning. Ultimately, mapping the current evidence landscape can guide more equitable and pedagogically grounded integration strategies, ensuring that GenAI adoption supports meaningful learning rather than superficial task completion or uncritical dependence on generated outputs.

Method

Search strategy

A comprehensive search was conducted using two major academic databases: Scopus and Web of Science. The search targeted empirical studies published between 2022 and June 5, 2024, with the intention to capture the most relevant research following the launch of ChatGPT in November 2022. To ensure a focused and efficient search, specific keywords were selected, such as generative AI, artificial intelligence, ChatGPT, student, learner, performance, outcomes, and achievement. Boolean operators (e.g., AND, OR) were used to combine these key terms in different configurations, allowing for a more precise search that covered a wide range of relevant studies. For example, the search was structured to include studies on generative AI tools and their impact on educational outcomes, including performance, achievements, and student experiences. The search strategy was developed iteratively and refined through pilot searches to ensure sensitivity and relevance.

Table 1

Inclusion and exclusion criteria

Criterion	Inclusion Criteria	Exclusion Criteria
Study type	Empirical studies reporting original data	Reviews, conceptual / theoretical papers, opinion pieces, editorials, theses, book chapters, conference proceedings, and other grey literature
Focus of study	Studies examining the use of GenAI tools in educational settings	Studies not involving generative AI tools or not conducted in educational contexts
Intervention	Studies involving a GenAI-related intervention, instructional activity, or structured learner engagement with a GenAI tool.	Studies without a GenAI-related intervention, instructional activity, or structured learner-facing use of a GenAI tool
Outcomes	Studies measuring learning-related outcomes using identifiable instruments (e.g., achievement, knowledge, skills, learning experience, perceptions related to learning processes)	Studies not reporting learning outcomes or not measuring outcomes using identifiable instruments
Participants	Learners or students engaged in educational contexts	Studies not involving learners (e.g., system development papers without learner data)
Language	Publications written in English	Publications written in languages other than English
Publication year	Studies published from 2022 onwards, following the public release of ChatGPT	Studies published before 2022

Study screening, extraction and analysis

The retrieved records were uploaded to Covidence, a web-based platform for managing systematic and scoping reviews. Screening was conducted in two stages: title-and-abstract screening followed by full-text screening (see Table 1 for the screening criteria). Both stages were undertaken by two members of the research team (CN and HF), and disagreements were resolved through discussion until consensus was reached. Studies were included if they reported original empirical data on learner-facing GenAI use in educational contexts and examined at least one identifiable learning-related outcome, including achievement, knowledge, skills, learning processes, dispositions, or learning experiences. Studies were excluded if they were conceptual papers, reviews, editorials, theses, book chapters, conference proceedings, or did not involve learners, GenAI tools, educational contexts, or identifiable outcome measures. Consistent with the purpose of scoping reviews, formal methodological quality appraisal was not used as an exclusion criterion. Instead, the synthesis focused on mapping the characteristics of the

evidence base, identifying recurring outcome domains, and describing patterns across study contexts, designs, and reported findings.

Findings

Characteristics of the included studies

Following the two-stage screening process, comprising title-and-abstract screening and full-text screening, 31 studies were included in the final review (see Figure 1). Collectively, the evidence base reflects increasing empirical attention to GenAI-enabled learning support, but it also reveals clear concentrations in particular geographical regions, educational sectors, and disciplinary domains, which should be considered when interpreting the scope and transferability of the findings.

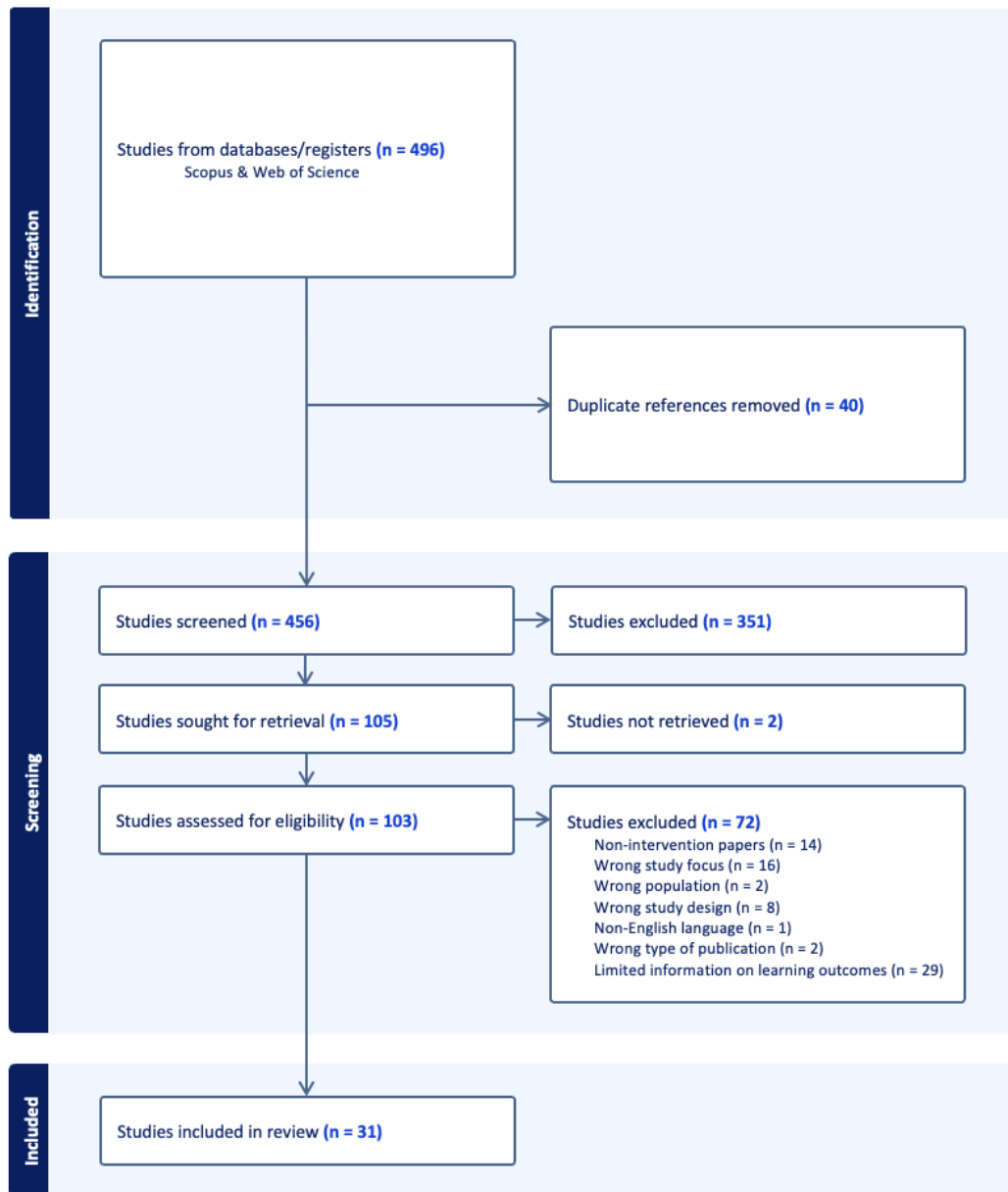
In terms of geographical distribution, the included studies were situated across a diverse but uneven range of national and regional contexts. Based on the available study-setting information, the evidence base represented 17 named countries or territories. China accounted for the largest share of studies ($n = 6$; 19.4%), followed by Taiwan ($n = 5$; 16.1%) and the United States ($n = 4$; 12.9%). Smaller clusters were identified in Saudi Arabia and Malaysia ($n = 2$ each; 6.5% each). The remaining studies were conducted in single-study contexts, including the United Arab Emirates, Croatia, Cyprus, Slovenia, Mexico, Israel, Norway, the Czech Republic, Turkey, and South Korea. One further study involved participants from both Finland and New Zealand ($n = 1$; 3.2%), while one study did not provide sufficiently clear information to determine the country setting ($n = 1$; 3.2%). This distribution suggests that the emerging GenAI-in-education evidence base remains geographically imbalanced, with research activity disproportionately centred in East Asian higher education contexts and selected Western university settings.

Consistent with this geographical pattern, the included studies primarily investigated GenAI use in tertiary education, indicating that current evidence is largely shaped by university-level instructional conditions. Twenty-nine studies (93.5%) were conducted in higher education contexts, encompassing undergraduate, postgraduate, and discipline-specific cohorts (e.g., nursing students, law students, and English language majors). By contrast, only two studies (6.5%) focused on secondary/K–12 learners, Chen and Chang (2024) in seventh-grade science and Alneyadi and Wardat (2024) in Grade 12 quantum theory, highlighting a clear gap in school-based research on GenAI-supported learning. Taken together, these features suggest that most current knowledge about GenAI's learning impact is derived from higher

education environments, where learners often have greater autonomy, digital access, and flexibility in tool use.

Figure 1

PRISMA flowchart of study screening and selection



Parallel to this concentration in higher education, the disciplinary distribution of studies indicates that GenAI applications are being explored widely but not equally across fields. Studies in Arts and Humanities accounted for the largest share (n = 10; 32.3%), largely driven by research in English language learning (EFL), academic writing, literature appreciation, design creativity, and music education. STEM disciplines represented nine studies (29.0%), including programming and computational thinking, quantum theory, game-based science learning, chemistry, and engineering-

related knowledge development. Social sciences ($n = 4$; 12.9%) included business, international relations, and legal education contexts, while Medical/Health sciences ($n = 4$; 12.9%) spanned nursing, pharmacy, dentistry, and healthcare management. A further four studies (12.9%) were situated in education-related contexts, encompassing teacher preparation, flipped learning, information literacy, and instructional design. This diversity reflects the cross-disciplinary appeal of GenAI tools; however, it also suggests that language learning and writing remain particularly prominent domains in which GenAI affordances, especially feedback, revision support, and ideation, are perceived as pedagogically relevant.

Methodologically, the evidence base was characterised by a strong orientation toward designs that can capture measurable effects of GenAI integration within short instructional timeframes. Using consistent design coding for synthesis, the studies comprised quantitative, qualitative, and mixed methods approaches, with quantitative designs being particularly common in performance-oriented research. Many quantitative studies adopted experimental or quasi-experimental structures involving treatment and comparison conditions and/or pre–post testing, which enabled examination of outcomes such as achievement, learning effectiveness, writing performance, or domain knowledge acquisition (e.g., Chen & Chang, 2024; Kosar et al., 2024; Tsai et al., 2024). Alongside this, qualitative studies typically focused on learner perceptions, experiences, and engagement dynamics, often through interviews, reflective accounts, or qualitative analyses of learning processes (e.g., Aggrawal & Magana, 2024; AlGhamdi, 2024). Mixed methods designs combined these approaches, integrating performance measures with qualitative evidence about student experiences, tool-use patterns, and engagement mechanisms (e.g., Kavarella et al., 2024; Naamati-Schneider & Alt, 2024).

With regard to the type of GenAI tools, ChatGPT was the most frequently examined tool, often in versions such as GPT-3.5 or ChatGPT-4, reflecting its accessibility and rapid adoption in education. Several studies expanded beyond ChatGPT to include other GenAI-enabled systems such as ERNIE Bot, AI companions, domain-specific chatbots, large-language-model-supported tutoring designs, and text-to-image generators (e.g., Chandrasekera et al., 2024; Hu et al., 2024; Liu et al., 2024). The learning functions addressed were similarly varied, spanning real-time feedback, guided revision, ideation and planning support, self-regulation scaffolding, conceptual clarification, creativity enhancement, and problem-solving assistance. Importantly, the reviewed interventions rarely framed GenAI as an autonomous instructional substitute; rather, it was typically positioned as a scaffold embedded in tasks,

activities, or feedback loops, underscoring the centrality of pedagogical design in shaping educational outcomes.

Finally, the direction of GenAI effect on learning outcomes across the included studies was predominantly positive, though not uniformly so. Overall, 27 studies (87.1%) reported positive impacts in at least one learning domain, such as improvements in academic performance, domain knowledge, writing quality, cognitive skills (e.g., critical thinking), process outcomes (e.g., debugging or revision behaviours), and affective-motivational factors (e.g., self-efficacy, confidence, reduced perceived burden). At the same time, two studies (6.5%) reported mixed effects, where benefits appeared in certain dimensions but not others, for example, improvements in structural coherence and language quality without consistent gains across all feedback dimensions (AlGhamdi, 2024), and gains in evaluative skills without broader improvements in digital literacy-related domains (Naamati-Schneider & Alt, 2024). Moreover, two studies (6.5%) reported negative or null outcomes, indicating no significant performance benefit attributable to GenAI use in comparison conditions (Bašić et al., 2023; Kosar et al., 2024).

Impact of GenAI on learning outcomes

Across the 31 included empirical studies, the educational impact of GenAI was examined through four outcome domains: (1) learning dispositions, (2) skills, (3) processes, and (4) achievements. Based on explicit measurement, outcomes assessed through an identifiable instrument, scale, rubric, or scored test, achievement was the most frequently operationalised domain (71.0%; 22/31 studies), followed equally by dispositions and skills (58.1% each; 18/31), while learning processes were the least commonly formally measured (32.3%; 10/31). Overall, the evidence suggests that GenAI-related benefits are documented across achievement, dispositional, and skill outcomes, while process-level evidence remains comparatively sparse and methodologically varied. The findings generally converge on a key implication: GenAI is not inherently beneficial; its educational value depends on whether learning tasks position AI output as material for critical evaluation, revision, and reflective judgement, rather than as a shortcut that reduces cognitive effort.

Achievement outcomes: Academic performance, assessment results, and contextual variability in learning gains

The most frequently reported learning outcome is students' learning achievement, which mostly involves test scores, post-test gains, course grades, and rated product quality. The evidence on achievement

outcomes is largely positive; however, a few studies reported null or mixed results. These findings suggest that improvements in learners' performance when using GenAI tools are sensitive to factors such as task demands, disciplinary context, and the design of instructional activities.

Several controlled studies reported significant achievement gains. In medical education, Hsu (2024) found that nursing students using AI-supported terminology learning tools (ChatGPT and Termbot) achieved significantly higher post-test outcomes than those using textbooks. In K–12 science learning, Chen and Chang (2024) reported improved science learning outcomes in AI-assisted game-based conditions relative to controls, particularly when scaffolding examples were included. In grade 12 quantum theory learning contexts, Alneyadi and Wardat (2024) found that ChatGPT-supported instruction improved post-test achievement across knowing, applying, and reasoning sub-skills, with evidence of a substantial effect size. Studies in professional or discipline-specific training, such as civil engineering and music education, also reported achievement gains in targeted knowledge or task performance (e.g., Uddin et al., 2023, 2024; Zhou & Kim, 2024).

However, achievement outcomes were not consistently positive. Bašić et al. (2023) reported no improvement in essay grades or writing efficiency in ChatGPT-assisted conditions, alongside modest concerns regarding text authenticity. Likewise, Kosar et al. (2024) found that novice programmers' performance and assessment scores were not significantly influenced by ChatGPT usage, suggesting that access alone does not guarantee improved learning products. These null findings indicate that achievement benefits are not automatic and may depend on whether GenAI support aligns with the cognitive demands of assessment tasks, particularly when assessments require independent reasoning, synthesis, or transfer rather than assisted production.

Dispositional outcomes: Motivation, self-efficacy, and affective engagement in GenAI-supported learning

The second most frequently assessed outcomes were dispositional, learners' motivational and affective orientations toward learning, including confidence, self-efficacy, motivation, flow, satisfaction, and perceived task value. The dominance of this category suggests that researchers commonly conceptualise GenAI not only as a cognitive tool, but also as a psychological facilitator that reshapes learners' willingness to engage with learning tasks.

A recurring finding is that GenAI-supported environments can increase learners' confidence and perceived competence, particularly when the tool functions as a responsive support system during problem-solving or performance tasks. For example, Aggrawal and Magana (2024) reported that most participants increased their confidence in addressing teamwork conflict after simulated conflict-resolution practice with generative AI, supported by structured reflection. In academic skill contexts, Gao et al. (2024) found positive associations between GenAI integration and motivational variables such as learning desire, self-efficacy, and future beliefs, linking these to improved performance outcomes. Similarly, Shi et al. (2024) showed that GenAI-supported situational interactive teaching enhanced learners' flow experiences alongside improvements across learning outcomes.

Beyond increased confidence, GenAI was often described as lowering psychological barriers to participation by making learning feel more manageable and supported. Chen and Chang (2024), for instance, showed that AI-assisted game-based learning enhanced intrinsic motivation and reduced reported mental burden compared with game-only conditions. In writing-related contexts, Song and Song (2023) found that GenAI-assisted instruction improved writing motivation, although qualitative accounts simultaneously indicated concerns about over-reliance and contextual accuracy. These patterns suggest that dispositional benefits are commonly produced when learners experience GenAI as an immediate, non-judgemental, and personalised source of support, which can increase persistence and engagement. At the same time, several studies indicate that dispositional gains may coexist with uncertainty around credibility, consistency, and ethical boundaries (e.g., AlGhamdi, 2024; Hamid et al., 2023), implying that motivational uplift does not automatically translate into deeper learning without pedagogical scaffolding.

Skill development: Transferable competencies, disciplinary reasoning, and cognitive displacement

A substantial proportion of studies assessed GenAI impact on learning skills, defined as transferable competencies that extend beyond immediate test performance. In this dataset, skills frequently included critical thinking, creativity, writing competence, disciplinary reasoning, and AI literacy-related capabilities (e.g., Huang et al., 2023; Li, 2023; Song & Song, 2023; Urban et al., 2024; Yilmaz & Yilmaz, 2023). This emphasis indicates that researchers often treat GenAI as a tool that may shift what learners become capable of doing, particularly in composing, reasoning, evaluating, and revising.

Several studies suggest that GenAI can strengthen skill development when it is embedded as a partner for iterative improvement rather than a substitute for learner production. In teacher education, Li et al. (2024) found that student teachers paired with ChatGPT demonstrated stronger critical thinking systematicity and task completion efficiency than those paired with in-service teachers, although the human-expert group slightly outperformed in final proposal quality. Yilmaz and Yilmaz (2023) explored the impact of ChatGPT on students' computational thinking skills in an experimental study and found that those in the treatment group experienced significantly higher levels of skill, as well as improved self-efficacy and motivation. This pattern is instructive: GenAI support may enhance procedural rigour and structured reasoning but does not necessarily guarantee superior holistic quality when expert human judgement and contextual nuance are central to the task.

In higher education contexts, Michalon and Camacho-Zuñiga (2023) reported that structured use of ChatGPT in a Future Studies course strengthened transversal competencies such as communication, critical thinking, and logical-methodical reasoning. However, the literature also includes skill-related cautionary evidence: Naamati-Schneider and Alt (2024) found that problem-based learning with ChatGPT (AI-PBL) conditions led students to utilise fewer digital literacy skills than non-AI PBL groups, suggesting that GenAI can sometimes displace rather than cultivate foundational competencies. Alhammad (2024) also reported improvements in the ability to interpret themes, symbols and character developments in literary texts among 28 Saudi Arabian EFL learners following three weeks of engagement with ChatGPT. In short, the skill-focused findings imply that GenAI may support competency growth most effectively when learners are required to evaluate, select, justify, and refine outputs, rather than accept them as final products.

Learning processes: Engagement patterns, revision cycles, and self-regulated learning

Although the least commonly operationalised of the four domains, learning processes, encompassing observable engagement patterns, interaction behaviours, self-regulation practices, and strategic cycles of work, offer some of the most conceptually illuminating evidence in this review, capturing not merely what learners achieve but how GenAI shapes the mechanisms through which learning unfolds.

A key insight is that GenAI can change learning processes in ways that differentiate higher- and lower-performing learners. Liu et al. (2024), using learning analytics, identified contrasting collaboration patterns between high- and low-performance groups, with high-performing students following a

structured sequence of ‘generating, monitoring, applying and evaluating’, accompanied by communication strategies that emphasised cognitive agency. This suggests that GenAI becomes educationally productive when learners maintain an active monitoring stance toward output quality. Similarly, Nguyen et al. (2024) demonstrated that doctoral writers who engaged in iterative, interactive collaboration patterns with a GAI writing tool achieved better writing performance than those using linear, minimally interactive approaches, highlighting the pedagogical importance of interaction design.

Other studies show GenAI influencing micro-processes of task completion. Sun et al. (2024) found that GenAI-supported programming modes produced more frequent debugging behaviours and feedback-reading cycles, although this did not necessarily yield significantly higher programming performance. Chen and Chang (2024) reported learning behavioural analytics indicating that GenAI-assisted game-based learning promoted more effective learning behaviours and reduced cognitive load, particularly when support included examples. Collectively, process evidence suggests that GenAI often functions as an external support that restructures learners’ task pathways, sometimes increasing strategic iteration but also raising the risk of superficial engagement if learners shift from generative effort to uncritical adoption.

Discussion

Overview of the findings

The reviewed evidence indicates that GenAI's educational influence is most frequently measured in achievement outcomes, though effects in this domain are also the most context-dependent, with both significant gains and null findings reported across the evidence base. Substantial measurement attention is also directed toward dispositional and skill outcomes, where findings more consistently suggest that GenAI functions as a motivational and confidence-supporting resource, and as a catalyst for iterative refinement, structured reasoning, and competency development when appropriately scaffolded. Learning processes were the least commonly operationalised domain, yet process-level evidence provides some of the strongest explanatory leverage for understanding why outcomes vary: how learners engage with GenAI, whether through active monitoring, iterative revision, and critical evaluation, or through cognitive offloading and uncritical adoption, appears to mediate whether achievement and dispositional gains materialise. Overall, the cross-study pattern suggests that the strongest benefits emerge when GenAI is embedded into pedagogical designs that sustain learners' agency and critical judgement,

positioning AI output as an artefact for evaluation, revision, and justification, rather than as a mechanism for bypassing cognitive work.

An important highlight in the findings of the review is concerned with the boundary between productive GenAI scaffolding and counterproductive over-reliance. When GenAI is used as a tool for explanation, feedback, and idea generation within guided tasks, it may strengthen learning through supported reasoning and iterative revision. However, when learners treat AI output as authoritative answers and minimise evaluative processing, GenAI can facilitate cognitive delegation rather than cognitive development. This distinction helps explain why some studies report limited or null effects relative to traditional instruction: the presence of GenAI alone does not guarantee learning gains; rather, gains appear contingent on pedagogical design features such as reflective prompting, requirement to justify choices, and structured opportunities to critique AI outputs.

Equity considerations remain under-examined in the current evidence base. Few included studies (e.g., Alneyadi & Wardat, 2024; Song & Song, 2023; Tsai et al., 2024) reported subgroup analyses by prior achievement, socioeconomic access to devices/internet, language proficiency, or learners' digital/AI literacy. Nevertheless, these factors are theoretically likely to shape who benefits most from GenAI-supported learning. For example, students with stronger prior knowledge or higher academic self-regulation may leverage GenAI more effectively as a strategic resource, whereas students with limited AI literacy may be more vulnerable to over-reliance, misinformation, or superficial learning (Liu et al., 2024; Nguyen et al., 2024). Therefore, future GenAI research should explicitly incorporate equity-sensitive designs, including subgroup analyses and measurement of access, literacy, and baseline proficiency.

A cognitive–affective–behavioural lens for interpreting GenAI's impact on learning outcomes

Drawing on the review findings, a heuristic cognitive–affective–behavioural (CAB) lens is proposed to explain how GenAI shapes learning outcomes through interacting mechanisms, rather than producing isolated effects (Figure 2). Across the included studies, GenAI was frequently embedded as a mediational resource within learning activities, providing explanation, feedback, suggestions, and task support. Yet the empirical evidence also indicates that the same GenAI tool can generate divergent outcomes depending on how learners engage with it. Figure 2 therefore conceptualises GenAI-supported learning

as unfolding through two interconnected pathways: a productive scaffolding cycle that strengthens learning through active processing and strategic engagement, and a risk cycle in which affective relief and efficiency shift into cognitive offloading, over-reliance, and weakened generative processing. In other words, CAB offers an explanatory account of why GenAI sometimes supports learning development and, in other cases, produces mixed, null, or ethically problematic outcomes.

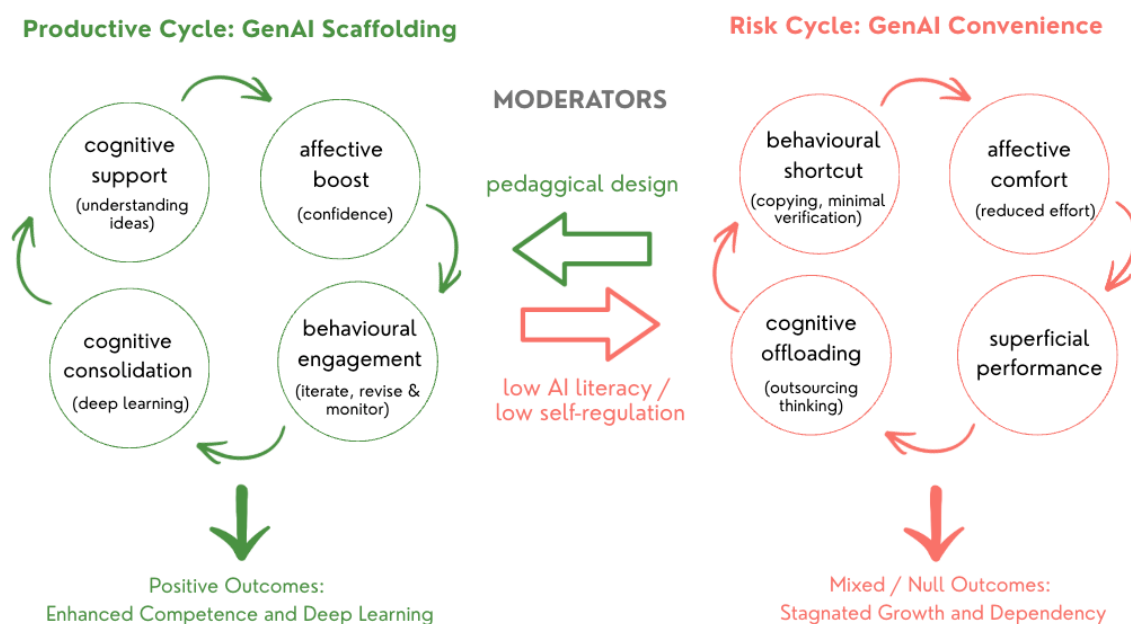
From the cognitive perspective, GenAI commonly operates as a scaffold that can accelerate comprehension and support information processing, particularly through responsive explanation, examples, and alternative reasoning routes. Studies in STEM and professional learning contexts frequently align this scaffold with improved knowledge acquisition or application (e.g., Alneyadi & Wardat, 2024; Uddin et al., 2023, 2024; Hsu, 2024). However, Figure 2 emphasises that the cognitive affordances of GenAI become educationally meaningful only when students actively process AI output, through verifying, integrating, and revising, rather than adopting responses as final answers (e.g., Rojas, 2024). This nuance helps clarify why achievement gains are inconsistent in parts of the evidence base: when tasks demand deeper synthesis, original production, or domain reasoning, access to GenAI alone may not yield measurable cognitive improvement (Bašić et al., 2023; Kosar et al., 2024). In Figure 2, this is represented by the cognitive node feeding into either deeper processing (productive pathway) or cognitive offloading (risk pathway).

These cognitive mechanisms intersect with the affective dimension in ways that are central to the model. A large proportion of studies reported increases in confidence, self-efficacy, and motivation, especially when GenAI functioned as an on-demand tutor providing immediate, non-judgemental support (Aggrawal & Magana, 2024; Chen & Chang, 2024; Gao et al., 2024). In CAB terms, affective shifts do not merely accompany learning, they shape learners' willingness to persist in cognitively demanding tasks. When GenAI reduces uncertainty and makes tasks feel manageable, learners may experience stronger expectancy of success and greater readiness to attempt difficult work. This is consistent with studies reporting reduced mental burden or cognitive load in GenAI-assisted conditions (Chen & Chang, 2024; Chandrasekera et al., 2024). In Figure 2, this affective uplift becomes a hinge mechanism: it can amplify strategic engagement (productive cycle), but it can also create vulnerability if students misinterpret AI-supported success as personal mastery. This risk is reflected in studies raising concerns about inflated indicators of competence or superficial improvement (Tsai et al., 2024), and in accounts of over-reliance (Nguyen et al., 2024).

The behavioural / learning-process dimension provides particularly strong explanatory leverage for understanding these divergent effects. A substantial proportion of studies emphasised that GenAI changes how learners work: prompting cycles of revision, monitoring, troubleshooting, feedback uptake, and reflective decision-making. This evidence suggests that outcomes are often determined not by the presence of GenAI but by the interaction patterns it induces. In writing contexts, for example, stronger learning-oriented outcomes tended to occur when learners used GenAI in a dialogic, iterative manner, treating AI suggestions as material for critique and refinement rather than text to accept uncritically (Nguyen et al., 2024). In instructional design contexts, Liu et al. (2024) showed that higher-performing students engaged in systematic cycles of ‘generating, monitoring, applying and evaluating’, illustrating behaviours consistent with self-regulation and epistemic agency. In Figure 2, these processes represent the behavioural engine of the productive scaffolding pathway: iterative engagement strengthens cognitive learning over time.

Figure 2

Two-pathway model of GenAI use in education: Scaffolding vs. risk cycles



Crucially, Figure 2 positions CAB as a two-pathway explanatory model that accounts for contradictory findings across studies. In the productive scaffolding cycle, cognitive scaffolding reduces uncertainty and mental burden, which strengthens affective engagement (confidence and motivation). This increased engagement supports iterative behaviours (revision, monitoring, troubleshooting), which then feed back into cognitive development through repeated evaluation and refinement. Over time, this cycle increases the probability that GenAI supports genuine learning rather than mere performance support. However,

the risk cycle is also plausible and visible in the evidence base. When affective relief turns into dependence, learners may reduce generative effort, defer judgement to AI output, and shift from strategic engagement toward shortcut completion. In such cases, behavioural patterns become less cognitively productive even if products appear improved. This is consistent with concerns about AI-assisted writing producing higher scores without clear evidence of competence development (Tsai et al., 2024), and with findings suggesting no learning advantage or potential confusion among inexperienced users (Bašić et al., 2023). Thus, CAB does not assume GenAI is inherently beneficial or harmful; rather, it explains how GenAI becomes educationally productive only when cognitive, affective, and behavioural mechanisms align toward active learning, and why misalignment can generate ethically and pedagogically problematic outcomes.

Educational implications

From a pedagogical perspective, the findings suggest that the educational value of GenAI depends less on tool availability than on *how learning tasks position AI outputs within students' thinking and decision-making*. Across the included studies, positive outcomes were more evident when GenAI was used as a scaffolding and feedback resource to support knowledge clarification, revision, and exploration, rather than as a source of ready-made answers. This suggests that AI-mediated learning tasks should require students to interpret, justify, and refine AI-generated suggestions, for example by comparing AI explanations with course concepts, revising outputs with explicit rationales, or documenting what was accepted, rejected, and why. Such activities create opportunities for learners to make their thinking visible through explanation, justification, and dialogue, whether in written reflections, peer discussion, or teacher-student interaction. This is consistent with Vygotsky's (1987) view that social interaction, in various direct and indirect forms, plays a foundational role in the development of higher psychological functions. In addition, the pedagogical framework proposed in this study can therefore support educators in distinguishing between productive scaffolding and cognitive offloading pathways in GenAI-assisted learning. It should inform lesson and task design by foregrounding active processing, iterative engagement, and accountable use of AI, while reducing the likelihood of over-reliance and superficial task completion.

Ethical and assessment-related concerns in several studies further imply that GenAI integration should be accompanied by transparent expectations regarding authorship, acceptable assistance, and academic integrity (e.g., Rojas, 2024). In particular, the evidence suggests that GenAI can produce improvements

in written products that may not fully reflect underlying competence, potentially creating fairness issues in assessment (e.g., Tsai et al., 2024). This suggests the need for a process-oriented approach to assessment, where learners' thinking traces, including revision rationales, decision-making reflections, and process documentation, are valued alongside final outputs.

In addition, the review indicates that many challenges in GenAI-assisted learning are pedagogical, ethical, and evaluative rather than purely technical. Teacher development therefore needs to go beyond prompting skills to include (i) designing tasks that sustain critical thinking and active learning, (ii) supporting students' AI literacy and verification habits, and (iii) aligning assessment with transparency and integrity expectations. Professional learning opportunities can help teachers develop context-sensitive strategies for integrating GenAI as a mediational tool, one that supports reasoning and revision, while managing risks such as over-reliance, superficial learning, and credibility errors.

At the institutional level, the findings point to the need for explicit institutional guidance on ethical GenAI use, particularly in relation to disclosure, attribution, verification, and the boundary between support and substitution. Rather than prescribing uniform rules across all learning contexts, such guidance should establish shared principles while allowing departments and instructors to adapt implementation to disciplinary norms, assessment purposes, and learner needs. In response to concerns about hallucination, bias, privacy, and assessment integrity, institutions should update academic integrity policies, promote responsible verification practices, and ensure equitable access to AI tools. These efforts should be supported by adequate infrastructure and AI literacy training for both educators and students, so that GenAI adoption does not exacerbate inequalities associated with prior knowledge, self-regulation, or AI literacy (Cong-Lem et al., 2025).

Limitations and future research directions

This scoping review has several limitations that should be considered when interpreting its findings. First, the inclusion criteria restricted the review to empirical studies published in English from 2022 onward, with a focus on GenAI tools such as ChatGPT. While this approach ensures that the review captures recent developments following the launch of prominent GenAI models, it may have excluded relevant studies in other languages or older foundational studies on related AI tools. Expanding the language and date range in future reviews could provide a more comprehensive view of GenAI's evolution in educational contexts and its impact across diverse linguistic and cultural settings.

Additionally, the review's reliance on studies indexed solely in Scopus and Web of Science may limit the breadth of perspectives included. Although these databases are widely used for high-quality academic research, valuable insights from educational conferences, emerging journals, or regional studies may not be captured. Including a wider range of sources, such as conference proceedings or grey literature, could deepen understanding of how GenAI is applied in various educational settings and capture experimental or case-based applications that may not yet be widely published.

The current evidence base is primarily characterised by short-term interventions and immediate post-intervention outcomes, limiting conclusions about sustainability of learning gains. As such, the present findings should be interpreted as indicating short-term potential rather than durable impact. Future research should incorporate longitudinal designs (e.g., delayed post-tests, follow-up measures across semesters) to evaluate retention, transfer to new tasks, and the possibility of unintended consequences such as dependency or reduced independent problem solving over time. Consequently, recommendations for widespread adoption should be approached cautiously and remain conditional on robust longitudinal evidence. Longitudinal studies would be essential in future research to examine how sustained use of GenAI influences long-term academic performance, self-regulation, and attitudes toward AI in learning. Furthermore, investigating indirect effects, such as the influence of GenAI on teachers' instructional practices and student-instructor interactions, would provide a more holistic view of its role in educational environments.

Future research should also focus on examining the ethical and practical implications of integrating GenAI in various educational levels and contexts. For instance, studies on the use of GenAI tools in primary and secondary education remain limited, yet the developmental and ethical considerations in these settings differ substantially from those in higher education. Research exploring age-appropriate, ethically sound implementations could help identify best practices for safely introducing GenAI to younger learners. Additionally, cross-disciplinary studies could reveal how GenAI might be tailored to different fields, such as humanities versus STEM, offering insights into field-specific applications and adaptations.

Lastly, the potential for GenAI to foster or hinder essential skills like critical thinking, creativity, and collaboration warrants further exploration. While initial findings suggest GenAI can support these skills through interactive feedback and scaffolding, the risk of over-reliance on AI-generated content poses a

challenge. Future studies might investigate methods to balance AI assistance with traditional learning approaches to ensure that GenAI enhances rather than replaces human-driven inquiry and problem-solving. These research directions would contribute to a nuanced understanding of GenAI's educational potential, guiding educators and policymakers in making informed decisions about its integration in diverse learning contexts.

Conclusion

This scoping review synthesised evidence from 31 empirical studies to map how GenAI tools shape learning across educational contexts. The findings suggest a clear but conditional conclusion: GenAI can support learning, but its value is not inherent to the technology itself. Rather, its educational impact depends on the pedagogical conditions under which it is used. Achievement outcomes were the most frequently measured but also the most context-dependent, while dispositional and skill-related gains were more consistently reported, particularly when GenAI functioned as a responsive scaffold for explanation, feedback, and iterative refinement. The dual-pathway cognitive–affective–behavioural (CAB) model proposed here offers an explanatory account of this variability: the same tool can activate either a productive scaffolding cycle, in which cognitive support, affective engagement, and strategic behaviour reinforce one another, or a risk cycle characterised by over-reliance, cognitive offloading, and superficially improved outputs. This distinction carries a direct implication for practice: educational impact is shaped less by whether GenAI is present than by whether learning tasks require students to evaluate, justify, and refine AI output, preserving the epistemic agency that drives genuine learning. Realising this potential will require research designs that move beyond simple with/without comparisons toward theoretically grounded, longitudinal investigations that trace process mechanisms, account for learner differences in prior knowledge, self-regulation, and AI literacy, and extend inquiry into the educational levels and cultural contexts currently underrepresented in the evidence base. Advancing this agenda will equip educators and policymakers with the nuanced, context-sensitive evidence needed to integrate GenAI in ways that are both pedagogically purposeful and equitable in their reach.

Appendix. Summary of Included Studies in This Review

No	Study	Aim	Research Design	Participants	Intervention Detail	Learning Outcomes	Key Findings
1	Aggrawal (2024)	To determine changes in confidence, conflict resolution styles, and enacted strategies after training and LLM-simulated practice.	Descriptive study using qualitative triangulation.	46 undergraduate systems analysis students.	Conflict management training followed by simulated practice sessions using an LLM to resolve team conflicts.	Dispositions (Confidence); Skills (Conflict resolution styles and strategies).	Significant shifts in confidence levels; identification of collaborative vs. avoiding styles; enacted strategies often differed from self-reported ones.
2	AlGhamdi (2024)	To understand the nuanced impacts of ChatGPT-generated feedback on the writing skills of computing students.	Qualitative research design (blinded study).	111 first-year technical writing male computing students.	Integration of ChatGPT-generated feedback (blinded from students) for weekly writing assignments submitted via blog entries.	Dispositions (Emotional responses); Skills (Writing development); Processes (Perceived quality and usefulness).	Positive perception of feedback quality; significant progress in skill development; identification of key emotional/psychological responses to AI feedback.
3	Alhammad (2024)	To investigate the impact of ChatGPT on literary appreciation and critical thinking skills in foreign language literature.	Experimental (one-group pre-post-test design).	28 sophomore female EFL learners, Saudi Arabia.	Interaction with ChatGPT to enhance appreciation and critical analysis of foreign language literature.	Achievements (Literature appreciation test scores); Skills (Critical thinking).	Significant improvement in literary appreciation (17.96 to 21.53, $p=0.000$); positive contribution to critical thinking skills.
4	Alneyadi (2024)	To assess the influence of chatbot applications on achievement, engagement, and sub-skills in Quantum Theory.	Mixed-methods quasi-experimental (pre-test, post-test, control group).	112 third-semester students (55 experimental, 57 control).	Customized ChatGPT-supported activities integrated into the Quantum Theory curriculum for an experimental group.	Achievements (Quantum Theory test: Knowing, Applying, Reasoning); Dispositions (Engagement, Perceived relevance).	Significant post-test improvements in all cognitive sub-skills (Knowing, Applying, Reasoning); large effect size ($\eta^2 > 0.14$).
5	Bašić (2023)	To compare traditional essay writing with ChatGPT-assisted writing among forensic science students.	Experimental control vs. intervention group design.	18 second-year master's students, Forensic Sciences.	Experimental group utilized ChatGPT version 2.1.0 assistance for essay writing; control group wrote traditionally.	Achievements (Essay scores); Processes (Duration, text authenticity).	No significant effect found for ChatGPT assistance on overall scores ($P=0.184$); students with higher GPAs in control group outperformed AI group.

6	Chandrasekera (2024)	To explore how text-to-image AI affects cognitive load and creativity metrics.	Independent samples experimental laboratory design.	40 participants (20 AI group, 20 Non-AI group).	Integration of text-to-image AI (assisted by ChatGPT conceptual statements) into a design task.	Dispositions (Cognitive Load); Skills (Creativity: Novelty, Resolution, Elaboration).	Significant differences in cognitive load and creativity; AI group showed a "residual learning effect" in a second task without AI.
7	Chen & Chang (2024)	To investigate the effectiveness of AI-assisted game-based learning on science outcomes and motivation.	Quasi-experimental (three-group: Game only, GameGPT, GameGPT_examples).	202 7th-grade science learners (Mainland China).	Integration of ChatGPT API (GPT-3.5-turbo) into a DGBL environment ("Summon of Magicrystal") providing feedback and examples.	Achievements (Science knowledge); Dispositions (Motivation, Cognitive load); Processes (Learning behavior patterns).	GameGPT_examples group significantly outperformed Game only; reduced mental burden and higher perceived competence in ChatGPT groups.
8	Gao et al. (2024)	To explore technology integration's impact on business students' motivation and performance.	Purposive sampling survey with PLS-SEM analysis.	376 Chinese international business students (Malaysia).	Integration of ChatGPT 3.5 into learning activities based on the ICAP framework (interactive, constructive, active, passive).	Achievements (Academic performance); Dispositions (Self-efficacy, desire for learning, beliefs about future).	Desire for learning and future beliefs mediated performance; epistemic curiosity moderated the impact of GenAI integration.
9	Hamid et al. (2023)	To explore ChatGPT's role in enhancing self-efficacy and preparedness in pharmacy Problem-Based Learning (PBL).	Exploratory mixed-methods study.	18 Bachelor of Pharmacy students (2nd, 3rd, 4th years).	Students integrated ChatGPT into pharmacy-specific PDPBL sessions using triggers from chemistry and clinical pharmacy.	Dispositions (Self-efficacy, Preparedness perceptions).	ChatGPT effectively enhanced students' perceptions of self-efficacy and preparedness for complex problem-solving.
10	Hsu (2024)	To compare ChatGPT and Termbot effectiveness in mastering medical terminology.	Randomized experimental design.	60 nursing students (3rd/4th year), Taiwan.	Experimental groups used ChatGPT or Termbot for terminology acquisition versus a control group using traditional textbooks.	Achievements (Medical terminology test scores); Processes (Weekly study time).	AI groups significantly outperformed control ($p < 0.05$); ChatGPT group showed highest study engagement time (10 hours/week).
11	Hu et al. (2024)	To evaluate an AI learning companion and Mandala Scaffold system for literacy and self-regulation.	Randomized three-group experimental design.	93 university freshmen, Creative Thinking course (Taiwan).	System using TABot/Master Socrates virtual librarians and Mandala Chart scaffolding for project ideation.	Dispositions (Self-efficacy, Self-regulated learning); Skills (Information retrieval, Critical thinking).	Group using AI + Mandala Scaffold significantly outperformed others in self-regulated learning and critical thinking.
12	Huang et al. (2023)	To improve programming review performance through AI-driven	Experimental design (intervention vs. traditional tutoring).	78 university students (Python programming course), Taiwan.	BERT and GPT-2 models generated personalized remedial content based on	Achievements (Programming test); Processes (Self-regulated learning strategies).	Significant improvement in performance and SRL strategies (rehearsal, critical thinking,

		personalized intervention.			individual conceptual mastery gaps.		metacognition) for AI intervention group.
13	Kavadella et al. (2024)	To evaluate ChatGPT's real-life implementation in undergraduate dental education.	Mixed-methods randomized group design.	77 dental students, European University Cyprus.	Intervention group used ChatGPT to develop learning assignments/presentations while literature group searched internet.	Achievements (MCQ examination grades); Dispositions (Student perceptions).	ChatGPT group performed significantly better (mean 7.54 vs. 6.94, $P=.045$); recognized benefits like immediate responses but cited false citations.
14	Kosar et al. (2024)	To investigate the influence of ChatGPT on object-oriented programming exams and assignments.	Controlled experimental between-subjects design.	198 computer science students (99 treatment, 99 control).	Treatment group encouraged to use ChatGPT for all lab work and practical assignments throughout the semester.	Achievements (Midterm exams, practical assignments, final grades); Dispositions (Perceived complexity).	No statistically significant difference in final grades or midterm performance between groups ($p > 0.05$).
15	Li (2023)	To test the effects of ChatGPT-based flipped learning (ChatGPT-FLGA) on project performance and perceptions.	Quasi-experiment (intact classes randomly assigned).	81 students (Modern Educational Technology course), China.	Experimental group utilized ChatGPT-FLGA guidance within Tencent QQ for a courseware design project.	Achievements (Project performance); Dispositions (Self-efficacy, Attitudes, Motivation); Skills (Creative thinking).	Significant improvement in all areas: performance, self-efficacy, motivation, and creative thinking ($p < 0.05$).
16	Li et al. (2024)	To compare "Human-Human" and "Human-Machine" collaboration in iSTEM teacher training.	Exploratory quasi-experimental randomized grouping.	23 master's student teachers, China.	Comparison of pairing students with in-service teachers (TSPL) vs. ChatGPT (CSPL) for iSTEM proposal design.	Achievements (Task performance); Dispositions (Cognitive load); Skills (Critical thinking systematicity).	ChatGPT group (CSPL) showed higher task efficiency and lower cognitive load; human teachers slightly better on final proposal design.
17	Liu et al. (2024)	To investigate the collaborative process of refining instructional design with GAI.	Qualitative learning analytics (LSA and ENA).	22 prospective chemistry educators.	Students collaborated with ERNIE Bot (GAI) to refine Technology-Enhanced Learning (TEL) chemistry instructional designs.	Achievements (TPACK-based design quality); Processes (Behavior sequences).	Significant improvement in TPACK scores; High-performing students demonstrated a structured generate-monitor-apply-evaluate cycle.
18	Michalon (2023)	To analyze prompt quality and interaction effectiveness through successive ChatGPT-assisted activities.	Longitudinal experimental activities design.	~19 university students, Mexico context.	Four successive assignments involving six distinct activities using ChatGPT for prospective methods.	Skills (Critical thinking, logical reasoning); Processes (Prompt quality, interaction productivity).	Significant improvement in communication via prompts; success required mobilizing written expression and critical thinking skills.

19	Naamati-Schneider & Alt (2024)	To measure the impact of AIPBL (ChatGPT + PBL) on digital literacy skills relative to traditional PBL.	Quasi-experimental (AIPBL vs. PBL vs. control LBL).	223 healthcare management students, Israel.	AIPBL group used ChatGPT-mediated guided learning for healthcare quality dilemma tasks.	Skills (Digital Literacy: Access, Analyze, Evaluate); Processes (Reflective journals).	AIPBL bolstered evaluation skills but did not significantly improve access/analysis skills compared to traditional PBL.
20	Nguyen et al. (2024)	To identify strategies and patterns in human-AI collaboration for academic writing.	Qualitative learning analytics (HMM and clustering).	10 doctoral students (Finland/New Zealand).	Use of ChatGPT as an AI-driven assistant for advanced academic writing tasks.	Achievements (Writing performance scores); Processes (Micro-processes/hidden states).	Identified two types: Structured Adaptivity (Type 1 - high performance) and Unstructured Streamline (Type 2 - low performance).
21	Rojas (2024)	To investigate the efficacy of ChatGPT for curation and drafting in elemental chemistry reports.	Mixed-methods multipart writing assignment study.	53 undergraduate inorganic chemistry students (Georgia Tech).	Multi-stage element report: topic generation with ChatGPT, literature verification, AI-assisted drafting/editing.	Achievements (Writing proficiency scores); Dispositions (Student perception).	High average score of 96%; students noted ease of report organization and curation with AI assistance.
22	Shaikh et al. (2023)	To assess usability and task performance of ChatGPT for formal English Language Learning.	Mixed-methods usability assessment (SUS/USE).	10 graduate students, NTNU (Norway).	Participants performed conversation, writing, grammar, and vocabulary tasks with ChatGPT.	Dispositions (Usability: Satisfaction, Ease of Use); Skills (Language competence).	ChatGPT demonstrated 'very good' performance (3.90-4.30/5 ratings); high usability scores across all student demographics.
23	Shi et al. (2024)	To investigate Situational Interactive Teaching (SIT) with GAI on legal outcomes and flow.	Controlled experiments comparing SIT+GAI vs. Traditional SIT.	128 Chinese law university students.	GAI-supported legal SIT application providing coherent context-driven interactive legal scenarios.	Achievements (Knowledge test); Dispositions (Flow experience, Affective); Skills (Cognitive/Technical).	Significant improvements in cognitive, skill, and affective domains; markedly enhanced flow experience ($p < 0.01$).
24	Song & Song (2023)	To examine the impact of AI-assisted instruction on EFL global writing performance and motivation.	Pre-post test experimental design.	50 Chinese EFL students (Bachelor's level).	12-week AI-assisted writing intervention versus traditional instruction.	Achievements (Global writing performance); Dispositions (Writing motivation); Skills (Organization, Content).	Significant difference in global performance ($M=59.12$ exp vs 45.18 control); medium to large effect sizes for content and organization.
25	Sun et al. (2024)	To compare ChatGPT-facilitated programming (CFP) with self-directed programming (SDP).	Comparative behavioral analytics and T-test.	82 college students (Python programming).	CFP group used gpt-3.5-turbo platform for real-time feedback and assistance during a radar chart coding task.	Achievements (Programming performance); Dispositions (User	CFP group had higher average scores (84.11 vs 78.36), though not statistically significant ($p=0.204$); moderate

						experience); Processes (Sequence analytics).	impact on perceived usefulness.
26	Tsai et al. (2024)	To verify the impacts of ChatGPT-assisted revisions on EFL composition and fairness.	Double-blind prospective experimental crossover design.	44 ChatGPT-naive EFL English major students, Taiwan.	Comparative analysis of student original essays versus revisions produced after ChatGPT strategy instruction.	Achievements (Composition scores: Total, Contents, Grammar, Vocabulary).	Significant total score increase (+6.36, $p=0.007$); greatest gains found in grammar and vocabulary for students with lower baseline scores.
27	Uddin et al. (2024)	To assess ChatGPT's effectiveness as an aid for introductory civil engineering topics.	Repeated-measure within-subject experimental design.	44 junior-level civil engineering students (US).	Three-phase intervention: pre-test, ChatGPT experimentation/guidance, and post-test assessment.	Achievements (Knowledge scores); Dispositions (Satisfaction, Sentiment analysis).	Statistically significant score improvements ($p < .01$); 89% positive sentiment regarding ChatGPT as a learning aid.
28	Uddin et al. (2023)	To leverage ChatGPT for improving construction hazard recognition ability.	Repeated-measures within-subject design.	Civil engineering students (North Carolina context).	75-minute session: pre-intervention baseline, ChatGPT capability demonstration, and post-intervention hazard recognition with AI.	Achievements (Hazard recognition performance scores); Dispositions (Agreement survey).	Significant increase in recognition ability (34.87% to 60.26%, $p < 0.01$); 95% of students agreed ChatGPT was helpful for safety learning.
29	Urban et al. (2024)	To examine ChatGPT's impact on actual ill-defined problem-solving performance and metacognition.	Randomized experimental laboratory design.	145 university students (Czech Republic).	Resolution of PIT (Product Improvement Task) ill-defined problems with or without ChatGPT assistance.	Achievements (Problem-solving: Quality, Elaboration, Originality); Dispositions (Self-efficacy, interest, difficulty); Processes (Mental effort).	Significant improvements in PIT quality, elaboration, and originality ($p < 0.001$); higher self-efficacy and lower invested mental effort.
30	Yilmaz & Yilmaz (2023)	To test GenAI tools on computational thinking skills, self-efficacy, and motivation.	Pretest-posttest control group experimental design.	45 undergraduate computer science students, Turkey.	Experimental group used ChatGPT for lab assignments in Java programming for 5 weeks.	Dispositions (Self-efficacy, Motivation); Skills (Computational Thinking: algorithmic thinking, problem-solving).	Experimental group scored significantly higher in all computational thinking factors, self-efficacy, and motivation ($p < .05$).
31	Zhou & Kim (2024)	To examine undergraduate benefits from ChatGPT-4 for music learning experiences.	Experimental control vs. ChatGPT-4 group design.	74 first-year music students, Chugye University (Korea).	Use of ChatGPT-4 technology to facilitate music learning and assignment tasks.	Achievements (Accuracy in test responses); Dispositions (Perceived music learning experience).	ChatGPT-4 group demonstrated significantly enhanced learning experiences and higher response accuracy.

Authors' contributions

CN led the study conceptualization, methodology, investigation, formal analysis, and preparation of the original draft, and also managed manuscript revision and project administration. HF contributed to conceptualization, formal analysis, validation, supervision, and critical review and editing of the manuscript. Both authors read and approved the final manuscript.

Authors' information

Dr. Ngo Cong-Lem earned his PhD from the Faculty of Education at Monash University, Australia. He is currently a Lecturer in the Faculty of Foreign Languages at Dalat University, Vietnam. His research focuses on critical thinking, academic integrity in AI-mediated learning and research, the role of emotions in learning and development, and evidence-informed policy and practice.

Hongming Fan, PhD, is a Lecturer in the School of Foreign Studies at Changsha University of Science and Technology. She obtained both her Master's degree and Ph.D. in TESOL from Monash University, Australia. Her teaching focus and research interests are situated within the fields of language education, learner identity, and educational technology. Her current projects focus on imagined identity and investment of English learners in China.

Funding

This research was supported by grants awarded to the first author by Dalat University and to the second author by the Hunan African Cultural Research & Exchange Center.

Availability of data and materials

No new data were generated for this review. All sources analysed are publicly available and cited in the reference list.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used ChatGPT (OpenAI) to support language refinement and improve the clarity and coherence of the manuscript. The authors reviewed, edited, and verified all AI-assisted content and take full responsibility for the integrity, originality, and accuracy of the final manuscript.

Declarations**Competing interests**

The authors declare that they have no competing interests.

Author details

Ngo Cong-Lem, Faculty of Foreign Languages, Dalat University, Vietnam
Hongming Fan, School of Foreign Studies, Changsha University of Science and Technology, China

Received: 28 February 2025 Accepted: 7 May 2026

Published online: 1 January 2027 (Online First: 26 June 2026)

References

- Aggrawal, S., & Magana, A. J. (2024). Teamwork conflict management training and conflict resolution practice via large language models. *Future Internet*, *16*(5), 1–25. <https://doi.org/10.3390/fi16050177>
- AlGhamdi, R. (2024). Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study. *Education and Information Technologies*, *29*, 18901–18926. <https://doi.org/10.1007/s10639-024-12594-2>
- Alhammad, A. I. (2024). The impact of ChatGPT in developing Saudi EFL learners' literature appreciation. *World Journal of English Language*, *14*(2), 331–338. <https://doi.org/10.5430/wjel.v14n2p331>
- Alneyadi, S., & Wardat, Y. (2024). Integrating ChatGPT in grade 12 quantum theory education: An exploratory study at Emirate school (UAE). *International Journal of Information and Education Technology*, *14*(3), 398–410. <https://doi.org/10.18178/ijiet.2024.14.3.2061>
- Alsofyani, A., & Barzanji, A. (2024). The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: A mixed-methods study. *Journal of Educational Computing Research*, *63*(2), 431–463. <https://doi.org/10.1177/07356331241307297>
- Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications*, *10*(1), 1–5. <https://doi.org/10.1057/s41599-023-02269-7>
- Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, *8*(1), 1–17. <https://doi.org/10.29140/ajal.v8n1.102418>
- Chandrasekera, T., Hosseini, Z., & Perera, U. (2024). Can artificial intelligence support creativity in early design processes? *International Journal of Architectural Computing*, 1–14. <https://doi.org/10.1177/14780771241254637>

- Chen, C. H., & Chang, C. L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Education and Information Technologies*, 29, 18621–18642. <https://doi.org/10.1007/s10639-024-12553-x>
- Cong-Lem, N. (2026). "Is this really your work?": A qualitative study of teacher-led interviews and student accountability in the age of generative AI. *Journal of Academic Ethics*, 24, 1–19. <https://doi.org/10.1007/s10805-025-09704-1>
- Cong-Lem, N., Van Bui, C. N., Nguyen, N. P. N., & Huynh, M. Q. (2025). Bridging or breaking? A systematic review of how generative AI shapes equity in foreign language education. *International Journal of Applied Linguistics*, 1–12. <https://doi.org/10.1111/ijal.70052>
- Einarsson, H., Lund, S. H., & Jónsdóttir, A. H. (2024). Application of ChatGPT for automated problem reframing across academic domains. *Computers and Education: Artificial Intelligence*, 6, 1–17. <https://doi.org/10.1016/j.caeai.2023.100194>
- Gao, Z., Cheah, J.-H., Lim, X.-J., & Luo, X. (2024). Enhancing academic performance of business students using generative AI: An interactive-constructive-active-passive (ICAP) self-determination perspective. *The International Journal of Management Education*, 22(2), 1–18. <https://doi.org/10.1016/j.ijme.2024.100958>
- Hamid, H., Zulkifli, K., Naimat, F., Che Yaacob, N. L., & Ng, K. W. (2023). Exploratory study on student perception on the use of chat AI in process-driven problem-based learning. *Currents in Pharmacy Teaching and Learning*, 15(12), 1017–1025. <https://doi.org/10.1016/j.cptl.2023.10.001>
- Hsu, M.-H. (2024). Mastering medical terminology with ChatGPT and TermBot. *Health Education Journal*, 83(4), 352–358. <https://doi.org/10.1177/00178969231197371>
- Hu, Y.-H., Hsieh, C.-L., & Salac, E. S. N. (2024). Advancing freshman skills in information literacy and self-regulation: The role of AI learning companions and Mandala Chart in academic libraries. *The Journal of Academic Librarianship*, 50(3), 1–11. <https://doi.org/10.1016/j.acalib.2024.102885>
- Huang, A. Y. Q., Chang, J. W., Yang, A. C. M., Ogata, H., Li, S. T., Yen, R. X., & Yang, S. J. H. (2023). Personalized intervention based on the early prediction of at-risk students to improve their learning performance. *Educational Technology & Society*, 26(4), 69–89. <https://www.jstor.org/stable/48747521>
- Kavadella, A., Dias Da Silva, M. A., Kaklamanos, E. G., Stamatopoulos, V., & Giannakopoulos, K. (2024). Evaluation of ChatGPT's real-life implementation in undergraduate dental education: Mixed methods study. *JMIR Medical Education*, 10, 1–14. <https://doi.org/10.2196/51344>
- Kosar, T., Ostojčić, D., Liu, Y. D., & Mernik, M. (2024). Computer science education in ChatGPT era: Experiences from an experiment in a programming course for novice programmers. *Mathematics*, 12(5), 1–22. <https://doi.org/10.3390/math12050629>
- Li, H. (2023). Effects of a ChatGPT-based flipped learning guiding approach on learners' courseware project performances and perceptions. *Australasian Journal of Educational Technology*, 39(5), 40–58. <https://doi.org/10.14742/ajet.8923>
- Li, T., Ji, Y., & Zhan, Z. (2024). Expert or machine? Comparing the effect of pairing student teacher with in-service teacher and ChatGPT on their critical thinking, learning performance, and cognitive load in an integrated-STEM course. *Asia Pacific Journal of Education*, 44(1), 45–60. <https://doi.org/10.1080/02188791.2024.2305163>
- Liu, J., Li, S., & Dong, Q. (2024). Collaboration with generative artificial intelligence: An exploratory study based on learning analytics. *Journal of Educational Computing Research*, 62(5), 1234–1266. <https://doi.org/10.1177/07356331241242441>
- Michalon, B., & Camacho-Zuñiga, C. (2023). ChatGPT, a brand-new tool to strengthen timeless competencies. *Frontiers in Education*, 8, 1–13. <https://doi.org/10.3389/feduc.2023.1251163>
- Naamati-Schneider, L., & Alt, D. (2024). Beyond digital literacy: The era of AI-powered assistants and evolving user skills. *Education and Information Technologies*, 29, 21263–21293. <https://doi.org/10.1007/s10639-024-12694-z>
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847–864. <https://doi.org/10.1080/03075079.2024.2323593>
- Rojas, A. J. (2024). An investigation into ChatGPT's application for a scientific writing assignment. *Journal of Chemical Education*, 101(5), 1959–1965. <https://doi.org/10.1021/acs.jchemed.4c00034>
- Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, 13(9), 1937–1960. <https://doi.org/10.3390/ejihpe13090140>
- Shi, S. J., Li, J. W., & Zhang, R. (2024). A study on the impact of Generative Artificial Intelligence supported Situational Interactive Teaching on students' 'flow' experience and learning effectiveness—A case study of legal education in China. *Asia Pacific Journal of Education*, 44(1), 112–138. <https://doi.org/10.1080/02188791.2024.2305161>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1–14. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, 21(1), 1–22. <https://doi.org/10.1186/s41239-024-00446-5>

- Tsai, C. Y., Lin, Y. T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, 29, 22427–22445. <https://doi.org/10.1007/s10639-024-12722-y>
- Uddin, S. M. J., Albert, A., Ovid, A., & Alsharif, A. (2023). Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability*, 15(9), 1–22. <https://doi.org/10.3390/su15097121>
- Uddin, S. M. J., Albert, A., Tamanna, M., Ovid, A., & Alsharif, A. (2024). ChatGPT as an educational resource for civil engineering students. *Computer Applications in Engineering Education*, 32, 1–18. <https://doi.org/10.1002/cae.22747>
- Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, 215, 1–15. <https://doi.org/10.1016/j.compedu.2024.105031>
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky: Problems of general psychology, including the volume thinking and speech*. Plenum Press.
- Wu, C.-H., & Ho, V. T. (2025). Critical factors for why ChatGPT enhances learning engagement and outcomes. *Education and Information Technologies*, 30(10), 14415–14446. <https://doi.org/10.1007/s10639-025-13346-6>
- Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 4, 1–14. <https://doi.org/10.1016/j.caeai.2023.100147>
- Zhou, W., & Kim, Y. (2024). Innovative music education: An empirical assessment of ChatGPT-4's impact on student learning experiences. *Education and Information Technologies*, 29, 20855–20881. <https://doi.org/10.1007/s10639-024-12705-z>

Publisher's Note

The Asia-Pacific Society for Computers in Education (APSCE) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Research and Practice in Technology Enhanced Learning (RPTEL)
is an open-access journal and free of publication fee.