Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute
InformingScience.org

JITEResearch.org

Volume 25, 2026

DUAL NATURE OF GENAI: CATALYST OR BARRIER FOR CRITICAL THINKING AND PROBLEM-SOLVING SKILLS? A SYSTEMATIC REVIEW

Tomáš Bendl	Technical University of Liberec, Liberec, Czechia	bendlt@natur.cuni.cz
Lenka Krajňáková*	Charles University, Prague, Czechia	lenka.krajnakova@natur.cuni.cz
Hilde Storrøsæter	Norwegian University of Science and Technology, Trondheim, Norway	hilde.storrosather@ntnu.no
Emil Drápela	Technical University of Liberec, Liberec, Czechia	emil.drapela@tul.cz
Martin Hanus	Charles University, Prague, Czechia	martin.hanus@natur.cuni.cz

* Corresponding author

ABSTRACT

Aim/Purpose	This systematic literature review aims to comprehensively and critically synthesize the impacts of generative artificial intelligence (GenAI) tools on students' critical thinking (CT) and problem-solving (PS) skills. In addition, it seeks to identify key research gaps and contribute to the theoretical discussion of GenAI's dual nature in education.
Background	GenAI is increasingly integrated into educational contexts, prompting ongoing debate about its influence on CT and PS – essential 21st-century skills. Despite the rapid growth in interest in GenAI-supported learning, a clear synthesis of empirical evidence on its impact on these skills remains limited.
Methodology	Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, we conducted a systematic review of empirical studies published between 2022 and 2025 that examined the use of GenAI in relation to students' CT and PS. In total, 173 records were included in the review.

Accepting Editor Torsten Reiners │ Received: January 14, 2026 │ Revised: April 30, June 13, 2026 │ Accepted: June 28, 2026.

Cite as: Bendl, T., Krajňáková, L., Storrøsæter, H., Drápela, E., & Hanus, M. (2026). Dual nature of GenAI: Catalyst or barrier for critical thinking and problem-solving skills? A systematic review. *Journal of Information Technology Education: Research, 25,* Article 30. https://doi.org/10.28945/5834

Contribution	This review advances a nuanced understanding of GenAI's dual role in education by synthesizing current empirical evidence and providing a foundation for informed pedagogical, policy, and research decisions.
Findings	Our systematic review has demonstrated that GenAI holds considerable potential to act as both a catalyst and a constraint in the development of students' CT and PS. On one hand, GenAI can scaffold reflective inquiry, foster diverse perspectives, and support structured engagement with complex tasks. On the other hand, its widespread and often uncritical use raises concerns around dependency, superficial reasoning, and the erosion of independent PS and CT capacities.
Recommendations for Practitioners	The findings highlight the importance of designing learning environments that cultivate AI literacy, understood not merely as technical proficiency, but as the capacity to interrogate, contextualize, and critically engage with AI-generated content. Embedding AI literacy across educational contexts may help protect CT and PS while enabling responsible use of GenAI.
Recommendations for Researchers	The review identified three major research gaps: a strong regional imbalance favoring high-income contexts; a predominance of self-reported and cross-sectional designs over experimental approaches; and a disproportionate focus on university students compared to earlier educational stages.
Impact on Society	The findings suggest that the key issue is not whether GenAI will shape education, but how educational systems will shape its use. Evidence-based and deliberate integration of GenAI is essential if it is to support, rather than undermine, the development of students' CT and PS.
Future Research	Future studies should employ robust longitudinal and experimental designs to examine the long-term cognitive and developmental impacts of sustained use of GenAI, particularly in underexplored educational contexts.
Keywords	generative artificial intelligence, GenAI, critical thinking, problem-solving, systematic review, education

INTRODUCTION

The advent of artificial intelligence (AI) has precipitated profound transformations across contemporary society, with education emerging as one of the most affected domains (Gunawan et al., 2024). This technological progress has reshaped how information is created, consumed, and shared, with AI tools being increasingly integrated into daily learning practices (Cheung et al., 2025; Rios et al., 2025; Su & Yang, 2022). This rapid diffusion raises pressing questions concerning AI's role in shaping core human capacities, particularly critical thinking (CT) and problem-solving (PS), which are widely recognized as essential skills in the 21st century (Gunawan et al., 2024).

AI broadly denotes computational systems that can perform tasks traditionally requiring human intelligence, such as learning, judgment, and decision-making (L. D. Xu et al., 2021). A more recent development is generative artificial intelligence (GenAI), which can produce novel content without explicit programming (Bandi et al., 2023). Indeed, GenAI anchored in large language models can generate human-like outputs – including texts, variable sorts of visualizations, audios, etc. – across a wide range of prompts, enabling applications such as translation, summarization, and interactive dialogue (Buripakdi & An, 2024; Kasneci et al., 2023; Xia et al., 2026).

GenAI's supporters argue that this technology can enrich education in multiple ways. It can personalize instruction, provide immediate feedback, scaffold complex tasks, and adapt content to individual students' needs, thereby enhancing engagement and performance (Farrokhnia et al., 2024; Gunawan et al., 2024; J. Lee et al., 2025). GenAI has also been shown to support interactive learning environments, formative assessment, and innovative teaching strategies that foster inclusivity and student autonomy (Ahmed et al., 2024; Antón-Sancho et al., 2021). From this perspective, GenAI is not merely a technical innovation but a catalyst for more equitable and engaging educational practices.

Yet, alongside these promises, a growing body of research warns of risks associated with the use of such tools. One major concern stems from the reliability and accuracy of AI-generated content: Because GenAI tools lack genuine understanding, they may provide misleading or superficial responses (Ahmed et al., 2024; Farrokhnia et al., 2024). Researchers and educators have also raised questions about the erosion of disciplinary integrity and the threats to students' independent reasoning and higher-order thinking skills, including CT and PS, should they come to rely on AI tools (Buripakdi & An, 2024). Critics further highlight ethical challenges and issues of privacy (Cotton et al., 2024; Dimeli & Kostas, 2025; Rudolph et al., 2023). This paradox has been conceptualized as the dual nature of GenAI in education (Buripakdi & An, 2024). While GenAI can scaffold engagement and foster CT and PS, overreliance on it may, conversely, inhibit students' independent CT and PS (Gunawan et al., 2024).

Against this backdrop, CT and PS remain non-negotiable priorities in primary, secondary, and tertiary education. These skills enable students to reason effectively, evaluate information, and navigate uncertainty, capacities that are indispensable in professional and civic life (González-Pérez & Ramírez-Montoya, 2022; Rios et al., 2025). Consequently, the accelerating integration of GenAI into educational contexts demands urgent inquiry into whether these technologies act as allies in cultivating these skills or as barriers that displace them.

Despite widespread debate of this topic having rapidly increased since 2022, systematic syntheses of the evidence remain fragmented (Andreucci-Annunziata et al., 2023; Farrokhnia et al., 2024; E. Xu et al., 2023). Educators, policymakers, and technology developers require a clear, evidence-based understanding to make informed decisions about the responsible integration of GenAI in classrooms. Still, such an understanding has not yet been established. More specifically, the existing literature has not yet provided a sufficiently focused and empirically grounded synthesis of how GenAI impacts students' CT and PS, under what conditions these impacts become supportive or constraining, and where the gaps in the evidence base remain.

This lack of evidence is especially significant, given that prior syntheses have often addressed GenAI or AI in education more broadly, examined a wide range of outcomes at once, or not focused specifically and systematically on the empirical evidence concerning CT and PS. As a result, the specific consequences of GenAI for CT and PS have remained insufficiently consolidated.

This study directly addresses the lack of evidence by conducting a systematic review focused on the relationship among GenAI, CT, and PS in education. Unlike previously published syntheses, it deliberately focuses on these two closely related higher-order skills rather than on GenAI in education more broadly. In addition, it adopts a methodologically rigorous PRISMA-based approach and includes only empirical studies, thereby grounding the synthesis in data-based findings rather than in purely conceptual claims. Because self-reported evidence remains prevalent in this emerging field, the review also distinguishes between subjective and performance-based forms of empirical evidence.

More specifically, this review addresses three main research questions:

- **RQ1:** What impact does GenAI have on CT and PS in the educational context?
- **RQ2:** What key gaps remain in the empirical evidence on GenAI impacts on CT and PS in the educational context?
- **RQ3:** Under what conditions does GenAI support or constrain students' CT and PS?

By addressing these questions, this review makes three specific contributions. First, it consolidates current empirical knowledge on how GenAI impacts students' CT and PS across educational contexts. Second, it identifies key gaps in the existing evidence base and offers a foundation for future research, pedagogy, and policy. Third, it provides a more differentiated insight into GenAI's dual nature by identifying the pedagogical, cognitive, and contextual conditions under which GenAI supports, constrains, or produces mixed impacts on CT and PS. In so doing, this review seeks to advance scholarly understanding and offer practical guidance for leveraging GenAI responsibly while safeguarding the development of fundamental human capacities.

DUAL NATURE OF GENAI: THEORETICAL PERSPECTIVES

The rapid integration of GenAI into educational contexts has intensified scholarly interest in its impact on CT (e.g., Liu & Wang, 2024; Qawqzeh, 2024) and PS (e.g., Lim et al., 2025; Sharma, 2024; Urban et al., 2024). Building on this growing body of research, the following section situates our systematic review within existing research and further examines its implications through relevant learning and technology theories.

Recent research syntheses provide an important foundation for understanding the educational implications of GenAI and highlight the complexity of its impacts. Across a growing number of reviews (e.g., Adejumo et al., 2026; Celik et al., 2024; Daniel et al., 2025; Dimeli & Kostas, 2025), there is a consistent recognition that GenAI can contribute to the development of CT and PS, particularly through scaffolding, feedback, interactive engagement, and iterative learning processes. These contributions are, however, often discussed in relation to a broad range of learning outcomes rather than systematically focused on CT and PS. At the same time, these systematic reviews converge on important tensions, including the risks of overreliance, superficial engagement, and variability in outcomes across pedagogical designs and contexts. Yet, they frequently approach these tensions at a general level without consistently examining the specific conditions under which GenAI supports or constrains the development of these skills.

Other review-oriented contributions similarly emphasize the context-dependent nature of these impacts. At the same time, complementary evidence suggests that factors such as self-regulation (Zhou et al., 2024), AI dependency (Duenas & Ruiz, 2024), and the changing nature of human-AI PS processes (Joksimovic et al., 2023) mediate these mixed outcomes.

Taken together, this body of work highlights the dual, context-dependent role of GenAI and indicates the need for a more focused synthesis that explicitly centers on CT and PS. The present review addresses this need by providing a systematic synthesis of empirical studies investigating the impacts of GenAI on students' CT and PS skills, as well as the conditions shaping their development in GenAI-supported learning environments.

To better understand why GenAI may produce both beneficial and detrimental learning outcomes, it is useful to situate its use within established learning and technology theories. One relevant perspective is Cognitive Load Theory, which suggests that learning is constrained by the limited capacity of working memory (Sweller, 1988). From this viewpoint, GenAI can reduce extraneous cognitive load by structuring information, providing explanations, and guiding students through complex tasks. However, it may also inadvertently reduce germane cognitive load if students rely on AI-generated outputs without engaging in deeper processing, thereby limiting the development of higher-order thinking skills such as CT and PS (Gerlich, 2025).

A complementary lens is provided by sociocultural learning theory, which conceptualizes learning as a socially mediated process (Vygotsky, 1978). GenAI can serve as a cognitive and dialogic scaffold, supporting students within their zone of proximal development by simulating interaction, prompting reflection, and offering alternative perspectives (e.g., Martínez-Peláez et al., 2025). At the same time, this mediation may displace authentic social interaction or internalization processes if overused, potentially weakening the development of independent reasoning (Barcaui, 2025).

Finally, technology integration frameworks such as TPACK (Technological Pedagogical Content Knowledge) highlight that the educational impact of technologies depends on how they are pedagogically embedded (Mishra & Koehler, 2006). From this perspective, GenAI does not inherently enhance or inhibit CT and PS; rather, its effects are contingent on instructional design, teacher guidance, and the alignment among technology, pedagogy, and content (e.g., Celik et al., 2024; Martínez-Peláez et al., 2025). These theoretical perspectives collectively suggest that the dual nature of GenAI is not paradoxical but conditional, emerging from the interaction between technological affordances, student engagement, and educational context.

Building on the theoretical perspectives outlined above, the educational implications of GenAI depend on how the key thinking skills (CT and PS) are defined. The following section, therefore, clarifies the conceptualization of CT and PS adopted in this study.

Critical Thinking and Problem-Solving

Although extensive literature has already conceptualized CT and PS from multiple perspectives, their definitions vary considerably across contexts (compare, for example, Calma and Davies (2025) with Dhakal et al. (2023)). For this review, clear and consistent framing is essential, as it directly guides the selection of empirical studies and the interpretation of findings.

Both CT and PS are widely recognized as core competencies in 21st-century education (e.g., Baucal et al., 2023), yet both constructs are characterized by substantial conceptual diversity across disciplines and research traditions as suggested by several comprehensive reviews and meta-analyses (Andreucci-Annunziata et al., 2023; Lai, 2011; Lorencová et al., 2019; E. Xu et al., 2023). Rather than reflecting a lack of clarity, this diversity points to their multifaceted nature and the range of cognitive, metacognitive, and contextual processes they encompass. Existing scholarship generally converges on the view that CT involves reflective and reasoned judgment (encompassing analysis, evaluation, and synthesis of information) while PS emphasizes goal-directed action in situations where solutions are not immediately apparent (Ennis, 1989; Facione, 1990; Mayer & Wittrock, 2006; Paul & Elder, 2006).

Importantly, prior research consistently highlights a strong interrelationship between CT and PS. CT is often conceptualized as a foundational component of PS, providing the evaluative and analytical processes necessary to identify, assess, and refine potential solutions (Halpern, 1998). At the same time, PS extends beyond CT by incorporating strategic, procedural, and context-sensitive dimensions, including the implementation and adaptation of solutions in real-world scenarios (Jonassen, 2000). Both CT and PS, therefore, rely on active cognitive engagement, metacognitive regulation, and iterative reflection, making them particularly sensitive to changes in how learning processes are structured and supported.

At the level of empirical research, this conceptual overlap is reflected in the way CT and PS are operationalized and measured. Studies frequently examine overlapping indicators but differ in emphasis depending on disciplinary context (compare, for example, Bendl et al. (2023) with Jäder and Johansson (2025)). This variability contributes to the fragmented nature of findings discussed above, as differences in definitions and operationalizations can lead to divergent conclusions.

For our systematic review and in line with the theoretical foundations described above, CT is understood as a *reflective process of analyzing, evaluating, and synthesizing information to guide judgment and decision-making*, while PS is defined as *a structured, goal-oriented process of identifying, exploring, and resolving complex or non-routine problems*. This integrated framing allows us to capture both the cognitive and procedural dimensions of higher-order thinking and provides a consistent analytical lens for examining how GenAI may support or constrain their development across different educational contexts. Building on the theoretical foundations outlined above, the following section presents the methodological approach used to synthesize the empirical evidence systematically.

METHODOLOGY

To examine the impact of GenAI use on students' CT and PS skills and, therefore, to address the three research questions, we conducted a systematic review by following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 guidelines, which provide an explicit systematic method for identifying, evaluating, and synthesizing relevant results (Page et al., 2021).

LITERATURE SEARCH

Given its global and multidisciplinary coverage, we used the Scopus database to search for relevant English-language literature. Scopus provides extensive coverage of peer-reviewed research in education, psychology, and computer science, aligning with the interdisciplinary nature of this study. Its broad scope and substantial overlap with other major citation databases (e.g., Web of Science) support comprehensive identification of relevant literature. Given the novelty of the topic, the search was limited to records published between 2022 (the year ChatGPT was released) and 2025 and included not only articles but also conference papers and book chapters to ensure coverage of both established and emerging research outputs.

The relevant terms for the search query were defined based on the aim and key concepts of our review, namely, GenAI, CT, and PS skills. To minimize the risk of omitting relevant literature in this stage, relevant synonyms and wildcards were used. Although the primary focus of this review is GenAI, the terminology used in the literature remains inconsistent, and studies often refer to broader concepts such as artificial intelligence or large language models (LLMs) without explicitly using the term "generative AI" or "GenAI" (as reflected in the titles of the included records, see Appendix). Therefore, broader AI-related terms were included to ensure comprehensive coverage of relevant studies.

As a result, we used the following combination of keywords in the search query:

> ("*AI" OR "artificial intelligence" OR "llm" OR "large language model") AND ("enquiry" OR "inquiry" OR "problem-solv*" OR "problem solv*" OR "problem-based learning" OR "problem based learning" OR "PBL" OR "critical thinking" OR "thinking critically") AND "education"

Irrelevant subject areas in the Scopus database were deselected during the search. The search process conducted on April 24, 2025, resulted in 1,930 potentially relevant records (see Figure 1).

SCREENING PROCESS

In the first stage, the titles and abstracts of the identified records were inspected. To enhance the reliability of the screening process, approximately 5% of the records (100) were first randomly selected and screened by all five authors. The record eligibility criteria were further specified (see Figure 1) due to unsatisfactory inter-coder reliability (73.8%; Cohen's kappa = 0.47). Moreover, the criteria were ordered from the most general to the most specific, and abstract screening was thus a step-by-step process (i.e., if a record met the first three criteria but not the fourth and fifth, the fourth criterion was cited as the reason for exclusion). Thanks to this specification, the agreement between the coders on another random sample was sufficiently high for five coders (81.2%, Cohen's kappa 0.61). The coders resolved the remaining disagreements and continued individually on the remaining records. The first stage resulted in 400 potentially relevant records (Figure 1).

In the second stage, we sought the relevant records for retrieval. For 45 records, the full text was not found or was not freely available; thus, we excluded them from the full-text screening.

In the third stage, the full texts were subjected to the same criteria as the abstracts, since it was not possible to verify from each abstract whether the records met all the stated criteria. Additionally, to ensure the quality of the research, we had to verify that the record contained sufficient information on the applied research methodology and, in particular, empirical evidence of the impact of GenAI on students' CT and/or PS skills.

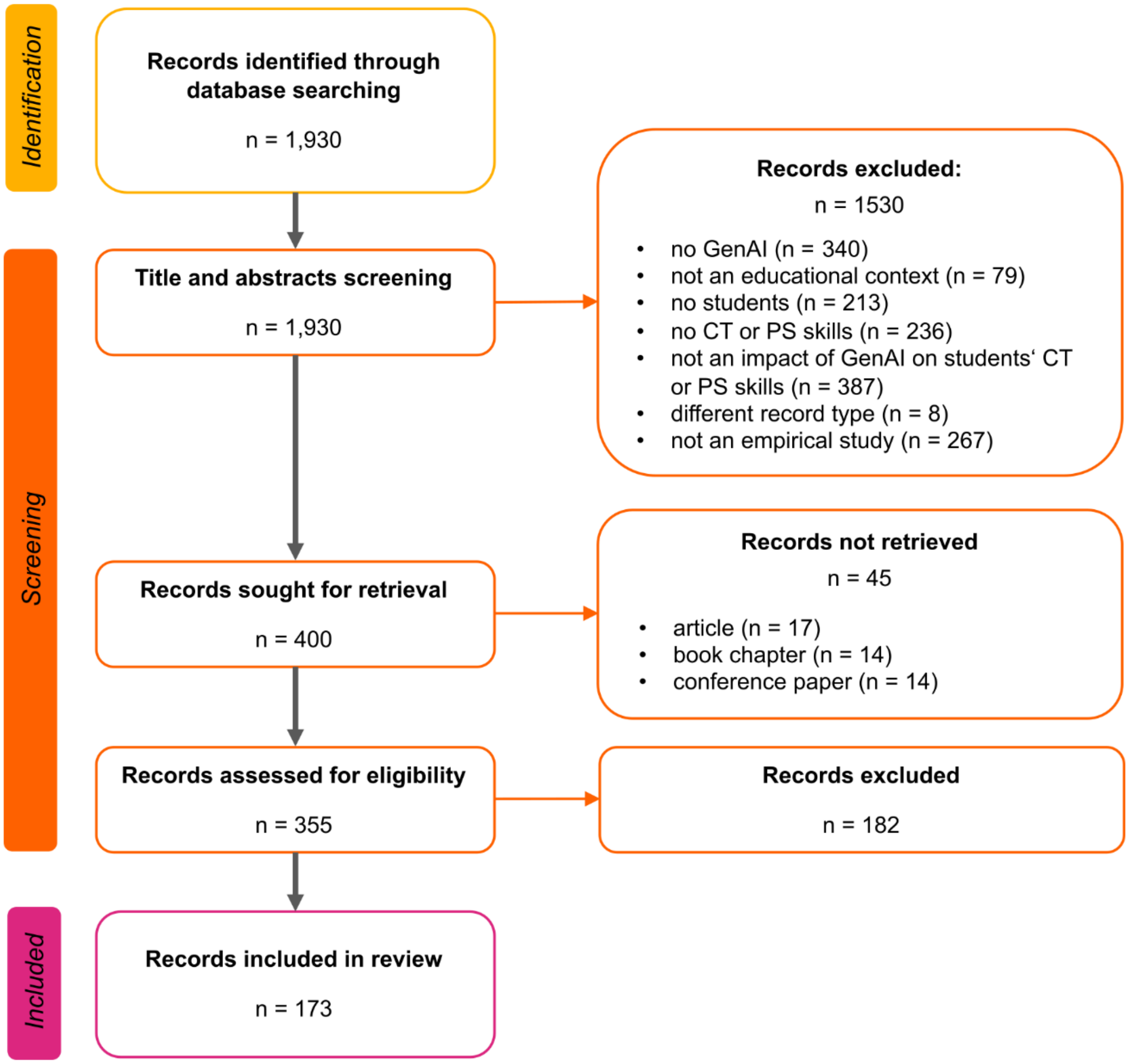

Figure 1. PRISMA flow diagram for the conducted systematic review

In the first step of the full-text screening, approximately 5% of the potentially eligible records (18) were randomly selected and evaluated by four coders (the first, second, third, and fourth authors). Due to the identified coding ambiguities (inter-coder reliability: 74.1%; Cohen's kappa: 0.51), the coding procedure was further specified and refined. Subsequently, the other 18 randomly selected records were evaluated, yielding a sufficient average agreement of 83.3% (Cohen's kappa = 0.67). After resolving the disagreements, the remaining full texts were divided among the four coders. In the end, 173 records were included in the review. In line with PRISMA guidelines (Page et al., 2021), we provide the full list of these studies in the Appendix, as it was not feasible to cite each individually in the Results section.

Data Extraction and Analysis

We developed a database grounded in a structured data-extraction framework, into which the relevant data from each included study were recorded. The framework defined in advance the variables recorded from each study, specifically:

- author's name
- study title
- publication year
- record type
- country
- research aim
- methodological approach
- method of data collection
- type of data collected
- participant type (educational level)
- sample size
- school subject
- GenAI tool
- way the GenAI tool was used
- identified positive impacts on students' CT skills
- identified negative impacts on students' CT skills
- identified positive impacts on students' PS skills
- identified negative impacts on students' PS skills.

An inductive thematic analysis, following Braun and Clarke (2006), was conducted to identify themes across the included studies. Prior to coding, all data recorded in the data extraction framework related to GenAI impacts were reviewed by the third and fourth authors to ensure familiarity with the dataset. Where necessary, the authors returned to the original records to verify and clarify relevant information, ensuring consistency between the extracted data and source documents. Initial coding was performed independently by the third and fourth authors. The third author coded reported positive and negative impacts of GenAI on students' CT skills, while the fourth focused on PS skills. This was followed by a collaborative review process in which both sets of codes were compared, discussed, and integrated. Through iterative discussion, the authors reached consensus on coding granularity and ensured conceptual alignment between CT and PS constructs. Codes were subsequently organized into higher-order themes based on conceptual similarity (their relationship). The analysis focused on semantic themes, capturing explicitly reported impacts rather than their latent interpretations. The preliminary codes and themes were then reviewed by all authors to ensure their coherence and consistency. Any discrepancies were resolved through discussion and consensus among the authors. In addition to a summary narrative text, we visualized the data using ArcGIS Pro, Gephi, Inkscape, and the Flourish web platform.

Results

This section presents the main findings in three steps. First, it provides an overview of the general findings derived from our PRISMA-based systematic review. It then offers a detailed analysis of how GenAI influences the development of CT and PS skills. Finally, it synthesizes the cross-cutting patterns that connect these findings.

In recent years, there has been a sharp increase in the number of publications addressing the impact of GenAI on CT and PS skills. Although this review focused on the short-term context of the past three years, there is clear and accelerating growth in research activity during this period (see Figure 2).

Geographically, this research is highly concentrated in east, south, and southeast Asian countries, particularly in China (28 studies), South Korea (9), Taiwan (9), and Indonesia (8), in the United States of America (15 studies), and in European countries (mainly Spain and United Kingdom, both with 9 studies), while studies from Africa, South America, and Central Asia remain largely underrepresented. This uneven distribution highlights a research gap that warrants further exploration, as shown in Figure 3.

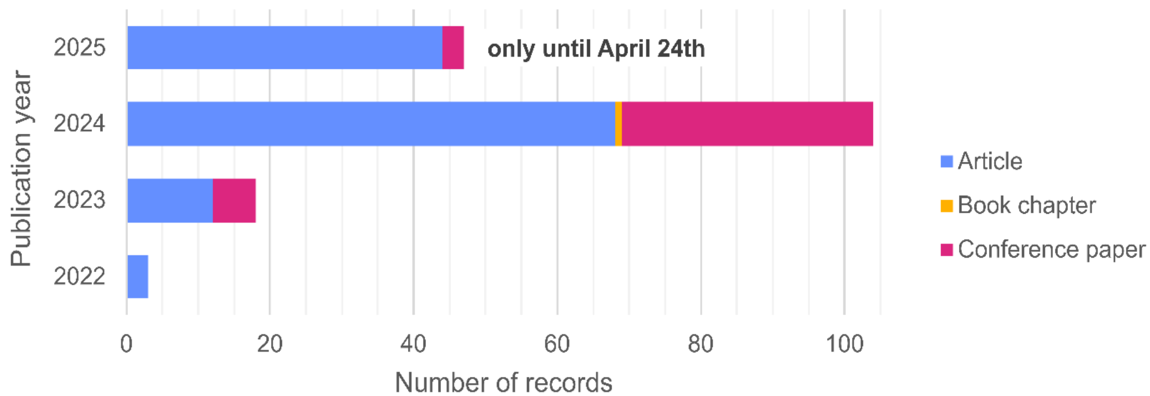

Figure 2. Research activity on GenAI in relation to CT and PS skills development

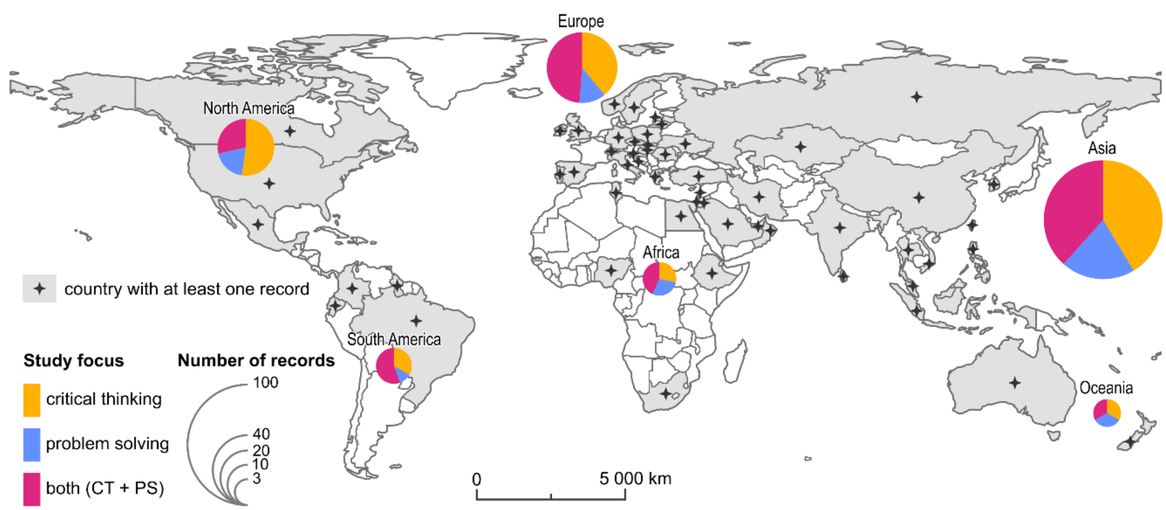

Figure 3. Spatial distribution of identified records

In terms of educational level, the literature is dominated by studies focusing on higher/university education. University students are by far the most common sample in research on GenAI's impacts on CT and PS skills (of 173 studies, 137 used university students as respondents). In contrast, research in primary and secondary education remains strikingly limited, despite growing concerns about how early exposure to GenAI tools may shape younger students' cognitive development. This imbalance suggests that the current understanding of GenAI's role in education is primarily framed through the lens of adult or near-adult students, leaving critical developmental stages largely unexplored.

A similar imbalance emerged regarding the research strategies and methods used to research this topic. Most studies relied on cross-sectional survey designs or qualitative interviews, with far fewer employing longitudinal or experimental approaches that could provide deeper causal insights into the relationship between GenAI and CT and PS skills.

Among the 173 identified empirical studies, we further distinguished between "objective" and "subjective" approaches (Figure 4). Objective studies employed experimental or quasi-experimental designs, using measurable interventions (e.g., standardized testing) to determine whether GenAI tools directly enhanced or hindered CT and PS skills. In contrast, subjective studies relied on self-reported data, such as surveys or interviews, in which participants provided their perceptions of how GenAI affected these skills. Notably, subjective studies were far more prevalent, underscoring the need for more robust, evidence-based evaluations in this emerging field.

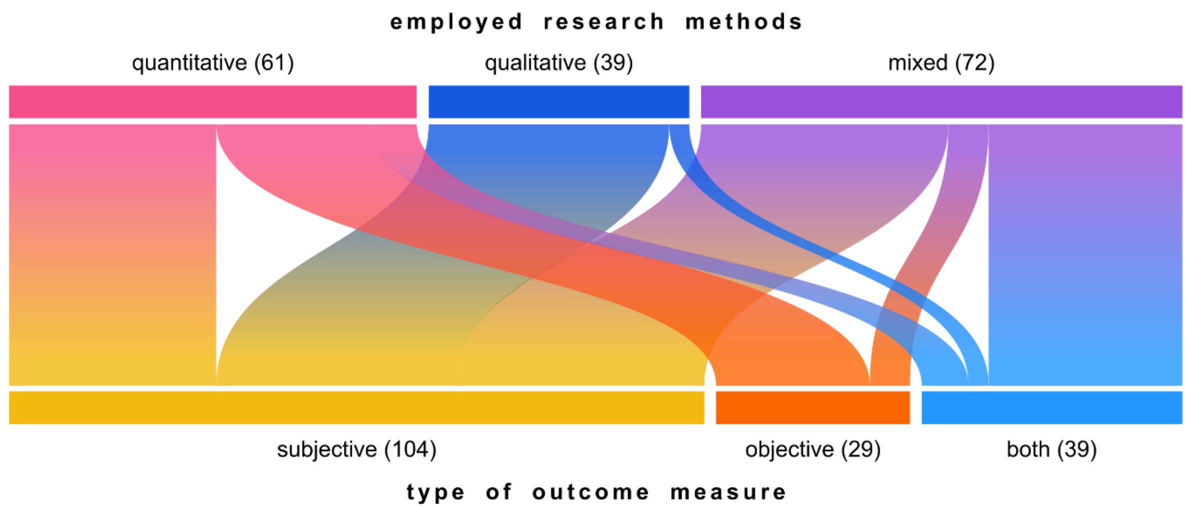

Figure 4. Methodological summary of analyzed studies

To specifically address RQ1, the following two sections present empirical evidence on how GenAI impacts the development of students' CT and PS skills

Impacts of GenAI on Critical Thinking Skills Development

Among the 139 studies addressing CT, 67 reported only positive impacts, 33 reported only negative impacts, and 39 reported both, suggesting that GenAI is more often associated with support for CT than with harm. Yet it still frequently produces mixed, context-dependent outcomes shaped by pedagogical design and patterns of use. The most frequently reported positive impacts were improved CT (n = 84), improved transversal skills (n = 63), and improved learning outcomes (n = 28), with fewer studies highlighting higher motivation (n = 18), enhanced learning experience (n = 12), and increased self-reflection (n = 12). The most frequently reported negative impacts were overreliance on GenAI (n = 37), erosion of CT (n = 36), and ethical concerns (n = 15), while conceptual misunderstanding (n = 12), reduced human interaction (n = 6), no improvement in CT (n = 3), and limitations of basic tool versions (n = 2) were less common. While a comprehensive overview of all these results is provided in Figure 5, the following subsections offer a more detailed insight into the most prevalent thematic categories.

Positive impacts

Improved CT. The reviewed studies indicate that GenAI can support the development of CT, particularly when students are required to evaluate, verify, and refine ideas rather than simply accept generated outputs. Lin et al. (2025) found that GenAI-supported writing tasks encouraged students to critically compare their own work with alternative perspectives, while Wu et al. (2025) showed that AI-assisted peer assessment strengthened students' evaluative capacity in relation to both their own and others' work. Several studies further suggested that because GenAI outputs may contain inaccuracies or bias, students are often prompted to assess reliability, verify claims, and triangulate information, thereby engaging in productive forms of analytical and evaluative thinking (Akolekar et al., 2025; Youssef et al., 2024). Relatedly, prompt refinement and iterative interaction with GenAI were described as forms of critical digital engagement that draw on higher-order information literacy skills (Fenske & Otts, 2024; Lijie et al., 2025).

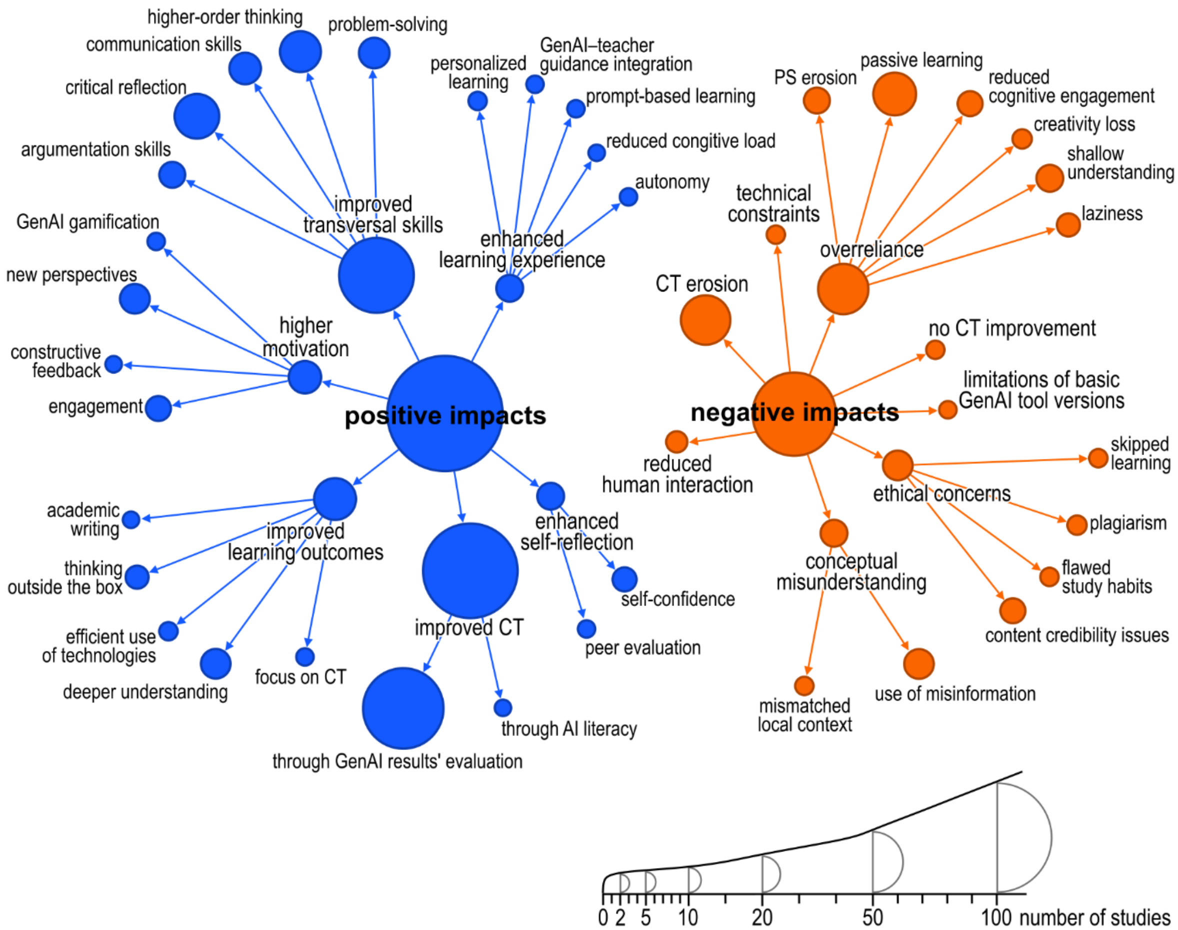

Figure 5. Categorization of the identified impacts of GenAI on CT

Improved transversal skills. A second prominent theme concerned the development of transversal skills. Studies suggested that GenAI can support broader intellectual capacities such as perspective-taking, independent thought, and engagement with multiple viewpoints, particularly when integrated into multimodal, collaborative, and project-oriented learning designs (Baek, 2025; Lin et al., 2025). Lin et al. (2025) and Suriano et al. (2025) observed that GenAI encouraged students to consider alternative perspectives, thereby broadening conceptual horizons and supporting more independent thinking. Within problem-based learning contexts, GenAI was also linked to the development of longer-term transversal competences relevant to CT (Baek, 2025).

Higher motivation, learning outcomes, and self-reflection. The reviewed studies also showed that GenAI may support motivation, learning outcomes, and self-reflection. GenAI appeared to increase engagement and students' willingness to apply CT, especially in collaborative and project-based tasks (du Plessis, 2025). When used as a scaffold through hints, questioning, and guided feedback rather than as a direct answer provider, it also helped students refine their reasoning and engage more actively in reflection (H.-Y. Lee et al., 2024; Sam et al., 2024; Wu et al., 2025). Additional studies suggested that GenAI supported creativity, personalized learning, and more complex reasoning (e.g., Aladsani, 2025; Nusivera et al., 2025; Suriano et al., 2025). Across contexts, these findings suggest that GenAI may support not only CT directly, but also broader conditions that enable students to engage more confidently and reflectively in learning.

Negative impacts and concerns

Overreliance and erosion of CT. Despite these benefits, the reviewed studies also highlighted substantial risks associated with overreliance and the erosion of CT. Several studies cautioned that habitual dependence on AI-generated outputs may weaken students' capacity or willingness to engage in independent reasoning and deeper analytical work (Aladsani, 2025; Hughes et al., 2025; Wu et al., 2025). Suriano et al. (2025) described how passive interaction with ChatGPT may encourage heuristic shortcuts and superficial processing, while Krupp et al. (2024) linked such use to "copy-and-paste" practices that bypass deliberate critical engagement.

Similarly, Akolekar et al. (2025) and Inderjeet and Bhardwaj (2024) found that dependence on GenAI explanations could discourage deeper analysis. More broadly, these concerns converged around the erosion of CT, as frequent use of GenAI was associated with weaker independent performance, confirmation bias, and difficulty distinguishing students' own ideas from AI-generated contributions (Gregory & Narang, 2024; Jošt et al., 2024; Kartal, 2024; Lijie et al., 2025).

Ethical concerns and conceptual misunderstanding. Another cluster of findings concerned ethical concerns and conceptual misunderstanding. Ethical concerns mainly involved academic integrity and the weakening of learning practices, as the use of GenAI could encourage plagiarism, bypass active engagement with learning materials, and promote study habits that prioritize convenience over critical inquiry (Akolekar et al., 2025; Lin et al., 2025; Nusivera et al., 2025). Conceptual misunderstanding referred to students' misinterpretation or uncritical acceptance of GenAI outputs, which could foster misconceptions, overtrust in hallucinated or inaccurate responses, and increased exposure to misinformation and bias (Akolekar et al., 2025; du Plessis, 2025; Nusivera et al., 2025; Styve et al., 2024; Suriano et al., 2025). Some studies further warned that such reliance may distort scholarly judgment and narrow students' epistemic horizons (Aladsani, 2025; Fenske & Otts, 2024).

Less frequently reported concerns included reduced human interaction, no observable improvement in CT, and limitations associated with basic versions of GenAI tools. Although these themes were only marginally represented, they suggest that both the quality of the tool and the conditions of use may shape the extent to which GenAI supports critical engagement.

IMPACTS OF GENAI ON PROBLEM-SOLVING SKILLS DEVELOPMENT

Among the 103 studies addressing PS, 57 reported only positive impacts, 23 reported only negative impacts, and 23 reported both, suggesting that GenAI is more often associated with support for PS than with harm, while still producing mixed and context-dependent outcomes in a substantial number of studies.

The most frequently reported positive impacts were enhanced learning outcomes (n = 60), social support and collaborative learning (n = 28), and improved cognition and strategic thinking (n = 22). Less frequently, studies reported positive affective responses (n = 17), expanded problem-solving opportunities (n = 15), and other benefits that did not fit the main thematic categories (n = 2). The most frequently reported negative impacts were cognitive risks and skill erosion (n = 28), quality and academic integrity issues (n = 8), and implementation barriers and technical constraints (n = 6), while social and engagement risks (n = 2) and other negative impacts (n = 2) were less common. A comprehensive summary of these results is presented in Figure 6.

The following sections offer a more detailed insight into the most prevalent thematic categories.

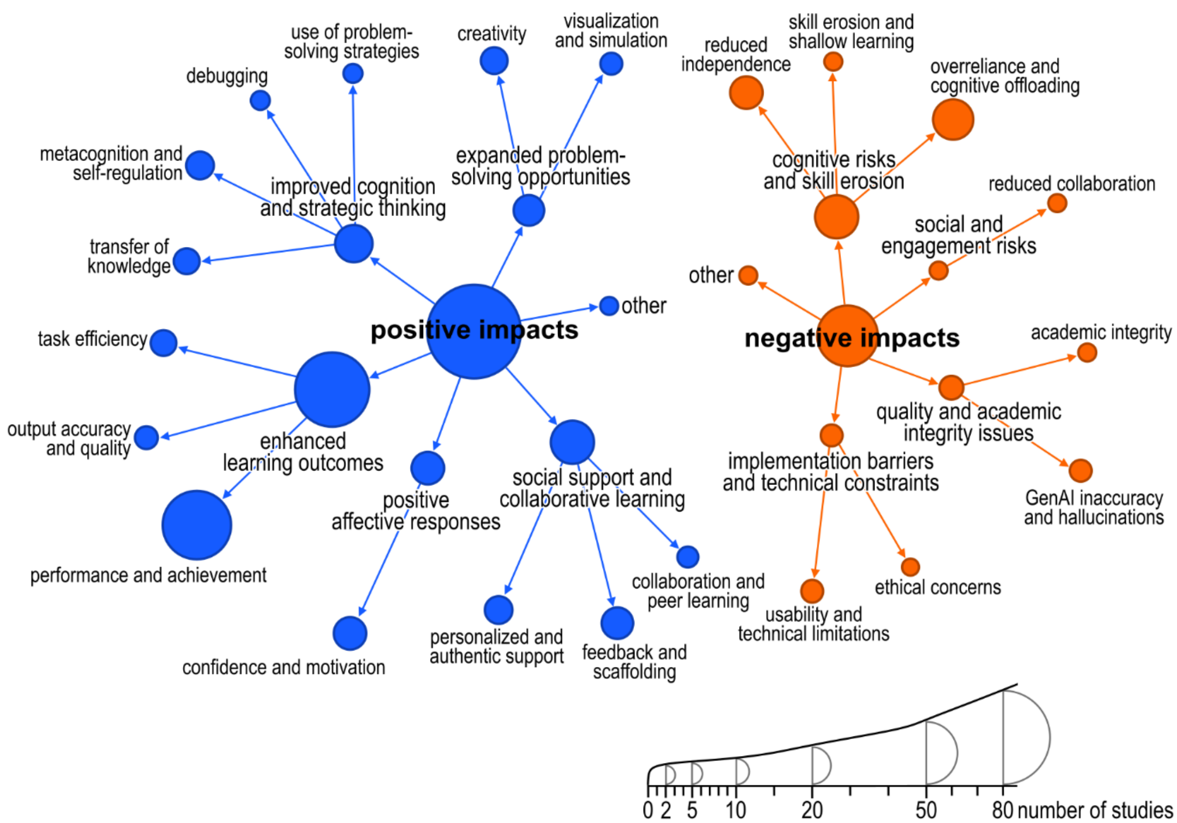

Figure 6. Categorization of the identified impacts of GenAI on PS

Positive impacts

Enhanced learning outcomes. The studies analyzed consistently documented improvements in observable student PS outcomes when GenAI was strategically integrated into learning activities. Okada et al. (2025), for instance, found that projects utilizing GenAI successfully supported student engagement in addressing both local and global challenges. Similarly, Hwang et al. (2025) documented how the structured use of GenAI-supported inquiry processes enabled students to develop more complete solutions while requiring significantly less time investment. Multiple studies also reported enhanced efficiency in PS workflows, with students demonstrating reductions in unnecessary procedural steps and an overall simplification of complex processes (Gonsalves, 2026; Hwang et al., 2025; Mike et al., 2024). Furthermore, when students used GenAI to verify their logical reasoning or refine their solutions, researchers observed more precise, error-free final products (Gonsalves, 2026; Maya & Valdes, 2024).

Social support and collaborative learning. The reviewed studies showed that GenAI frequently functions as an adaptive scaffold, providing just-in-time support through hints, explanations, and step-by-step guidance that students can utilize independently or integrate into collaborative discussions (Hwang et al., 2025; Maya & Valdes, 2024). Related to the latter, Rodriguez et al. (2024) emphasized that GenAI serves as a catalyst for group PS activities, facilitating more productive team interactions. The personalization capabilities of GenAI emerged as a particularly valuable component, with students leveraging GenAI-generated adaptations to enhance group-work productivity and accommodate diverse learning needs (Filiz & Gür, 2025; Gonsalves, 2026; Gregory & Narang, 2024; Okada et al., 2025).

Improved cognitive processes and strategic thinking. The identified studies showed that students demonstrate enhanced metacognitive awareness when using GenAI, employing the technology to plan solution steps, monitor progress, reflect on decision-making processes, and transfer successful strategies to novel problem contexts (Hwang et al., 2025; Okada et al., 2025). In technical and design-oriented disciplines, GenAI functions effectively as a "debug partner," systematically identifying errors and suggesting targeted remedial actions (Wang et al., 2023). GenAI also supports knowledge transfer and application by enabling students to engage with concepts through realistic, interactive scenarios. Relatedly, Zheng et al. (2024) demonstrated that GenAI enhances practical skill development through role-play activities, simulations, and collaborative PS experiences while fostering higher-order thinking in game-based learning contexts.

Positive affective responses and expanded problem-solving opportunities. Positive affective responses and expanded problem-solving opportunities were also reported across the reviewed studies. GenAI appeared to increase students' willingness to engage with challenging problems, sustain effort during complex tasks, and reduce anxiety by providing immediate guidance and high-quality examples (French et al., 2023; Rodríguez-Rivera et al., 2025). At the same time, it expanded problem-solving opportunities by generating alternative ideas, scenarios, and perspectives, as well as by offering visual and simulated representations that supported reasoning about abstract relationships and system dynamics (Del Mundo et al., 2024; Kishore et al., 2023; Zheng et al., 2024).

Negative impacts and concerns

Cognitive risks and skill erosion. Despite the benefits outlined above, several studies have raised concerns that GenAI may undermine essential cognitive processes (see Figure 6). The theme of cognitive risks and skill erosion was by far the most prominent among the negative PS findings. Studies particularly expressed apprehension that students may become less inclined to engage in productive struggle, potentially weakening their independent PS capabilities over time (Gerlich, 2025; Vieriu & Petrea, 2025). One participant in Gerlich's (2025, p. 20) study articulated this concern: "*The more I use AI, the less I feel the need to solve problems on my own. It's like I'm losing my ability to think critically*." Baláž et al. (2024) provided empirical evidence for these concerns, explicitly warning of skill erosion when GenAI substitutes for hands-on practice.

Quality and academic integrity issues, implementation barriers and technical constraints, and social and engagement risks. Quality and academic integrity issues, implementation barriers and technical constraints, and social and engagement risks were also identified, although less frequently. Several studies have shown that students sometimes accept plausible-sounding but incorrect GenAI outputs, which could misdirect their problem-solving efforts (Akolekar et al., 2025; Filiz & Gür, 2025). Other studies pointed to practical limitations, particularly in complex or abstract tasks, where GenAI could reproduce flawed reasoning, fabricate data, or otherwise provide unreliable support (Juáres & Guzmán, 2024). Concerns were also raised that prolonged reliance on GenAI in demanding tasks such as programming may foster unhealthy dependency and weaken long-term skill development (Palacios & Rivadeneira, 2025). In addition, a smaller but notable set of findings suggested risks for peer collaboration, as students may turn to GenAI as a private assistant rather than engaging with classmates, thereby weakening the social dimensions of problem-solving (Wang et al., 2023).

Synthesis of the Findings

Across the reviewed studies, positive impacts of GenAI on CT and PS were reported more frequently than negative impacts, although a substantial number of studies identified both positive and negative impacts simultaneously (see Figure 7). This pattern highlights that while GenAI was often associated with support for CT and PS, its impacts remained contingent upon conditions of use and pedagogical design.

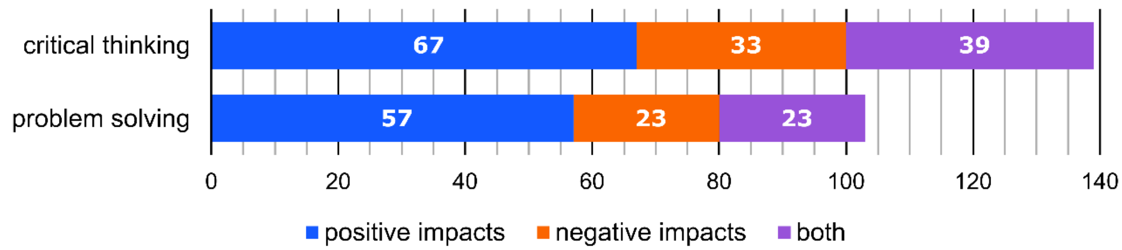

Figure 7. Distribution of GenAI impacts on CT and PS across identified records

While CT and PS were analyzed as distinct domains in this systematic review, the mechanisms through which GenAI affects them often overlapped (see Figures 5 and 6). Across both skill sets, metacognitive scaffolding, prompt-based reflection, and collaborative engagement appeared to shape whether GenAI functioned primarily as a cognitive support or as a constraint. When GenAI was used to prompt students to question, refine, and justify their decisions in CT-related tasks, similar processes also seemed to support strategic planning and adaptive reasoning in PS. Likewise, GenAI's role in facilitating structured dialogue, whether through peer assessment or collaborative PS, was associated with both reflective judgment and shared inquiry.

Taken together, the findings suggest that the educational impacts of GenAI are shaped not only by the technology itself but also by how it is embedded in pedagogical design and taken up by students. Purposeful integration through active engagement, reflective use, and alignment with collaborative or project-based pedagogies appeared to support motivation, creativity, communication, and self-reflection, thereby reinforcing both CT and PS. By contrast, passive use, limited AI literacy, and overreliance on automated responses were more often associated with ethical concerns, conceptual misunderstanding, and reduced higher-order reasoning.

Overall, the findings indicate that GenAI is most supportive when students are encouraged to treat its outputs as starting points for further exploration rather than as authoritative answers. In this sense, the implications of GenAI are not tied only to the targeted skills but also to the pedagogical and cognitive context in which the technology is used, where it may either support reflective inquiry and analytical growth or hinder the productive struggle central to learning.

DISCUSSION

While the preceding sections presented empirical evidence on how GenAI influences CT and PS, this section seeks to go further by addressing RQ2. We begin by reflecting on key research gaps identified by our systematic review. We then turn to a more nuanced discussion of the dual nature of GenAI and the conditions under which it supports or constrains the development of CT and PS, thereby revisiting RQ3. In doing so, we situate the results within the "bigger picture" of rapidly expanding GenAI adoption. Lastly, we reflect on the limitations of our review that shape the scope and interpretation of these findings.

IDENTIFYING KEY RESEARCH GAPS AND THEORETICAL IMPLICATIONS

To address RQ2, our systematic review highlights several critical research gaps that limit a comprehensive understanding of how GenAI influences the development of CT and PS skills. These gaps point not only to areas of ambiguous or even contradictory evidence but also to opportunities to advance the field through more rigorous and targeted research.

First, there is an apparent imbalance in the spatial distribution of the analyzed literature. Most studies are concentrated in high-income regions, seemingly because these are areas with strong research and innovation capacity, whereas research from low- and middle-income regions remains scarce. This lack of balance in global representation risks reinforcing a narrow perspective of GenAI's educational impact, potentially overlooking challenges related to infrastructure, cultural factors, and pedagogical traditions in less-represented contexts.

The methodological gap is equally striking. The majority of existing studies relied on cross-sectional surveys or qualitative interviews, with very few adopting experimental or longitudinal designs that could provide more robust evidence of causal relationships. Moreover, the limited number of empirical studies, particularly "objective" ones that measure skill development through standardized assessments, weakens the evidence base. As a result, many reported impacts reflect perceived changes rather than observable cognitive development, making it difficult to determine whether GenAI genuinely enhances higher-order thinking or merely alters students' experiences of it. To address this, future research should prioritize experimental and longitudinal approaches, integrate objective measures of CT and PS, and triangulate findings using mixed-methods designs to better capture both the measurable and experiential dimensions of GenAI's influence on key 21st-century skills.

Lastly, our review shows that the current research has overwhelmingly focused on university students, leaving primary and secondary education largely unexplored. Yet, the increasing availability of GenAI tools for younger students raises important questions about how early exposure to these tools might shape the development of CT and PS skills during the formative stages of education. Consequently, we recommend expanding research efforts to include primary and (lower- and upper-) secondary educational levels, where the pedagogical and developmental implications of GenAI may differ substantially from higher-education/university contexts.

Taken together, the evidence base resembles a mosaic of promising case studies rather than a coherent picture of GenAI's educational impact. While this mosaic confirms that GenAI can support CT and PS, it tells us far less about the conditions under which such support is sustainable or about the risks of cumulative negative impacts over time. Effectively addressing the dual nature of GenAI and its impact on the development of CT and PS will require not only a deliberate effort to include underrepresented educational settings and regions but, above all, the implementation of more diverse and empirically grounded methodological approaches. By adopting such methods, future studies can move beyond the fragmented evidence to date and offer a more nuanced and globally relevant understanding of GenAI's dual role in fostering or constraining these essential skills.

Beyond identifying empirical gaps, the findings of this review also point to several important theoretical implications, which are broadly consistent with emerging research on the impacts of AI on our cognition and skill development (e.g., Du et al., 2025). First, the impact of GenAI on CT and PS should not be understood as inherently positive or negative, but rather as conditional and context-dependent. This suggests that the "dual nature" of GenAI is better explained by how students engage with AI-generated outputs and how these tools are embedded within specific learning tasks, rather than by the technology itself. Second, the development of CT and PS appears to be mediated not by the presence of GenAI but by factors such as students' self-regulation, the design of learning activities, the quality of scaffolding, and the depth of cognitive engagement they elicit. Third, these findings support a reconceptualization of GenAI in educational settings – not as a mere answer-providing tool, but as a scaffold, a dialogic partner, and a potential cognitive amplifier, whose impact depends on whether it is used to support or replace students' thinking processes. These theoretical implications provide a basis for a more nuanced interpretation of the findings, which is further elaborated in the following discussion of GenAI's dual role in shaping CT and PS.

A More Complex View of the Dual Nature of GenAI in Relation to Critical Thinking and Problem-Solving

In addressing RQ3, this section examines the conditions under which GenAI supports or constrains the development of students' CT and PS skills, building on the empirical patterns identified in the preceding analysis. Our findings suggest that the integration of GenAI into education holds considerable promise for enhancing CT and PS. Across the empirical studies examined, the dominant trend points toward positive outcomes. The mechanisms behind these impacts are consistent with existing cognitive theories of scaffolding and guided discovery learning (e.g., Jonassen, 2000). GenAI tools, when designed or implemented as supportive "scaffolds" rather than direct answer-providers, can prompt students both to think critically and solve problems. This effect likely occurs because such interactions require students to actively evaluate, compare, and refine information, thereby engaging the core processes of analysis and judgment that are central to CT and PS. From a sociocultural perspective (Vygotsky, 1978), these findings can be understood as evidence that GenAI functions as a form of mediated support within students' zone of proximal development, enabling the co-construction of knowledge through guided interaction and dialogic engagement. This also aligns with both Facione's (1990) and Halpern's (1998) conceptualizations of CT as a skill that develops most robustly in environments that provide both cognitive challenge and structured support. In this sense, GenAI can act as a novel form of cognitive apprenticeship, offering timely feedback, simulating complex scenarios, and enabling students to test their reasoning in a low-stakes environment.

However, these findings must be interpreted with caution. Although the reviewed literature demonstrated a preponderance of positive results, this trend may reflect a partial picture. Most included studies examined GenAI use in narrowly defined educational contexts. Under such conditions, positive impacts are perhaps unsurprising: If GenAI is integrated into well-structured problem-based learning or inquiry-driven environments, it can indeed amplify existing pedagogical strengths. Thus, the supportive role of GenAI is highly contingent upon intentional pedagogical design and active student engagement. Yet, the broader reality of students' everyday use of GenAI extends far beyond controlled classroom settings. Outside carefully designed interventions, the risk of overdependence, superficial engagement, or uncritical reliance on AI-generated content is significantly greater. From the perspective of Cognitive Load Theory (Sweller, 1988), such patterns may be explained by the externalization of cognitive processes, where excessive reliance on GenAI reduces opportunities for germane cognitive load associated with deep processing, thereby limiting the development of CT and PS. In other words, when students outsource key steps of reasoning to GenAI, they bypass the very mental effort through which higher-order thinking skills are formed. This highlights the constraining conditions under which GenAI may undermine CT and PS, particularly when use becomes passive, uncritical, or habitual.

Herein lies the crux of GenAI's "dual nature". While it can strengthen CT and PS development under optimal conditions, it can equally weaken these very skills when employed uncritically or excessively. Even if GenAI integration in classrooms successfully promotes CT and PS in specific tasks, one must consider what happens when students rely on GenAI tools across the full spectrum of daily cognitive challenges. For instance, a student may practice critical evaluation when comparing AI-generated essays to their own writing in a controlled activity. Yet, the same student might outsource everyday decision-making to personal AI assistants, delegating tasks such as choosing a computer, organizing finances, or even drafting personal communications. These are precisely the kinds of everyday situations that historically nurtured informal PS and CT skills. Increasingly mediated these decisions with GenAI may represent a "silent erosion" of cognitive effort. In other words, while individual studies have highlighted micro-level gains, the macro-level reality could involve a net loss of opportunities to engage in authentic, self-directed thinking. This paradox has been noted in broader philosophical debates on automation and cognitive offloading (Gerlich, 2025), but it remains underexplored in empirical educational research.

This tension becomes even clearer when considering the extraordinary speed with which GenAI has become embedded in everyday life. ChatGPT, for example, was released on November 30, 2022, and within only two months was estimated to have reached 100 million active users, with around 13 million unique visitors daily, making it the fastest-growing application in history (Peters et al., 2024). Such unprecedented adoption illustrates not only the enthusiasm for GenAI but also the rapid normalization of these tools in education, work, and daily decision-making. As the social theorist McLuhan (1994, p. 9) observed, "we shape our tools, and thereafter our tools shape us." The current trajectory suggests that GenAI is no longer a peripheral experiment but a central element of how we think, solve problems, and interact with knowledge. The crucial question, therefore, is not whether GenAI will reshape CT and PS development; rather, it concerns how we can maintain agency over how it shapes us. From the perspective of RQ3, this underscores that the conditions shaping GenAI's impact extend beyond formal educational settings and increasingly include everyday patterns of use. As a result, the conditions under which GenAI supports CT and PS are closely tied to structured guidance, reflective use, and pedagogical alignment, whereas unstructured or unguided use is more likely to constrain their development.

This situation points toward a fundamental challenge for educators and policymakers: how to preserve and cultivate CT and PS in an era when many of the tasks that once required them can now be outsourced to AI systems. Our review suggests that the answer does not lie in resisting GenAI altogether, but in equipping students with the competence to use GenAI critically and responsibly. Furthermore, this aligns with the TPACK framework (Mishra & Koehler, 2006), underscoring that the educational impact of technologies depends not on their mere presence but on how effectively they are aligned with pedagogical goals and content knowledge. This is where the concept of AI literacy becomes central (Ng et al., 2021; Zhong & Liu, 2025).

AI literacy, in this sense, is not simply about knowing how AI systems function, but about developing the metacognitive skills to interrogate their outputs, recognize their limitations, and decide when (and when not) to rely on them. This includes fostering dispositions such as skepticism, reflection, and intellectual humility, attitudes long associated with CT (Ennis, 1989; Facione, 1990). Equally important, however, is the educator's intentional role. Our review revealed that the most positive impacts of GenAI on CT and PS occur when it is carefully embedded in well-designed teaching and learning activities. This means that the quality of educators' planning, i.e., how they structure tasks, frame prompts, and anticipate moments of reflection, largely determines whether GenAI becomes a catalyst for deeper thinking or a shortcut to superficial answers. By embedding GenAI literacy as a transversal competence, educational systems can help ensure that GenAI becomes a catalyst rather than a crutch for CT and PS. Without such deliberate cultivation, there is a real risk that the short-term gains highlighted in current empirical studies will be overshadowed by long-term cognitive atrophy in everyday reasoning and decision-making.

Furthermore, while concerns about overreliance, diminished originality, and shallow engagement are valid, several positive outcomes identified in this review suggest pathways to mitigate these risks. Notably, many of the most promising benefits, such as increased evaluative reasoning, metacognitive awareness, and collaborative PS, directly counteract these negative impacts when GenAI is embedded within reflective, guided, and supported learning contexts. Rather than ignoring the limitations of GenAI, effective pedagogical designs can turn these limitations into productive triggers for critical engagement. In this sense, the positive mechanisms do not merely coexist with the negative; they actively provide a means to neutralize or reverse them, provided that GenAI use is intentional, scaffolded, and aligned with the broader goals of cognitive autonomy and ethical, responsible AI use.

Limitations

To situate our results in a broader context, it is important to acknowledge several limitations inherent in the chosen theoretical and methodological framework. The first significant limitation of this review is the predominance of self-reported data across the included studies. Much of the research on GenAI's influence on CT and PS relies on surveys and interviews, which provide valuable insights into students' and educators' perceptions but offer limited evidence of actual skill development. As such, reported benefits may reflect perceived improvements rather than demonstrable gains in higher-order thinking, while reported concerns may not always correspond to measurable declines. To mitigate this limitation, we distinguished between "objective" studies, which employed measurable interventions or standardized assessments, and "subjective" studies, which were based on selfreporting. Although this allowed us to evaluate findings from both groups separately, the overall evidence base in current academic literature remains weighted toward subjective perceptions, limiting the strength of causal interpretations and underscoring the need for more robust, performance-based research designs. In addition, we believe that the relative scarcity of experimental and longitudinal studies further constrains the ability to draw firm conclusions about causality and long-term impacts. Most existing studies capture short-term outcomes within specific instructional contexts, making it difficult to determine whether observed improvements or declines in CT and PS are sustained over time or transferable across different learning environments. At the same time, it should be noted that the review did not employ a formal quality assessment tool to evaluate the methodological rigor of the included studies systematically. These studies were included based on predefined eligibility criteria and their empirical relevance to the research questions, but not weighted by methodological quality. While differences in study design were considered in the analysis, the absence of a standardized assessment of study quality means that variations in methodological robustness were not formally weighted in the synthesis.

Second, our search strategy was limited to the Scopus database, which might have restricted the comprehensiveness of the review. Although Scopus provides broad coverage of peer-reviewed literature across disciplines relevant to this study, other databases (such as ERIC) and grey literature (e.g., dissertations and preprints) may include relevant studies not captured by Scopus. In particular, databases focused on education, psychology, and technology in education research (e.g., ERIC, Web of Science, PsycINFO, IEEE, or ACM) may contain relevant studies that were not indexed in Scopus. In addition, the search relied on the keywords “critical thinking” and “problem-solving.” This focus may have excluded research addressing specific subcomponents of these skills (e.g., inference, evaluation, metacognition) without explicitly using the broader terms. As a result, our findings should be viewed as a synthesis of the literature explicitly framed in CT and PS, rather than a comprehensive mapping of all related cognitive skills.

Third, a further limitation relates to the geographic and contextual distribution of the included studies. The existing evidence base is heavily concentrated in high-income regions and higher education settings, with limited representation from low-income and middle-income contexts and earlier stages of education. This imbalance may restrict the generalizability of the findings, as the impact of GenAI on CT and PS is likely shaped by factors such as educational infrastructure, cultural norms, and pedagogical traditions. This pattern likely reflects disparities in access to GenAI technologies and research capacity across regions, highlighting the need for future studies that more systematically include diverse educational contexts, underrepresented regions, and a broader range of student populations.

Fourth, both CT and PS are complex constructs with multiple definitions. Our decision to adopt specific definitions, while grounded in well-established scholarship, inevitably shaped how we interpreted and categorized the literature. Studies using alternative conceptualizations may have emphasized different skills or relationships, a possibility that underscores the need to view our conclusions within these specific theoretical frameworks.

CONCLUSION

This systematic review set out to examine the impact of GenAI on two foundational 21st-century skills: CT and PS. Drawing on a PRISMA-based synthesis of 173 empirical studies published between 2022 and 2025, our analysis demonstrated that GenAI holds considerable potential to act as both a catalyst and a constraint in the development of higher-order cognitive skills mentioned. On one hand, GenAI can scaffold reflective inquiry, foster diverse perspectives, and support structured engagement with complex tasks. On the other hand, its widespread and often uncritical use raises concerns about dependency, superficial reasoning, and the erosion of independent PS and CT capacities. These findings confirm the dual nature of GenAI in education: It can be either a contributor or an obstacle to students' intellectual growth, depending on how it is embedded in pedagogical practices and educational settings.

Three key insights emerged from our review. First, the benefits of GenAI are most pronounced when it is integrated into structured, inquiry-based, or problem-oriented instructional designs that deliberately engage students in evaluating, questioning, and refining GenAI outputs. In such contexts, GenAI serves less as an answer provider and more as a cognitive scaffold, enabling students to cultivate the evaluative and metacognitive habits essential for CT and PS. Second, risks arise primarily in unstructured or passive use, where the temptation to offload cognitive effort to GenAI tools can undermine students' independent reasoning and originality. Third, the current evidence base is highly uneven, dominated by self-reported studies and largely focused on higher-educational levels, highlighting substantial gaps that limit our ability to generalize findings across contexts and student populations.

The implications of these findings are twofold. For educational practice, our review underscores the urgent need to design learning environments that cultivate AI literacy, understood not merely as technical proficiency, but as the capacity to interrogate, contextualize, and critically engage with AI-generated content. In practical terms, this may involve practices such as teaching students to verify GenAI outputs against other sources, identify possible hallucinations and bias, justify when and why GenAI was used, and reflect on which parts of a task should remain cognitively their own. Likewise, pedagogical design should, for example, position GenAI as a scaffold rather than as a substitute for thinking by asking students to compare AI-generated and self-generated responses, critique and revise GenAI outputs, and use GenAI as a starting point for further reasoning rather than as a final answer. These implications are consistent with emerging literature emphasizing that AI should complement rather than replace students' CT and PS, and that pedagogical scaffolding and assessment design matter for preserving cognitive engagement and academic integrity. By embedding AI literacy across disciplines, educators and policymakers may help safeguard CT and PS while enabling students to leverage GenAI responsibly. For research, there is a clear demand for more robust, longitudinal, and experimental studies that examine not only short-term outcomes but also the long-term cognitive and developmental consequences of sustained GenAI use, particularly in underexplored contexts.

More specifically, future research should prioritize experimental and quasi-experimental designs capable of testing causal effects, extend the evidence base beyond higher education into primary and (lower and upper) secondary education contexts, and include a stronger representation of underrepresented regions. There is also a need for more performance-based assessments of CT and PS, so that future studies can move beyond predominantly self-reported perceptions and provide stronger evidence of actual CT and PS levels and development.

Ultimately, the question is not whether GenAI will shape education, but how education can shape its use. By providing a focused and empirically grounded synthesis of 173 studies, this PRISMA-based review presents a more differentiated understanding of the conditions under which GenAI supports, constrains, or ambiguously influences students' CT and PS. It therefore provides a foundation for future, more informed pedagogical, policy, and research decisions. Only through deliberate, evidence-based integration can GenAI serve as an aid rather than a barrier to the cultivation of the CT and PS capacities that remain indispensable for students in a rapidly evolving world.

REFERENCES

Adejumo, A. A., Oyelere, S. S., Sanusi, I. T., & Suhonen, J. (2026). A systematic review of the impact of GenAI on learning performance, AI hallucinations, and problem-solving in computer science education. *Computers and Education: Artificial Intelligence*, *10*, Article 100570. https://doi.org/10.1016/j.caeai.2026.100570

Ahmed, Z., Shanto, S. S., Rime, M. H. K., Morol, Md. K., Fahad, N., Hossen, Md. J., & Abdullah-Al-Jubair, Md. (2024). The generative AI landscape in education: Mapping the terrain of opportunities, challenges, and student perception. *IEEE Access*, *12*, 147023–147050. https://doi.org/10.1109/ACCESS.2024.3461874

Akolekar, H., Jhamnani, P., Kumar, V., Tailor, V., Pote, A., Meena, A., Kumar, K., Challa, J. S., & Kumar, D. (2025). The role of generative AI tools in shaping mechanical engineering education from an undergraduate perspective. *Scientific Reports*, *15*, Article 9214. https://doi.org/10.1038/s41598-025-93871-z

Aladsani, H. K. (2025). Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: A participatory action research. *Interactive Learning Environments*, *33*(10), 5747–5765. https://doi.org/10.1080/10494820.2025.2487517

Andreucci-Annunziata, P., Riedemann, A., Cortés, S., Mellado, A., del Río, M. T., & Vega-Muñoz, A. (2023). Conceptualizations and instructional strategies on critical thinking in higher education: A systematic review of systematic reviews. *Frontiers in Education*, *8*, Article 1141686. https://doi.org/10.3389/feduc.2023.1141686

Antón-Sancho, Á., Vergara, D., Lamas-Álvarez, V. E., & Fernández-Arias, P. (2021). Digital content creation tools: American university teachers' perception. *Applied Sciences*, *11*(24), Article 11649. https://doi.org/10.3390/app112411649

Baek, J. (2025). PBL education plan for video content production class based on generative AI technology. *Edelweiss Applied Science and Technology*, *9*(4), 83–87. https://doi.org/10.55214/25768484.v9i4.5940

Baláž, N., Porubän, J., Horváth, M., & Kormaník, T. (2024). Using ChatGPT during implementation of programs in education. In A. L. Santos & M. Pinto-Albuquerque (Eds.), *Proceedings of the 5th International Conference on Programming Education for Children:* OASIcs (Vol. 122, pp. 18:1–18:9). Dagstuhl Publishing. https://doi.org/10.4230/OASICS.ICPEC.2024.18

Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet*, *15*(8), Article 260. https://doi.org/10.3390/fi15080260

Barcaui, A. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. *Social Sciences & Humanities Open*, *12*, Article 102287. https://doi.org/10.1016/j.ssaho.2025.102287

Baucal, A., Jošić, S., Stepanović Ilić, I., Videnović, M., Ivanović, J., & Krstić, K. (2023). What makes peer collaborative problem solving productive or unproductive: A qualitative systematic review. *Educational Research Review*, *41*, Article 100567. https://doi.org/10.1016/j.edurev.2023.100567

Bendl, T., Marada, M., & Havelková, L. (2023). Preservice geography teachers' exposure to problem solving and different teaching styles. *Journal of Geography*, *122*(3), 66–76. https://doi.org/10.1080/00221341.2023.2220114

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buripakdi, A., & An, L. T. (2024). "Why can't I use ChatGPT for my academic learning?" Voices from Ph.D. students in a Thai university. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(2), 299–319. https://doi.org/10.70730/VRVS6900

Calma, A., & Davies, M. (2025). Assessing students' critical thinking abilities via a systematic evaluation of essays. *Studies in Higher Education*, *51*(2), 422–437. https://doi.org/10.1080/03075079.2025.2470969

Celik, I., Gedrimiene, E., Siklander, S., & Muukkonen, H. (2024). The affordances of artificial intelligence-based tools for supporting 21st-century skills: A systematic review of empirical research in higher education. *Australasian Journal of Educational Technology*, *40*(3), 19–38. https://doi.org/10.14742/ajet.9609

Cheung, K. K. C., Long, Y., Liu, Q., & Chan, H.-Y. (2025). Unpacking epistemic insights of artificial intelligence (AI) in science education: A systematic review. *Science & Education*, *34*(2), 747–777. https://doi.org/10.1007/s11191-024-00511-5

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Daniel, K., Msambwa, M. M., & Wen, Z. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education*, *60*(1), Article e70036. https://doi.org/10.1111/ejed.70036

Del Mundo, M. A., Delos Reyes, E. F., Gervacio, E. M., Manalo, R. B., Book, R. J. B., Chavez, J. V., Espartero, M. M., & Sayadi, D. S. (2024). Discourse analysis on experience-based position of science, mathematics, and Tech-Voc educators on generative AI and academic integrity. *Environment and Social Psychology*, *9*(8), 1–16. https://doi.org/10.59429/esp.v9i8.3028

Dhakal, K. R., Watson Todd, R., & Jaturapitakkul, N. (2023). Unpacking the nature of critical thinking for educational purposes. *Educational Research and Evaluation*, *28*(4–6), 130–151. https://doi.org/10.1080/13803611.2023.2262447

Dimeli, M., & Kostas, A. (2025). The role of ChatGPT in education: Applications, challenges: Insights from a systematic review. *Journal of Information Technology Education: Research*, *24*, Article 2. https://doi.org/10.28945/5422

Du, X., Du, M., Zhou, Z., & Bai, Y. (2025). Facilitator or hindrance? The impact of AI on university students' higher-order thinking skills in complex problem solving. *International Journal of Educational Technology in Higher Education*, *22*, Article 39. https://doi.org/10.1186/s41239-025-00534-0

Duenas, T., & Ruiz, D. (2024). *The risks of human overreliance on large language models for critical thinking* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.26002.06082

du Plessis, E. D. (2025). Embracing project-based assessments in the age of AI in open distance e-learning. *International Journal of Information and Education Technology*, *15*(2), 372–381. https://doi.org/10.18178/ijiet.2025.15.2.2249

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, *18*(3), 4–10. https://doi.org/10.3102/0013189X018003004

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* (ED315423). ERIC. https://eric.ed.gov/?id=ED315423

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Fenske, R. F., & Otts, J. A. A. (2024). Incorporating generative AI to promote inquiry-based learning: Comparing elicit AI research assistant to PubMed and CINAHL Complete. *Medical Reference Services Quarterly*, *43*(4), 292–305. https://doi.org/10.1080/02763869.2024.2403272

Filiz, A., & Gür, H. (2025). Students' perceptions and applications of metacognitive awareness levels in problem solving with ChatGPT. *Educational Process: International Journal*, *14*, Article e2025063. https://doi.org/10.22521/edupij.2025.14.63

French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: Case studies from game development. *Multimodal Technologies and Interaction*, *7*(8), Article 81. https://doi.org/10.3390/mti7080081

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

Gonsalves, C. (2026). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. *Journal of Marketing Education*, *48*(1), 4–19. https://doi.org/10.1177/02734753241305980

González-Pérez, L. I., & Ramírez-Montoya, M. S. (2022). Components of education 4.0 in 21st century skills frameworks: Systematic review. *Sustainability*, *14*(3), Article 1493. https://doi.org/10.3390/su14031493

Gregory, R. W., & Narang, S. (2024). AI for learning unleashed: Pioneering generative AI in education at the University of Miami. *Journal of Information Technology Teaching Cases, 16*(1), 3–12. https://doi.org/10.1177/20438869241266258

Gunawan, Palupi, S., Hardi, R., Pribadi, A. S., Zulkarnain, R., Setiawan, M. N., Sari, N. W. W., Sulistiyanto, & Utomo, D. T. (2024). Enhancing computational strategies to decode ChatGPT's influence on the critical thinking abilities of university students. *Proceedings of the 4th International Conference of Science and Information Technology in Smart Administration, Balikpapan, Indonesia*, 27–32. https://doi.org/10.1109/ICSINTESA62455.2024.10748139

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, *53*(4), 449–455. https://doi.org/10.1037/0003-066X.53.4.449

Hughes, L., Malik, T., Dettmer, S., Al-Busaidi, A. S., & Dwivedi, Y. K. (2025). Reimagining higher education: Navigating the challenges of generative AI adoption. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-025-10582-6

Hwang, G.-J., Cheng, P.-Y., & Chang, C.-Y. (2025). Facilitating students' critical thinking, metacognition and problem-solving tendencies in geriatric nursing class: A mixed-method study. *Nurse Education in Practice*, *83*, Article 104266. https://doi.org/10.1016/j.nepr.2025.104266

Inderjeet, & Bhardwaj, R. (2024). AI-integrated model design for the education system for analytical computation (mathematics). *Proceedings of the 4th International Conference on Advance Computing and Innovative Technologies in Engineering, Greater Noida, India,* 1412–1417. https://doi.org/10.1109/ICACITE60783.2024.10616503

Jäder, J., & Johansson, H. (2025). Exploring students' conceptual understanding through mathematical problem solving: Students' use of and shift between different representations of rational numbers. *Research in Mathematics Education*, 1-18. https://doi.org/10.1080/14794802.2025.2456840

Joksimovic, S., Ifenthaler, D., Marrone, R., De Laat, M., & Siemens, G. (2023). Opportunities of artificial intelligence for supporting complex problem-solving: Findings from a scoping review. *Computers and Education: Artificial Intelligence*, *4*, Article 100138. https://doi.org/10.1016/j.caeai.2023.100138

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85. https://doi.org/10.1007/BF02300500

Jošt, G., Taneski, V., & Karakatič, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences*, *14*(10), Article 4115. https://doi.org/10.3390/app14104115

Juáres, E. C., & Guzmán, D. S. (2024). Implementing the LSEESC methodology with AI-guided support for active learning in engineering physics courses. *Proceedings of the IEEE Frontiers in Education Conference, Washington, DC, USA*, 1–7. https://doi.org/10.1109/FIE61694.2024.10893032

Kartal, G. (2024). The influence of ChatGPT on thinking skills and creativity of EFL student teachers: A narrative inquiry. *Journal of Education for Teaching*, *50*(4), 627–642. https://doi.org/10.1080/02607476.2024.2326502

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kishore, S., Hong, Y., Nguyen, A., & Qutab, S. (2023). Should ChatGPT be banned at schools? Organizing visions for generative artificial intelligence (AI) in education. *Proceedings of the ICIS 2023 Conference* (pp. 1–17). AIS. https://aisel.aisnet.org/icis2023/learnandiscurricula/learnandiscurricula/5

Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K. E., Lukowicz, P., Kuhn, J., Küchemann, S., & Karolus, J. (2024). Challenges and opportunities of moderating usage of large language models in education. *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia* (pp. 249–254). Association for Computing Machinery. https://doi.org/10.1145/3701571.3701590

Lai, E. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*, 40–41.

Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, *21*, Article 16. https://doi.org/10.1186/s41239-024-00447-4

Lee, J., Cimová, T., Foster, E. J., France, D., Krajňáková, L., Moorman, L., Rewhorn, S., & Zhang, J. (2025). Transforming geography education: The role of generative AI in curriculum, pedagogy, assessment, and fieldwork. *International Research in Geographical and Environmental Education*, *34*(3), 237–253. https://doi.org/10.1080/10382046.2025.2459780

Lijie, H., Mat Yusoff, S., & Mohamad Marzaini, A. F. (2025). Influence of AI-driven educational tools on critical thinking dispositions among university students in Malaysia: A study of key factors and correlations. *Education and Information Technologies*, *30*(6), 8029–8053. https://doi.org/10.1007/s10639-024-13150-8

Lim, J., Lee, U., Koh, J., Jeong, Y., Lee, Y., Byun, G., Jung, H., Jang, Y., Lee, S., & Moon, J. (2025). Development and implementation of a generative artificial intelligence-enhanced simulation to enhance problem-solving skills for pre-service teachers. *Computers & Education*, *232,* Article 105306. https://doi.org/10.1016/j.compedu.2025.105306

Lin, C.-H., Zhou, K., Li, L., & Sun, L. (2025). Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms. *Computers and Composition*, *75*, Article 102895. https://doi.org/10.1016/j.compcom.2024.102895

Liu, W., & Wang, Y. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*, *59*(4), Article e12804. https://doi.org/10.1111/ejed.12804

Lorencová, H., Jarošová, E., Avgitidou, S., & Dimitriadou, C. (2019). Critical thinking practices in teacher education programmes: A systematic review. *Studies in Higher Education*, *44*(5), 844–859. https://doi.org/10.1080/03075079.2019.1586331

Martínez-Peláez, R., Mena, L. J., Toral-Cruz, H., Ochoa-Brust, A., Potes, A. G., Flores, V., Ostos, R., Pacheco, J. C. R., Félix, R. A., & Félix, V. G. (2025). Can large language models foster critical thinking, teamwork, and problem-solving skills in higher education? A literature review. *Systems*, *13*(11), Article 1013. https://doi.org/10.3390/systems13111013

Maya, H. R. R., & Valdes, A. B. S. (2024). Empowering critical thinking through AI: The PAIR model's impact on higher education excellence. In *Proceedings of the 8th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (pp. 1–6). IEEE. https://doi.org/10.1109/ISMSIT63511.2024.10757186

Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 287–303). Routledge.

McLuhan, M. (1994). *Understanding media: The extensions of man*. MIT Press.

Mike, N., Karsai, K., Orbán, G., Bubelényi, A., Nagy, C. N., & Polyák, G. (2024). The impact of ChatGPT on student performance in higher education. In J. Ubacht, J. Crompvoets, C. Csáki, L. Danneels, M. Janssen, M. R. Johannessen, T. Lampoltshammer, H. Lee, I. Lindgren, A. Nikiforova, S. Hofmann, P. Parycek, G. V. Pereira, G. Schwabe, I. Susha, & E. Tambouris (Eds.), *Proceedings of Ongoing Research, Practitioners, Posters, Workshops, and Projects of the International Conference EGOV-CeDEM-ePart 2024*. https://ceur-ws.org/Vol-3737/

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Nusivera, E., Hikmat, A., & Ghani, A. R. A. (2025). Integration of Chat-GPT usage in language learning model to improve argumentation skills, complex comprehension skills, and critical thinking skills. *International Journal of Learning, Teaching and Educational Research*, *24*(2), 375–390. https://doi.org/10.26803/ijlter.24.2.19

Okada, A., Sherborne, T., Panselinas, G., & Kolionis, G. (2025). Fostering transversal skills through open schooling supported by the CARE-KNOW-DO pedagogical model and the UNESCO AI competencies framework. *International Journal of Artificial Intelligence in Education*, *35*(4), 1953–1998. https://doi.org/10.1007/s40593-025-00458-w

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*(71). https://doi.org/10.1136/bmj.n71

Palacios, D., & Rivadeneira, L. (2025). Influence of ChatGPT on programming code generation: A case study of the Technical University of Manabí. In T. Guarda, F. Portela, & G. Gatica (Eds.), *Advanced research in technologies, information, innovation and sustainability* (pp. 276–289). Springer. https://doi.org/10.1007/978-3-031-83210-9_21

Paul, R. W., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education*, *30*(2), 34–35.

Peters, M. A., Jackson, L., Papastephanou, M., Jandrić, P., Lazaroiu, G., Evers, C. W., Cope, B., Kalantzis, M., Araya, D., Tesar, M., Mika, C., Chen, L., Wang, C., Sturm, S., Rider, S., & Fuller, S. (2024). AI and the future of humanity: ChatGPT-4, philosophy and education – Critical responses. *Educational Philosophy and Theory*, *56*(9), 828–862. https://doi.org/10.1080/00131857.2023.2213437

Qawqzeh, Y. (2024). Exploring the influence of student interaction with ChatGPT on critical thinking, problem solving, and creativity. *International Journal of Information and Education Technology*, *14*(4), 596–601. https://doi.org/10.18178/ijiet.2024.14.4.2082

Rios, T. C.-D. L., Solis-Trujillo, B., Perez-Ruiz, J., & Aquije-Mansilla, M. (2025). Systematic review of critical thinking using artificial intelligence. *Edelweiss Applied Science and Technology*, *9*(3), 990–1001. https://doi.org/10.55214/25768484.v9i3.5405

Rodriguez, D., Mourad, F., Vatus, C., & Blavy, E. (2024). Uses and impacts of LLMs on the learning experience of future engineers: Case studies in the context of a French engineering school. *Proceedings for the 52nd Annual Conference of the European Society for Engineering Education* (pp. 2063–2075). SEFI.

Rodríguez-Rivera, P., Rodríguez-Ferrer, J. M., & Manzano-León, A. (2025). Designing digital escape rooms with generative AI in university contexts: A qualitative study. *Multimodal Technologies and Interaction*, *9*(3), Article 20. https://doi.org/10.3390/mti9030020

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*, 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Sam, K.-L., Ieong, U.-S., & da Costa, M. G. (2024). PeerSynergy: An innovative application applying large language models to peer instruction. Proceedings of the *IEEE International Conference on Teaching, Assessment and Learning for Engineering, Bengaluru, India,* 1–7. https://doi.org/10.1109/TALE62452.2024.10834348

Sharma, D. (2024). Critical thinking and problem-solving in the age of ChatGPT: Experiential-bibliotherapy-blogging project. *Business and Professional Communication Quarterly*, *87*(4), 630–653. https://doi.org/10.1177/23294906241254780

Styve, A., Virkki, O. T., & Naeem, U. (2024). Developing critical thinking practices interwoven with generative AI usage in an introductory programming course. *Proceedings of the IEEE Global Engineering Education Conference, Kos Island, Greece*, 1–8. https://doi.org/10.1109/EDUCON60312.2024.10578746

Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. *Computers and Education: Artificial Intelligence*, *3*, Article 100049. https://doi.org/10.1016/j.caeai.2022.100049

Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction*, *95*, Article 102011. https://doi.org/10.1016/j.learninstruc.2024.102011

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285. https://doi.org/10.1207/s15516709cog1202_4

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*, Article 105031. https://doi.org/10.1016/j.compedu.2024.105031

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, *15*(3), Article 343. https://doi.org/10.3390/educsci15030343

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, Article 104703. https://doi.org/10.1016/j.compedu.2022.104703

Wu, T.-T., Lee, H.-Y., Chen, P.-H., Lin, C.-J., & Huang, Y.-M. (2025). Integrating peer assessment cycle into ChatGPT for STEM education: A randomised controlled trial on knowledge, skills, and attitudes enhancement. *Journal of Computer Assisted Learning*, *41*(1), Article e13085. https://doi.org/10.1111/jcal.13085

Xia, Q., Liu, Q., Tlili, A., & Chiu, T. K. F. (2026). A systematic literature review on designing self-regulated learning using generative artificial intelligence and its future research directions. *Computers & Education*, *240*, Article 105465. https://doi.org/10.1016/j.compedu.2025.105465

Xu, E., Wang, W., & Wang, Q. (2023). The effectiveness of collaborative problem solving in promoting students' critical thinking: A meta-analysis based on empirical literature. *Humanities and Social Sciences Communications*, *10*, Article 16. https://doi.org/10.1057/s41599-023-01508-1

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. *IEEE Internet of Things Journal*, *8*(13), 10452–10473. https://doi.org/10.1109/JIOT.2021.3060508

Youssef, E., Medhat, M., Abdellatif, S., & Al Malek, M. (2024). Examining the effect of ChatGPT usage on students' academic learning and achievement: A survey-based study in Ajman, UAE. *Computers and Education: Artificial Intelligence*, *7*, Article 100316. https://doi.org/10.1016/j.caeai.2024.100316

Zheng, K., Shen, Z., Chen, Z., Che, C., & Zhu, H. (2024). Application of AI-empowered scenario-based simulation teaching mode in cardiovascular disease education. *BMC Medical Education*, *24*, Article 1003. https://doi.org/10.1186/s12909-024-05977-z

Zhong, B., & Liu, X. (2025). Evaluating AI literacy of secondary students: Framework and scale development. *Computers & Education*, *227*, Article 105230. https://doi.org/10.1016/j.compedu.2024.105230

Zhou, X., Teng, D., & Al-Samarraie, H. (2024). The mediating role of generative AI self-regulation on students' critical thinking and problem solving. *Education Sciences*, *14*(12), Article 1302. https://doi.org/10.3390/educsci14121302

APPENDIX

General findings of the reviewed studies

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Acuña, O. L., Santos Carvajal, D. M., Bolanos-Barbosa, A. D., Torres-Vanegas, J. D., Alvarez Solano, O. A., Cruz, J. C., & Reyes, L. H. (2025). Fostering technical proficiency and professional skills: A multifaceted PO-PBL strategy for unit operations education. *Education for Chemical Engineers, 51*, 64–78. https://doi.org/10.1016/j.ece.2025.01.001	✓		✓	
Akolekar, H., Jhamnani, P., Kumar, V., Tailor, V., Pote, A., Meena, A., Kumar, K., Challa, J. S., & Kumar, D. (2025). The role of generative AI tools in shaping mechanical engineering education from an undergraduate perspective. *Scientific Reports*, *15*, Article 9214. https://doi.org/10.1038/s41598-025-93871-z	✓	✓	✓	✓
Al Yakin, A., Muthmainnah, M., Al-Matari, A. S., de Barros Cardoso, L. M. O., Yunus, N. H., Hajar, S., & Harianti, H. (2025). Transferability cybersocialization as a digital civility through artificial intelligence-based design thinking mindset to modern pedagogy. In V. Goar, M. Kuri, R. Kumar, & T. Senjyu (Eds), *Advances in information communication technology and computing* (pp. 237–250). Springer. https://doi.org/10.1007/978-981-97-6106-7_13	✓			
Aladini, A., Ismail, S. M., Ahmad Saleem Khasawneh, M., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports, 17,* 100566. https://doi.org/10.1016/j.chbr.2024.100566	✓			
Aladsani, H. K. (2025). Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: A participatory action research. *Interactive Learning Environments*, *33*(10), 5747–5765. https://doi.org/10.1080/10494820.2025.2487517	✓	✓		
Alarcón, C. M. M., Lasekan, O. A., & Godoy Pena, M. T. (2025). AI concerns among educators: A thematic analysis. *International Research Journal of Multidisciplinary Scope*, *6*(1), 1031–1045. https://doi.org/10.47857/irjms.2025.v06i01.01929		✓		✓
Alarcón-López, C., Krütli, P., & Gillet, D. (2024). Assessing ChatGPT's influence on critical thinking in sustainability oriented activities. *Proceedings of the IEEE Global Engineering Education Conference, Kos Island, Greece* (pp. 1–10). https://doi.org/10.1109/EDUCON60312.2024.10578790		✓		
Almahasees, Z., Khalil, M., & AMinzadeh, S. (2024). Students' perceptions of the benefits and challenges of integrating ChatGPT in higher education. *Pakistan Journal of Life and Social Sciences, 22*(2). https://doi.org/10.57239/PJLSS-2024-22.2.00256		✓		
Almulla, M., & Ali, S. I. (2024). The changing educational landscape for sustainable online experiences: Implications of ChatGPT in Arab students' learning experience. *International Journal of Learning, Teaching and Educational Research, 23*(9), Article 9. https://ijlter.org/index.php/ijlter/article/view/10586	✓	✓		

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
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	✓		✓	
Alwaqdani, M. (2025). Investigating teachers' perceptions of artificial intelligence tools in education: Potential and difficulties. *Education and Information Technologies, 30*(3), 2737–2755. https://doi.org/10.1007/s10639-024-12903-9		✓		✓
Al-Zahrani, A. M. (2024). Unveiling the shadows: Beyond the hype of AI in education. *Heliyon, 10*(9), e30696. https://doi.org/10.1016/j.heliyon.2024.e30696		✓		
Alzubi, A. A. F., Nazim, M., & Alyami, N. (2025). Do AI-generative tools kill or nurture creativity in EFL teaching and learning? *Education and Information Technologies, 30*, 15147–15184. https://doi.org/10.1007/s10639-025-13409-8		✓		✓
Araujo, S. M., & Cruz-Correia, R. (2024). Incorporating ChatGPT in medical informatics education: Mixed methods study on student perceptions and experiential integration proposals. *JMIR Medical Education, 10*, e51151. https://doi.org/10.2196/51151	✓	✓		
Aristovnik, A., Umek, L., Brezovar, N., Keržič, D., & Ravšelj, D. (2024). The role of ChatGPT in higher education: Some reflections from public administration students. In S. K. S. Cheung, F. L. Wang, N. Paoprasert, P. Charnsethikul, K. C. Li, & K. Phusavat (Eds.), *Technology in education. Innovative practices for the new normal* (pp. 254–263). Springer Nature. https://doi.org/10.1007/978-981-99-8255-4_22		✓		
Ayanwale, M. A., Adelana, O. P., Molefi, R. R., Adeeko, O., & Ishola, A. M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Computers and Education Open*, *6*, 100179. https://doi.org/10.1016/j.caeo.2024.100179			✓	
Ayman, S. E., El-Seoud, S. A., Nagaty, K., & Karam, O. H. (2023). The influence of ChatGPT on student learning and academic performance. *Procedings of the 2023 International Conference on Computer and Applications* (pp. 1–5). IEEE. https://doi.org/10.1109/ICCA59364.2023.10401713	✓			
Baek, J. (2025). PBL education plan for video content production class based on generative AI technology. *Edelweiss Applied Science and Technology*, *9*(4), 83–87. https://doi.org/10.55214/25768484.v9i4.5940	✓		✓	
Baláž, N., Porubän, J., Horváth, M., & Kormaník, T. (2024). Using ChatGPT during implementation of programs in education. In A. L. Santos & M. Pinto-Albuquerque (Eds.), *Proceedings of the 5th International Conference on Programming Education for Children* (Vol. 122, pp. 18:1–18:9). Dagstuhl Publishing. https://doi.org/10.4230/OASICS.ICPEC.2024.18		✓		✓
Barana, A., Marchisio, M., & Roman, F. (2023). Fostering problem solving and critical thinking in mathematics through generative artificial intelligence. *Proceedings of the 20th International Conference on Cognition and Exploratory Learning in the Digital Age* (pp. 377–385). IADIS Press. https://doi.org/10.33965/CELDA2023_202306L046	✓		✓	✓
Biberman-Shalev, L. (2025). Prompting theory into practice: Utilizing ChatGPT-4 in a curriculum planning course. *Education Sciences, 15*(2), Article 2. https://doi.org/10.3390/educsci15020196	✓	✓		

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Bokan, M., Abdimalikkyzy, J., Babaeva, A., Keldibekova, A., Makpyr, S., & Smagulov, Y. (2025). Enhancing logical thinking skills of future informatics teachers through artificial intelligence. *Qubahan Academic Journal*, *5*, 543–551. https://doi.org/10.48161/qaj.v5n1a1329			✓	
Briedienė, S., & Pranckevičiūtė, V. (2024). Transforming maritime education: Impact of artificial intelligence on learning. *Proceedings of the 28th International Scientific Conference. Transport Means* (pp. 198–202). KTU Leidykla Technologija. https://doi.org/10.5755/e01.2351-7034.2024.P198-202	✓	✓		
Buripakdi, A., & An, L. T. (2024). "Why can't I use ChatGPT for my academic learning?" Voices from Ph.D. students in a Thai university. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(2), 299–319. https://doi.org/10.70730/VRVS6900		✓		✓
Buselic, V., & Rajkovic, I. (2024). Teaching generic skills with/or against Chat GPT. *Proceedings of the 47th MIPRO ICT and Electronics Convention, Opatija, Croatia,* 707–712. https://doi.org/10.1109/MIPRO60963.2024.10569431	✓			
Canonigo, A. (2024). Levering AI to enhance students' conceptual understanding and confidence in mathematics. *Journal of Computer Assisted Learning*, *40*, 3215–3229. https://doi.org/10.1111/jcal.13065			✓	
Chang, C.-Y., Yang, C.-L., Jen, H.-J., Ogata, H., & Hwang, G.-H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology & Society*, *27*(1), 215–230. https://doi.org/10.30191/ETS.202401_27(1).TP02	✓		✓	
Chavez, J. V., Cuilan, J. T., Mannan, S. S., Ibrahim, N. U., Carolino, A. A., Radjuni, A., Albani, S. E., & Garil, B. A. (2024). Discourse analysis on the ethical dilemmas of using AI in academic settings from ICT, science, and language instructors. *Forum for Linguistic Studies*, *6*(5), 349–363. https://doi.org/10.30564/fls.v6i5.6765		✓		✓
Chung, J. Y., & Jeong, S.-H. (2024). Exploring the perceptions of Chinese pre-service teachers on the integration of generative AI in English language teaching: Benefits, challenges, and educational implications. *Online Journal of Communication and Media Technologies, 14*(4), e202457. https://doi.org/10.30935/ojcmt/15266		✓		
Conde, M. Á., & Román-Gallego, J.-Á. (2024). Do engineering students know how to use generative artificial intelligence? A case study. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies* (pp. 241–254). Springer. https://doi.org/10.1007/978-3-031-61691-4_16			✓	
Cong-Lem, N., Tran, T. N., & Nguyen, T. T. (2024). Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers. *Teaching English with Technology, 24*(1), 28–47. https://doi.org/10.56297/FSYB3031/MXNB7567		✓		
Costa, A. R., Lima, N., Viegas, C., & Caldeira, A. (2024). Critical minds: Enhancing education with ChatGPT. *Cogent Education, 11*(1), Article 2415286. https://doi.org/10.1080/2331186X.2024.2415286	✓			
Dai, Y., Xiao, J.-Y., Huang, Y., Zhai, X., Wai, F.-C., & Ming, Z. (2025). How generative AI enables an online project-based learning platform: An applied study of learning behavior analysis in undergraduate students. *Applied Sciences, 15*(5), 2369. https://doi.org/10.3390/app15052369	✓		✓	

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
de la Puente, M., Torres, J., Troncoso, A. L. B., Meza, Y. Y. H., & Carrascal, J. X. M. (2024). Investigating the use of ChatGPT as a tool for enhancing critical thinking and argumentation skills in international relations debates among undergraduate students. *Smart Learning Environments, 11*, Article 55. https://doi.org/10.1186/s40561-024-00347-0	✓			
de la Puente Pacheco, M. A., Torres, J., Blanco Troncoso, A. L., Guzmán Murrilo, H. J., & Carrascal, J. X. M. (2025). Enhancing critical thinking and argumentation skills in Colombian undergraduate diplomacy students: ChatGPT-assisted and traditional debate methods. *Journal of Political Science Education, 21*(4), 728–738. https://doi.org/10.1080/15512169.2025.2449936	✓			
Dobre, S.-C., & Popescu, E. (2024). Exploring students' perception and experience with ChatGPT and critical thinking in a higher education context. *Proceedings of the 21st International Conference on Information Technology Based Higher Education and Training, Paris, France,* 1–6. https://doi.org/10.1109/ITHET61869.2024.10837650	✓			
Drushlyak, M., Lukashova, T., Sabadosh, Y., Melnikov, I., & Semenikhina, O. (2024). Using ChatGPT for the development of critical thinking in youth: Example of inequality proof. *Proceedings of the 47th MIPRO ICT and Electronics Convention, Opatija, Croatia*, 334–339. https://doi.org/10.1109/MIPRO60963.2024.10569759	✓	✓		
du Plessis, E. D. (2025). Embracing project-based assessments in the age of AI in open distance e-learning. *International Journal of Information and Education Technology*, *15*(2), 372–381. https://doi.org/10.18178/ijiet.2025.15.2.2249	✓	✓	✓	✓
Eltahir, M. E., & Babiker, F. M. E. (2024). The influence of artificial intelligence tools on student performance in e-learning environments: Case study. *Electronic Journal of E-Learning, 22*(9), Article 9. https://doi.org/10.34190/ejel.22.9.3639	✓			
Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (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	✓	✓		
Fadillah, M. A., Usmeldi, U., & Asrizal, A. (2024). The role of ChatGPT and higher-order thinking skills as predictors of physics inquiry. *Journal of Baltic Science Education, 23*(6), 1178–1192. https://doi.org/10.33225/jbse/24.23.1178			✓	
Famaye, T., Abimbade, O., & Adisa, O. (2025). "AI is not just a robot; it's more than a robot.": Understanding children's AI competencies in an AI literacy workshop. *Journal of Interactive Learning Research, 36*(1), 71–82.	✓			
Farhi, F., Jeljeli, R., Aburezeq, I., Dweikat, F. F., Al-shami, S. A., & Slamene, R. (2023). Analyzing the students' views, concerns, and perceived ethics about chat GPT usage. *Computers and Education: Artificial Intelligence, 5*, 100180. https://doi.org/10.1016/j.caeai.2023.100180	✓			
Farinetti, L., & Cagliero, L. (2025). A critical approach to ChatGPT: An experience in SQL learning. *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1* (pp. 318–324). Association for Computing Machinery. https://doi.org/10.1145/3641554.3701932	✓			

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Fenske, R. F., & Otts, J. A. A. (2024). Incorporating generative AI to promote inquiry-based learning: Comparing elicit AI research assistant to PubMed and CINAHL Complete. *Medical Reference Services Quarterly*, *43*(4), 292–305. https://doi.org/10.1080/02763869.2024.2403272	✓	✓		
Filiz, A., & Gür, H. (2025). Students' perceptions and applications of metacognitive awareness levels in problem solving with ChatGPT. *Educational Process International Journal*, *14*(1), Article e2025063. https://doi.org/10.22521/edupij.2025.14.63			✓	✓
Fontao, C. B., Lopéz Santos, M., & Lozano, A. (2024). ChatGPT's role in the education system: Insights from the future secondary teachers. *International Journal of Information and Education Technology, 14*(8), 1035–1043. https://doi.org/10.18178/ijiet.2024.14.8.2131		✓		
French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: Case studies from game development. *Multimodal Technologies and Interaction*, *7*(8), Article 81. https://doi.org/10.3390/mti7080081	✓	✓	✓	
Gattupalli, S., Lee, W., Allessio, D., Crabtree, D., Arroyo, I., & Woolf, B. (2023). Exploring pre-service teachers' perceptions of large language models-generated hints in online mathematics learning. *Proceedings of AIED 2023: Empowering Education with LLMs – The Next-Gen Interface and Content Generation* (pp. 1–12). CEUR-WS. https://ceur-ws.org/Vol-3487/paper10.pdf			✓	✓
Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), Article 6. https://doi.org/10.3390/soc15010006		✓		✓
Gervacio, A. P. (2024). Exploring how generative AI contributes to the motivated engagement and learning production of science-oriented students. *Environment and Social Psychology, 9*(11), Article 11. https://doi.org/10.59429/esp.v9i11.3194	✓		✓	
Gonsalves, C. (2026). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. *Journal of Marketing Education*, *48*(1), 4–19. https://doi.org/10.1177/02734753241305980	✓		✓	
Govindarajan, R., & Christuraj, G. (2024). Opportunities and challenges of using ChatGPT in the ELT scenario of Utas, Nizwa, Oman. *Journal of Teaching English for Specific and Academic Purposes, 11*(3), 593–605. https://doi.org/10.22190/JTESAP230529046G		✓		
Gregory, R. W., & Narang, S. (2024). AI for learning unleashed: Pioneering generative AI in education at the University of Miami. *Journal of Information Technology Teaching Cases, 16*(1), 3–12. https://doi.org/10.1177/20438869241266258	✓	✓	✓	
Gunawan, Palupi, S., Hardi, R., Pribadi, A. S., Zulkarnain, R., Setiawan, M. N., Sari, N. W. W., Sulistiyanto, & Utomo, D. T. (2024). Enhancing computational strategies to decode ChatGPT's influence on the critical thinking abilities of university students. *Proceedings of the 4th International Conference of Science and Information Technology in Smart Administration, Balikpapan, Indonesia,* 27–32. https://doi.org/10.1109/ICSINTESA62455.2024.10748139	✓			

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Guo, Q., Zhen, J., Wu, F., He, Y., & Qiao, C. (2025). Can students make STEM progress with large language models (LLMs)? An empirical study of LLMs integration within middle school science and engineering practice. *Journal of Educational Computing Research*, *63*(2), 372–405. https://doi.org/10.1177/07356331241312365			✓	
Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education, 100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505	✓			
Gupta, S., Dharamshi, R. R., & Kakde, V. (2024). An impactful and revolutionized educational ecosystem using generative AI to assist and assess the teaching and learning benefits, fostering the post-pandemic requirements. *Proceedings of the Second International Conference on Emerging Trends in Information Technology and Engineering, Vellore, India*, 1–4. https://doi.org/10.1109/ic-ETITE58242.2024.10493370			✓	
He, Z., Chai, S., Zhou, F., Wu, H., Lai, N., & Jiang, T. (2024). Enhancing computational thinking in knowledge building community: Analyzing ChatGPT's role and impact among undergraduates. *Proceedings of the 8th APSCE International Conference on Computational Thinking and STEM Education, Beijing*, 141-143. https://doi.org/10.6084/m9.figshare.26008282			✓	
Herdliska, A., & Zhai, X. (2023). Artificial intelligence-based scientific inquiry. In X. Zhai & J. Pellegrino (Eds.), *Uses of artificial intelligence in STEM education* (pp. 201–218). Oxford University Press. https://doi.org/10.1093/oso/9780198882077.003.0009			✓	
Hieu, H. H., & Thao, L. T. (2024). Exploring the impact of ai in language education: Vietnamese EFL teachers' views on using ChatGPT for fairy tale retelling tasks. *International Journal of Learning, Teaching and Educational Research, 23*(3), 486–503. https://doi.org/10.26803/ijlter.23.3.24	✓			
Hijón-Neira, R., Pizarro, C., Borrás-Gené, O., & Cavero, S. (2024). AI-generated context for teaching robotics to improve computational thinking in early childhood education. *Education Sciences, 14*(12), 1401. https://doi.org/10.3390/educsci14121401			✓	
Hong, J., & Kim, K. (2025). Impact of AIoT education program on digital and AI literacy of elementary school students. *Education and Information Technologies*, *30*(1), 107–130. https://doi.org/10.1007/s10639-024-12758-0			✓	
Huang, X. (2024). Research on the construction of English intelligent teaching mode in colleges and universities facilitated by artificial intelligence technology. *Applied Mathematics and Nonlinear Sciences, 9*(1), 1–11. https://doi.org/10.2478/amns-2024-0657	✓		✓	
Hughes, L., Malik, T., Dettmer, S., Al-Busaidi, A. S., & Dwivedi, Y. K. (2025). Reimagining higher education: Navigating the challenges of generative AI adoption. *Information Systems Frontiers*. Advance online publication. https://doi.org/10.1007/s10796-025-10582-6		✓		✓
Hussein, R., Zhang, Z., Amarante, P., Hancock, N., Orduna, P., & Rodriguez-Gil, L. (2024). Integrating personalized AI-Assisted instruction into remote laboratories: Enhancing engineering education with OpenAI's GPT models. *Proceedings of the IEEE Frontiers in Education Conference, Washington, DC, USA*, 1–7. https://doi.org/10.1109/FIE61694.2024.10892918	✓		✓	

Record reference	Critical thinking		Problem-solving	
	Positive	**Negative**	**Positive**	**Negative**
Hwang, G.-J., Cheng, P.-Y., & Chang, C.-Y. (2025). Facilitating students' critical thinking, metacognition and problem-solving tendencies in geriatric nursing class: A mixed-method study. *Nurse Education in Practice*, *83*, Article 104266. https://doi.org/10.1016/j.nepr.2025.104266	✓		✓	
Inderjeet, & Bhardwaj, R. (2024). AI-integrated model design for the education system for analytical computation (mathematics). *Proceedings of the 4th International Conference on Advance Computing and Innovative Technologies in Engineering, Greater Noida, India,* 1412–1417. https://doi.org/10.1109/ICACITE60783.2024.10616503	✓	✓	✓	
Jang, H., & Choi, H. (2025). A double-edged sword: Physics educators' perspectives on utilizing ChatGPT and its future in classrooms. *Journal of Science Education and Technology*, *34*(2), 267–283. https://doi.org/10.1007/s10956-024-10173-1		✓	✓	✓
Jang, J., Jeon, J., & Jung, S. K. (2022). Development of a STEM-based AI education program for sustainable improvement of elementary learners. *Sustainability, 14*, 15178. https://doi.org/10.3390/su142215178	✓		✓	
Jayasinghe, S. (2024). Promoting active learning with ChatGPT: A constructivist approach in Sri Lankan higher education. *Journal of Applied Learning and Teaching, 7*(2), 141–154. https://doi.org/10.37074/jalt.2024.7.2.26	✓		✓	
Ji, Y., Zhan, Z., Li, T., Zou, X., & Lyu, S. (2025). Human–machine cocreation: The effects of ChatGPT on students' learning performance, AI awareness, critical thinking, and cognitive load in a STEM course toward entrepreneurship. *IEEE Transactions on Learning Technologies, 18*, 402–415. https://doi.org/10.1109/TLT.2025.3554584		✓		
Jia, X.-H., & Tu, J.-C. (2024). Towards a new conceptual model of AI-enhanced learning for college students: The roles of artificial intelligence capabilities, general self-efficacy, learning motivation, and critical thinking awareness. *Systems, 12*(3), 74. https://doi.org/10.3390/systems12030074		✓		
John, A. (2025). Revolutionizing STEAM education: Harnessing the power of AI and digital technology to deliver personalized learning experiences. In N. D. Son (Ed.), *Integrating Personalized Learning Methods into STEAM Education* (pp. 143–168). IGI Global. https://doi.org/10.4018/979-8-3693-7718-5.ch007	✓			
Jošt, G., Taneski, V., & Karakatič, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences*, *14*(10), Article 4115. https://doi.org/10.3390/app14104115	✓	✓	✓	✓
Juáres, E. C., & Guzmán, D. S. (2024). Implementing the LSEESC methodology with AI-guided support for active learning in engineering physics courses. *Proceedings of the IEEE Frontiers in Education Conference, Washington, DC, USA*, 1–7. https://doi.org/10.1109/FIE61694.2024.10893032			✓	✓
Jung, D., & Suh, S. (2024). Enhancing soft skills through generative AI in sustainable fashion textile design education. *Sustainability, 16*, 6973. https://doi.org/10.3390/su16166973			✓	
Kakhki, M., Oguz, A., & Gendron, M. (2024). exploring the affordances of chatbots in higher education: A Framework for understanding and utilizing ChatGPT. *Journal of Information Systems Education, 35*(3), 284–302. https://doi.org/10.62273/UIRX9922	✓			

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Kalniņa, D., Nīmante, D., & Baranova, S. (2024). Artificial intelligence for higher education: Benefits and challenges for pre-service teachers. *Frontiers in Education, 9*, 1501819. https://doi.org/10.3389/feduc.2024.1501819	✓	✓	✓	✓
Kamoun, F., El Ayeb, W., Jabri, I., Sifi, S., & Iqbal, F. (2024). Exploring students' and faculty's knowledge, attitudes, and perceptions towards ChatGPT: A cross-sectional empirical study. *Journal of Information Technology Education: Research*, *23,* 4. https://doi.org/10.28945/5239	✓	✓	✓	✓
Kartal, G. (2024). The influence of ChatGPT on thinking skills and creativity of EFL student teachers: A narrative inquiry. *Journal of Education for Teaching*, *50*(4), 627–642. https://doi.org/10.1080/02607476.2024.2326502	✓	✓		
Khadragy, S., Selim, N., Hassan, D., Hegazy, A., & Ali, S. (2024). Exploring the effects of using AI technologies in higher education institutions in UAE. *Journal of Curriculum and Teaching, 13*(5), 10–21. https://doi.org/10.5430/jct.v13n5p10	✓		✓	
Kilz, L. (2024). Student perspective on the use of AI in higher education: Analysis of the students' podcast. *IDIMT-2024: Changes to ICT, Management, and Business Processes through AI, 53*, 133–140. https://doi.org/10.35011/IDIMT-2024-133		✓		✓
Kim, J., & Cho, Y. H. (2025). My teammate is AI: Understanding students' perceptions of student-AI collaboration in drawing tasks. *Asia Pacific Journal of Education, 45*(3), 1013–1027. https://doi.org/10.1080/02188791.2023.2286206	✓			
Kishore, S., Hong, Y., Nguyen, A., & Qutab, S. (2023). Should ChatGPT be banned at schools? Organizing visions for generative artificial intelligence (AI) in education. *Proceedings of the ICIS Conference* (pp. 1–17). https://aisel.aisnet.org/icis2023/learnandiscurricula/learnandiscurricula/5	✓	✓	✓	✓
Klimova, B., & De Campos, V. P. L. (2024). University undergraduates' perceptions on the use of ChatGPT for academic purposes: Evidence from a university in Czech Republic. *Cogent Education*, *11*(1), 2373512. https://doi.org/10.1080/2331186X.2024.2373512	✓	✓		✓
Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K. E., Lukowicz, P., Kuhn, J., Küchemann, S., & Karolus, J. (2024a). Challenges and opportunities of moderating usage of large language models in education. *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia* (pp. 249–254). Association for Computing Machinery. https://doi.org/10.1145/3701571.3701590	✓			✓
Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K. E., Lukowicz, P., Kuhn, J., Küchemann, S., & Karolus, J. (2024b). Unreflected acceptance: Investigating the negative consequences of ChatGPT-assisted problem solving in physics education. *Proceedings of the 33rd International Conference on Frontiers in Artificial Intelligence and Applications* (pp. 199–212). IOS Press. https://doi.org/10.3233/FAIA240195		✓		✓
Lai, C.-H., & Lin, C.-Y. (2025). Analysis of learning behaviors and outcomes for students with different knowledge levels: A case study of intelligent tutoring system for coding and learning (ITS-CAL). *Applied Sciences*, *15*(4). https://doi.org/10.3390/app15041922	✓	✓	✓	✓

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Lamanauskas, V., & Makarskaitė-Petkevičienė, R. (2025). Pre-service preschool and primary school teachers' position on artificial intelligence. In M. Turčáni (Ed.), *15th International Scientific Conference on Distance Learning in Applied Informatics* (pp. 183–197). Springer. https://doi.org/10.1007/978-3-031-81261-3_14	✓			✓
Lee, C., Myung, J., Han, J., Jin, J., & Oh, A. (2024). Learning from teaching assistants to formulate subgoals for programming tasks: Exploring the potential for AI teaching assistants. *Proceedings of the L3MNGET Workshop at NeurIPS 2024: CEUR Workshop Proceedings* (Vol. 3840, pp. 1–9). CEUR. https://ceur-ws.org/Vol-3840/L3MNGET24_paper5.pdf			✓	✓
Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, *21*(1), Article 16. https://doi.org/10.1186/s41239-024-00447-4	✓	✓	✓	✓
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	✓		✓	✓
Li, W., Ji, J.-W., Tseng, J. C. R., Liu, C.-Y., Huang, J.-Y., Liu, H.-Y., & Zhou, M. (2025). A ChatGPT-AIPSS approach for SDGs: Classroom engagement and collective efficacy of students with different levels of critical thinking. *Educational Technology & Society, 28*(1), 301–318.	✓			
Liao, J., Zhong, L., Zhe, L., Xu, H., Liu, M., & Xie, T. (2024). Scaffolding computational thinking with ChatGPT. *IEEE Transactions on Learning Technologies*, *17*, 1628–1642. https://doi.org/10.1109/TLT.2024.3392896	✓			✓
Lijie, H., Mat Yusoff, S., & Mohamad Marzaini, A. F. (2025). Influence of AI-driven educational tools on critical thinking dispositions among university students in Malaysia: A study of key factors and correlations. *Education and Information Technologies*, *30*(6), 8029–8053. https://doi.org/10.1007/s10639-024-13150-8	✓	✓		
Lim, J., Lee, U., Koh, J., Jeong, Y., Lee, Y., Byun, G., Jung, H., Jang, Y., Lee, S., & Moon, J. (2025). Development and implementation of a generative artificial intelligence-enhanced simulation to enhance problem-solving skills for pre-service teachers. *Computers & Education*, *232,* Article 105306. https://doi.org/10.1016/j.compedu.2025.105306			✓	
Lin, C.-H., Zhou, K., Li, L., & Sun, L. (2025). Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms. *Computers and Composition*, *75*, Article 102895. https://doi.org/10.1016/j.compcom.2024.102895	✓	✓		
Liu, W., & Cui, X. (2024). Generative AI-assisted collaborative argumentation: Implications for the argumentation process and outcome. *Proceedings of the 4th International Conference on Educational Technology, Wuhan, China,* 220–225. https://doi.org/10.1109/ICET62460.2024.10868171	✓			
Liu, W., & Wang, Y. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*, *59*(4), Article e12804. https://doi.org/10.1111/ejed.12804	✓			

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Mahrous, A., Botsko, D. L., Elgreatly, A., Tsujimoto, A., Qian, F., & Schneider, G. B. (2023). The use of artificial intelligence and game-based learning in removable partial denture design: A comparative study. *Journal of Dental Education, 87*(8), 1188–1199. https://doi.org/10.1002/jdd.13225	✓			
Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open, 5*, 100296. https://doi.org/10.1016/j.ijedro.2023.100296	✓	✓		
Maqbali, S. B. A., & Tahir, N. M. (2024). AI Opportunities and risks for students' decision-making skill. *Proceedings of the IEEE 14th International Conference on Control System, Computing and Engineering, Penang, Malaysia,* 321–326. https://doi.org/10.1109/ICCSCE61582.2024.10696240		✓		
Markos, A., Prentzas, J., & Sidiropoulou, M. (2024). Pre-service teachers' assessment of ChatGPT's utility in higher education: SWOT and content analysis. *Electronics, 13*(10), 1985. https://doi.org/10.3390/electronics13101985		✓		
Maya, H. R. R., & Valdes, A. B. S. (2024). Empowering critical thinking through AI: The PAIR model's impact on higher education excellence. *Proceedings of the 8th International Symposium on Multidisciplinary Studies and Innovative Technologies, Ankara, Turkiye,* 1–6. https://doi.org/10.1109/ISMSIT63511.2024.10757186	✓		✓	
Michalon, B., & Camacho-Zuñiga, C. (2023). ChatGPT, a brand-new tool to strengthen timeless competencies. *Frontiers in Education, 8*, 1251163. https://doi.org/10.3389/feduc.2023.1251163	✓			
Mike, N., Karsai, K., Orbán, G., Bubelényi, A., Nagy, C. N., & Polyák, G. (2024). The impact of ChatGPT on student performance in higher education. In J. Ubacht, J. Crompvoets, C. Csáki, L. Danneels, M. Janssen, M. R. Johannessen, T. Lampoltshammer, H. Lee, I. Lindgren, A. Nikiforova, S. Hofmann, P. Parycek, G. V. Pereira, G. Schwabe, I. Susha, & E. Tambouris (Eds.), *Proceedings of Ongoing Research, Practitioners, Posters, Workshops, and Projects of the International Conference EGOV-CeDEM-ePart 2024* (Vol. 3737). CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3737/	✓	✓	✓	✓
Mohamed, A. M., Shaaban, T. S., Bakry, S. H., Guillén-Gámez, F. D., & Strzelecki, A. (2025). Empowering the faculty of education students: Applying AI's potential for motivating and enhancing learning. *Innovative Higher Education*, *50*(2), 587–609. https://doi.org/10.1007/s10755-024-09747-z	✓		✓	
Moskovich, L., & Rozani, V. (2025). Health profession students' perceptions of ChatGPT in healthcare and education: Insights from a mixed-methods study. *BMC Medical Education*, *25*(1), 98. https://doi.org/10.1186/s12909-025-06702-0		✓		✓
Del Mundo, M. A., Delos Reyes, E. F., Gervacio, E. M., Manalo, R. B., Book, R. J. B., Chavez, J. V., Espartero, M. M., & Sayadi, D. S. (2024). Discourse analysis on experience-based position of science, mathematics, and Tech-Voc educators on generative AI and academic integrity. *Environment and Social Psychology*, *9*(8), 1–16. https://doi.org/10.59429/esp.v9i8.3028		✓	✓	

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Naamati-Schneider, L., & Alt, D. (2024). Beyond digital literacy: The era of AI-powered assistants and evolving user skills. *Education and Information Technologies*, *29*(16), 21263–21293. https://doi.org/10.1007/s10639-024-12694-z	✓			✓
Nazri, I. S. B. M., Rodzi, Z. M., Razali, I. N. B., Rahman, H. A., Rahman, A. B. A., & Al-Sharqi, F. (2023). Unraveling the factors influencing the adoption of artificial intelligence (AI) in education. *Proceedings of the 4th International Conference on Artificial Intelligence and Data Sciences, Ipoh, Malaysia,* 186–193. https://doi.org/10.1109/AiDAS60501.2023.10284698	✓		✓	
Ng, P. H. F., Chen, P. Q., Wu, A. C. H., Tai, K. S. K., & Li, C. (2024). Reimagining STEM learning: A comparative analysis of traditional and service learning approaches for social entrepreneurship. *IEEE Transactions on Learning Technologies*, *17*, 2212–2226. https://doi.org/10.1109/TLT.2024.3492352			✓	
Nusivera, E., Hikmat, A., & Ghani, A. R. A. (2025). Integration of ChatGPT usage in language learning model to improve argumentation skills, complex comprehension skills, and critical thinking skills. *International Journal of Learning, Teaching and Educational Research*, *24*(2), 1–16. https://doi.org/10.26803/ijlter.24.2.19	✓	✓		
Okada, A., Sherborne, T., Panselinas, G., & Kolionis, G. (2025). Fostering transversal skills through open schooling supported by the CARE-KNOW-DO pedagogical model and the UNESCO AI competencies framework. *International Journal of Artificial Intelligence in Education*, *35*(4), 1953–1998. https://doi.org/10.1007/s40593-025-00458-w	✓		✓	
Palacios, D., & Rivadeneira, L. (2025). Influence of ChatGPT on programming code generation: A case study of the Technical University of Manabí. In T. Guarda, F. Portela, & G. Gatica (Eds.), *Advanced research in technologies, information, innovation and sustainability* (pp. 276–289). Springer. https://doi.org/10.1007/978-3-031-83210-9_21	✓		✓	✓
Pan, A.-J., Huang, Y.-C., & Lai, C.-F. (2024). Constructing hands-on distance labs: The development and implementation of an intelligent learning management system (ILMS-d) in undergraduate IoT courses. *Interactive Learning Environments*, *32*(10), 6413–6429. https://doi.org/10.1080/10494820.2023.2263061			✓	
Pantazatos, D., Grammatikou, M., & Maglaris, V. (2024). Enhancing soft skills in network management education: A study on the impact of GenAI-based virtual assistants. *Proceedings of the IEEE Global Engineering Education Conference, Kos Island, Greece,* 1–5. https://doi.org/10.1109/EDUCON60312.2024.10578597	✓		✓	
Parveen, M., & Alkudsi, Y. (2024). Graduates' perspectives on AI integration: Implications for skill development and career readiness. *International Journal of Educational Research and Innovation*, *22*, 1–17. https://doi.org/10.46661/ijeri.10651	✓		✓	
Pisica, A. I., Popa, A., Ioan, R., Zaharia, R. M., & Bucur, L.-M. (2024). Romanian students' opinions on implementing artificial intelligence in higher education: A qualitative approach. *Transformations in Business & Economics,* *23*(2), 21–35.		✓		

Record reference	Critical thinking		Problem-solving	
	Positive	**Negative**	**Positive**	**Negative**
Pitso, T. (2023). Post-COVID-19 higher learning: Towards telagogy, a web-based learning experience. *IAFOR Journal of Education*, *11*(2), 39–59. https://doi.org/10.22492/ije.11.2.02			✓	✓
Puig-Ortiz, J., Pamies-Vila, R., & Nebot, L. J. (2023). Exploring the application of ChatGPT in mechanical engineering education. *Proceedings of the 51st Annual Conference of the European Society for Engineering Education.* https://doi.org/10.21427/PP64-KG63			✓	
Qawqzeh, Y. (2024). Exploring the influence of student interaction with ChatGPT on critical thinking, problem solving, and creativity. *International Journal of Information and Education Technology*, *14*(4), 596–601. https://doi.org/10.18178/ijiet.2024.14.4.2082	✓		✓	
Razali, S. (2024). Lecturers' and students' perspectives on using ChatGPT in academics for creative problem solving: A dilemma or an opportunity for improvement*? Proceedings of the 5th International Conference on Artificial Intelligence and Data Sciences, Bangkok, Thailand,* 30–34. https://doi.org/10.1109/AiDAS63860.2024.10730665		✓		✓
Robledo-Rella, V., & Toh, B.-Y. (2024). Artificial intelligence in physics courses to support active learning. *Proceedings of the 2024 10th International Conference on E-Society, e-Learning and e-Technologies* (pp. 68–75). Association for Computing Machinery. https://doi.org/10.1145/3678610.3678631			✓	✓
Rodríguez-Rivera, P., Rodríguez-Ferrer, J. M., & Manzano-León, A. (2025). Designing digital escape rooms with generative AI in university contexts: A qualitative study. *Multimodal Technologies and Interaction*, *9*(3), Article 20. https://doi.org/10.3390/mti9030020			✓	
Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability, 16*(13), 5367. https://doi.org/10.3390/su16135367	✓	✓	✓	
Sachyani, D., & Gal, A. (2025). artificial intelligence tools in environmental education: Facilitating creative learning about complex interaction in nature. *European Journal of Educational Research, 14*(2), 395–413. https://doi.org/10.12973/eu-jer.14.2.395	✓	✓		
Said, N. A., Osman, N. A. Q., Mohamed, T. I., & Al-Said, K. (2025). The role of artificial intelligence in enhancing problem-solving skills: The mediating role of adaptive learning environments. *International Journal of Innovative Research and Scientific Studies*, *8*(2), Article 2. https://doi.org/10.53894/ijirss.v8i2.5961			✓	
Sam, K.-L., Ieong, U.-S., & da Costa, M. G. (2024). PeerSynergy: An innovative application applying large language models to peer instruction. *Proceedings of the IEEE International Conference on Teaching, Assessment and Learning for Engineering, Bengaluru, India,* 1–7. https://doi.org/10.1109/TALE62452.2024.10834348	✓		✓	
Santamaria-Velasco, J., Núñez-Naranjo, A., & Morales-Urrutia, X. (2025). Critical thinking and AI: Enhancing history teaching through ChatGPT simulations. *International Journal of Innovative Research and Scientific Studies, 8*(1), 564-575. https://doi.org/10.53894/ijirss.v8i1.4403	✓			

Record reference	**Critical thinking**		**Problem-solving**	
	Positive	**Negative**	**Positive**	**Negative**
Sarwanti, S., Sariasih, Y., Rahmatika, L., Islam, M. M., & Riantina, E. M. (2024). Are they literate on ChatGPT? University language students' perceptions, benefits and challenges in higher education learning. *Online Learning, 28*(3), 105–130. https://doi.org/10.24059/olj.v28i3.4599		✓		
Shalong, W., Zuo, Y., Zou, B., Liu, G., Zhou, J., Zheng, Y., Zhang, Z., Yuan, L., & Feng, R. (2025). Enhancing self-directed learning with custom GPT AI facilitation among medical students: A randomized controlled trial. *Medical Teacher, 47*(7), 1126–1133. https://doi.org/10.1080/0142159X.2024.2413023	✓			
Sharma, D. (2024). Critical thinking and problem-solving in the age of ChatGPT: Experiential-bibliotherapy-blogging project. *Business and Professional Communication Quarterly*, *87*(4), 630–653. https://doi.org/10.1177/23294906241254780	✓			
Silitonga, L. M., Wiyaka, Suciati, S., & Prastikawati, E. F. (2024). The impact of integrating AI chatbots and microlearning into flipped classrooms: Enhancing students' motivation and higher-order thinking skills. In Y.-P. Cheng, M. Pedaste, E. Bardone, & Y.-M. Huang (Eds.), *Innovative Technologies and Learning* (pp. 184–193). Springer. https://doi.org/10.1007/978-3-031-65884-6_19	✓		✓	
Slimi, Z., Benayoune, A., & Alemu, A. E. (2025). Students' perceptions of artificial intelligence integration in higher education. *European Journal of Educational Research*, *14*(2), 471–484. https://doi.org/10.12973/eu-jer.14.2.471	✓		✓	
Stampfl, R., Geyer, B., Deissl-O'Meara, M., & Ivkic, I. (2024). Revolutionising role-playing games with ChatGPT. *Advances in Artificial Intelligence and Machine Learning*, *4*(2), 2244–2257. https://doi.org/10.54364/AAIML.2024.42129	✓		✓	
Styve, A., Virkki, O. T., & Naeem, U. (2024). Developing critical thinking practices interwoven with generative AI usage in an introductory programming course. *Proceedings of the IEEE Global Engineering Education Conference, Kos Island, Greece,* 1–8. https://doi.org/10.1109/EDUCON60312.2024.10578746	✓	✓		
Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction*, *95*, Article 102011. https://doi.org/10.1016/j.learninstruc.2024.102011	✓	✓	✓	
Sykes, E. R. (2024). AI-assisted academic writing: A comparative study of student-crafted and ChatGPT-enhanced critiques in ubiquitous computing. *International Journal of Advanced Computer Science and Applications, 15*(9), 43–53. https://doi.org/10.14569/IJACSA.2024.0150905	✓	✓		
Šedlbauer, J., Činčera, J., Slavík, M., & Hartlová, A. (2024). Students' reflections on their experience with ChatGPT. *Journal of Computer Assisted Learning, 40*(4), 1526–1534. https://doi.org/10.1111/jcal.12967	✓	✓		
Takerngsaksiri, W., Warusavitarne, C., Yaacoub, C., Hou, M. H. K., & Tantithamthavorn, C. (2024). Students' perspectives on AI code completion: Benefits and challenges. *Proceedings of the IEEE Computer Software and Applications Conference, Osaka, Japan,* 1606–1611. https://doi.org/10.1109/COMPSAC61105.2024.00252				✓

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Tlais, S., Alkhatib, A., Hamdan, R., HajjHussein, H., Hallal, K., & Malti, W. E. (2025). Artificial intelligence in higher education: Early perspectives from Lebanese STEM faculty. *TechTrends, 69*, 598–606. https://doi.org/10.1007/s11528-025-01061-5		✓		
Torres-Peña, R. C., Peña-González, D., Chacuto-López, E., Edwan, A. A., & Vergara, D. (2024). Updating calculus teaching with AI: A classroom experience. *Education Sciences, 14*(9), 1019. https://doi.org/10.3390/educsci14091019			✓	✓
Toscano, R., Guerra, M. A., Durán-Ballén, S., & Valarezo, B. M. (2024). WIP - Development of critical thinking in AEC students aided by artificial intelligence. *Proceedings of the IEEE Frontiers in Education Conference, Washington, DC, USA,* 1–6. https://doi.org/10.1109/FIE61694.2024.10893092	✓			
Tseng, Y.-C., & Lin, Y.-H. (2024). Enhancing English as a foreign language (EFL) learners' writing with ChatGPT: A university-level course design. *Electronic Journal of E-Learning, 22*(2), 78–97. https://doi.org/10.34190/ejel.21.5.3329	✓			
Uwosomah, E. E., & Dooly, M. (2025). It is not the huge enemy: Preservice teachers' evolving perspectives on AI. *Education Sciences*, *15*(2), Article 2. https://doi.org/10.3390/educsci15020152		✓		✓
Vasconcelos, M. A. R., & Santos, R. P. dos. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), em2296. https://doi.org/10.29333/ejmste/13313	✓		✓	
Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, *15*(3), Article 343. https://doi.org/10.3390/educsci15030343	✓	✓		✓
Villarino, R. T. (2025). Artificial intelligence (AI) integration in rural Philippine higher education: Perspectives, challenges, and ethical considerations. *International Journal of Educational Research and Innovation,* 23. https://doi.org/10.46661/ijeri.10909		✓		✓
Waluyo, B., & Kusumastuti, S. (2024). Generative AI in student English learning in Thai higher education: More engagement, better outcomes? *Social Sciences & Humanities Open, 10*, 101146. https://doi.org/10.1016/j.ssaho.2024.101146	✓	✓		
Wang, M., Sun, Z., Jia, M., Wang, Y., Wang, H., Zhu, X., Chen, L., & Ji, H. (2022). Intelligent virtual case learning system based on real medical records and natural language processing. *BMC Medical Informatics and Decision Making, 22*, Article 60. https://doi.org/10.1186/s12911-022-01797-7	✓	✓		
Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, Article 104703. https://doi.org/10.1016/j.compedu.2022.104703			✓	✓
Weng, J. (2024). Critical thinking with AI: Navigating ChatGPT in engineering education. *Proceedings of the 52nd Annual Conference of SEFI, Lausanne, Switzerland.* https://doi.org/10.5281/zenodo.14256737	✓			

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities, 11*(1), 2392388. https://doi.org/10.1080/23311983.2024.2392388		✓		
Woerner, J. H. R., Turtova, A. P., & Lang, A. S. I. D. (2024). Transformative potentials and ethical considerations of AI tools in higher education: Case studies and reflections. *Proceedings of the SoutheastCon, Atlanta, GA, USA*, 510–515. https://doi.org/10.1109/SoutheastCon52093.2024.10500042	✓	✓		
Wu, T.-T., Lee, H.-Y., Chen, P.-H., Lin, C.-J., & Huang, Y.-M. (2025). Integrating peer assessment cycle into ChatGPT for STEM education: A randomised controlled trial on knowledge, skills, and attitudes enhancement. *Journal of Computer Assisted Learning*, *41*(1), Article e13085. https://doi.org/10.1111/jcal.13085	✓	✓	✓	✓
Xiao, J., Bozkurt, A., Nichols, M., Pazurek, A., Stracke, C. M., Bai, J. Y. H., Farrow, R., Mulligan, D., Nerantzi, C., Sharma, R. C., Singh, L., Frumin, I., Swindell, A., Honeychurch, S., Bond, M., Dron, J., Moore, S., Leng, J., van Tryon, P. J. S., … Themeli, C. (2025). Venturing into the unknown: Critical insights into grey areas and pioneering future directions in educational generative AI research. *TechTrends, 69*, 582–597. https://doi.org/10.1007/s11528-025-01060-6		✓		
Xu, J., & Liu, Q. (2025). Uncurtaining windows of motivation, enjoyment, critical thinking, and autonomy in AI-integrated education: Duolingo Vs. ChatGPT. *Learning and Motivation, 89*, 102100. https://doi.org/10.1016/j.lmot.2025.102100	✓			
Xu, X. S., Wang, X. B., Zhang, Y. F., & Zheng, R. (2024). Applying ChatGPT to tackle the side effects of personal learning environments from learner and learning perspective: An interview of experts in higher education. *PLoS ONE*, *19*(1), e0295646. https://doi.org/10.1371/journal.pone.0295646	✓		✓	
Xu, Y., Zhu, J., Wang, M., Qian, F., Yang, Y., & Zhang, J. (2024). The impact of a digital game-based AI chatbot on students' academic performance, higher-order thinking, and behavioral patterns in an information technology curriculum. *Applied Sciences*, *14*(15), 6418. https://doi.org/10.3390/app14156418				✓
Yang, A. C. M., Lin, J.-Y., Lin, C.-Y., & Ogata, H. (2024). Enhancing Python learning with PyTutor: Efficacy of a ChatGPT-based intelligent tutoring system in programming education. *Computers and Education: Artificial Intelligence*, *7*, 100309. https://doi.org/10.1016/j.caeai.2024.100309				✓
Yang, C., Wei, M., & Liu, Q. (2025). Intersections between cognitive-emotion regulation, critical thinking and academic resilience with academic motivation and autonomy in EFL learners: Contributions of AI-mediated learning environments. *British Educational Research Journal.* Advance online publication. https://doi.org/10.1002/berj.4140	✓			
Yang, Q.-F., Lian, L.-W., & Zhao, J.-H. (2023). Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students. *International Journal of Educational Technology in Higher Education*, *20*(1), 18. https://doi.org/10.1186/s41239-023-00391-9			✓	

Record reference	Critical thinking		Problem-solving	
	Positive	Negative	Positive	Negative
Ye, W., Li, S., Liu, S., & Zhou, Y. (2022). Application of artificial intelligence technology in martial arts education governance. *Discrete Dynamics in Nature and Society*, *2022*(1), 5606280. https://doi.org/10.1155/2022/5606280			✓	
Youssef, E., Medhat, M., Abdellatif, S., & Al Malek, M. (2024). Examining the effect of ChatGPT usage on students' academic learning and achievement: A survey-based study in Ajman, UAE. *Computers and Education: Artificial Intelligence*, *7*, Article 100316. https://doi.org/10.1016/j.caeai.2024.100316	✓	✓		
Zeng, C. (2024). Innovation and implementation of the teaching mode of a higher vocational advertising planning course in the context of AI smart marketing. *Applied Mathematics and Nonlinear Sciences*, *9*(1). https://doi.org/10.2478/amns-2024-1679	✓		✓	
Zhang, H., & Li, M. (2024). A study on the impact of junior high school students' programming learning effect based on generative artificial intelligence. *Proceedings of the 4th International Conference on Educational Technology, Wuhan, China*, 106–109. https://doi.org/10.1109/ICET62460.2024.10869172			✓	
Zhang, X. (2025). Incorporating AI tools into database classes. *Journal of Information Systems Education, 36*(1), 37–52. https://doi.org/10.62273/GKZI2477	✓		✓	
Zheng, K., Shen, Z., Chen, Z., Che, C., & Zhu, H. (2024). Application of AI-empowered scenario-based simulation teaching mode in cardiovascular disease education. *BMC Medical Education*, *24*, Article 1003. https://doi.org/10.1186/s12909-024-05977-z	✓		✓	
Zhou, X., Teng, D., & Al-Samarraie, H. (2024). The mediating role of generative AI self-regulation on students' critical thinking and problem solving. *Education Sciences*, *14*(12), Article 1302. https://doi.org/10.3390/educsci14121302	✓	✓	✓	
Zhu, G., Sudarshan, V., Kow, J. F., & Soon Ong, Y. (2024). Human-generative AI collaborative problem solving: Who leads and how students perceive the interactions. *Proceedings of the IEEE Conference on Artificial Intelligence, Singapore,* 680–686. https://doi.org/10.1109/CAI59869.2024.00133			✓	✓

Authors

Tomáš Bendl (0000-0003-1533-097X) is an Assistant Professor at the Faculty of Science, Humanities and Education of the Technical University of Liberec (FP TUL), where he focuses on geography education and education as both his professional specialization and personal passion. His research centers on the reconceptualization of higher-order thinking skills, the impact of artificial intelligence on pupils' and students' thinking, and discipline-specific thinking, particularly geographical thinking. He regularly publishes in high-quality, peer-reviewed education journals and actively contributes to international academic discourse.

He is also involved in selected geography education projects at the Faculty of Science of Charles University and serves in several educational bodies and initiatives, including as co-leader of IGU CGE Special Interest Group on Thinking Development in (Geography) Education.

Lenka Krajňáková (0000-0002-6266-4801) is an Assistant Professor at the Faculty of Science, Charles University, Prague, Czechia. Her research focuses on cartographic education, particularly students' misconceptions and the enhancement of map-based learning. In recent research projects, she has explored the role of generative artificial intelligence and other digital technologies in geography education. She co-leads the IGU CGE Special Interest Group on AI and Technologies.

In addition to her research activities, she is involved in the Czech Geography Olympiad. She teaches courses for pre-service geography teachers, including cartography, field-based geography education, and methodological seminars for bachelor's and master's theses.

Hilde Storrøsæter (0009-0008-0642-5970) is an Assistant Professor at the Norwegian University of Science and Technology (NTNU), where she also serves as Head of Practicum and coordinator for the third year of the Master's program in Primary and Lower Secondary Teacher Education (years 5–10). Her work is strongly grounded in education, particularly in geography education research, with a focus on didactics. She focuses on approaches that foster mastery, curiosity, exploration, and active learning, while also engaging with multiple perspectives in the classroom. A key area of her interest is the use of AI in education and sustainability education, where she emphasizes a balanced, non-indoctrinating approach that integrates environmental, social, and economic dimensions. In addition to her research-oriented work, she contributes to the ongoing development and reorganization of practicum within teacher education at NTNU, linking academic research with practical training in schools.

Emil Drápela (0000-0002-6537-3962) is an Assistant Professor at the Technical University of Liberec whose research focuses on the regional development of peripheral and border areas, with a long-term specialization in cross-border cooperation and the socio-economic transformation of border regions. His work further examines geography education, tourism, and transport as key drivers of spatial development in these territories. Methodologically, he combines quantitative and qualitative approaches, works with diverse data sources, and applies advanced spatial and GIS-based analytical methods. In addition to his scientific activities, he is actively engaged in popularising geographical topics through public lectures, school educational programs, and the development of educational games and competitions.

Martin Hanus (ORCID: 0000-0002-9873-5106) is an Associate Professor at the Faculty of Science, Charles University, Prague, Czechia. His research focuses on geography education, with particular attention to curriculum, teaching strategies, spatial thinking, map-based learning, GeoCapabilities, world-mindedness, and the integration of digital technologies into teaching and learning. Across his work, he examines how learners develop geographical understanding, solve problems, interpret representations, and engage critically with information in educational contexts. His current interests include the educational potential and challenges of generative artificial intelligence, particularly regarding problem-solving, critical thinking, teacher education, and the evolving role of technology in geography education. Martin teaches preservice teacher-training programs and professional development courses for in-service teachers. He is also active in international and professional service, including as co-chair of the Steering Committee of the IGU Commission on Geographical Education for the 2024–2028 term.