https://doi.org/10.17323/jle.2025. 27387

Enhancing Critical Thinking Skills in ChatGPT-Human Interaction: A Scoping Review

Lilia Raitskaya [1], **Elena Tikhonova** [2, 3]

[1] MGIMO University, Moscow, Russia
[2] HSE University, Moscow, Russia
[3] RUDN University, Moscow, Russia

ABSTRACT

Introduction: The rapid integration of generative artificial intelligence (GenAI) technologies, including ChatGPT, into educational environments has introduced both opportunities and challenges for learners and educators. While GenAI can support advanced learning practices, it also raises concerns about critical engagement and the accuracy of generated content. Previous systematic reviews have explored GenAI's relationship with critical thinking (CT) and self-regulated learning, but a focused synthesis of recent empirical evidence on GenAI's impact on university students' CT skills remains lacking.

Method: This scoping review followed the PRISMA-ScR guidelines and applied the Arksey and O'Malley framework alongside the Population – Concept – Context (PCC) model. Studies were identified via the Scopus database, using inclusion criteria limited to the years 2024–2025, English language, and the Social Sciences subject area. Thirty eligible empirical studies were analysed and visualised using VOSviewer to identify thematic clusters and categories in the literature.

Results: The reviewed studies were grouped into seventeen thematic clusters by the VOSviewer and then manually synthesized into six categories based on semantic interpretation: cognitive and metacognitive development, pedagogical innovation and learning design, academic writing and language learning, AI literacy and learner perception, evaluation and assessment technologies, global and ethical dimensions of GenAI use. The findings were analysed as (1) direct enhancement of CT, (2) metacognitive and reflective gains, (3) contextual factors shaping CT, (4) risks of cognitive offloading, and (5) instructional strategies mediating AI's effect. 21 publications showed predominantly positive impact of GenAI on CT (idea generation, conceptual understanding, construction of arguments, literature review, academic writing, etc.) whereas reported found mixed impact.

Conclusion: The review concludes that GenAI holds substantial potential to support CT development, particularly when pedagogically integrated to promote active reasoning, metacognitive monitoring, and critical autonomy. However, the evidence base is still emerging and is limited by its short temporal scope, narrow database coverage, and reliance on self-reported data. Future research should focus on long-term effects, discipline-specific instructional models, and robust theoretical frameworks linking AI use to cognitive outcomes.

KEYWORDS

development of critical thinking skills, generative AI impact on higher education, ChatGPT and student cognition, AI-supported academic writing, metacognitive engagement with GenAI, GenAI and argumentation, prompt engineering and critical thinking

Citation: Raitskaya, L., & Tikhonova, E. (2025). Enhancing critical thinking skills in chatgpt-human interaction: A scoping review. *Journal of Language and Education, 11*(2), 5-19. https://doi.org/10.17323/jle.2025.27387

Correspondence:
Elena Tikhonova,
etihonova@hse.ru

Received: May 23, 2025
Accepted: June 10, 2025
Published: June 30, 2025

INTRODUCTION

The recent advancements in generative artificial intelligence (GenAI), particularly the release of ChatGPT 3.5, have catalyzed a surge of interest in the educational potential and cognitive implications of human-AI interaction (Fabio et al., 2025). As educational institutions across the globe begin integrating GenAI technologies into teaching and learning processes, researchers have turned to examining how these tools shape students' skills and learning behaviours. Among these skills, critical thinking (CT) has received considerable attention, as it is widely regarded as both a prerequisite for and an outcome of meaningful engagement with AI (Tikhonova & Raitskaya, 2023; Jiang et al., 2024).

Critical thinking plays a central role in the digital age, not only as a measure of academic development but also as a safeguard against the epistemic risks posed by algorithmically generated content. When students engage in tasks such as refining prompts, interpreting AI-generated responses, verifying information accuracy, or reflecting on the ethical implications of machine-produced language, they are actively employing critical thinking strategies (Babin et al., 2024; Gonsalves, 2024). As Darwin et al. (2024) note, CT becomes a mediating function in student-AI interaction, supporting both metacognitive awareness and epistemological vigilance.

While the concept of critical thinking is well-established, with over 36,000 documents indexed in Scopus as of May 2025 and roots stretching back to the early 20th century, the arrival of GenAI technologies has redefined its operationalization in education. Since the 1990s, interest in CT has grown steadily, with significant acceleration in the 2000s. The foundational framework for CT in educational research remains Bloom's taxonomy and its revisions, which distinguish between levels of cognitive complexity, from basic recall to synthesis and creative problem-solving (Krathwohl, 2002).

The proliferation of GenAI tools challenges educators and researchers to reconsider how critical thinking is taught, assessed, and applied. On one hand, GenAI offers powerful means of enhancing CT through interactive, scaffolded, and personalized learning environments. On the other, there are growing concerns about cognitive offloading, information naivety, and uncritical dependence on AI systems (Risko & Gilbert, 2016; Gerlich, 2025). These tensions underscore the need for a more precise understanding of how GenAI influences different dimensions of critical thinking, especially in higher education contexts where autonomy, reflection, and analytical skills are paramount.

To date, only a limited number of systematic reviews have addressed the intersection of GenAI and critical thinking in higher education. Notably, Melisa et al. (2025) explored how ChatGPT use correlates with students' ability to evaluate information and develop independent judgments, while Sardi et al. (2025) focused on the interplay between CT and self-regulated learning in AI-supported environments. Both reviews emphasized generally positive or neutral outcomes, though they also pointed to a lack of consensus regarding the mechanisms through which GenAI supports or inhibits higher-order thinking.

This scoping review builds on that emerging body of scholarship by focusing specifically on empirical studies published between 2024 and 2025, a period that coincides with a marked expansion in the use of GenAI in academic settings. By concentrating on recent peer-reviewed research indexed in the Scopus database, the review aims to synthesize current knowledge on how generative AI influences the development of critical thinking skills among university students. Unlike earlier reviews that cast a broad net, the present study narrows its focus to uncover the patterns, strategies, and pedagogical conditions under which CT is most effectively cultivated through GenAI interaction.

Accordingly, this review addresses the following research questions:

RQ1. What are the key directions of empirical research on generative AI impact on critical thinking?
RQ2. How are critical thinking skills influenced in human-ChatGPT interaction?

By answering these questions, the review seeks to provide educators, policymakers, and researchers with a more nuanced understanding of the opportunities and limitations associated with GenAI in higher education. It also aims to identify promising avenues for further inquiry and to support the development of evidence-based pedagogical frameworks that foreground critical thinking in the age of artificial intelligence.

METHOD

Protocol

Getting down to the present scoping review, we meticulously developed a research protocol. The reviewers hereby certify that this scoping review report constitutes a faithful, precise, and transparent description of the conducted review. No deviations from the protocol were registered. Any departures from the original study design have been duly elucidated. This scoping review stick to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for Scoping Reviews (Tricco et al., 2018) and the framework proposed by Arksey and O'Malley (2005).

Eligibility Criteria

In this review, the population, concept, and context (PCC) framework was applied to devise an effective search strategy where each criterion was justified (Table 1).

Information Sources and Search Strategies

The Scopus database was searched on May 25, 2025. The search string used was: TITLE-ABS-KEY("critical thinking") AND TITLE-ABS-KEY("ChatGPT" OR "generative AI" OR "generative artificial intelligence"). Only documents published in

Table 1

Eligibility Criteria

Criterion	Inclusion	Exclusion	Rationale
Population	Students, EFL learners and teachers in the higher education institutions	Other users of GenAI and population at other educational levels	The introduction of generative AI into higher education requires research of the challenges and barriers that may be overcome via fostering a complex of competencies, including higher order thinkings skills at that educational level
Concept	Critical thinking, including Generative AI impact on critical thinking	Other concepts	The aim of the review is to identify the scope and recent trends of research on critical thinking in the context of appliances of GenAI
Context	Higher education	Other contexts	The review dwells upon studies in the higher education environment
Language	English	Other languages	The object of all research in focus is scholarly publications in English. The language choice is also identified by its status as a lingua franca of international science.
Time period	2024-2025	Publications before 2024	The period is selected due to the breakthrough in the generative AI, starting from ChatGPT 3.5. Only the recent publications were considered
Types of sources	In the Scopus database: full texts of articles and book chapters	Unavailable sources, unavailable full texts	This review aims to get a comprehensive understanding of the field
Geographical location	Any location	N/A	Getting international perspective
Database	Scopus	Other bases	Scopus was selected as a comprehensive database indexing top-ranking publications on higher education and technological innovations in the society
Areas of Research	Social Sciences	Other areas	The research in Social Sciences were chosen as the context is limited to higher education

2024 or 2025 were included. Pre-protocol test searches were conducted to assess alternative keywords, but no additional relevant studies were identified.

Document Identification and Screening

Both reviewers independently identified documents (empirical articles and book chapters) subject to the eligibility criteria enumerated in Table 1. The searches base in the Scopus data brought 566 titles. After the Scopus filters (research field, language and types of sources) had been applied, the total decreased to 290. Each author independently screened the titles and abstracts of the 290 documents. Only empirical research papers were selected. Special focus was made on the context (higher education). 207 documents were subsequently eliminated after the screening. Full texts of 36 documents out of 83 were retrieved from the journal sites and the publishers' sites for further screening. Each full text was downloaded, thoroughly read and independently analysed by each author. After they had been thoroughly analysed, six documents were found irrelevant as they were not based on empirical research, or the research was placed in a school environment. When occasional disagreements arose, they were settled by mutual consent. Thirty documents were ultimately included into the review (Figure 1).

Data Charting

A data-charting form was cooperatively developed. Both authors independently charted the data extracted from five identified documents as a pilot calibration. The data-charting form was discussed in an iterative process. The ultimate data included in the form are enumerated in Table 2. All the extracted data were double-checked by the authors.

Data Analysis

Following the data charting process, we employed a combination of descriptive synthesis and thematic analysis to interpret the extracted information and answer the research questions. The analysis was structured in alignment with the review's two guiding questions: (1) the identification of key directions in empirical research on the impact of generative AI (GenAI) on critical thinking; and (2) the examination of how specific aspects of human - ChatGPT interaction influence the development of critical thinking skills among university students.

Initially, we conducted within-case analysis of each study, focusing on contextual variables such as research design, population, mode of GenAI use, and specific indicators of

Figure 1

Selection of Publications for the Review

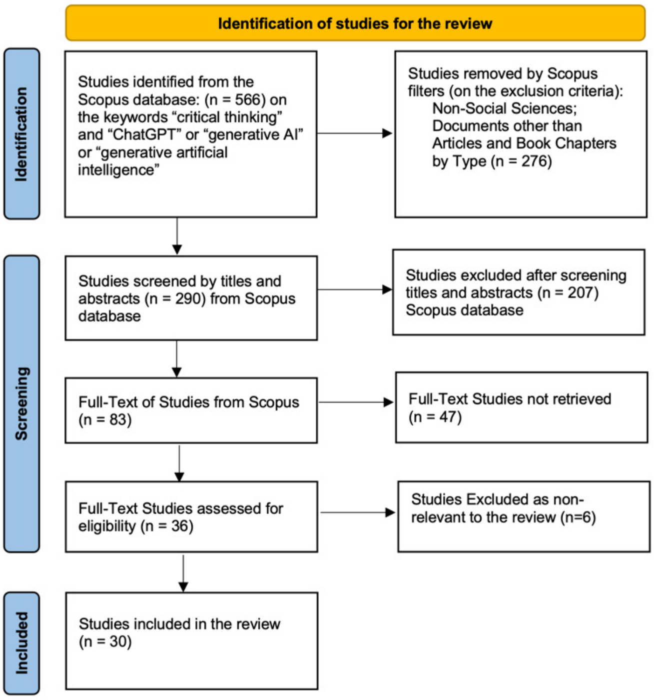


Table 2

Data-Charting Form

Data to be extracted	Population	Main findings related to CT
Title of study	Sample size	Type of GenAI impact
Author(s)	Data collection methods	Reported benefits of GenAI
Country	Main findings related to CT	Reported drawbacks
Type of publication	Type of GenAI impact	Limitations noted by authors
Study design	Sample size	Keywords
Context of GenAI use	Data collection methods	Discipline

critical thinking. This was followed by a cross-study comparison, during which patterns, convergences, and divergences across the publications were systematically coded. The coding was carried out independently by both reviewers and then iteratively refined through discussion to ensure analytical consistency.

Studies were grouped thematically based on the nature of GenAI application (e.g., academic writing support, feedback provision, argument development, or reflection tasks), and the type of critical thinking engagement reported (e.g., evaluation, reasoning, synthesis, or metacognitive awareness). These groupings were not predefined but emerged inductively from the data through repeated reading and clustering of findings. In addition, we paid close attention to how critical thinking was operationalised and measured across studies. Where applicable, we noted whether frameworks such as Bloom's taxonomy, self-regulated learning models, or other cognitive scaffolds were employed to define and assess critical thinking.

To support visual interpretation of trends, thematic clusters were validated using VOSviewer for co-occurrence analysis of keywords. The integration of keyword mapping allowed us to triangulate inductively identified trends with bibliometric signals from the literature set. The final synthesis was structured around the emergent thematic clusters, each supported by illustrative examples from the reviewed studies. These clusters served as the basis for answering the two research questions and outlining research directions in the concluding section of the review.

Data Validation

To ensure the trustworthiness and consistency of the review findings, we implemented a multi-step data validation process at both the extraction and synthesis stages. This process was designed to minimise bias, enhance inter-reviewer reliability, and support analytical transparency.

During the data charting phase, both reviewers independently extracted data from a randomly selected subset of five studies to establish a common interpretive framework and calibrate the data-charting form. Discrepancies were discussed and resolved through consensus, resulting in refinements to the extraction categories and clarification of ambiguous fields. Once alignment was reached, the remaining studies were divided equally between the reviewers, with a second round of cross-checking performed to confirm the consistency of coding and categorisation decisions.

To validate the thematic analysis, each emergent category or cluster was reviewed against the original full texts to confirm that the assigned labels accurately reflected the underlying content and empirical focus of the studies. Particular attention was paid to studies that overlapped thematically or contained findings relevant to multiple categories; these were flagged and jointly reviewed to ensure appropriate placement without duplication or over-interpretation. Additionally, to reduce the risk of interpretive bias, we employed a form of analyst triangulation: each reviewer independently interpreted the findings within each thematic group and then collaboratively compared interpretations. This process helped ensure that the identified trends and conclusions were not the result of individual analytical preferences but rather emerged through iterative dialogue and critical discussion.

Finally, keyword co-occurrence analysis conducted via VOSviewer served as a supplementary validation tool, offering a bibliometric lens to reinforce or question patterns observed in the thematic synthesis. Convergences between manually identified clusters and automatically generated keyword networks strengthened the internal coherence of the review's analytical framework.

RESULTS

Search and Selection Results

The searches in the Scopus database were made as of April 24, 2025. A total of 566 documents were found on a combination of the previously identified keywords: "critical thinking" AND "ChatGPT" OR "generative AI" OR "generative artificial intelligence".

After the Scopus filters, including the eligibility criteria of language, types of publications, and research field had been applied, the total decreased to 290. The authors screened the titles and abstracts of 290 documents. 207 documents were eliminated after the screening. Full texts of 36 documents out of 83 were retrieved for further screening. After they had been thoroughly analysed, six documents were found irrelevant as they were not based on empirical research, or the research was placed in a school environment. Thirty documents were ultimately included into the review (Figure 2).

A Bibliometric Analysis

The 30 documents ultimately included in the review were analysed. The publications indexed in the Scopus database included 29 research articles and one book chapter. The review entailed 19 and 11 documents published in 2024 and in 2025 respectively. Three publications appeared in *Thinking Skills and Creativity*. Two articles were published in each of the following four journals: *Cogent Education, Education and Information Technologies, Frontiers in Education, and International Journal of Educational Technology in Higher Education.* The remaining 18 articles came out in another 18 journals. The sampling entailed one book chapter. It was part of the book entitled "Studies in Systems, Decision and Control" (Emran et al., 2024).

Geographically, the reviewed documents were distributed as follows: China, Indonesia, the United Kingdom, and Taiwan accounted for four publications each. Two documents came from Educator and Malaysia each. The other countries included Italy (2 documents), Spain (2 documents), Switzerland (2 documents), and Bahrain (1 document).

The thirty publications were authored by 91 researchers, nearly 3 authors per record on average. The authors had 63 affiliations, with King's College London, Università degli Studi di Messina, National Yunlin University of Science and Technology, and National Cheng Kung University as the frontrunners (two authors from each). The other 59 affiliations were represented by one researcher each.

Included Studies

After identifying and screening the relevant documents, the ultimate 30 documents were included into the review. The total population in the reviewed documents randomly, purposively or otherwise selected for 29 out of the 30 studies amounted to 4785 participants, including university students, EFL learners, and university teachers. One of the studies that did not report population dwells upon the transformative impact of AI in educational settings, focusing on the necessity for AI literacy, prompt engineering proficiency, and enhanced critical thinking skills (Walter, 2024).

The data on the population and sample size are included in Appendix 1.

The study designs entail qualitative, quantitative, and mixed. Though, in describing study designs, we focused on the authors' wordings of the type of design in the reviewed documents. In Table 3, we also outlined quasi-experimental design, randomised controlled tests, and other designs, covering "a naturalistic inquiry approach" (Gonsalves, 2024), "a cross-sectional survey design" (Lijie et al., 2025), "an explanatory study" (Panit, 2024), "a case-study" (Walter, 2024), and "a true-experimental design" (Xu & Liu, 2025).

Thematic Clusters

The co-occurrence analysis of author keywords from the 30 included publications was conducted using VOSviewer, based on full counting and minimum threshold set to two occurrences per keyword. The resulting visualization (Figure 2) reveals 17 initial clusters, automatically generated via VOSviewer's clustering algorithm based on keyword prox-

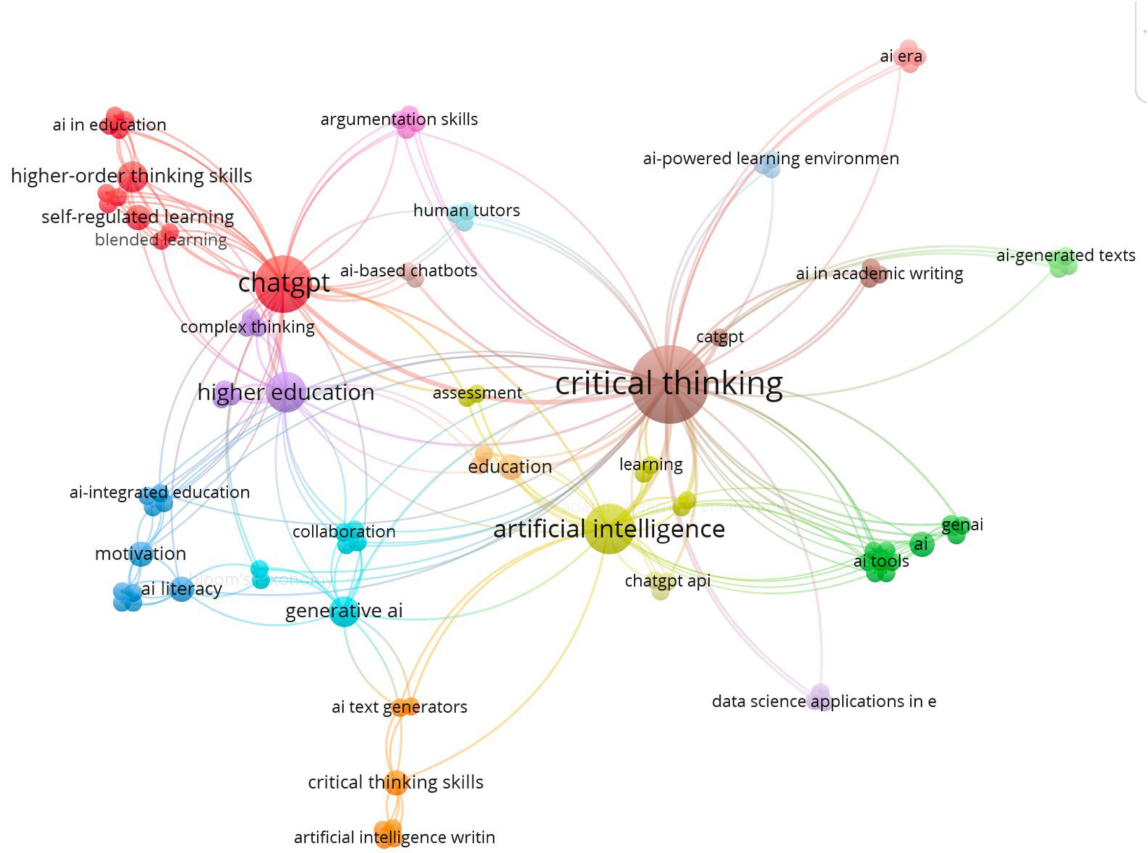
Table 3

Study Designs in the Reviewed Documents

Study design	Documents in the review
Qualitative design	Darwin et al., 2024; Octaberlina et al., 2024; Santamaria-Velasco et al., 2025
Quantitative design	Fadillah et al., 2024; Naatonis et al., 2024; Zhou et al., 2024
Mixed design	Essien et al., 2024; Fakour & Imani, 2025; Gerlich, 2025; Hwang et al., 2025;
Quasi-experimental design	Emran et al., 2024; George-Reyes et al., 2024; Liu & Tu, 2024;
Randomized controlled trial	Lee et al., 2024; Wang et al., 2025
Other designs	Boers et al., 2025; de la Puente ae al., 2024; Fabio et al., 2025; Chaparro-Banegas et al., 2024; Gonsalves, 2024; Lijie et al., 2025; Oates & Johnson, 2025; Panit, 2024; Shen & Teng, 2024; Suriano et al., 2025; Tang et al., 2024; Walter, 2024; Xiao et al., 2025; Xu & Liu, 2025; Yasuf et al., 2024

Figure 2

VOSviewer Visualization of Keyword Co-occurrence and Thematic Clusters



Note. Circle size indicates keyword frequency; spatial distance reflects co-occurrence strength; colors represent cluster groupings by VOSviewer's modularity-based algorithm. Thematic categories were manually synthesized based on semantic interpretation.

imity and frequency. However, due to thematic overlaps and conceptual affinities among clusters, we consolidated them into six broader thematic categories, described below.

Cognitive and Metacognitive Development (Clusters 1, 2, 5, 6, 13, 17)

This group includes studies exploring how GenAI tools support critical thinking, self-regulation, cognitive complexity, dual-process reasoning, and reflective thinking. Bloom's taxonomy, problem-based gamification learning (PBGL), and dual-process theory frequently appear in this set. These studies investigate how AI interaction fosters higher-order thinking skills (HOTS), metacognitive strategies, and deep learning.

Pedagogical Innovation and Learning Design (Clusters 4, 12, 15)

This thematic area focuses on instructional models integrating GenAI, such as guidance-based ChatGPT scaffolding, dialogic learning, and AI-mediated tutoring. Emphasis is placed on human - AI collaboration, Socratic methods, and blended feedback strategies that enhance reasoning, argumentation, and learner autonomy.

Academic Writing and Language Learning (Clusters 7, 8, 10)

Here, the focus is on academic discourse and EFL learners, examining how ChatGPT and other GenAI tools support writing pedagogy, improve argumentative skills, and challenge originality. Concerns about overreliance and intellectual autonomy are also raised, especially for students in postgraduate and EFL contexts.

AI Literacy and Learner Perception (Clusters 3, 9, 11)

This category examines students' and educators' perceptions of GenAI tools, their motivation, trust, and AI-literacy levels. Key concepts include attitudes toward AI, ethical reasoning, critical assessment of AI-generated content, and equity of access. A sub-focus is placed on learners' engagement across geographic and socio-economic contexts.

Evaluation and Assessment Technologies (Clusters 14, 16)

Studies in this group explore automated feedback systems, peer review tools, and educational data mining techniques. Emphasis is on using GenAI to evaluate critical thinking through scalable models (e.g., BERT, LLM4HA), epistemic network analysis, and evaluation metrics for learning analytics.

Global and Ethical Dimensions of GenAI Use (Clusters 9, 16)

This group overlaps with AI-literacy research but uniquely addresses macro-level implications: global disparities, digital inequality, responsible innovation, and sustainable development goals (SDGs). These studies treat AI not only as a pedagogical tool but as a social phenomenon shaping educational ethics and access.

Taken together, the thematic clusters highlight not only research directions but also the wide spectrum of competencies being shaped through human - ChatGPT interaction. Studies grouped under cognitive and metacognitive clusters emphasize the development of critical reasoning, reflective thinking, and self-regulation - core components of higher-order thinking skills. Clusters focused on academic writing, argumentation, and AI literacy foreground academic and ethical competencies, particularly among EFL learners, with attention to intellectual autonomy and authorship.

The clusters related to pedagogical innovation and evaluation technologies address more instrumental and analytical competencies, such as data analysis, gamification, and feedback design. Finally, clusters engaging with sustainability, equity, and global challenges point to the formation of value-oriented and civic competencies, including ethical judgment, teamwork, and global citizenship.

This competency-oriented lens reinforces the idea that generative AI tools do more than support skill acquisition: they participate in reshaping the educational paradigm, where flexible, integrative, and critically aware thinking becomes central to student preparedness for uncertain, technology-rich futures.

Generative AI Influence on Critical Thinking in Human-ChatGPT Interaction

This section synthesizes the findings of 30 empirical studies addressing how generative AI tools, especially ChatGPT, influence students' critical thinking (CT) skills in higher education. To respond to RQ2, the studies were analysed thematically and categorised into five analytical groups: (1) direct enhancement of CT, (2) metacognitive and reflective gains, (3) contextual factors shaping CT, (4) risks of cognitive offloading, and (5) instructional strategies mediating AI's effect. A structured summary of the reported benefits and drawbacks of AI impact on CT is provided in Appendix 1. 21 publications reported positive influence whereas the remaining nine showed mixed influence.

Direct Enhancement of CT Through AI-Supported Learning Tasks

A growing body of empirical studies has shown that generative AI enhances critical thinking (CT) when integrated into authentic, cognitively demanding tasks such as debate, argumentation, and analytical writing. These interventions appear most effective when AI tools are positioned not as information providers but as reasoning partners. For instance, de la Puente et al. (2024) demonstrated that incorporating

ChatGPT into structured debate sessions significantly improved both argumentation and CT skills, as confirmed by structural equation modelling. Emran et al. (2024) similarly found that students used ChatGPT to explore diverse perspectives and to practise logic-based evaluation, which in turn stimulated fact-checking behaviour and deeper cognitive engagement. Fabio et al. (2025) reported statistically significant gains across multiple CT dimensions (cognitive complexity, reasoning style, and openness) particularly among students who approached ChatGPT with caution and intent. Other studies support these findings across various disciplinary contexts: Santamaria-Velasco et al. (2025) linked ChatGPT use to improved evaluative judgement in historical and ethical analysis; Yusuf et al. (2024) showed that a scaffolded framework for AI-generated texts fostered synthesis and critique in academic writing; and Wang et al. (2025) observed CT improvement via AI-supported feedback mechanisms. Oates & Johnson (2025) further reported increases in students' critical evaluation scores during AI-assisted assessment tasks. Together, these findings highlight the potential of ChatGPT to reinforce CT when its use is grounded in well-designed learning scenarios that prioritize interpretation, justification, and reasoning.

Metacognitive and Reflective Engagement

Beyond direct skill acquisition, generative AI also plays a notable role in cultivating learners' metacognitive awareness and reflective thinking. This refers to students' ability to monitor, regulate, and evaluate their own cognitive processes in interaction with AI tools. Studies included in the review show that when learners are prompted not just to consume AI outputs, but to critically interrogate them, deeper layers of cognition are activated.

A useful conceptual distinction is offered between two complementary forms of critical engagement: one that focuses on the AI itself and another oriented toward the task at hand. Gonsalves (2024) refers to these as "critical thinking toward the AI", including practices such as refining prompts and evaluating potential biases, and "critical thinking for the assignment," which involves the application of AI-generated content to solve real academic problems. Such dual engagement fosters both reflective scepticism and problem-solving reasoning.

Reflective thinking is further supported when students become aware of AI's limitations. Darwin et al. (2024) reported that learners demonstrated increased awareness and caution when confronting the inability of AI systems to recognize irony or emotional nuance. Other studies revealed that this awareness becomes more acute when students encounter the misinformation produced by AI (plausible but incorrect outputs) which stimulate the need for verification and information control. Essien et al. (2024) and Emran et al. (2024) observed that these moments of AI failure led students to fact-check, triangulate sources, and engage more cautiously with content-behaviours closely tied to metacognitive regulation.

In addition to these spontaneous reflective responses, several pedagogical interventions were designed to scaffold metacognition intentionally. For example, Lee et al. (2024) introduced a guidance-based ChatGPT-assisted learning tool that provided indirect prompts rather than direct answers, thereby encouraging learners to articulate their own reasoning. Similarly, Hwang et al. (2025) demonstrated that a prompt-based learning model increased students' ability to generate questions and reflect on their learning paths. Both studies emphasize that structured interaction with AI (especially when mediated through carefully designed instruction) can enhance the quality and depth of students' reflective thought.

These findings suggest that generative AI, when embedded in learning environments that foreground interpretation and self-regulation, has the capacity to develop metacognitive dimensions of critical thinking, rather than merely facilitating content delivery.

Individual and Contextual Moderators of AI's Influence on Critical Thinking

While generative AI offers various opportunities to enhance critical thinking (CT), its actual impact is not uniform across learners or learning environments. Multiple studies suggest that individual learner characteristics and contextual factors substantially shape how AI tools affect cognitive engagement. Motivation, prior digital competence, and self-regulatory behaviours emerged as key variables influencing the effectiveness of AI-assisted learning. Lijie et al. (2025) demonstrated that motivation, AI literacy, and perceived usefulness of generative tools positively predicted students' disposition toward CT. However, ease of use was paradoxically associated with weaker CT development, likely due to reduced cognitive investment. This suggests that overly intuitive tools may inadvertently discourage deeper engagement. Similarly, Shen and Teng (2024) found that students with stronger self-directed learning (SDL) skills benefited more from AI-assisted writing tasks in terms of CT growth. SDL functioned as a moderator in the relationship between AI-supported writing and critical reasoning, indicating that autonomous learning behaviours amplify AI's educational potential.

Platform type also appeared less decisive than quality of engagement. Xu and Liu (2025) compared ChatGPT with Duolingo and found no significant differences in outcomes for CT and learner autonomy, implying that interactive depth, rather than platform architecture, drives cognitive gains. In a related vein, Suriano et al. (2025) showed that students' level of engagement and trust in ChatGPT were more powerful predictors of CT development than tool availability itself.

Other studies highlighted how design features and perceived reliability influence higher-order thinking skills (HOTS). Fadillah et al. (2024) emphasized that students' perceptions of convenience, responsiveness, and output accuracy significantly affected their cognitive engagement, with these factors positively correlating with HOTS development. Context-specific practices also matter: Chaparro-Banegas et al. (2024) in science education, and George-Reyes et al. (2024) in entrepreneurship training, reported notable CT improvements when AI was integrated into active, student-centred learning formats.

AI's educational impact is mediated not only by its functionality but also by learners' motivation, autonomy, and instructional context. These factors must be strategically addressed to unlock the full potential of AI-enhanced critical thinking (Table 4).

Risks of Over-Reliance and Cognitive Offloading

While many studies acknowledge the potential of generative AI to support critical thinking, growing evidence points to notable risks associated with its unmoderated use. A recurring concern across several studies is cognitive offloading - the tendency of students to delegate analytical tasks to AI systems instead of engaging with them independently. Gerlich (2025) identified a significant negative correlation between frequent use of AI tools and critical thinking performance, showing that reliance on automated outputs reduced students' cognitive effort and led to poorer outcomes. Octaberlina et al. (2024) and Panit (2024) reinforced this concern, noting that habitual dependence on AI tools can diminish active engagement and foster a passive learning stance.

More worryingly, this behavior was shown to compromise ethical judgment, particularly in academic writing, where the boundary between support and substitution becomes blurred. In several cases, students were found to accept AI-generated information without sufficient scrutiny, bypassing the deeper cognitive processes necessary for synthesis and evaluation (Panit, 2024). Gerlich (2025) further demonstrated that such patterns were associated with lower critical thinking scores and reduced long-term retention. These findings collectively suggest that, without careful scaffolding and pedagogical oversight, the use of GenAI may undermine the very cognitive capacities it is expected to cultivate.

Instructional Strategies Mediating the Impact of GenAI on Critical Thinking

A growing body of evidence suggests that the pedagogical design of GenAI integration plays a decisive role in shaping its influence on students' critical thinking (CT). Rather than functioning as a neutral tool, GenAI becomes pedagogically meaningful when embedded within instructional frameworks that prioritize active engagement, cognitive challenge, and reflective processing.

Table 4

Moderating Variables Influencing the Effect of GenAI on Critical Thinking Development

Moderator Type	Specific Factor	Effect on Critical Thinking (CT)	Representative Sources
Individual	Motivation (MO)	Strong positive predictor of critical thinking disposition	Lijie et al., 2025
	AI Literacy (AIL)	Promotes metacognitive awareness and more deliberate engagement with GenAI	Lijie et al., 2025
	Self-Directed Learning (SDL)	Enhances the impact of AI-assisted writing on CT development	Shen & Teng, 2024
	Trust and Engagement	High engagement levels correlate with greater CT improvement	Suriano et al., 2025
	Educational Attainment	Higher education levels associated with stronger CT outcomes	Gerlich, 2025
Technological	Perceived Ease of Use (PEOU)	Paradoxically lowers CT due to reduced cognitive effort	Lijie et al., 2025
	Output Reliability and Accuracy	Reinforces analytical behavior and user trust	Fadillah et al., 2024
	Platform Structure (e.g., ChatGPT vs Duolingo)	No significant difference if user engagement is comparable	Xu & Liu, 2025
Pedagogical	Active Learning Strategies (e.g., debates, inquiry tasks)	Amplify CT when integrated with GenAI-supported activities	Chaparro-Banegas et al., 2024; George-Reyes et al., 2024
	Interaction Format (e.g., scaffolding, feedback, mentorship)	Strengthens meaningful application of GenAI and metacognitive growth	Lee et al., 2024; Walter, 2024

One of the most promising instructional strategies involves prompt engineering[1]. As Walter (2024) argues, this practice not only requires students to think analytically about how they interact with GenAI but also fosters deeper understanding of the technology's limitations and ethical dimensions. His findings also emphasize the value of instructional scaffolding in AI literacy courses, where guided questioning, explicit feedback, and modeling of critical evaluation help students move from surface-level interaction to meaningful reflection and ethical reasoning.

Other studies underline the effectiveness of problem-based and gamified learning formats**.** For example, Naatonis et al. (2024) demonstrated that embedding ChatGPT into inquiry-driven modules enhanced students' CT by activating iterative feedback loops and engaging learners in self-directed problem-solving. Similarly, Xiao et al. (2025) evaluated the performance of large language models (LLMs), such as BERT and LLM4HA, in automatically assessing higher-order thinking skills (HOTS). Their results suggest that such tools can serve as efficient complements to human-led instruction by providing scalable, formative evaluation aligned with cognitive complexity.

Instructional strategies that promote self-regulation have been shown to significantly strengthen the impact of GenAI on critical thinking. Tools that prompt learners to plan, monitor, and reflect on their thinking processes encourage deeper engagement with complex problems and foster higher-order reasoning. When these self-regulatory mechanisms are embedded into digital learning environments, students not only solve tasks more effectively but also develop a stronger disposition toward critical inquiry. This approach was demonstrated to be particularly effective in a recent study involving AI-assisted learning design (Zhou et al., 2024). The findings are consistent with broader research emphasizing that autonomy-supportive environments enhance metacognitive awareness and sustain cognitive engagement over time (Ryan & Deci, 2017).

The reviewed evidence affirms that generative AI, particularly in the form of ChatGPT, can meaningfully support the development of critical thinking skills. This contribution is achieved not by substituting instructional efforts, but by enhancing them through intentional and well-structured pedagogical integration. When situated within guided, dialogic, or inquiry-based learning environments, generative AI serves as a catalyst for analytical reasoning, metacognitive development, and reflective judgment. Nevertheless, the effectiveness of such technologies depends substantially on the presence of appropriate instructional scaffolding. In its absence, learners are more likely to interact with AI tools in superficial or utilitarian ways, thereby limiting their transformative potential for education.

DISCUSSION

This scoping review synthesised empirical evidence on the influence of generative AI, particularly ChatGPT, on the development of critical thinking (CT) among university students. The overall trend across the included studies points to a predominantly positive effect, provided that AI use is pedagogically embedded and critically mediated. Key reported benefits include support for idea generation, conceptual understanding, construction of arguments, literature review, academic writing, and engagement in iterative reasoning processes (Fabio et al., 2025; Yusuf et al., 2024; Wang et al., 2025; Santamaria-Velasco et al., 2025). Students also benefited from time saved on routine academic tasks, which allowed greater cognitive focus on higher-order thinking. The reviewed literature consistently emphasised that engaging critically with AI through practices such as fact-checking, prompt refinement, and bias identification stimulates analytical reasoning and fosters critical thinking skills (Emran et al., 2024; Babin et al., 2024; Gonsalves, 2024).

Although benefits were prominent, several studies reported potential challenges. These include cognitive offloading, over-reliance on AI, superficial engagement with generated content, and diminished awareness of authorship or originality (Gerlich, 2025; Panit, 2024; Octaberlina et al., 2024). Darwin et al. (2024) warned against the formation of "echo chambers" and the inability of AI systems to grasp affective nuances. These risks were particularly salient when AI was used without adequate scaffolding or reflection, highlighting the need for structured instructional environments. Nevertheless, pedagogical frameworks such as scaffolded prompting, guided evaluation, and dialogic feedback demonstrated mitigating effects, enabling learners to maintain agency and reflective engagement (Lee et al., 2024; Walter, 2024; Hwang et al., 2025).

While the present findings generally align with the conclusions of earlier systematic reviews (Sardi et al., 2025; Melisa et al., 2025), which reported predominantly positive or neutral outcomes of GenAI use in education, this review extends prior research by offering a more granular synthesis of how specific instructional conditions, user dispositions, and AI design features interact to mediate the development of critical thinking. Unlike previous studies that focused on general outcomes, the current review maps out nuanced moderators and pedagogical mechanisms that shape the quality of human-GenAI interaction. Although Sardi et al. (2024) reported gains in higher-order CT skills and Erlich

[1] The deliberate crafting of inputs to elicit specific cognitive responses from AI systems.

(2025) identified improvements primarily at the lower levels of Bloom's taxonomy, the present review contributes a more differentiated perspective by systematically integrating these diverging outcomes. It further substantiates the ambivalence observed in broader research on AI's cognitive effects (Royer, 2024; Raitskaya & Lambovska, 2024; Crompton & Burke, 2023; Fuchs, 2023; Tikhonova & Raitskaya, 2023; Ivakhnenko & Nikolskiy, 2023), emphasizing the critical role of instructional context, task design, and learner characteristics in determining whether GenAI use promotes or undermines critical thinking.

Limitations in the Literature

The reviewed studies faced methodological limitations that affect the strength of inferences. Most relied on outcome-based indicators of CT, with few offering process-based or behavioural analyses (Boers et al., 2025; Suratmi & Sopandi, 2022). Instruments often included self-report questionnaires or task-based proxies, rather than longitudinal or mixed-method approaches (Essien et al., 2024; Fakour & Imani, 2025; Pan et al., 2025). Furthermore, the cognitive processes underpinning CT remain difficult to isolate in AI-mediated settings, and the influence of disciplinary differences was not systematically addressed.

The Role of Moderating Factors

Several moderating factors were identified that condition the impact of GenAI on CT. Individual factors such as motivation, AI literacy, self-directed learning, and critical disposition were found to correlate with better CT outcomes (Lijie et al., 2025; Shen & Teng, 2024; Suriano et al., 2025). Technological characteristics such as ease of use, perceived reliability, and platform structure played a nuanced role. For instance, Lijie et al. (2025) observed that while ease of use facilitated adoption, it paradoxically reduced cognitive effort. Pedagogical variables such as gamified learning, scaffolded prompts, and instructor feedback were consistently associated with improved CT (Chaparro-Banegas et al., 2024; George-Reyes et al., 2024; Walter, 2024; Lee et al., 2024).

Practical and Theoretical Implications

Instructional strategies such as prompt engineering, metacognitive scaffolding, and reflective tasks appear essential for maximising the educational value of GenAI. Walter (2024) and Gonsalves (2024) emphasised that teaching students to generate purposeful prompts cultivates both analytical reasoning and epistemic awareness. The reviewed studies highlight that GenAI can serve as a catalyst for critical thinking, provided that its use is guided, intentional, and aligned with pedagogical goals. These findings support a constructivist understanding of AI-enhanced learning and underscore the importance of integrating digital literacy and CT into academic curricula (Babin et al., 2024; Chen et al., 2025; Wong, 2024).

This review highlights that the influence of generative AI on critical thinking is shaped by a combination of pedagogical design, learner characteristics, and the nature of AI-supported tasks. While structured and reflective integration of GenAI tends to support the development of analytical and metacognitive skills, uncritical or instrumental use may lead to cognitive passivity and reduced engagement. The evidence points to the need for more targeted instructional strategies that align AI use with specific cognitive goals. These findings call for a reorientation of future research toward identifying the conditions under which GenAI fosters meaningful learning and for refining pedagogical models accordingly.

CONCLUSION

This scoping review examined how generative artificial intelligence (GenAI), particularly in educational contexts involving tools like ChatGPT, contributes to the development of critical thinking among university students. The analysis of recent empirical studies revealed a broad spectrum of approaches to integrating GenAI into teaching and learning. These included strategies such as prompt engineering, task-based learning, self-regulated and self-directed learning, AI-assisted writing, scaffolded instruction, and reflective evaluation of AI-generated content. The reviewed literature confirmed that, under appropriate pedagogical conditions, GenAI can serve as a meaningful catalyst for the development of critical thinking skills. Students who engage actively and consciously with AI tools tend to demonstrate increased metacognitive awareness, improved argumentation, and more effective reasoning.

At the same time, the review identified factors that may inhibit the positive effects of GenAI. Passive use of AI, lack of instructional support, and over-reliance on automatically generated outputs were associated with reduced cognitive engagement and a decline in students' ability to independently evaluate information. The outcomes of GenAI use were also shaped by individual learner variables such as motivation, AI literacy, and the ability to self-regulate learning activities.

Several limitations of the present review must be acknowledged. The analysis was restricted to publications indexed in the Scopus database and covered a narrow timespan, focusing only on the years 2024 and 2025. This may have resulted in the omission of relevant studies published earlier or indexed in other databases. Moreover, the majority of included research relied on short-term observations or self-reported data, which limits the generalisability of the findings and precludes conclusions about the long-term impact of GenAI use on cognitive development. The lack of consistent theoretical frameworks across the reviewed studies also complicates efforts to compare results and synthesise evidence in a cumulative way.

Future research should address these gaps in several directions. First, longitudinal studies are needed to investigate the enduring effects of GenAI-supported instruction on critical thinking across diverse educational contexts and learner populations. Second, there is a clear need to develop more robust theoretical models that explicitly connect the dimensions of critical thinking, as defined in Bloom's taxonomy, with the specific learning mechanisms that are activated through the use of generative AI. Third, further work should focus on evaluating how different forms of instructional support, including scaffolded AI use and targeted AI literacy training, influence students' cognitive outcomes and their ability to engage with AI critically and productively. By advancing these research priorities, scholars and educators can better understand how to integrate GenAI tools into higher education in ways that support intellectual autonomy, foster analytical thinking, and reinforce the development of critical competencies in an increasingly digital learning environment.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

Lilia Raitskaya: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – review & editing; other contribution.

Elena Tikhonova: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – review & editing; other contribution.

REFERENCES

Ajlouni, A. O., Wahba, F. A. A., & Almahaireh, A. S. (2023). Students' attitudes towards using ChatGPT as a learning tool: The case of the university of Jordan. *International Journal of Interactive Mobile Technologies*, *17*(18), 99-117. https://doi.org/10.3991/ijim.v17i18.41753

Al-Mamary, Y.H., & Abubakar, A.A. (2025). Empowering ChatGPT adoption in higher education: A comprehensive analysis of university students' intention to adopt artificial intelligence using self-determination and technology-to-performance chain theories. *The Internet and Higher Education*, *66*, Article 101015. https://doi.org/10.1016/j.iheduc.2025.101015

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/101080/1364557032000119616

Babin, J.I., Raber, H., & Mattingly II, T.J. (2024). Prompt pattern engineering for test question mapping using ChatGPT: A cross-sectional study. *American Journal of Pharmaceutical Education*, *88*, Article 101266. https://doi.org/10/1016/j.ajpe.2024.101116

Bertocchi, F. M., De Oliveira, A. C., Lucchetti, G., & Lucchetti, A.L.G. (2022). Smartphone use, digital addiction and physical and mental health in community-dwelling older adults: A population-based survey. *Journal of Medical Systems, 46*(8), 53. https://doi.org/10.1007/s10916-022-01839-7

Bloom, B. S., Krathwohl, D. R., Engelhart, M.B.., Furst, E.J., & Hill, W.H. (1956). *Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. Handbook I: Cognitive domain.* Longman.

Boers, J., Etty, T., Baars, M., & van Boekhoven, K. (2025). Exploring cognitive strategies in human-AI interaction: ChatGPT's role in creative tasks. *Journal of Creativity*, *35*(1), Article 100095. http://doi.org/ 10.1016/j.yjoc.2025.100095

Chaparro-Banegas, N., Alicia, M.-T., & Roig-Tierno, N. (2024). Challenging critical thinking in education: New paradigms of artificial intelligence. *Cogent Education*, *11*(1), Article 2437899. https://doi.org/10.59429/esp.v9i11.3141

Chen, B., Zhang, Z., Langere, N., & Zhu, S. (2025). Unleashing the potential of prompt engineering for large language models. *Patterns*, *6*(6), Article 101260. https://doi.org/10.1016/j.patter.2025.101260

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), 22. https://doi.org/10.1186/s41239-023-00392-8

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E.D., & Marzuki (2024). Critical thinking in the AI era: An exploration of EFLbstudents' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), Article 2290342. https://doi.org/10.1080/2331186X.2023.2290342

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*, 55. https://doi.org/10.1186/s40561-024-00347-0

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, *227*, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Emran, A.Q.M., Aldallal, A., & Nadheer, A. (2024). Investigating the impact of ChatGPT on enhancing university students' critical thinking skills. In A. Hamdan & A. Harraf (Eds.), *Business development via AI and digitalization* (vol. 1, pp. 567-574). Springer. https://doi.org/10.1007/978-3-031-62102-4_47

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

Fabio, R.A., Plebe, A., & Suriano, R. (2025). AI-based chatbot interactions and critical thinking skills: An explanatory study. *Current Psychology*. https://doi.org/10.1007/s12144-024-06795-8

Fadillah, M.A., Usmeldi, U., & Asrizal, A. 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

Fakour, H., & Imani, M. (2025). Socratic wisdom in the age of AI: a comparative study of ChatGPT and human tutors in enhancing critical thinking skills. *Frontiers in Education*. https://doi.org/10.3389/feduc.2025.1528603

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, *6*(1), 57-63. https://doi.org/10.37074/jalt.2023.6.1.22

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, *8*, 1166682. https://doi.org/10.3389/feduc.2023.1166682

George-Reyes, C.E., Vilhunen, E., Avello-Martinez, R., & Lopez-Caudana, E. (2025). Developing scientific entrepreneurship and complex thinking skills; Creating narrative scripts using ChatGPT. *Frontiers in Education*, *9,* Article 1378564. https://doi.org/10.3389/feduc.2024.1378564

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

Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting bloom's taxonomy. *Journal of Marketing Education*, 1-16. https://doi.org/10.1177/02734753241305980

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

Ivakhnenko, E. N., Nikolskiy, V. S. (2023). ChatGPT in higher education and science: A threat or a valuable resource? *Higher Education in Russia*, *32*(4), 9-22. https://doi.org/10.31992/0869-3617-2023-32-4-9-22

Jiang, T., Sun, Z., Fu, S., & Lv, Y. (2024). Human-AI interaction research agenda: A user-centered perspective. *Data and Information Management*, *8,* Article 1000078. https://doi.org/10.1016/j.dim.2024.100078

Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. *Theory in to Practice*, *41*(4), 37-41. https://doi.org/10.1207/s15430421tip4104_2

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*, 16. https://doi.org/10.org/10.1186/s41239-024-00447-4

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

Lim, W. M., A. Gunasekara, J. L. Pallant, J. I. Pallant, & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or Reformation? A paradoxical perspective from management educators*. The International Journal of Management Education*, *21*(2), Article 100790. https://doi.org/10.1016/j.ijme.2023.100790

Liu, Q., & Tu, C.C. (2024). Improving critical thinking through AI-supported socio-scientific issues instruction. *Journal of Logistics, Informatics and Service Science*, *11*(3), 52-65. https://doi.org/10.33168/JLISS.2024.0304

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin., & Fuad, Z. A. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal*, *14*, Article e2025031. https://doi.org/10.22521/edupij.2025.14.31

Naatonis, R.N., Rusijono, Jannah, M., & Malahina, E.A.U. (2024). Evaluation of Problem Based Gamification Learning (PBGL) model on critical thinking ability with Artificial Intelligence approach integrated with ChatGPT API: An experimental study. *Qubahan Academic Journal*, *4*(3), 485-520. https://doi.org/10.58429/qaj.v4n3a919

Oates, A., & Johnson, D. (2024). ChatGPT in the classroom: Evaluating its role fostering critical evaluation skills. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007.s40593-024-00452-8

Octaberlina, L.R., Muslimin, A.I., Chamidah, D., Surur, M., & Mustikawan, A. (2024). Exploring the impact of AI threats on originality and critical thinking in academic writing. *Edelweiss Applied Science and Technology*, *8*(6), 8805-8814. https://doi.org/10.55214/25768484.v8i6.3878

Pan, Z., Moore, O.A., Papadimitriou, A., & Zhu, J. (2025). AI literacy and trust: a multi-method study of Human-GAI team collaboration. *Computers in Human Behavior: Artificial Humans, 4*, Article 100162. https://doi.org/10.106/j.chbah.2025.100162

Panit, N.M. (2024). Can critical thinking and AI work together? Observations of science, mathematics, and language instructors. *Environment and Social Psychology, 9*(11), Article 3141. https://doi.org/10.59429/esp.v9i11.3141

Promma, W., Imjai, N., Usman, B., & Aujirapongpan, S. (2025). The influence of AI literacy on complex problem-solving skills through systematic thinking skills and intuition thinking skills: An empirical study in Thai gen Z accounting students. *Artificial Intelligence*, *8*, Article 100382. https://doi.org/10.1016/j.caeai.2025.100382

Raitskaya, L., & Lambovska, M. (2024). Prospects for ChatGPT application in higher education: A scoping review of international research. *Integration of Education*, *28*(1), 10-21. https://doi.org/10.15507/1991-9468.114.028.202401.010-021

Risko, E.F., & Gilbert, S.J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20,* 676-688. htto://doi.org/10.1016/j.tics.2016.07.002

Royer, C. (2024). Outsourcing humanity? ChatGPT, critical thinking, and the crisis in higher education. *Studies in Philosophy and Education*, *43*, 479-497. https://doi.org/10.1007/s11217-024-09946-3

Santamaria-Velasco, J., Nunez-Naranjo, A., & Morales-Urrutia, X. (2025). Critical thinking and AI: Enhancing history teaching through ChatGPT simulations. *Internation Journal of Innovative Research and Scientific Studies*, *8*(1), 564-575. https://doi.org/10.53894/ijirss.v8i1.4403

Sardi, J., Darmansyah, Candra, O., Devi Faizah, Y., Habibullah, Yanto, D.T.P., & Fivia, E. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? A systematic review. *International Journal of Engineering Pedagogy*, *15*(1), 94-108. https://doi.org/10.3991/ijep.v15i1.53379

Shen, X., & Teng, M.F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity*, *52*, Article 101524. https://doi.org/10.1016/j.tsc.2024.101524

Steele, J.L. (2023). To GPT or not GPT? Empowering our students to learn with AI. *Computers and Education: Artificial Intelligence*, *5*, 100160. https://doi.org/10.1016/j.caeai.2023.100160

Suratmi, S., & Sopandi, W. (2022). Knowledge, skills, and attitudes of teachers in training critical thinking of elementary school students. *Journal of Education and Learning (EduLearn), 16*(3), 291–298. https:// doi. org/ 10. 11591/ edule arn.v16i3. 20493

Suriano, R., Plebe, A., Accial, A., & Fabio., R.A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction*, *95*, Article 102011. https://doi.org/9510.1016/j.learninstruc.2024.102011

Tang, T., Sha, J., Zhao, Y., Wang, S., Wang, Z., & Shen, S. (2024). Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria. *Thinking Skills and Creativity*, *52*, Article 101607. https://doi.org/10.1016/j.tsc.2024.101607

Tikhonova E., & Raitskaya L. (2023). ChatGPT: Where Is a silver lining? Exploring the realm of GPT and large language models. *Journal of Language and Education*, *9*(3), 5-11. https://doi.org/10.17323/jle.2023.18119

Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horseley, T., Weeks, L., Hempel, S., & Akl, E.A. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–73. https://doi.org/107326/M18-0850

van Rensburg, J.J. (2024). Artificial human thinking: ChatGPT's capacity to be a model for critical thinking when prompted with problembased writing activities. *Discover Education*, *3*(1), 42. https://doi.org/10.1007/s44217-024-00113-x

Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Technology in Higher Education, 21*, Article 15. https://doi.org/10.1186/s41239-024-00448-3

Wang, W.-S., Lin, C.-J., Lee, H.-Y., Huang, Y.-M., & Wu, T.-T. (2025). Enhancing self-regulated learning and higher-order thinking skills in virtual reality: The impact of ChatGPT-integrated feedback aids. *Education and Information Technologies.* https://doi.org/10.1007/s10639-025-13557-x

Wong, R. S.-Y. (2024). ChatGPT in medical education: Promoting learning or killing critical thinking? *Education in Medicine Journal*, *16*(2), 177-183. https://doi.org/10.21315/eimj2024.16.2.13

Xiao, X., Li, Y., He, X., Fang, J., Yan, Z., & Xie, C. (2025). An assessment framework of higher-order thinking skills based on fine-tuned large language models. *Expert Systems with Applications*, *272*, Article 126531. https://doi.org/10.1016/j.eswa.2025.126531

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*, Article 102100. https://doi.org/10.1016/j.lmot.2025.102100

Yusuf, A., Bello, S., Pervin, N., & Tukur, A.K. (2024). Implementing a proposed framework for enhancing critical thinking skills in synthesizing AI-generated texts. *Thinking Skills and Creativity*, *53*, Article 101619. https://doi.org/10.1016/j.tsc.2024.101619

Zhai, X., Nyaaba, M., & Ma, W. (2025). Can generative AI and ChatGPT outperform humans on cognitivedemanding problem-solving tasks in science? *Science & Education*, *34*, 649-670. https://doi.org/10.1007/s11191-024-00496-1

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*, Article 1302. https://doi.org/10.3390/educsci14121302