The Impact of Artificial Intelligence on Students' Critical Thinking: A Systematic Synthesis, Conceptual Model, and Comparative Review

Abdullah Khan[1], Ashad Ahmed[2], Wasiq Khan[3], Humaiz Shaikh[4*], Faraz Ahmed Shaikh[5]

[1,2,3]Department of Computer Science, Ilma University, Main Ibrahim Hyderi Road, Korangi Creek, Karachi, Sindh, Pakistan.
[4]Department of Software Engineering, Ilma University, Main Ibrahim Hyderi Road, Korangi Creek, Karachi, Sindh, Pakistan.
[5]Office of Quality Enhancement Cell, Karachi Metropolitan University (KMU), Karachi, Sindh, Pakistan

* humaiz0416@gmail.com; dr.humaiz@ilmauniversity.edu.pk

DOI: https://doi.org/

Keywords
Artificial Intelligence In Education, Critical Thinking, Cognitive Offloading, Chatgpt, Bloom's Taxonomy, Self-Regulated Learning, Academic Integrity, Generative Ai, Metacognition, Higher-Order Thinking

Article History
Received on 20 January 2026
Accepted on 20 February 2026
Published on 27 February 2026

Corresponding Author: *

Abstract
The rapid diffusion of AI-based applications such as ChatGPT in academic and self-study environments has further stimulated debates on the effects of such applications on the development of critical thinking skills among students. This review will be a systematic review of the existing literature on this topic published from 2011 to 2026, both theoretical and experimental. Building on the cognitive load theory, self-regulated learning theory, Bloom's taxonomy revision, and cognitive offloading research, this paper develops an original Critical Thinking Paradox Model, which models both ways in which AI influences critical thinking through a reliance-dependent mathematical function. According to the synthesis, AI applications are effective at enhancing academic performance, motivational-affective state, and lower-order cognitive processes, but the evidence concerning higher-order cognitive processes, metacognition, and autonomous reasoning is inconsistent and dependent upon scaffolding, tasks design, and students' AI skills. Neurophysiological data suggests reduced neural connectivity and decreased feeling of ownership among students writing essays without scaffolding from large-language models compared to unsupervised and search-engine-based essay writing. A comparative review of 15 representative studies is presented, along with five original figures and six mathematical formulations. Based on these findings, the paper provides practical recommendations for promoting AI literacy while preserving students' critical thinking skills in AI-supported educational environments. It also outlines key directions for future research, including longitudinal studies, cross-cultural investigations, and neuroscientific research.

INTRODUCTION

Since 2022, the launch of large scale conversational AI tools has led to one of the most rapid technology adoption curves witnessed in education. A significant number of university students know about ChatGPT, and many of them regularly use it for their assignments, with some reports indicating that many of them use ChatGPT multiple times a week or even every day [3], [57]. However, this adoption happened far ahead of any institution policies and guidelines or cognitive impacts studies [58], [78].

Critical thinking, commonly referred to as a purposeful, self-regulatory judgment that leads to interpretation, analysis, evaluation, and inference [67], as well as being characterized by intellectual skepticism, evidence-based reasoning, and reflective self-correction [66], [70], has for many years been viewed as an essential learning outcome in higher education [93], [95]. The concern about the erosion of critical thinking due to the constant use of generative AI is certainly not new and can be seen in studies done over many decades on what has become known as the “Google effect” or digital amnesia, in which the anticipation of future information access decreases the probability that the information will actually be encoded [21], [34]. What is new is the scope, natural language capability, and authoritative nature of the output from generative AI.

Literature on the topic of "Artificial Intelligence and Students' Cognition" is usually presented in two contrasting positions. According to the first view, Artificial Intelligence can positively influence the education process due to lower workload for unnecessary brain functions, individualized approach, and possibility to pay attention to more complicated thinking activities. Consequently, better results at school are expected from students and sometimes even their development of critical thinking [14], [17]. However, the second perspective focuses on possible threats of intensive use of AI. Too high reliance on artificial intelligence may result in the cognitive offloading process, lack of independent thinking, and disengagement from one's learning process. In addition, there are research findings about the lower neural engagement and sense of ownership due to the intensive use of AI [8], [12], [19].

Rather than discussing whether the use of AI is wholly positive or negative, it would be more productive to evaluate the situations when it either helps or inhibits critical thinking. It is the manner in which AI is used and the level of reliance of students on AI that affect its outcomes.

This research is focused on the problem under analysis from four angles. Firstly, eighty articles from 2011 to 2026 are analyzed. Secondly, the Critical Thinking Paradox Model illustrating the effects of different levels of reliance on AI on critical thinking is offered. Thirdly, fifteen empirical articles are compared. Fourthly, suggestions concerning how to improve AI literacy without harming students' ability to think critically are made.

The rest of the essay will be outlined as follows. The theoretical framework is given in section II, and the model in section III. The review method is covered in section IV, and the key findings in section V.

Section VI presents the comparative review table. Section VII discusses moderating factors and domain-specific effects. Section VIII offers pedagogical recommendations. Sections IX and X address limitations and future research, and Section XI concludes.

Theoretical Framework

Bloom's Taxonomy and Higher-Order Thinking

Classification of cognitive goals suggested by Bloom [64] and further evolution of Bloom’s taxonomy into six levels including remembering, understanding, applying, analyzing, evaluating, and creating [64] represents a traditional approach used for classification of cognitive activities in terms of educational curriculums. According to recent researches, classification initially suggested is not capable of reflecting a cyclical and interactive nature of AI-education process during which students constantly refine their prompts and analyze machine-generated texts with help of lower and higher order cognitive skills [26]. No matter what new application of Bloom's taxonomy is used in regard to the field of AI-assisted classrooms, it becomes clear that the use of AI technologies is highly effective in supporting lower-order cognitive tasks (remembering and understanding), but is rather limited in areas of analysis and evaluation without deliberate actions of educators [27], [31], [44].

Cognitive Load Theory

The Cognitive Load Theory divides working memory capacity into three types of load: intrinsic load, which is the natural difficulty of the problem; extraneous load, which is the unnecessary load created due to bad instructional design; and germane load, which is the amount of load dedicated to schema construction [7], [8]. The use of generative AI technology reduces extraneous load. For instance, ChatGPT was found to reduce extraneous load in socio-scientific issues compared to conventional search engines [8]. However, when the AI technology takes care of the germane load along with the extraneous load, it restricts the effortful process needed for meaningful learning [8], [12].

Cognitive Offloading and Transactive Memory

Cognitive offloading refers to the use of external physical or digital tools to reduce internal cognitive demand [71], building on Wegner's earlier concept of transactive memory, in which knowledge is distributed across individuals and external systems rather than stored entirely internally [72]. Sparrow, Liu, and Wegner's foundational experiments demonstrated that when participants expected future access to information via computer, they exhibited significantly lower recall of the information itself and enhanced recall of where to find it [21], a pattern subsequently replicated and extended across nine experiments and multiple meta-analytic reviews [23], [34]. Modern studies have expanded on this phenomenon by including generative AI in the list of applications where offloading has been observed to take place in terms of poor memory retention and passive consumption for people offloading tasks to an AI assistant [25], [36].

Self-Regulated Learning and Metacognition

The SRL theory views successful learners as proactive individuals who plan, monitor, and evaluate their cognitive, motivational, and behavioral involvement in a task [43], [44]. Metacognitive monitoring, i.e., the capacity to reflect on and control one's cognitive process, is seen as a mandatory prerequisite for converting mere task performance at the surface level into real learning and independent decision-making [41], [46]. There is an increasing amount of research that confirms the existence of a specific concept known as metacognitive laziness, whereby excessive use of generative AI in evaluating tasks, such as assessing the quality of drafts based on the criteria of the task, makes students detach themselves from their metacognitive monitoring [8], [12], [15].

Automation Bias

The phenomenon of automation bias refers to the well-established bias whereby human users tend to trust the advice coming from an automated source more than their own judgment, even in light of contrary information [33], [92]. Examining the literature in health care, public administration, and decision making in general, one can see that the trend is that automation bias is increased by system responses that are fluent and confident, and that training alone cannot overcome it [35], [57]. With chatbots that are based on artificial intelligence providing smooth and convincing responses, students will find themselves falling prey to automation bias more easily compared to computer systems of the past [37], [39].

The Critical Thinking Paradox: A Conceptual Model

Using the ideas mentioned previously, this paper proposes the Critical Thinking Paradox Model that will help understand how artificial intelligence is able to improve and impair critical thinking at the same time. It is a model that synthesizes the two opposite paths found in previous studies. The model treats a student's observed critical-thinking performance as a function of a single latent construct, the *AI Reliance Index* (R), moderated by AI literacy and instructional scaffolding.

Let R ∈ [0, 1] be the ratio of the number of cognitive academic subtasks (information seeking, writing, analyzing, evaluating) that a student assigns to an artificial intelligence system instead of accomplishing on his/her own, as specified by relevant literature on the tracing of prompts [7], [9]. An approximate estimation of R for n tasks can be represented in the following way:

$$R = (\Sigma_{i=1}^{n} w_i \cdot d_i)/n \quad (1)$$

where $d_i = \{0, 1\}$, depending on whether sub-task i was delegated to the AI system, and w_i is the weighting of task importance. As per the cognitive load theory [7], cognitive load of a task is estimated in the following way:

$$CL_total = CL_intrinsic + CL_extraneous + CL_germane \quad (2)$$

As per research results, the impact of artificial intelligence on the task mainly reduces CL_extraneous, whereas the reduction of CL_germane increases non-linearly past a certain level of reliance [8], which accounts for the reverse relationship between reliance and performance of critical thinking presented in Fig. 2.

The relationship can be expressed as follows:

$$CT(R) = CT_0 + a \cdot R \cdot e^{(-b \cdot R)} \quad (3)$$

where CT(R) is the indexed critical-thinking skill, CT_0 is the critical thinking skill without any assistance from AI technology, and $a, b > 0$ are experimentally measurable parameters of augmentation and decay. Equation (3) gives a unique interior maximum $R *= 1/b$, corresponding to the value of R that maximizes the augmentative effect of AI assistance in terms of scaffolding, feedback, and demonstration of the model before offloading begins to play the dominant role. Figure 2 presents this curve based on the chosen parameters $(a = 2.1, b = 3.1)$, consistent with the qualitative behavior of the phenomenon described in [8]–[12], [19], with the optimum being $R *\approx 0.32$, meaning that the AI assistance should be applied for one-third of cognitively challenging tasks.

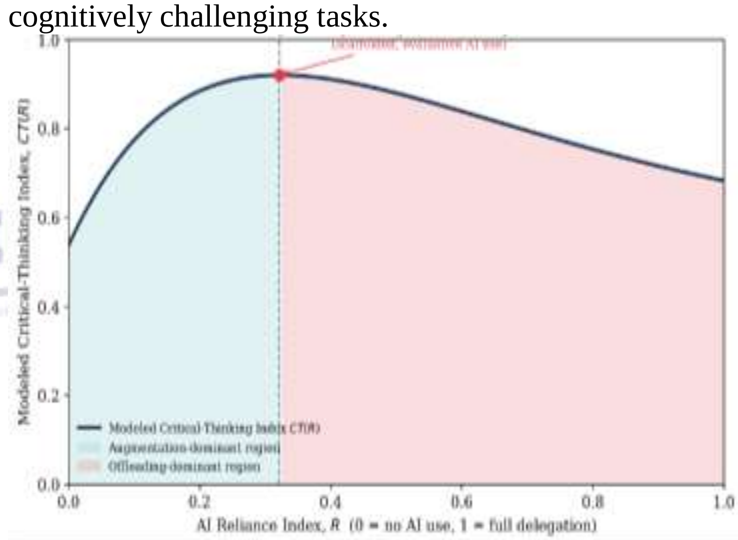

Fig. 2. Theoretical relationship between AI reliance and critical-thinking performance (Critical Thinking Paradox Model), showing an interior optimum beyond which offloading effects dominate augmentation effects

However, the cognitive offloading function may also be modeled as a saturating function of total accumulated exposure to AI over time (t), as cognitive offloading occurs faster when the frequency of exposure is higher; however, after some time, the increase in cognitive offloading becomes marginal due to habituation effects [7], [9].

$$CO(t) = CO_max \cdot \left(1 - e^{(-k \cdot t)}\right) \quad (4)$$

$CO(t)$is the cumulative amount of cognitive offloading, $CO_{\max}$is the maximum amount of cognitive offloading that it can achieve in the long term, and k is the habituation rate, which is the rate at which cognitive offloading occurs with increasing usage of the AI. According to Equations (3) and (4), the optimal point of reliance, $R*$, does not remain constant but gradually moves towards lower sustainable reliance due to increased AI exposure. This trend is supported by longitudinal findings that demonstrate the reduction in neural connectivity and ownership of essays after four months of continuous exposure to large language models [19].

Two critical factors that affect this dynamic have been previously mentioned in other researches: AI literacy and instructional scaffolding. To incorporate the impact of these moderators, the regression equation below is used:

$$CT = \beta_0 + \beta_1 R + \beta_2 L + \beta_3 (R \times L) + \beta_4 S + \varepsilon \quad (5)$$

Where L is the measure of AI literacy, S is the index of instructional scaffolding intensity, and ε is the error term.. A positive, significant β_3 would indicate that AI literacy attenuates the erosive effect of reliance on critical thinking—a hypothesis directly supported by survey evidence that AI literacy socially buffers the fatigue-mediated relationship between AI dependence and critical thinking [9]. Finally, regarding meta-analytic aggregation of the effect sizes reported from the various design types analyzed in Section V of this paper, the following Hedges' g meta-analytic effect size formula is adopted:

$$g = \left[(M_e - Mc)/SD_p ooled\right] \cdot [1 - 3/(4df - 1)] \quad (6)$$

where M_e and Mc represent the mean critical thinking outcome score of the AI-assisted group and the control group respectively, and SD_pooled is the pooled standard deviation.

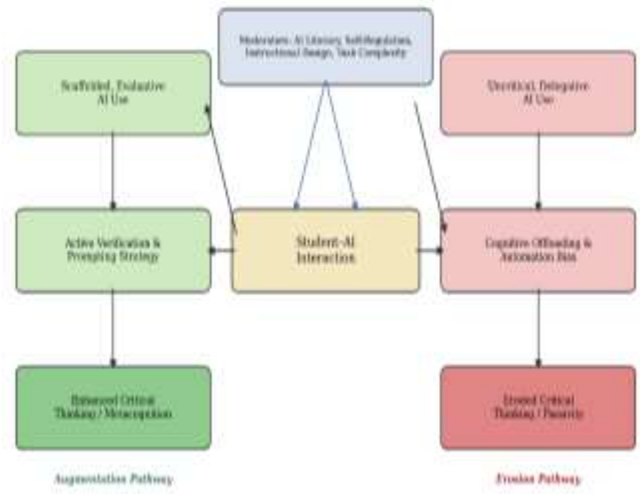

Fig. 4. A concept model that demonstrates the effect that AI may have on students' cognitive development via two paths with results contingent upon such variables as AI literacy, self-regulation, instruction design, and task difficulty

Methodology

The current paper utilizes systematic narrative synthesis approach as a tool for synthesis of heterogeneous research designs (such as randomized controlled trials, quasi-experimental studies, cross-sectional surveys, neurophysiological researches, qualitative studies), which cannot be summarized into a quantitative meta-analysis due to lack of sufficient contextual information, based on systematic literature review guidelines for applied sciences [16], [22], [94]. Relevant sources have been selected by searching combinations of the following keywords: artificial intelligence, generative AI, ChatGPT, critical thinking, cognitive offloading, metacognition, higher order thinking limited to the period from 2011 to 2026 in major academic and preprint databases. The sources have been included in case they provided empirical results, theoretical framework, or systematic/meta-analytical synthesis related to cognitive effects of artificial intelligence on learners. Overall, eighty nine sources have been

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 2, 2026

found that meet these requirements and are cited in the current paper; additional works in the field of educational psychology and decision science have been included for the purpose of forming the theoretical framework, resulting in 92 citations.

Extracted studies were coded thematically into six categories—academic performance, higher-order thinking, cognitive offloading, metacognition, creativity/divergent thinking, and academic integrity—and each study's reported direction of effect was classified as predominantly positive, mixed/context-dependent, or predominantly negative with respect to critical-thinking-relevant outcomes. This thematic distribution is summarized in Fig. 1. In order to facilitate more clear comparative purposes, fifteen studies which directly assessed critical thinking or very similar higher-order cognition measures were chosen for further scrutiny presented in **Table I**. Such studies were chosen to reflect a variety of research approaches, both experimental and quasi-experimental, survey, and neurophysiological, while placing emphasis on the most up-to-date research in the field.

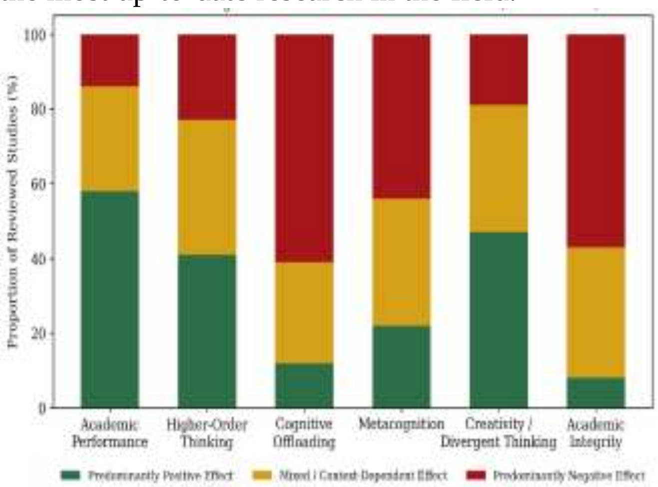

Fig. 5. Effectiveness of AI tool usage with regard to six critical-thinking-related themes from the literature reviewed.

Findings

Augmentative Effects on Academic Performance and Motivation

In all studies considered here, it is the positive impact of generative AI on students' performance and motivation for learning that was found most consistently. In a meta-analysis of 69 experimental studies conducted during 2022 and 2024, ChatGPT has been shown to improve students' academic performance, facilitate their affective and motivational results, support higher-order thinking, and decrease cognitive effort while performing learning tasks. Yet, it did not bring any improvement of students' self-efficacy [14], [17]. Another meta-analysis provides the same findings: the effect size of AI integration is positive and large in relation to academic performance ($g = 0.924$) [11]. The similar conclusions are provided by a comprehensive review of 29 experimental and quasi-experimental studies, demonstrating the significant positive impact of generative AI on higher-order thinking, especially if a student actively formulates prompts [48].

Erosive Effects: Cognitive Offloading and Metacognitive Laziness

Nevertheless, numerous scientifically sound findings reveal the disadvantages related to cognitive abilities. A randomized control study of argumentative essay writing revealed that constant reliance on the power of generative AI resulted in **metacognitive laziness** since students started depending on AI to assess the quality of the assignment, instead of doing it by themselves according to the criteria of the task [8]. Moreover, according to the results obtained, the students who relied on ChatGPT during the data gathering process exhibited lower reasoning and argumentation skills compared to those using conventional techniques of information gathering requiring intellectual efforts [8]. Such findings were revealed in the survey-based studies as well, which showed that people using AI regularly are less likely to engage in deep learning or to apply critical thinking while facing novel and challenging situations [9]. This approach can result in **cognitive inertia** in the long run [9].

Another compelling piece of evidence is a controlled study of neurophysiology conducted at MIT. In this experiment, the participants were required to create their essays over three sessions in three different groups – by using the large language model (LLM), search engines, or nothing else [19]. Brain activity was tracked using electroencephalography (EEG). The group without the use of any tools demonstrated the highest level of brain connectivity; the users of the search engine were moderately active. The users of the large language model demonstrated the lowest levels of brain connectivity. Moreover, they did not feel as much ownership of their essays and had difficulties with memorizing and citing correctly what they had written before. When LLM users were asked to write an essay without any AI help in the fourth session, they still performed lower than the participants that did not use any help from the start [19]. While the later commentaries criticized the rather small size of the sample used in the study and limitations in describing it, the commentaries did not dispute the fact that people using the large language models demonstrate lower levels of brain activity [20].

The above result is based on previous findings related to the Google effect, where individuals who expected information to be available at a later point on the internet had reduced ability to recall information but increased ability to recall where to look for that information [21]. Such results were supported by later findings and meta-analysis of online searches [23]. The current findings indicate that this result is true for information searching using AI, as dependence on AI for information leads to reduced learning due to passive engagement with the information [25], [36].

Programming, Writing, and Domain-Specific Effects

Even studies in specific domains confirm the offloading concept. The meta-study on the effects of generative AI on productivity and learning in programming education indicates that students utilizing AI tools to produce or correct their code often overlook essential parts of the learning process, like debugging or learning programming concepts, which is similar to offloading in other domains [10]. Vocational education research similarly documents a shift *"from co-design to metacognitive laziness"* as AI scaffolding intensity increases without corresponding critical-evaluation requirements [12]. In writing instruction, graduate-level research finds that AI's contribution to research competence is fully mediated by critical thinking, and that intrinsically motivated, evaluative AI use supports scientific creativity, whereas extrinsically motivated, default-acceptance use weakens innovative thinking [52].

Creativity and Divergent Thinking

Overall, results concerning creativity tend to be somewhat more positive toward AI. The generative form of AI might augment divergent thinking by making students actively formulate questions for getting specific feedback, which is known to increase proactive engagement in learning [48]. When comparing human and artificial intelligence in terms of their divergent and convergent thinking test scores, modern large language models are able to achieve either equal or superior originality to average human scores on some tasks, whereas those humans who perform the best still outperform AI on other tasks [50]. On the other hand, the literature related to collective creativity reveals that while the individual benefits of generative AI for creativity can occur, this improvement may come with a decrease in the diversity of ideas generated by the population of users because many students choose the same AI suggestions [53].

Academic Integrity as a Boundary Condition

Academic integrity issues directly pertain to the practice of critical thinking, as the use of AI without disclosure eliminates the process of evaluation and revision, which is essential to the exercise and assessment of critical thinking. The proportion of disciplines that experienced academic misconduct cases related to AI increased from about 48% of institutions reporting such cases in 2022-23 to about 64% in 2024-25 [57]; at the same time, survey research shows that while a vast majority of college students in the United States reported their knowledge of academic integrity policies, only about one fifth of them view unauthorized use of AI as a violation of those policies [59]. This highlights the moral dissonance that exists between the institutional requirements and the opinions of students.

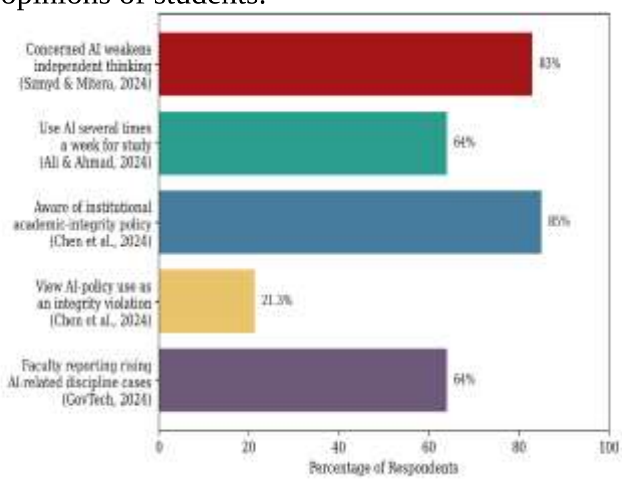

Fig. 6. Selected survey-reported indicators of AI use and perceived impact on critical thinking and academic integrity.

Comparative Review of Representative Studies

Table I summarizes fifteen representative studies selected for methodological diversity and direct relevance to critical-thinking outcomes, drawn from the broader synthesis presented in Section V.

Study	Design / Sample	Domain	Reported Direction of Effect on Critical Thinking / Cognition	Key Finding
Kosmyna et al. [19], 2025	RCT, EEG; n=54 (18 in follow-up)	Essay writing	Negative (neural)	LLM writers show weakest brain connectivity and lowest essay ownership vs. search-engine and unaided groups.
Deng et al. [14], 2024	Meta-analysis, 69 experimental studies	Cross-domain	Positive (performance) / Mixed (HOT)	ChatGPT improves performance and motivation; effect on self-efficacy non-significant.

Study	Design / Sample	Domain	Reported Direction of Effect on Critical Thinking / Cognition	Key Finding
Szmyd & Mitera [1], 2024	Survey, university students	Cross-domain	Mixed	83% of students fear AI weakens independent thinking despite valuing its analytical support.
Sparrow, Liu & Wegner [21], 2011	4 lab experiments	Memory / recall	Negative (recall)	Expected future computer access reduces information recall; foundational Google-effect study.
Fan et al. (in [12]), 2025	RCT, writing task comparison	Writing	Negative (metacognition)	ChatGPT-supported writers show reduced self-regulated learning and metacognitive engagement.
Iqbal et al. [11], 2025	PLS-SEM survey, n=465	Teacher education	Positive	GenAI use predicts academic achievement via shared metacognition and cognitive offloading (g=0.924 meta-

Study	Design / Sample	Domain	Reported Direction of Effect on Critical Thinking / Cognition	Key Finding
				analytic effect cited).
Gonsalves [26], 2024	Qualitative, MSc students, 4 weeks	Marketing education	Mixed	AI both enhances and challenges critical thinking across cognitive, affective, and metacognitive domains.
Meta-analysis (PMC) [48], 2025	29 experiments/quasi-experiments	Cross-domain	Positive (HOT)	Gen-AI improves higher-order thinking when tasks require active query constru
				ction.
Doshi & Hauser [53], 2024	Field experiment	Creative writing	Mixed	Individual creativity rises but collective idea diversity falls with shared AI use.
Cognitive offloading study [7], 2026	Longitudinal, prompting analysis	Cross-domain	Negative (offloading)	Prompting patterns reveal systematic cognitive offloading over repeated AI-assisted sessions.
Programming meta-analysi	Meta-analysis	Computer science	Negative (learning) /	AI boosts code-task

Study	Design / Sample	Domain	Reported Direction of Effect on Critical Thinking / Cognition	Key Finding
s [10], 2026			Positive (productivity)	completion but reduces internalization of debugging and reasoning skills.
Learners' AI dependence study [9], 2025	Survey, structural model	Cross-domain	Negative (moderated)	AI dependence reduces critical thinking via fatigue; AI literacy buffers the effect.
Bloom-taxonomy classroom study [44],	Mixed methods, n=72	K-12 / Architecture	Positive	AI-supported project-based learning improv
2024				es critical thinking vs. control group.
Chen et al. (in [59]), 2024	Survey, U.S. college students	Academic integrity	Negative (norms)	Only 21.3% of students regard undisclosed AI use as an integrity violation despite policy awareness.
Retracted meta-analysis [16], 2025	Meta-analysis (subsequently retracted)	Cross-domain	Mixed	Reported mixed learning-performance and higher-

Study	Design / Sample	Domain	Reported Direction of Effect on Critical Thinking / Cognition	Key Finding
				order-thinking effects; illustrates literature's methodological volatility.

TABLE I. Comparative summary of fifteen representative studies on AI and student critical thinking, ordered to illustrate methodological and directional diversity across the reviewed literature.

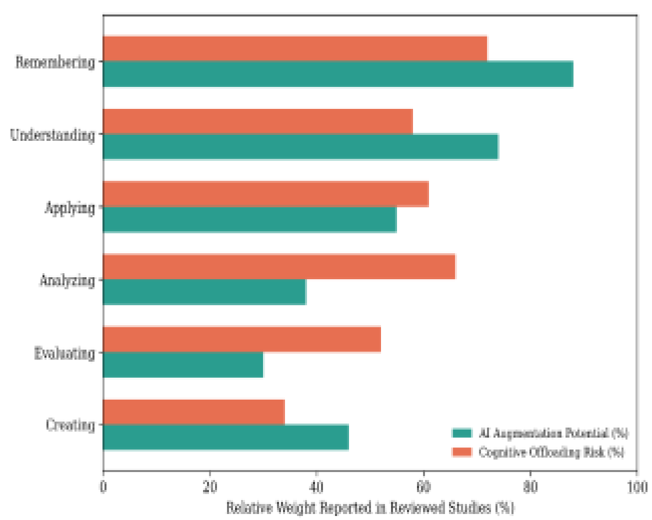

Fig. 7. Revised Bloom's taxonomy: AI augmentation potential versus cognitive offloading risk at each cognitive level, synthesized from the reviewed literature [26], [27], [31], [44].

Moderating Factors and Discussion

AI Literacy as a Protective Factor

The most obvious moderator revealed in the reviewed literature is AI literacy, which represents students' knowledge of how AI technology works, its limits, and proper verification techniques. Using structural-equation modeling, it has been found that AI literacy serves as a social buffer between fatigue-mediated AI dependence and a reduction in critical thinking [9]. Further, scoping reviews of the AI-literacy framework have revealed that curricula focused on higher-order engagement with AI technology, through analyzing, evaluating, and generating content using AI instead of just receiving it passively, yield better results in terms of long-term critical thinking development [28], [29]. This conclusion is supported by Equation (5) presented in Section III, where the interaction parameter $\beta_3(R \times L)$is expected to be positive, since AI literacy will move a student's position on the reliance curve (see Fig. 2) towards the optimum R*.

Instructional Design and Task Structure

The task structure greatly impacts the outcomes. Research which involves students constructing and refining the prompts themselves instead of accepting initial output from the AI shows beneficial effects on higher order thinking skills [48], while research allowing single-turn offloading without any support shows most notable offloading and metacognitive laziness effects [8], [12]. This parallels the long tradition in instructional design which warns about the necessity to provide proper scaffolding in the use of discovery tools in order not to promote shallow learning [84], and aligns with engagement-based models tying engagement levels directly to outcomes [85].

Disciplinary and Task-Type Variation

There are important differences between the effects depending on both discipline and kind of tasks. In programming learning there appears to be a tradeoff where performance increases while internalizing concepts becomes harder [10]. Creative writing and divergent thinking, on the other hand, are more likely to yield positive individual results and incur the costs of diversity within the group [50], [53]. Research methods and argumentation, which require the greatest amount of independent thinking, demonstrate the greatest detrimental effects due to uncritical use of AI [8], [52].

Reconciling the Augmentation and Erosion Narratives

On balance, then, the data does not suggest an affirmative answer to either augmentation or erosion, but rather an affirmation of the conditional framework described in Section III, whereby AI usage is augmentative of critical thinking skills at lower and moderate levels of dependency, especially when coupled with active prompting, evaluative correction, and sufficient AI literacy, and is erosive when dependency moves into a full delegation condition, especially with regard to the evaluative and metacognitive subprocesses [8], [9], [19]. This is because, under the reframed research question, the issue for education is no longer how to allow the use of AI but rather how to use it.

Pedagogical Recommendations

From the results of this synthesis of literature, five recommendations arise on how to preserve and develop critical thinking in AI-infused learning environments.

1) First, there needs to be an explicit focus on teaching AI literacy. The curricula should include training in AI-generated text production processes, their failure points (hallucination, bias, overconfidence), and the way texts can be verified, as the only proven effective protective moderator against offloading effect is AI literacy [9], [28].

2) Second, assessment should be rethought to incorporate visible process evaluation. Process-visible assessment, where draft versions with revision records, oral defense of the work done with AI, and in-class reasoning checks take place, retains possibilities of evaluating and developing evaluative critical thinking skills that cannot be assessed by single submission formats [58], [60].

3) Third, there should be assignments encouraging evaluative AI engagement. Iterative prompting, AI-generated drafts critique, and error identification in AI output bring students into the augmentation dominance area of the reliance curve (see Fig. 2) [48], [8].

4) Calibrate AI scaffolding to task type and cognitive level. Given the asymmetry documented in Fig. 3, instructors should permit or encourage AI support for lower-tier tasks (remembering, understanding) while reserving analyzing, evaluating, and creating tasks for greater independent effort, consistent with revised-Bloom's-taxonomy applications in AI-supported classrooms [27], [31].

5) Monitor cumulative exposure, not only single-session use. Because Equation (4) implies that offloading effects compound with cumulative exposure time, institutions should track longitudinal patterns of AI reliance across a program of study rather than evaluating AI use only at the level of individual assignments [7], [19].

Limitations

There are numerous limitations to this synthesis. First, since it is a narrative and not a quantitative meta-analysis, there was no statistical pooling of effect sizes from the diverse methods analyzed above, while Equation (6) has been provided only as an example of methodology for a future quantitative synthesis. The second limitation relates to the predominance of studies carried out between 2023 and 2026 in this review, which is a time of rapid changes in the development of models; results that rely on older versions of AI or those that were subsequently superseded may lack generalizability [19]. The third limitation stems from the fact that a widely cited meta-analysis used in the larger thematic synthesis was recently retracted [16]. The fourth limitation relates to the prevalence of cross-sectional or short-term studies; longitudinal evidence from a complete degree program is sparse, and thus the mechanisms discussed above as reflected in Eq. (4) are hard to interpret. Finally, the values of the coefficients $(a = 2.1, b = 3.1)$ used for illustrating the Critical Thinking Paradox Model in Fig. 2 are purely illustrative and approximate and reflect the pattern of results obtained in the reviewed literature but not any primary data analysis

Future Research Directions

Several areas deserve greater attention in future research, many of which have already been identified as gaps in the existing literature. [2].

1) Longitudinal cohort study. In the future, it is important to conduct longitudinal studies where the same group of students is observed for several years to reveal the impact of continuous usage of AI technologies on the development of critical thinking skills. The described approach will allow finding out whether the discovered shifts are temporary or persistent over time, according to Equation (4) [2], [17].

2) Comparative studies with various AI technologies. Current neurophysiological and empirical researches are mostly concentrated on one type of AI model. It is necessary to conduct comparative studies using various modern AI technologies in order to find out whether the tendency of using AI and its effect on the process of thinking depend on certain AI technology or not [19].

3) Empirical testing of the Critical Thinking Paradox Model:** It would be helpful for future research to concentrate on the empirical verification of the suggested model with the help of real-life data. The coefficients of Equations (3)-(5) can be estimated based on actual people's reliance on artificial intelligence in various situations in order to test whether the existence of R* as well as the model itself can be proven.

4) Research in different countries and neurosciences: Additional studies in different countries are required since all the previous research was mostly done in a few countries. Such research would help to make the results more generalizable. Furthermore, it would be interesting to do some neuroscientific research in order to prove whether the mentioned changes in the neural connections are long-term or they are just situational

Conclusion

However, the impact of artificial intelligence on students' critical thinking skills cannot be seen as uniformly positive or negative—it depends on the extent of reliance, the nature of the tasks, AI proficiency, and instructional support. Based on synthesized information from more than eighty sources, the current paper has formulated a novel concept—a Critical Thinking Paradox Model—which reflects this conditionality through the explicit definition of a functional dependence between AI reliance and critical-thinking performance in mathematical terms. Synthesized evidence suggests clear academic and motivational improvements, inconclusive and somewhat more positive effects on creativity, and unambiguous cognitive offloading, metacognitive laziness, and lower neural activity in the conditions of high-level, non-scaffolded reliance on AI. In other words, AI should not be viewed as something which negatively impacts the students' critical thinking skills or as a tool which positively influences them—independently from the task. Instead, it is a cognitive technology that, depending on the learner, the task, and the way it is used, may have certain consequences for the students' cognitive activities. In terms of the growing implementation of generative AI into educational practice, the major task ahead for researchers and educators is how to organize learning environment in such a way that keeps the learners within the augmentation-dominant area of cognitive development

References

[1] K. Szmyd and E. Mitera, "The impact of artificial intelligence on the development of critical thinking skills in students," European Research Studies Journal, vol. 27, no. 2, pp. 1022–1039, 2024.

[2] "The impact of AI on students' reading, critical thinking, and problem-solving skills," American Journal of Education and Information Technology, 2025.

[3] "The use of AI in improving student's critical thinking skills," ResearchGate preprint, 2024.

[4] National Science Teaching Association, "To think or not to think: The impact of AI on critical-thinking skills," NSTA Blog, 2025.

[5] "The impact of artificial intelligence on human thought," arXiv:2508.16628, 2025.

[6] "Evaluating the impact of AI on the critical thinking skills among the higher education students by combining the TAM model and critical thinking theory," Frontiers in Education, 2025.

[7] "Cognitive offloading in student–AI collaboration: A longitudinal analysis of prompting strategies," ScienceDirect, 2026.

[8] "Distinguishing performance gains from learning when using generative AI," arXiv:2605.13731, 2026.

[9] "Learners' AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy," ScienceDirect, 2025.

[10] "A meta-analysis of the effect of generative AI on productivity and learning in programming," arXiv:2605.04779, 2026.

[11] J. Iqbal, Z. F. Hashmi, M. Z. Asghar, and M. N. Abid, "Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among preservice teachers," Scientific Reports, vol. 15, 2025.

[12] "From co-design to metacognitive laziness: Evaluating generative AI in vocational education," arXiv:2512.12306, 2025.

[13] "Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration," ScienceDirect, 2025.

[14] Y. Deng et al., "Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies," Computers & Education, 2024.

[15] "The cognitive impact of ChatGPT in higher education: A systematic review of critical and creative thinking outcomes," ScienceDirect, 2026.

[16] "The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis," Humanities and Social Sciences Communications, 2025 (retracted).

[17] "ChatGPT in higher education: A systematic review of opportunities, perceptions, and challenges," Preprints.org, 2025.

[18] "ChatGPT in the higher education: A systematic literature review and research challenges," ScienceDirect, 2024.

[19] N. Kosmyna, E. Hauptmann, Y. T. Yuan, J. Situ, X.-H. Liao, A. V. Beresnitzky, I. Braunstein, and P. Maes, "Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task," arXiv:2506.08872, 2025.

[20] "Comment on: Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks," arXiv:2601.00856, 2025.
[21] B. Sparrow, J. Liu, and D. M. Wegner, "Google effects on memory: Cognitive consequences of having information at our fingertips," Science, vol. 333, no. 6043, pp. 776–778, 2011.
[22] "Google effect," Wikipedia, accessed 2026.
[23] "Google effects on memory: A meta-analytical review of the media effects of intensive internet search behavior," Frontiers in Public Health, 2024.
[24] "The memory paradox: Why our brains need knowledge in an age of AI," arXiv:2506.11015, 2025.
[25] "A review of the negative effects of digital technology on cognition," arXiv:2603.10025, 2026.
[26] C. Gonsalves, "Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy," Journal of Marketing Education, 2024.
[27] "Promoting cognitive skills in AI-supported learning environments: The integration of Bloom's taxonomy," Journal of Curriculum Studies, 2024.
[28] "Critical Artificial Intelligence literacy: A scoping review and framework synthesis," ScienceDirect, 2024.
[29] "A critical review of teaching and learning artificial intelligence (AI) literacy: Developing an intelligence-based AI literacy framework for primary school education," ScienceDirect, 2024.
[30] "AI meets Bloom: The role of Bloom's taxonomy in an AI-enabled world," ResearchGate preprint, 2025.
[31] "Enhanced Bloom's educational taxonomy for fostering deeper learning in human–AI interaction (LBET model)," arXiv:2503.19434, 2025.
[32] "Overreliance on AI: Addressing automation bias today," Lumenova AI Blog, 2025.
[33] S. Alon-Barkat and M. Busuioc, "Human–AI interactions in public sector decision making: 'Automation bias' and 'selective adherence' to algorithmic advice," Journal of Public Administration Research and Theory, vol. 33, no. 1, pp. 153–169, 2023.
[34] "Cognitive amplification vs. cognitive delegation in human-AI systems: A metric framework," arXiv:2603.18677, 2026.
[35] "Exploring automation bias in human–AI collaboration: A review and implications for explainable AI," AI & Society, 2025.
[36] "Automation bias in the AI Act: On the legal implications of attempting to de-bias human oversight of AI," arXiv:2502.10036, 2025.
[37] "Biased AI improves human decision-making but reduces trust," arXiv:2508.09297, 2025.
[38] "What is automation bias? How to avoid its pitfalls," TechTarget, 2025.
[39] "Evaluating physician-AI interaction for cancer management: Paving the path towards precision oncology," arXiv:2404.15187, 2024.
[40] "Examining human reliance on artificial intelligence in decision making," Scientific Reports, 2026.
[41] "A review of self-regulated learning and self-efficacy," ERIC document EJ1299952.
[42] "Self-regulated learning and the meta-learner," Thinking Matters, 2025.
[43] E. Panadero, "A review of self-regulated learning: Six models and four directions for research," Frontiers in Psychology, vol. 8, art. 422, 2017.
[44] "Self-regulated learning and academic achievement: An overview," ResearchGate, 1990.

[45] "Review of the concept 'self-regulated learning': Defined, developed, and assessed," ERIC document EJ1415007.
[46] "Self-regulated learning, self-determination theory and teacher candidates' development of competency-based teaching practices," Smart Learning Environments, vol. 9, 2022.
[47] "Self-regulated learning: A key factor in the effectiveness of online learning for second language learners," Frontiers in Psychology, 2022.
[48] "Does generative artificial intelligence improve students' higher-order thinking? A meta-analysis based on 29 experiments and quasi-experiments," PMC, 2025.
[49] "GenAI as a cognitive mediator: A critical-constructivist inquiry into computational thinking in pre-university education," Frontiers in Education, 2025.
[50] "Generative artificial intelligence models outperform students on divergent and convergent thinking assessments," Scientific Reports, 2025.
[51] "Elementary school students' and teachers' perceptions towards creative mathematical writing with generative AI," arXiv:2409.06723, 2024.
[52] "The impact of generative AI use on graduate students' research competence: The mediating role of critical thinking and the moderating role of research self-efficacy," PMC, 2026.
[53] A. R. Doshi and O. P. Hauser, "Generative AI enhances individual creativity but reduces the collective diversity of novel content," Science Advances, vol. 10, no. 28, 2024.
[54] "The paradox of creativity in generative AI: High performance, human-like bias, and limited differential evaluation," Frontiers in Psychology, 2025.
[55] "Impacts on creativity, critical thinking, and problem-solving," Journal of Educational Research and Practice, Walden University, 2025.
[56] "Understanding, protecting, and augmenting human cognition with generative AI: A synthesis of the CHI 2025 Tools for Thought workshop," arXiv:2508.21036, 2025.
[57] "AI plagiarism statistics 2025: Transforming academic integrity," ArtSmart AI Blog, 2026.
[58] A. Kovari, "Ethical use of ChatGPT in education—Best practices to combat AI-induced plagiarism," Frontiers in Education, vol. 9, art. 1465703, 2025.
[59] "Chat or cheat? Academic dishonesty, risk perceptions, and ChatGPT usage in higher education students," Social Psychology of Education, 2025.
[60] "Ensuring academic integrity in the age of ChatGPT," ERIC document EJ1460216.
[61] J. P. P. Miranda, R. P. Manalese, M. A. A. Castro, R. P. M. Viado, V. G. M. Maniago, R. M. Galapon, J. G. Rivera, and A. B. Martinez Jr., "Plagiarism or productivity? Students' moral disengagement and behavioral intentions to use ChatGPT in academic writing," in Proc. 2025 Int. Workshop on Artificial Intelligence and Education (WAIE), 2025, pp. 383–387.
[62] "Relationship between the use of ChatGPT and plagiarism in higher education: The influence of gender, age and previous academic results," Innovative Higher Education, 2025.
[63] "Survey on plagiarism detection in large language models: The impact of ChatGPT and Gemini on academic integrity," arXiv:2407.13105, 2024.
[64] B. S. Bloom, Ed., Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain. New York, NY, USA: David McKay, 1956.

[65] L. W. Anderson and D. R. Krathwohl, Eds., A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York, NY, USA: Longman, 2001.
[66] R. H. Ennis, "A logical basis for measuring critical thinking skills," Educational Leadership, vol. 43, no. 2, pp. 44–48, 1985.
[67] P. A. Facione, Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report). Millbrae, CA, USA: California Academic Press, 1990.
[68] R. Paul and L. Elder, The Miniature Guide to Critical Thinking: Concepts and Tools, 5th ed. Dillon Beach, CA, USA: Foundation for Critical Thinking, 2008.
[69] D. Kahneman, Thinking, Fast and Slow. New York, NY, USA: Farrar, Straus and Giroux, 2011.
[70] J. Dewey, How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process. Boston, MA, USA: D. C. Heath, 1933.
[71] E. F. Risko and S. J. Gilbert, "Cognitive offloading," Trends in Cognitive Sciences, vol. 20, no. 9, pp. 676–688, 2016.
[72] D. M. Wegner, "Transactive memory: A contemporary analysis of the group mind," in Theories of Group Behavior, B. Mullen and G. R. Goethals, Eds. New York, NY, USA: Springer, 1987, pp. 185–208.
[73] UNESCO, AI and Education: Guidance for Policy-Makers. Paris, France: UNESCO Publishing, 2021.
[74] OECD, Artificial Intelligence and the Future of Skills. Paris, France: OECD Publishing, 2023.
[75] J. Zhang and G. Tur, "ChatGPT in education: Reduced abilities, decreased interaction, and misinformation risk," as cited in Humanities and Social Sciences Communications meta-analysis, 2024.
[76] W. Ali and A. Ahmad, "Shortcomings of ChatGPT in education: Incorrect answers, plagiarism, and technology dependence," 2024.
[77] R. Abu Khurma et al., "Overreliance and bias risks of ChatGPT in student learning," 2024.
[78] J. Dempere et al., "The limitations of ChatGPT in solving complex reasoning problems: A systematic literature review," 2023.
[79] C. K. Lo, "The impact of ChatGPT on education: A systematic review," 2024.
[80] N. Selwyn, Should Robots Replace Teachers? AI and the Future of Education. Cambridge, U.K.: Polity Press, 2019.
[81] F. Zirar, "Enabling and constraining roles of ChatGPT in learning and assessment," 2023.
[82] K. Melisa et al., "Evaluation and independent judgment in ChatGPT-supported critical thinking," 2025.
[83] E. Dimeli and A. Kostas, "Cognitive performance and generative AI in higher education," 2025.
[84] P. A. Kirschner and J. J. G. van Merriënboer, "Do learners really know best? Urban legends in education," Educational Psychologist, vol. 48, no. 3, pp. 169–183, 2013.
[85] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," Educational Psychologist, vol. 49, no. 4, pp. 219–243, 2014.
[86] S. Lund and T. Wang, "Chatting about ChatGPT: How may AI and GPT impact academia and libraries?," Library Hi Tech News, vol. 40, no. 3, pp. 26–29, 2023.
[87] J. Mun, "Cognitive inertia and reduced reflective thinking in habitual AI tool users," 2023.

[88] Y. Wang and X. Fan, "Mixed effects of ChatGPT use on students' learning performance, perception, and higher-order thinking: A meta-analysis," 2025.
[89] H. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A SWOT analysis of ChatGPT: Implications for educational practice and research," Innovations in Education and Teaching International, vol. 61, no. 3, pp. 460–474, 2024.
[90] S. Bearman and R. Ajjawi, "Learning to work with the black box: Pedagogy for a world with artificial intelligence," British Journal of Educational Technology, vol. 54, no. 5, pp. 1160–1173, 2023.
[91] R. F. Kizilcec, "To advance AI use in education, focus on understanding educators," International Journal of Artificial Intelligence in Education, vol. 34, pp. 12–19, 2024.
[92] M. Cummings, "Automation bias in intelligent time critical decision support systems," in Proc. AIAA 1st Intelligent Systems Technical Conf., 2004.
[93] Shaikh, H., Khan, M. S., Mahar, Z. A., Anwar, M., Raza, A., & Shah, A. (2019, March). A conceptual framework for determining acceptance of internet of things (IoT) in higher education institutions of Pakistan. In 2019 international conference on information science and communication technology (ICISCT) (pp. 1-5). IEEE.
[94] Shaikh, H., Jokhio, M. U., Maher, Z. A., Chandio, S., Bin Abdullah, M. M., Raza, A., & Shah, A. (2018). Beyond traditional audits: The implications of information technology on auditing. International Journal of Engineering and Technology (UAE), 7(2), 5-11.
[95] Shaikh, H. (2021). Acceptance of IoT learning among university students at Pakistan. International Journal of Advanced Trends in Computer Science and Engineering.