



Conditional Effects of Generative AI on Metacognitive Learning

Martahani ^{1*}, Rita Retnowati ¹, Surti Kurniasih ¹, Romi Suarti ¹, Indriyani Rachman ²

¹Universitas Pakuan, Bogor, Indonesia

²University of Kitakyushu, Japan

martahani27@gmail.com, ritaretnowati@unpak.ac.id, surti_kurniasih@unpak.ac.id, romisuarti@gmail.com,
indriyani.rachman@unpak.ac.jp

corresponding author: martahani27@gmail.com

ABSTRACT

Generative artificial intelligence (GenAI) has been rapidly adopted in education, yet its impact on students' metacognitive learning remains poorly understood. Most existing reviews focus broadly on learning outcomes or technology acceptance, neglecting specific metacognitive processes, planning, monitoring, evaluation, and reflection that underpin self-regulated learning. This systematic review synthesizes empirical evidence on how GenAI influences metacognitive learning, which pedagogical approaches support it, and the risks that emerge. Following PRISMA 2020 guidelines, a systematic search of Scopus, ERIC, and Google Scholar identified 12 empirical studies published between 2020 and 2026. The findings reveal that GenAI's effect on metacognition is fundamentally conditional: structured scaffolding improves self-regulation, whereas unrestricted use leads to metacognitive laziness, cognitive offloading, and a decline in self-regulation. Four effective pedagogical approaches were identified: metacognitive prompting, guided inquiry with scaffolding, reflective journaling, and feedback literacy instruction. Significant risks include passive acceptance of AI answers, illusion of competence, and unequal benefits for students with weaker feedback literacy. This review presents a Conditional Metacognitive Support framework to guide GenAI integration. For educators, effective implementation requires structured prompts, reflective activities, and explicit feedback literacy instruction.

Keywords: generative AI, metacognition, self-regulated learning, systematic literature review, scaffolding, AI in education

Submitted	Accepted	Published
27 June 2026	19 July 2026	30 July 2026

Citation	:	Martahani., Retnowati , R., Kurniasih, S., Suarti, R., & Rachman, I. (2026). Conditional Effects of Generative AI on Metacognitive Learning. <i>Jurnal PAJAR (Pendidikan dan Pengajaran)</i> , 10(4), 707-721. DOI: http://dx.doi.org/10.33578/pjr.v10i4.1656 .
----------	---	---

INTRODUCTION

The arrival of generative artificial intelligence (GenAI) in education has sparked both enthusiasm and unease. However, the most fundamental question is not how to maximize test scores or prevent cheating, but whether students are still learning to think about how they think when machines can perform much of that thinking for them. GenAI, from ChatGPT to various large language model-based systems, offers unprecedented adaptive capabilities: explaining complex concepts in plain language, providing personalized feedback, and responding to student queries conversationally (Kasneci et al., 2023; Kong & Yang, 2024). Used judiciously, these capabilities could make GenAI a powerful learning partner. Used carelessly, however, they risk eroding the metacognitive capacities that have long underpinned independent learning.

This concern is not nostalgia for older methods. Metacognition, the ability to monitor and regulate one's own thinking processes (Flavell, 1979) has long been recognized as a strong predictor of academic success. Zimmerman (2002) extended this into a framework of self-regulated learning (SRL) comprising three cyclical phases: forethought (planning), performance (monitoring during learning), and self-reflection (evaluation). In traditional learning environments, these phases arise naturally through teacher-student interactions and through the cognitive struggle inherent in learning. GenAI, with its ease of access, removes much of that struggle. When a student can simply ask an AI and receive an instant answer, the opportunity to ask "Do I really understand this?" may never arise.

What is concerning is that recent research shows GenAI's effect on metacognition is far from uniform. On one hand, several studies report enhanced metacognitive engagement when AI is used with appropriate

scaffolding (Xu et al., 2025; Fütterer et al., 2026). On the other hand, findings by Fan et al. (2025) reveal "metacognitive laziness" and cognitive offloading among students who use AI freely without guidance. In extreme cases, students experience an illusion of competence, believing they understand, because the AI explained things convincingly, but when tested independently, their comprehension proves shallow (Yang & Park, 2025). This contradiction suggests that GenAI has no inherent effect on metacognition; its impact depends heavily on how students are guided to interact with it.

Unfortunately, this insight has not been systematically synthesized in the existing literature. Most reviews on AI in education treat metacognition as a secondary outcome or ignore it entirely, focusing instead on what AI can do: improve test scores, reduce cognitive load, or deliver personalized feedback (Zawacki-Richter et al., 2019; Wu & Yu, 2024). As a result, the critical trade-off between enhanced cognitive outcomes and weakened metacognitive monitoring remains invisible. A student might achieve high grades with AI assistance but lose the ability to assess their own understanding, a loss that remains unmeasured if we rely solely on conventional learning metrics.

From these gaps, three fundamental questions emerge. *First*, which metacognitive processes, planning, monitoring, evaluation, or reflection, are most likely to be supported by GenAI, and which are most vulnerable to disruption? *Second*, what pedagogical conditions make GenAI a support rather than a hindrance to metacognition? Scattered findings suggest that unguided use tends to foster passive acceptance, whereas use with metacognitive prompts or reflective scaffolding enhances self-regulation. However, this pattern has not been comparatively analyzed across studies with different designs and measures. *Third*, what are the actual risks of GenAI use for metacognitive development, and how significant are these risks in practice?

This systematic review addresses these questions with four distinct emphases that set it apart from previous reviews. First, it focuses specifically on metacognition, not merely on learning outcomes or SRL broadly. Second, its scope is limited to generative AI, recognizing that earlier generations of educational AI operated differently and may not generalize. Third, it specifically synthesizes evidence on pedagogical approaches, including instructional strategies, prompting techniques, and scaffolding mechanisms, that have been empirically shown to support metacognitive learning. Fourth, it critically examines risks and unintended consequences, a balanced approach rather than an overly optimistic one.

Through this synthesis, I hope to provide a more nuanced understanding of when and how GenAI supports or undermines students' metacognitive development. Accordingly, this review is guided by four research questions:

RQ1: What forms of metacognitive learning processes are facilitated through generative artificial intelligence in education?

RQ2: How does generative artificial intelligence influence students' reflective thinking, self-regulation, and metacognitive awareness?

RQ3: What pedagogical approaches are used to support metacognitive learning processes through generative artificial intelligence?

RQ4: What are the challenges and ethical concerns of using generative artificial intelligence for metacognitive learning?

METHOD

This study employed a Systematic Literature Review (SLR) design to synthesize empirical evidence regarding the relationship between generative artificial intelligence (GenAI) and metacognitive learning. An SLR was selected because the research questions required a structured synthesis of findings across multiple studies rather than evidence from a single investigation. To ensure transparency and methodological rigor, the review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (Page et al., 2021).

Literature searches were conducted between March and April 2026 across three databases: Scopus, ERIC, and Google Scholar. Scopus was selected for its international coverage, ERIC as a specialized education database, and Google Scholar as a supplementary source.

The search strategy combined terms related to generative artificial intelligence and metacognitive learning using Boolean operators. The primary search string was: ("generative AI" OR "GenAI" OR "ChatGPT" OR "large language model" OR "LLM" OR "AI chatbot") AND ("metacognition" OR "metacognitive" OR "self-regulated learning" OR "SRL" OR "reflective thinking" OR "metacognitive awareness").

The search was limited to studies published between 2020 and 2026 to capture research emerging after the rapid development and public adoption of generative AI technologies, particularly ChatGPT. In addition to database searches, backward and forward snowballing techniques were applied to identify additional relevant studies from reference lists and citation tracking. Studies were selected based on predefined inclusion and exclusion criteria presented in Table 1.

Table 1. Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Publication type	Peer-reviewed journal articles	Editorials, opinion papers, dissertations, book reviews
Topic focus	Studies examining GenAI and metacognitive learning or self-regulated learning	Studies focusing only on cognitive achievement or technology acceptance without metacognitive aspects
Participants	Students in formal educational settings	Non-educational or workplace contexts
Study design	Empirical qualitative, quantitative, or mixed-methods studies	Conceptual papers without empirical data
Publication year	2020–2026	Publications before 2020
Language	English or Indonesian language publications	Non-English publications

The study selection process followed the PRISMA stages of identification, screening, eligibility, and inclusion. Duplicate records were removed using Mendeley reference management software. Titles and abstracts were screened according to predefined inclusion and exclusion criteria. Articles considered relevant were further assessed through full-text reading to determine eligibility for inclusion in the review. After screening and eligibility, 12 empirical studies met all inclusion criteria and were included in the final synthesis.

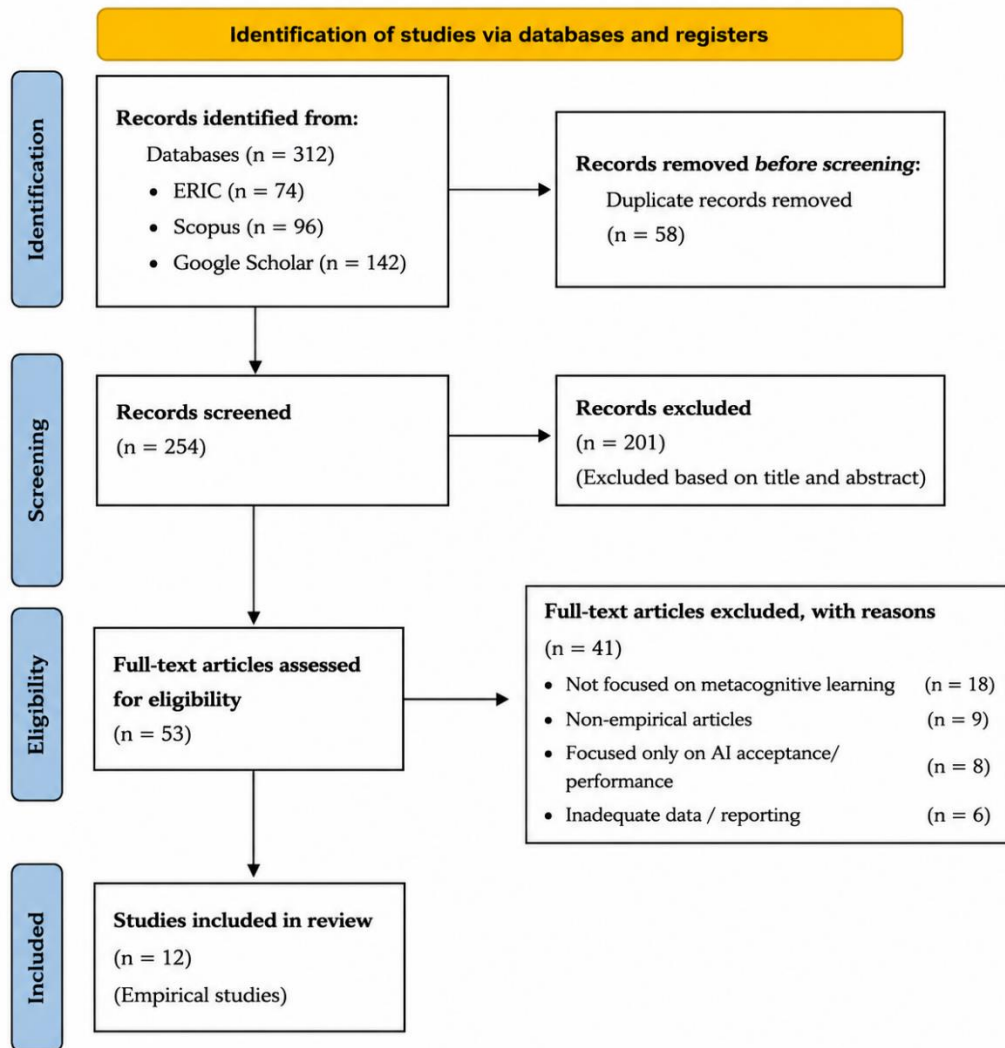


Figure 1. PRISMA Flow Diagram of Study Selection

The methodological quality of the included studies was appraised using the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018), which was selected because the 12 included studies spanned experimental, quasi-experimental, survey, qualitative, and mixed-methods designs that a single-design tool such as JBI or CASP could not accommodate uniformly. Each study was first screened against the two MMAT eligibility questions (clear research question; adequacy of collected data) and then rated against the five quality criteria corresponding to its specific design category (e.g., appropriateness of randomization for quasi-experimental designs, adequacy of qualitative data interpretation, or coherence between quantitative and qualitative components for mixed-methods designs). Two reviewers (the first and third authors) independently appraised each study; disagreements were resolved through discussion until consensus was reached. Studies meeting fewer than 60% of applicable MMAT criteria would have been excluded from the synthesis; all 12 included studies met this threshold, indicating acceptable methodological quality for narrative synthesis. This appraisal was conducted to support the interpretation of the evidence and to identify the overall credibility of the included studies.

A narrative thematic synthesis approach was employed because the included studies showed substantial heterogeneity in research designs, interventions, participant characteristics, and outcome measures, making meta-analysis unsuitable.

The analysis was conducted in three stages, following an operationalized coding protocol to support replicability. First, findings from the included studies were coded inductively through open coding, using a coding manual in which each metacognitive process (planning, monitoring, evaluation, reflection) was defined operationally against Zimmerman's (2002) three-phase SRL model (forethought, performance, self-reflection); for example, "monitoring" was coded only where a study reported learners tracking or checking their understanding during a task, as distinct from planning (goal-setting prior to a task) or evaluation (judging the correctness or quality of an AI-generated response). Coding was performed independently by two authors, and inter-coder agreement was calculated using Cohen's kappa ($\kappa = 0.81$), indicating substantial agreement; remaining discrepancies were resolved through discussion with a third author. Second, related codes were grouped into broader themes corresponding to the research questions. Third, cross-study comparisons were conducted to identify recurring patterns, contradictions, and emerging conceptual relationships regarding the role of GenAI in metacognitive learning.

The final synthesis generated four major thematic categories: (1) metacognitive processes facilitated by GenAI, (2) effects on metacognitive awareness and self-regulated learning, (3) pedagogical approaches supporting metacognitive learning, and (4) challenges and ethical concerns associated with GenAI use in education.

RESULTS AND DISCUSSION

Characteristics of Included Studies

A total of 12 empirical studies met the inclusion criteria and were included in this systematic review. The selected studies, published between 2024 and 2026, examined the relationship among generative artificial intelligence (GenAI), metacognitive learning, and self-regulated learning (SRL) in formal educational settings. Most studies employed experimental, quasi-experimental, mixed-methods, or qualitative approaches and primarily involved university students.

ChatGPT was the most frequently used GenAI tool, although several studies implemented customized GPT systems or AI-supported scaffolding platforms. Across the included studies, the most commonly investigated metacognitive dimensions were planning, monitoring, evaluation, reflection, and self-regulated learning behaviors.

Table 2. Characteristics of Included Studies

No	Authors (Year)	Research Design	GenAI Tool	Main Focus
1	Fan et al. (2025)	Experimental	ChatGPT 4.0	SRL processes and metacognitive monitoring
2	Ng et al. (2024)	Quasi-experimental	SRLbot (GPT-based)	Goal setting and self-evaluation
3	Futterer et al. (2026)	Quasi-experimental	ChatGPT 4.0	Task strategy and self-evaluation
4	Huda et al. (2025)	Mixed-methods	ChatGPT	Planning, monitoring, and evaluation
5	Xu et al. (2025)	Quasi-experimental	ChatGPT with metacognitive prompts	Metacognitive regulation
6	Mendoza et al. (2026)	Survey	ChatGPT	Feedback literacy and reflection
7	Yang and Park (2025)	Mixed-methods	ChatGPT with reflective journaling	Reflective thinking

No	Authors (Year)	Research Design	GenAI Tool	Main Focus
8	Nguyen and Nguyen (2024)	Qualitative	ChatGPT 3.5	Reflective practices and SRL
9	Weng et al. (2024)	Survey	ChatGPT	Forethought, performance, and self-reflection
10	Nazaretsky et al. (2026)	Mixed-methods	ChatGPT	Evaluative judgement and feedback literacy
11	Yilmaz et al. (2026)	Mixed-methods	ChatGPT-4	Feedback perception and SRL
12	Lowry et al. (2025)	Multi-method	Customized GPT	Metacognition and self-directed learning

Eight studies were conducted in higher education contexts, while four were conducted in secondary or K–12 educational settings. The studies demonstrated substantial variation in instructional design, intervention duration, and forms of AI-supported learning activities.

Metacognitive Processes Facilitated by GenAI

The included studies identified four major metacognitive processes facilitated by GenAI-supported learning environments: planning, monitoring, evaluation, and reflection. Table 3 summarizes the major metacognitive processes facilitated by GenAI across the included studies.

Table 3. Summary of Metacognitive Processes Facilitated by GenAI

Metacognitive Process	Reported Outcomes	Representative Studies
Planning and goal setting	Improved learning preparation and strategic planning	Ng et al. (2024); Weng et al. (2024)
Monitoring	Improved monitoring of learning progress	Huda et al. (2025); Xu et al. (2025)
Evaluation	Increased evaluative judgment toward AI-generated responses	Huda et al. (2025); Nazaretsky et al. (2026)
Reflection	Enhanced reflective engagement and metacognitive awareness	Yang & Park (2025); Lowry et al. (2025)

Planning and goal-setting processes were the most consistently reported outcomes. Studies using structured prompts and AI-supported SRL systems found that students became more capable of identifying learning objectives and organizing learning strategies before task completion. Ng et al. (2024) and Weng et al. (2024) reported that GenAI-assisted learning activities supported students' forethought and strategic preparation. Monitoring processes were facilitated through scaffolding and metacognitive prompting. Huda et al. (2025) found that students improved their ability to monitor learning progress and identify ineffective problem-solving strategies during AI-assisted inquiry activities. Similarly, Xu et al. (2025) reported that metacognitive prompts encouraged students to track their understanding while interacting with GenAI systems.

Evaluation processes were strengthened when students were encouraged to assess AI-generated responses critically. Several studies reported improvements in evaluative judgment, particularly when reflective questioning and guided feedback were integrated into learning activities. Reflection was also identified as an important metacognitive process supported by GenAI. Yang and Park (2025) and Nguyen and Nguyen (2024) found that reflective journaling and AI-supported reflection activities enhanced students' metacognitive awareness and reflective engagement. In addition, Lowry et al. (2025) reported significant improvements in students' metacognitive awareness after repeated reflective interactions with a customized GPT system designed to support self-directed learning.

Effects on Metacognitive Awareness and Self-Regulated Learning

The findings across studies indicated that the effects of GenAI on metacognitive awareness and self-regulated learning were conditional rather than uniformly positive. Positive effects were generally observed when AI interaction was accompanied by scaffolding, reflective questioning, or metacognitive guidance. Several studies reported improvements in self-regulated learning behaviors. Fütterer et al. (2026) found that structured AI interaction improved students' task strategies and self-evaluation processes. Similarly, Yilmaz et al. (2026) reported that AI-supported feedback activities enhanced students' perceptions of feedback usefulness and supported self-regulated learning behaviors.

Lowry et al. (2025) found statistically significant improvements in both metacognitive awareness and self-directed learning readiness following repeated interactions with a customized GPT model. Students also reported improvements in learning strategy evaluation, time management, and reflective thinking. Despite these positive outcomes, several studies also identified potential risks associated with unrestricted AI use. Fan et al. (2025) reported reduced reflective effort and increased reliance on AI-generated responses among some students. Other studies noted that students occasionally accepted AI-generated answers without sufficient critical evaluation.

Moderating factors were also identified. Mendoza et al. (2026) found that students with stronger feedback literacy gained greater metacognitive benefits from AI interaction. Weng et al. (2024) reported that learner characteristics such as openness and conscientiousness influenced self-regulated learning outcomes. Comparing across the 12 studies reveals a consistent pattern rather than isolated exceptions. Studies that paired GenAI with explicit metacognitive prompts or structured scaffolding (Xu et al., 2025; Fütterer et al., 2026; Huda et al., 2025) consistently reported gains in self-regulation and monitoring, whereas studies examining unrestricted or minimally guided GenAI use (Fan et al., 2025) consistently reported reduced reflective effort and passive acceptance of AI-generated answers. This pattern held across both higher-education and secondary-school contexts and across quantitative, qualitative, and mixed-methods designs, suggesting that the presence or absence of scaffolding rather than the specific GenAI tool, subject domain, or educational level is the primary driver of divergent outcomes. This cross-study comparison is corroborated by two independent 2025–2026 syntheses: Nguyen, Chen, and Asfani's (2026) comparative review of earlier AI versus GenAI scaffolding mechanisms, and Xia, Liu, Tlili, and Chiu's (2026) systematic review of GenAI-supported SRL design, both of which likewise conclude that structuring, not the technology itself, determines whether GenAI enhances or erodes self-regulation.

Pedagogical Approaches Supporting Metacognitive Learning

Four major pedagogical approaches supporting metacognitive learning through GenAI were identified across the reviewed studies: metacognitive prompting, guided inquiry with scaffolding, reflective journaling, and feedback literacy instruction. The pedagogical approaches identified across the reviewed studies are summarized in Table 4.

Table 4. Pedagogical Approaches Supporting Metacognitive Learning

Pedagogical Approach	Description	Representative Studies
Metacognitive prompting	Prompts supporting planning, monitoring, and evaluation	Xu et al. (2025); Fütterer et al. (2026)
Guided inquiry with scaffolding	Structured inquiry activities supported by AI	Huda et al. (2025)
Reflective journaling	Reflection activities based on AI interaction	Yang & Park (2025)
Feedback literacy instruction	Critical evaluation of AI-generated feedback	Mendoza et al. (2026); Yilmaz et al. (2026)
Reflective conversational scaffolding	GPT-supported reflective dialogue for SDL development	Lowry et al. (2025)

Metacognitive prompting was the most frequently implemented approach. Xu et al. (2025) and Fütterer et al. (2026) integrated prompts encouraging students to plan, monitor, and evaluate their learning during AI-supported activities. Guided inquiry with scaffolding was commonly used to structure AI-supported learning processes. Huda et al. (2025) implemented inquiry stages involving prior knowledge activation, guided exploration, analysis, evaluation, and reflection.

Reflective journaling activities were also frequently reported. Yang and Park (2025) found that repeated reflective interactions with ChatGPT increased students' awareness of their learning strategies and reflective practices. Feedback literacy instruction emerged as another important pedagogical approach. Mendoza et al. (2026), Nazaretsky et al. (2026), and Yilmaz et al. (2026) emphasized the importance of teaching students to interpret and evaluate AI-generated feedback critically. Lowry et al. (2025) also demonstrated that conversational interaction with customized GPT systems can serve as reflective scaffolding, supporting students' metacognitive development and readiness for self-directed learning.

Challenges and Ethical Concerns

Several challenges and ethical concerns associated with GenAI-supported metacognitive learning were identified across the reviewed studies. The most frequently reported concerns involved metacognitive laziness and cognitive offloading. Several studies indicate that excessive reliance on AI-generated responses can reduce students' reflective engagement and independent problem-solving. For example, Fan et al. (2025) reported that some students relied heavily on AI-generated answers and demonstrated lower levels of reflective effort during learning activities.

Over-reliance on AI systems was also identified as a potential risk. Several studies noted that unrestricted GenAI use could weaken students' independent problem-solving and self-regulated learning behaviors over time. Another recurring issue involved the illusion of competence, in which students perceived themselves as understanding content because AI-generated explanations appeared fluent and convincing. This issue was particularly evident among students with lower feedback literacy and weaker critical evaluation skills. Feedback literacy deficits further limited the effectiveness of GenAI-supported learning. Mendoza et al. (2026) found that students with weaker evaluative skills experienced greater difficulty critically assessing AI-generated responses. Additional concerns included data privacy, algorithmic bias, unequal access to AI technologies, and variations in students' readiness to engage in reflective learning practices.

DISCUSSION

The Conditional Nature of GenAI's Metacognitive Effects

The most striking finding across all 12 studies is that GenAI's effect on metacognition is fundamentally conditional. This is not a minor qualification; it challenges the dominant framing of AI in education discourse. When technology advocates proclaim that "AI will transform education," they typically assume that the technology itself drives improvement. The evidence from this review suggests the opposite: GenAI has no inherent pedagogical effect. Its impact depends entirely on how it is used, by whom, and under what conditions.

Consider the contrast between two sets of findings. In Xu et al. (2025) and Fütterer et al. (2026), students who received metacognitive prompts showed significant improvements in self-regulation, while control groups using identical AI tools without prompts showed declines. The same technology, the same tasks, the same duration, only the presence or absence of structured guidance differentiated the outcomes. This pattern cannot be explained by technological determinism. It requires a pedagogical explanation.

This finding resonates with, but also extends, earlier work on educational technology. Clark's (1983) famous claim that "media do not influence learning" was controversial precisely because it challenged the assumption that technology itself drives outcomes. Clark argued that instructional methods, not media, cause learning. The present review suggests a similar principle applies to GenAI: the medium (the AI tool) matters less than the method (how students are guided to interact with it).

However, GenAI introduces a new dimension that earlier technologies did not. Unlike static educational software or even rule-based intelligent tutoring systems, GenAI is conversational and adaptive. It can respond to student queries in unpredictable ways. This interactivity creates both opportunities and risks. The opportunity is that GenAI can engage students in genuine dialogue, potentially supporting deeper reflection. The risk is that the same interactivity can create an illusion, leading students to mistake the AI's coherence for their own comprehension. This concern is reinforced by García-López and Trujillo-Liñán's (2025) systematic review of GenAI ethics in education, which identifies epistemic over-trust in fluent AI output as a recurring risk factor across the reviewed literature, independent of subject domain.

Resolving the Contradiction: Why the Same Tool Produces Opposite Outcomes

How can the same AI tool improve metacognition in some studies and undermine it in others? The answer lies in the mechanism of metacognitive offloading. When students interact with GenAI, they face a decision point: whether to engage in reflective processing or to accept the AI's response as sufficient. This decision is influenced by at least three factors identified across the studies.

First, task demands. When tasks are perceived as difficult or time-pressured, students are more likely to offload cognitive work to the AI (Fan et al., 2025). This is not irrational; it is efficient. The problem is that efficiency in task completion may come at the cost of learning. Students who use AI to complete assignments quickly may perform well on those assignments but fail to develop the underlying knowledge or skills. This explains why Fan et al. found that ChatGPT improved essay scores but not knowledge transfer.

Second, metacognitive skill. Students who lack metacognitive awareness may not recognize when they are offloading. They may believe they understand because the AI explained it well, but when tested independently, their understanding is shallow, the "illusion of competence" documented by Yang and Park (2025). Huda et al.'s (2025) finding that 28% of students passively accepted AI answers without scrutiny illustrates this phenomenon. These students were not being lazy; they were being unaware. They did not know that they should evaluate the AI's responses because they had not developed the metacognitive habit of self-questioning.

Third, pedagogical scaffolding. When students received explicit prompts to plan, monitor, and evaluate, they engaged in more reflective processing (Xu et al., 2025; Fütterer et al., 2026). These prompts externalized the metacognitive questions that students should internalize: "What is my goal?" "Is this response relevant?" "Do I really understand?" Over time, external prompts may become internal habits. This is the classic Vygotskian mechanism of scaffolding moving from the social to the individual plane. The implication is that GenAI, when paired with appropriate scaffolding, can serve as a tool for developing metacognitive capacity rather than merely a crutch for completing tasks.

This analysis leads to a testable hypothesis: GenAI's effect on metacognition follows an inverted-U curve. At low levels of scaffolding, students offload excessively, and metacognition declines. At moderate levels of scaffolding, students learn to regulate their AI interactions, and metacognition improves. At very high levels of scaffolding, students may become dependent on external prompts and fail to internalize metacognitive strategies. The optimal level likely varies by individual, task, and context, and is an important direction for future research. It should be noted that this inverted-U hypothesis is our own theoretical interpretation of the cross-study pattern observed in this review, rather than a synthesized empirical finding; it is broadly consistent with the dose-response pattern reported in Tsakeni, Nwafor, Mosia, and Egara's (2025) bibliometric-systematic review of AI-scaffolded metacognition in STEM classrooms, but it has not yet been tested directly and warrants dedicated empirical investigation.

Extending the Theoretical Framework

The findings extend Zimmerman's (2002) SRL framework in three ways. First, the forethought phase appears most amenable to GenAI support. Across multiple studies, goal setting and planning were consistently improved by GenAI interactions (Ng et al., 2024; Weng et al., 2024). This may be because planning involves articulating explicit goals, a task that GenAI can readily scaffold through structured questioning. The AI asks,

"What do you want to achieve?" and the student must formulate an answer. This process, even if initiated by the AI, engages metacognitive processes.

Second, the performance phase shows the most variable outcomes. Monitoring tracking one's understanding during learning improved under guided conditions but declined under unrestricted use (Fütterer et al., 2026; Xu et al., 2025). This suggests that monitoring is easily disrupted by cognitive offloading. When students can instantly access answers, the perceived need to monitor their own understanding diminishes. The AI becomes a substitute for self-checking.

Third, the self-reflection phase is the most challenging to support. Only studies that explicitly structured reflection through journals (Yang & Park, 2025), structured prompts (Huda et al., 2025), or feedback literacy instruction (Mendoza et al., 2026) showed improvements. Reflection cannot be assumed; it must be engineered.

These findings suggest a new concept: metacognitive load. Just as cognitive load theory describes the limited capacity of working memory, metacognitive load describes the cognitive cost of monitoring and regulating one's own learning. GenAI reduces cognitive load by automating lower-level tasks, potentially freeing capacity for metacognitive processing. However, GenAI may also reduce *metacognitive load* to the point where monitoring no longer feels necessary. When answers come effortlessly, why check if you understand? The optimal level of metacognitive load, enough to engage monitoring, not so much that it overwhelms, is a critical design parameter for GenAI learning environments.

A Proposed Conceptual Framework

Based on the synthesis, we propose a Conditional Metacognitive Support (CMS) framework for GenAI integration. The framework specifies four conditions that must be met for GenAI to support rather than undermine metacognitive learning.

Condition	Description	Evidence
Explicit prompting	Students must receive explicit prompts to plan, monitor, and evaluate	Xu et al. (2025); Fütterer et al. (2026)
Reflective structure	Activities must include structured opportunities for reflection (journals, guided questions)	Yang & Park (2025); Huda et al. (2025)
Feedback literacy	Students must be taught to evaluate AI-generated feedback critically	Mendoza et al. (2026); Nazaretsky et al. (2026)
Graduated autonomy	Scaffolding should fade as students develop metacognitive skills	Implied by Huda et al. (2025); Nguyen & Nguyen (2024)

The CMS framework is not merely a checklist. It posits that these four conditions interact: prompting without reflection may produce compliance without understanding; reflection without feedback literacy may produce journals without insight. The whole is greater than the sum of the parts.

The proposed Conditional Metacognitive Support (CMS) framework contributes a conceptual explanation for why generative AI may simultaneously support and weaken metacognitive learning. Rather than positioning GenAI as inherently beneficial or harmful, the framework emphasizes that its educational impact depends on reflective engagement, pedagogical mediation, feedback literacy, and learners' active regulatory participation.

What This Review Contributes That Previous Reviews Have Not

Previous systematic reviews on AI in education have typically asked broad questions, such as "Is AI effective?" or "What are the applications of AI in education?" (e.g., Zawacki-Richter et al., 2019; Wu & Yu, 2024). They have synthesized evidence across diverse outcomes, often concluding that AI has positive but variable effects. This review differs in three important respects.

First, it focuses specifically on metacognition. By isolating metacognitive outcomes from cognitive ones, this review reveals trade-offs that broader reviews miss. For example, Fan et al. (2025) found that ChatGPT improved essay scores (a cognitive outcome) while reducing knowledge transfer (a metacognitive outcome). A review focused solely on academic achievement would have reported positive effects; one focused solely on metacognition would have reported negative effects. Neither would capture the trade-off. This review's specific focus makes trade-offs visible.

Second, it synthesizes evidence on *pedagogical mechanisms* rather than merely *effects*. Knowing that an intervention works is less useful than knowing *how* it works. This review identifies specific pedagogical approaches, such as prompting, scaffolding, journaling, and feedback literacy instruction, that have been empirically shown to support metacognitive learning. These mechanisms are actionable for educators in ways that general conclusions about AI's effectiveness are not.

Third, it treats risks and negative effects as central rather than peripheral. The dominant narrative in educational technology research is optimistic, emphasizing potential benefits. Negative findings are often relegated to limitations sections or discussed briefly. This review brings metacognitive laziness, cognitive offloading, and the illusion of competence to the foreground. A balanced evidence base requires attention to both benefits and harms.

Implications for Future Research

Several directions for future research emerge from this synthesis. First, longitudinal studies are needed. Most included studies measured outcomes immediately after brief interventions (ranging from a single session to four weeks). The effects of prolonged GenAI use on metacognitive development are unknown. Does metacognitive laziness accumulate over time? Can metacognitive skills trained with AI support transfer to non-AI contexts? These questions require longer-term designs.

Second, developmental studies across age groups are needed. The predominance of university students in the sample limits the generalizability of the findings. Younger students may be more susceptible to metacognitive offloading and may benefit more from structured scaffolding. Research comparing the effects of metacognition across primary, secondary, and tertiary education is urgently needed.

Third, research on individual differences is needed. Weng et al. (2024) showed that personality traits predict SRL outcomes in the context of GenAI. Mendoza et al. (2026) showed that feedback literacy moderates AI's effects. These findings suggest that GenAI interventions may not work equally well for all students. Identifying which students are most at risk for metacognitive laziness and which scaffolding strategies work for which students is a priority.

Fourth, research on the design of GenAI tools themselves is needed. Most studies used generic ChatGPT without modification. Studies that customized GPT prompts or built specialized tools (e.g., Ng et al.'s SRLbot; the customized GPT for reflective practice) showed promising results. Designing GenAI tools that actively scaffold metacognitive processes rather than passively respond represents an important direction for human-AI interaction research.

CONCLUSIONS AND RECOMMENDATION

This systematic review synthesized evidence from 12 empirical studies on the relationship between generative artificial intelligence and metacognitive learning. Unlike broader 2025–2026 syntheses of GenAI and self-regulated learning (Xia et al., 2026; Tsakeni et al., 2025; Romdhoni et al., 2025), which map the field's volume and topical spread, this review's novelty lies in isolating metacognition as its unit of analysis and translating the evidence into an actionable pedagogical framework.

Three main findings emerge. First, GenAI's effect on metacognition is fundamentally conditional: the same tool that improves self-regulation when paired with structured scaffolding can produce metacognitive laziness and declining self-regulation when used without guidance, challenging assumptions that AI automatically improves learning outcomes. Second, four pedagogical approaches support metacognitive learning with GenAI: metacognitive prompting, guided inquiry with scaffolding, reflective journaling, and

feedback literacy instruction, each addressing a specific phase of self-regulated learning (planning, monitoring, evaluation, or reflection). Third, significant risks require active management: passive acceptance of AI-generated answers, an illusion of competence, and unequal benefits for students with weaker feedback literacy were consistently reported across the included studies.

For educators, effective GenAI integration requires structured prompts, reflective activities, and explicit feedback literacy instruction. For instructional designers, GenAI tools should actively prompt metacognitive engagement rather than only respond passively. For policymakers, equity concerns require attention, as students with weaker feedback literacy benefit less from GenAI support. We recommend that future research prioritize longitudinal designs tracking whether AI-supported metacognitive gains persist and transfer to non-AI contexts, developmental comparisons across primary, secondary, and tertiary education, and direct empirical tests of the inverted-U scaffolding hypothesis proposed in this review.

Limitations include the predominance of English-language and university-student samples, potential publication bias, and the rapid evolution of GenAI, which may limit the durability of these findings. Ultimately, this review suggests that the educational value of generative AI lies not in the technology itself, but in the reflective and pedagogical conditions surrounding its use, a conclusion that distinguishes this review's actionable, mechanism-focused contribution from broader 2025–2026 mapping reviews of the same field.

Credit authorship contribution statement

First Author : Conceptualization, methodology, investigation, data collection, formal analysis, data interpretation, and writing original draft.

Second Author : Supervision, methodology validation, and academic guidance.

Third Author : Instrument validation, content review, data interpretation, and manuscript revision.

Fourth Author : Review and editing

Fifth Author : Critical review of the theoretical framework and final manuscript revision

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Ethical Declaration

All participants were informed of the study's purpose and procedures before their involvement. Student participation was conducted with the school's permission, and the data were used only for research purposes. The confidentiality of students' identities was maintained throughout the research process. The learning activities were conducted as part of classroom instruction and did not involve procedures that could harm the participants.

Acknowledgment

The authors would like to thank the school, science teacher, and students who participated in this study. The authors also appreciate the support and academic guidance provided during the preparation of this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. *Educational Technology Research and Development*, 72, 2401–2416. <https://doi.org/10.1007/s11423-024-10366-w>
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445–459. <https://doi.org/10.3102/00346543053004445>

- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530. <https://doi.org/10.1111/bjet.13544>
- Fitriyah, L., & Puspasari, D. (2021). The effect of self-regulated learning and self-efficacy on students' learning outcomes in the Arabic subject at SMK Negeri 10 Surabaya. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 5(5), 1321–1338. <https://doi.org/10.33578/pjr.v5i5.8414>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Futterer, T., Bardach, L., Kuhn, J., Keller, S. D., & Gerjets, P. (2026). Enhancing school students' self-regulated learning through generative AI support: A randomized controlled trial. *Educational Psychology Review*, 38, Article 42. <https://doi.org/10.1007/s10648-026-10133-8>
- García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of Generative AI in education: a systematic review. *Frontiers in Education*, 10, Article 1565938. <https://doi.org/10.3389/educ.2025.1565938>
- Hadian, T., Awaluddin, A. I., Rostini, D., & Rosa, A. T. R. (2023). Learning implementation using ChatGPT media at SMA Negeri 1 Sukabumi. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 7(4), 691–698. <https://doi.org/10.33578/pjr.v7i4.9453>
- Hartini, Chairilsyah, D., Kurnia, R., Anjana, C. F., Nizami, I., Noverman, & Warni. (2026). Transformation of gadget use as adaptive learning media based on deep learning at junior high school. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 10(4), 604–620. <https://doi.org/10.33578/pj.v10i4.1552>
- Hayati, N., Pulungan, N. H., Adhli, A., Saputra, E., & bin Kirin, A. (2026). A metacognitive approach in teaching Ulumul Qur'an and Ulumul Hadith: A study of students at STAIN Mandailing Natal. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 10(3), 414–426. <https://doi.org/10.33578/pj.v10i3.1502>
- Hong, Q. N., Pluye, P., Bujold, M., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT)*, version 2018. Registration of Copyright (#1148552), Canadian Intellectual Property Office, Industry Canada.
- Huda, N., Anwar, K., Novferma, & Kurniawan, W. (2025). ChatGPT scaffolding in supporting metacognition for limit concepts in guided inquiry mathematics learning. *Journal of Information Technology Education: Research*, 24, Article 2. <https://doi.org/10.28945/5422>
- Jumika, M., Erni, & Erlisnawati. (2026). Primary school teacher professionalism: Analysis based on principal managerial competencies and teacher TPACK in Rumbai District Pekanbaru City. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 10(4), 634–644. <https://doi.org/10.33578/pjr.v10i4.1562>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kong, S.-C., & Yang, Y. (2024). A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in K–12 settings. *IEEE Transactions on Learning Technologies*, 17, 1554–1568. <https://doi.org/10.1109/TLT.2024.3392830>
- Lowry, B., McGrath, S., Eitel, C., Hall, H., & Clapp, T. R. (2025). Leveraging generative AI to foster metacognition and self-directed learning. *Journal of Microbiology and Biology Education*, 27(1), e00153-25. <https://doi.org/10.1128/jmbe.00153-25>

- Majidah, Rullyana, G., & Triandari, R. (2025). Google Gemini as a learning assistant: Exploring student perceptions. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 9(2), 163–172. <https://doi.org/10.33578/pjr.v9i2.10008>
- Mendoza, N. B., Xiong, Y., & Yan, Z. (2026). Making sense of AI feedback: How students' feedback literacy moderates the link between ChatGPT acceptance and self-regulated learning. *Educational Psychology*, 46(1), 1–23. <https://doi.org/10.1080/01443410.2026.2641521>
- Molenaar, I. (2022). The concept of hybrid human–AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence*, 3, 100070. <https://doi.org/10.1016/j.caeai.2022.100070>
- Nazaretsky, T., Gabbay, H., & Käser, T. (2026). Can students judge like experts? A large-scale study on the pedagogical quality of AI and human personalized formative feedback. *Computers and Education: Artificial Intelligence*, 10, 100533. <https://doi.org/10.1016/j.caeai.2025.100533>
- Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328–1353. <https://doi.org/10.1111/bjet.13454>
- Nguyen, H., Chen, H.-L., & Asfani, K. (2026). Supporting self-regulated learning in the age of artificial intelligence: A comparative systematic review of earlier AI and generative AI. *Journal of Educational Computing Research*. <https://doi.org/10.1177/07356331261450647>
- Nguyen, H., & Nguyen, A. (2024). Reflective practices and self-regulated learning in designing with generative artificial intelligence: An ordered network analysis. *Journal of Science Education and Technology*, 33, 1–15. <https://doi.org/10.1007/s10956-024-10175-z>
- Nurjanah, S., Nadiroh, N., Rohmaniyah, N., & Cahyanti, E. (2022). Hubungan pembelajaran self-regulated dengan hasil belajar siswa MI di OKU Timur. *Primary: Jurnal Pendidikan Guru Sekolah Dasar*, 11(6), 1826–1834. <https://doi.org/10.33578/jpfkip.v11i6.9218>
- Pacheco, A. J., Boude Figueredo, O. R., Chiappe, A., & Fontán de Bedout, L. (2025). AI-powered learning analytics for metacognitive and socioemotional development: a systematic review. *Frontiers in Education*, 10, Article 1672901. <https://doi.org/10.3389/educ.2025.1672901>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Panggabean, N. N. A., & Turmudi. (2023). Analysis of students' self-regulated mathematical learning during the COVID-19 pandemic. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 7(2), 323–330. <https://doi.org/10.33578/pjr.v7i2.9139>
- Pramusti, R. A., Ulum, B., & Kumalasani, M. P. (2024). Exploration of TPACK in IPAS subjects in optimizing digital literacy of students with special needs. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 8(6), 555–563. <https://doi.org/10.33578/pjr.v8i6.10108>
- Ren, L., Stephens, J. M., & Lee, K. (2026). The impact of AI on learners' self-efficacy: A meta-analysis. *Behavioral Sciences*, 16(1), 158. <https://doi.org/10.3390/bs16010158>
- Romdhoni, R. D., Arrasyid, R., Widodo, S., & Elviani, U. (2025). AI-driven learning analytics for self-regulated and metacognitive learning: A systematic review. *Intellect: Indonesian Journal of Innovation Learning and Technology*, 4(2), 329–347. <https://doi.org/10.57255/intellect.v4i2.1657>
- Sharizan, S., Purwanti, I. T., & Eliwarti, E. (2022). Self-efficacy and self-regulated learning for English content courses in online learning. *Jurnal PAJAR (Pendidikan dan Pengajaran)*, 6(2), 608–619. <https://doi.org/10.33578/pjr.v6i2.8703>
- Tsakeni, M., Nwafor, S. C., Mosia, M., & Egara, F. O. (2025). Mapping the scaffolding of metacognition and learning by AI tools in STEM classrooms: A bibliometric–systematic review approach (2005–2025). *Journal of Intelligence*, 13(11), 148. <https://doi.org/10.3390/jintelligence13110148>

- Wu, R., & Yu, Z. (2024). Do AI chatbots improve students' learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, 55(1), 10–33. <https://doi.org/10.1111/bjet.13334>
- Xia, Q., Liu, Q., Tlili, A., & Chiu, T. K. F. (2026). A systematic literature review on designing self-regulated learning using generative artificial intelligence and its future research directions. *Computers & Education*, 240, Article 105465. <https://doi.org/10.1016/j.compedu.2025.105465>
- Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology*, 56(5), 1842–1863. <https://doi.org/10.1111/bjet.13599>
- Yang, J., & Park, M. (2025). Beyond the algorithm: How metacognitive reflection amplifies AI feedback in L2 writing development. *Multimedia-Assisted Language Learning*, 28(3), 174–193. <https://doi.org/10.15702/mall.2025.28.3.174>
- Yilmaz, M., Temur, H. B., Emmungil, L., Çelik, E., Gauthier, A., & Cukurova, M. (2026). Supporting self-regulated learning through generative AI feedback in online higher education: The importance of student perceptions of the source of feedback. *International Journal of Educational Technology in Higher Education*, 23, 16. <https://doi.org/10.1186/s41239-026-00592-y>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2