Contents lists available at International Journal Press

Journal of Interdisciplinary Learning and Knowledge Sciences

journal homepage: https://ijpress.com/jilks

Volume 1 · Issue 1 · 2026 pp. 23-33

Artificial Intelligence-Supported Metacognitive Learning in Higher Education: A Narrative Review

Ankita Shukla[1,*], Dr. Geetha C A[2]

[1] *Research Scholar, Department of Education, Central University of Kerala, Kerala, India*

[2] *Assistant Professor, Central University of Kerala, Kerala, India*

* Corresponding author: ankita.shukla.phd.9@gmail.com

ARTICLE INFO

Article history:
Received: 28-07-2026
Received in revised form: 06-08-2026
Accepted: 08-08-2026
Available online: 08-08-2026

Keywords:
AI Literacy
Artificial Intelligence
Higher Education
Metacognition
Self-Regulated Learning

ABSTRACT

Artificial intelligence (AI) is transforming how metacognitive skills such as planning, monitoring, evaluating, and regulating learning are developed in higher education. This narrative review synthesises 2021–2026 literature on AI-supported metacognitive learning, covering intelligent tutoring systems, generative AI chatbots, learning analytics dashboards, open learner models, and AI literacy frameworks. It addresses four questions: how AI supports the phases of self-regulated learning; the evidence for specific AI tools; the risks these tools pose to independent thinking; and the implications for teacher education and the Education 4.0 agenda. Well-structured AI scaffolds such as metacognitive prompts, negotiable open learner models, and reflective dashboards are associated with improved self-regulation, motivation, and achievement, yet research also documents risks of cognitive offloading and metacognitive laziness, where uncritical reliance on generative AI narrows engagement with reasoning itself. For pre-service teachers, AI literacy and metacognitive competence intersect with professional identity formation and the Education 4.0 agenda. The review identifies unresolved tensions between AI as a cognitive amplifier and a cognitive substitute, and gaps in long-term, discipline-specific, and culturally situated evidence, especially from South Asian higher education. It concludes that pedagogical design, not technology alone, determines whether AI acts as a metacognitive scaffold or a metacognitive shortcut.

1. INTRODUCTION

Higher education institutions are experiencing a swift technological shift, where artificial intelligence (AI) systems, from intelligent tutoring systems (ITS) to generative large language models (LLMs) like ChatGPT, have transitioned from peripheral tools to become commonplace elements of student learning [1]. A parallel shift of equal importance is metacognition: learners' awareness and regulation of their own cognitive processes [2]. Metacognition and self-regulated learning (SRL) have long been identified as key predictors of academic achievement, deep learning and lifelong learning capacity [3, 4]. The coming together of these two developments poses an urgent question for higher education research. Is AI helping learners to plan, monitor and evaluate their own thinking, or is it subtly eroding that ability by doing the cognitive work that learners would otherwise be doing?

This tension is not just theoretical. A global survey carried out in 2024 showed that the vast majority of university students are currently using AI tools in their studies, the most common purposes being to summarize texts, check grammar and produce first drafts (as cited in the ScienceDirect review of scaffolding critical thinking with generative AI). At the same time, empirical research has begun to record a phenomenon that researchers have come to term "metacognitive laziness", a tendency wherein frequent and unreflective use of AI is linked to lower levels of self-regulation, less critical assessment of AI outputs, and a growing dependence on external cognitive support [5, 6]. These results are part of a large body of evidence indicating that well-designed AI tools, pedagogical agents, metacognitive prompts, and open learner models are capable of effectively enhancing planning, monitoring, and self-evaluation [7, 8]. The literature does not simply express either optimism or pessimism regarding the role of AI in the development of metacognition; instead, it is split. The main factor appears to be the quality of the pedagogical design rather than the mere presence of AI itself.

This narrative review has compiled the research that was published between 2021 and 2026 on the use of AI in promoting metacognitive learning in higher education, and it addresses four questions: (a) in what way is it believed that AI supports the various stages of the self-regulated learning (SRL) cycle; (b) what do the studies reveal regarding the impact of particular AI tools such as intelligent tutoring systems, generative AI chatbots, learning analytics dashboards, and open learner models on metacognitive processes; (c) what risks to the development of metacognitive and critical thinking skills does the literature point out; and (d) what are the specific implications for teacher education and the wider Education 4.0 skills agenda? The review concludes by highlighting gaps that need more research, especially in non-Western and teacher education contexts.

2. THEORETICAL FRAMEWORK

Metacognition, first systematically theorised by Flavell, refers to an individual's knowledge of and control over their own cognitive processes [2]. Flavell distinguished metacognitive knowledge (what learners know about cognition) from metacognitive regulation (how learners plan, monitor, and evaluate their cognitive activity), a distinction later elaborated by Schraw and Moshman into declarative, procedural, and conditional knowledge components [9]. Metacognition is widely treated as the "engine" of self-regulated learning: without accurate monitoring of one's own understanding, planning and reflection have little basis [10].

SRL theory extends metacognition into a dynamic, cyclical process. Zimmerman's model describes three recursive phases: forethought (goal-setting and strategic planning), performance (self-monitoring and strategy execution), and self-reflection (self-evaluation and adaptation), while Pintrich frames SRL as an active, constructive process in which learners set goals and then monitor, regulate, and control their cognition, motivation, and behaviour within a given context. Winne and Hadwin's information-processing model similarly casts SRL as a sequence of monitoring-and-control loops operating across task definition, goal-setting, strategy enactment, and adaptation. These models converge on a key premise directly relevant to AI-supported learning: metacognitive and self-regulatory competence can, in principle, be scaffolded by external agents, including human tutors, peers, and, increasingly, computational systems, before being internalised by the learner [3, 4, 7, 10].

This review adopts a synthesis of these frameworks as its interpretive lens, treating AI tools as potential external regulators that can scaffold each phase of the SRL cycle: generative prompts and adaptive hints supporting forethought and planning; real-time feedback, open learner models, and dashboards supporting monitoring; and reflective prompts or performance summaries supporting self-evaluation [11]. The review also draws on the concept of shared or socially distributed metacognition [12], which recognises that in AI-mediated environments, metacognitive regulation is increasingly co-constructed between the learner and the system rather than residing solely within the individual, a shift that recent scholarship has begun to frame through posthumanist and distributed-cognition perspectives [13].

3. REVIEW APPROACH

This is a narrative rather than a systematic review, and its purpose is interpretive synthesis rather than exhaustive enumeration. Literature was identified through structured searches of academic databases and search engines for peer-reviewed journal articles, systematic reviews, and reputable institutional reports published between 2021 and 2026, using combinations of the terms "artificial intelligence," "generative AI," "metacognition," "self-regulated learning," "intelligent tutoring systems," "AI literacy," "higher education," and "pre-service teachers." Priority was given to studies published in established journals in educational technology and educational psychology (e.g., British Journal of Educational Technology, Computers and Education: Artificial Intelligence, Metacognition and Learning, Frontiers in Psychology/Education) and to recent systematic reviews and meta-analyses that themselves synthesise larger bodies of primary research. Sources were selected for thematic relevance and quality rather than through a fixed inclusion algorithm, consistent with the interpretive, argument-building purpose of a narrative review [14]. Approximately sixty sources are cited, spanning theoretical, empirical, and policy-oriented literature.

4. AI-SUPPORTED METACOGNITIVE LEARNING: A CONCEPTUAL LANDSCAPE

In the literature that has been reviewed, AI-supported metacognitive learning is not merely a single approach but rather a set of overlapping technologies all serving the same educational aim. According to Tsakeni et al., who carried out a bibliometric-systematic review of 135 articles published between 2005 and 2025, this field has been mapped and it has been found that intelligent tutoring systems, learning analytics platforms, and generative AI systems have gradually changed the way metacognition is discussed [13]. This change involves a shift from an emphasis on individual reflection to one involving system-level regulation and shared cognition. The review, which was based on Flavell's theory, General Systems Theory, and posthumanist perspectives, concluded that AI systems are now viewed as going beyond being simply tools; they function as co-regulators of learning and take on the responsibility of monitoring and adjusting the learning process together with the human learner.

Tankelevitch and colleagues in 2024 put forward a different view, claiming that generative AI introduces new metacognitive challenges and opportunities which are different from those associated with earlier educational technologies [11]. The reason for this is that generative AI is able to produce output that seems smooth and believable even when it is not accurate, so learners have to be more careful. They should not only evaluate their own understanding but also judge the reliability of the AI-generated material. This double layer of monitoring (self-monitoring plus AI-output monitoring) is a recurring theme across the literature and helps explain why some scholars treat generative AI as a potentially powerful metacognitive scaffold, while others treat it as a significant metacognitive risk, depending on whether learners are equipped with the literacy and disposition to perform this dual evaluation.

Figure 1 and Table 1 below provide an early summary of the framework used throughout this review.

Table 1 synthesises the tool categories reviewed below against the specific phase(s) of the self-regulated learning cycle they are theorised and empirically shown to support, together with the principal risk each category carries when poorly designed.

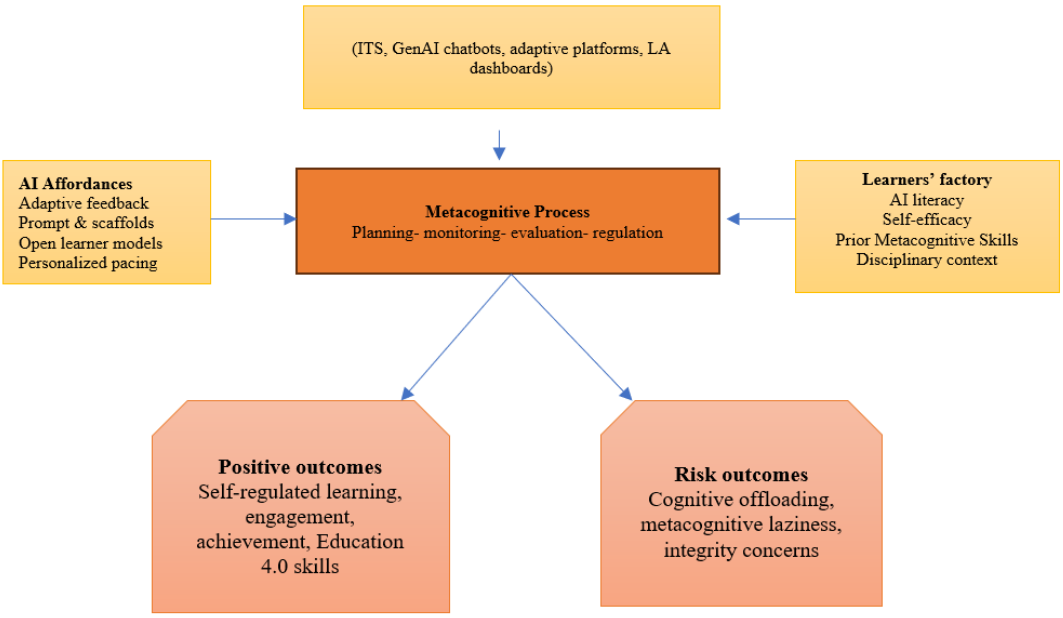

Figure 1. Conceptual Framework of AI-Supported Metacognitive Learning in Higher Education

Table 1. *AI Tool Categories Mapped to Self-Regulated Learning (SRL) Phases and Associated Risks*

AI Tool Category	Primary SRL Phase Supported	Representative Mechanism	Principal Risk if Poorly Designed
Intelligent tutoring systems (e.g., MetaTutor)	Monitoring; strategy enactment	Pedagogical-agent prompts triggered by multichannel learner data	Over-scaffolding delays learner self-initiation
Generative AI chatbots (e.g., ChatGPT-based tutors)	Forethought; performance	Adaptive, personalised dialogic feedback	Cognitive offloading; metacognitive laziness
Learning analytics	Monitoring	Visualisation of	Passive viewing

AI Tool Category	Primary SRL Phase Supported	Representative Mechanism	Principal Risk if Poorly Designed
dashboards		progress, performance, and (in emotion-aware designs) affect	without reflective prompting
Open learner models	Monitoring; self-evaluation	Negotiable, dialogic representation of the learner's own knowledge state	Low uptake if not integrated into course workflow
AI literacy curricula/frameworks	Cross-cutting (precondition)	Explicit instruction in evaluating and critically using AI output	Framework adoption without assessment or reinforcement
Adaptive/personalised learning platforms	Forethought; performance	Real-time adjustment of task difficulty and pacing	Invisible adaptivity reduces learner self-monitoring

5. INTELLIGENT TUTORING SYSTEMS AND METACOGNITIVE SCAFFOLDING

The most theoretically mature strand of this literature concerns intelligent tutoring systems (ITS) explicitly designed to scaffold SRL. MetaTutor, a hypermedia-based ITS built around a biology learning task, has served as a flagship research platform for over a decade. Reporting on lessons learned from this programme, Azevedo et al. describe how pedagogical agents embedded within MetaTutor prompt learners to engage in planning, monitoring, and strategy-use behaviours, leveraging multichannel data, including eye-tracking and log files, to detect when scaffolding is needed [7]. Building on this platform, Dever et al. applied a complex-systems lens to show that pedagogical agents' scaffolding interacts dynamically with learners' own self-initiated strategies rather than operating as a simple additive effect, while a related study found that agent-delivered prompts were most effective when content relevance to learners' subgoals was high, underscoring that learners do not respond uniformly to external metacognitive scaffolds [8, 15].

A subsequent study extended this line of inquiry outside the MetaTutor programme, examining how different types of computer-based scaffolding, conceptual, strategic, and metacognitive, influenced students' metacognitive monitoring accuracy and problem-solving efficiency within an ITS [16]. Their findings reinforced a now-common conclusion in this literature: metacognitive scaffolding (prompts that increase learners' awareness of their own progress) produces more durable gains in monitoring accuracy than scaffolding that simply supplies content or strategy hints, because it targets the regulatory process directly rather than the task outcome. Related work using AI-based scaffolding with structured pre-test or post-test designs has similarly found a strong positive relationship between exposure to AI-based scaffolds and learners' metacognitive awareness scores, particularly in problem-solving-heavy STEM domains.

A consistent design principle across this body of research is that scaffolding must fade over time. Systems that continue to prompt learners indefinitely risk the same dependency effects documented for generative AI (see Section 8); the more durable finding is that ITS scaffolding is most effective when framed as a temporary, reducible support that learners can eventually enact for themselves, consistent with the original Vygotskian premise underlying scaffolding theory.

6. GENERATIVE AI, CHATBOTS, AND SELF-REGULATED LEARNING

The emergence of conversational generative AI has produced a second, more recent strand of research examining chatbot-mediated SRL support. In an often-cited pilot study, Ng et al. compared a generative AI-based chatbot (SRLbot) against a rule-based chatbot (Nemobot) in a three-week secondary science learning context; although conducted with secondary rather than university students, the study is widely cited in higher-education scholarship because it isolated the specific contribution of generative flexibility to SRL outcomes [17]. SRLbot's personalised, adaptive responses were associated with greater behavioural engagement, motivation, and self-regulated learning than the static rule-based alternative, with the number of chatbot interactions significantly predicting variation in students' SRL scores.

Within higher education, Xu et al. conducted a quasi-experiment with 68 college students to explore the role of explicit metacognitive support in a generative AI environment. Their main finding was that metacognitive support, rather than just access to generative AI, was crucial for improving students' self-regulated learning and overall learning experience [18]. This has important implications for teaching. It indicates that simply providing access to tools like ChatGPT is not enough. Measurable improvements depend on intentionally including metacognitive prompts, reflection checkpoints, or structured guidance around the AI interaction. This matches other findings on generative AI and second-language writing. Students with higher pre-existing metacognitive awareness could use AI-generated feedback more effectively. In contrast, students with lower metacognitive awareness gained much less. This suggests that generative AI may highlight existing self-regulatory differences rather than level the playing field. Evidence about learning outcomes in general is mixed. A meta-analysis by Sun and Zhou looked at whether generative AI improves college students' academic achievement [19]. They found that the effects varied widely depending on context, type of tool, and outcome measures. They also warned against blindly claiming uniform benefits. Overall, this research suggests that generative AI chatbots can effectively support self-regulated learning in higher education, but mostly when metacognitive support is intentionally integrated into the interaction rather than assumed based on access to the tool alone.

7. AI LITERACY AS A PRECURSOR TO METACOGNITIVE ENGAGEMENT

A growing consensus in the literature holds that AI literacy is a necessary precondition for productive metacognitive engagement with AI tools [20]. In an early and widely cited exploratory review, Ng et al. conceptualised AI literacy along four competency dimensions: knowing and understanding AI, using and applying AI, evaluating and creating AI applications, and understanding the ethical issues surrounding AI, a framework later operationalised into a validated measurement instrument spanning affective, behavioural, cognitive, and ethical components [21]. Almatrafi et al., synthesising AI literacy research published between 2019 and 2023, found that most existing frameworks emphasise technical and evaluative competencies but under-specify the metacognitive dimension, that is, learners' capacity to reflect on when, why, and how they are relying on AI during a given task [22].

Several higher-education-specific frameworks attempt to close this gap. Kong et al. developed and evaluated an AI literacy curriculum built around project-based learning, finding measurable gains in students' evaluative and ethical reasoning competencies [23]. What unites these frameworks is the claim that AI literacy and metacognition are mutually reinforcing: learners with stronger AI literacy are better positioned to monitor when AI assistance is helping versus substituting for their own reasoning, while learners with stronger baseline metacognitive skill are better positioned to acquire AI literacy through reflective practice.

8. LEARNING ANALYTICS DASHBOARDS AND OPEN LEARNER MODELS

A further strand of research examines how AI-powered learning analytics dashboards (LADs) and open learner models (OLMs) externalise metacognitive information for students. A systematic review of 21 studies on AI-powered LADs found that most were used for predictive performance modelling, SRL support, and teacher-facing insights, but cautioned that dashboards only foster metacognitive awareness when they go beyond simple visualisation to actively prompt reflection and support learner agency, a design principle traced to an earlier framework by Jivet et al. for evaluating dashboard design against learning-science principles [24]. Consistent with this, a meta-synthesis of open learner models in higher education characterised OLMs along a continuum from inspectable (learners can view their model) to negotiable (learners can discuss and contest their model with the system), finding that negotiable OLMs, which transform static feedback into a dialogic exchange, were most strongly associated with strengthened self-efficacy and emotional engagement with the learning process.

Experimental evidence on dashboards' specific effects on metacognitive regulation is more mixed than for OLMs. A randomised study comparing an emotion-aware dashboard, a standard dashboard, and a no-dashboard control among undergraduates in a self-paced MOOC found that the emotion-aware dashboard produced the strongest gains in metacognitive regulation and engagement, while an earlier study of social-comparison dashboard features found meaningful gains in motivation and achievement but no statistically significant change in metacognitive ability suggesting that not all dashboard features are equally effective at building metacognitive skill, even when they succeed on other outcomes. This nuance is supported by qualitative research: university students using a learning-analytics dashboard reported that it supported metacognitive monitoring mostly in the performance phase of the SRL cycle (tracking progress towards a task), rather than in the forethought or self-reflection phases, suggesting that current dashboard designs may support different points in the SRL cycle unevenly.

9. PRE-SERVICE TEACHERS, TEACHER EDUCATION, AND METACOGNITIVE-AI INTERFACES

For teacher education specifically, the population most central to Education 4.0 skills development, the literature converges on the Technology Acceptance Model (TAM) and its extensions as the dominant analytic lens for understanding pre-service teachers' generative AI adoption. Studies applying extended TAM models to pre-service teacher populations have found that AI self-efficacy is among the strongest predictors of perceived ease of use, which in turn shapes perceived usefulness and behavioural intention to adopt generative AI tools in teaching practice. Beyond adoption intentions, however, a smaller but theoretically important body of work links generative AI use directly to shared metacognition and cognitive offloading among pre-service teachers: Iqbal et al. found that generative AI tool use was associated with enhanced academic achievement in sustainability-related coursework, specifically through pathways of shared metacognition and cognitive offloading, suggesting that these cognitive mechanisms can operate as productive rather than purely detrimental processes when the offloaded tasks are well matched to learning goals [12].

Qualitative and mixed-methods research on pre-service teachers' lived experience of generative AI paints a more ambivalent picture. Bae et al. found that pre-service teachers perceived generative AI as a genuinely useful instructional tool, while simultaneously expressing concern that it could limit their own creativity and professional agency, a duality that maps closely onto the broader tension between AI as amplifier versus AI as substitute [25]. A study looking at generative AI as a 'learning buddy and teaching assistant' for pre-service teachers also reported generally positive attitudes, but ongoing concerns about over-reliance undermining the development of independent pedagogical judgement [26]. For a population that is still developing both subject-matter expertise and pedagogical identity, these findings are particularly salient: pre-service teachers too willing to cede instructional planning or reflective judgement to AI risk graduating with underdeveloped metacognitive and pedagogical reasoning skills in the very domain of teaching where such skills are most consequential professionally.

10. RISKS: COGNITIVE OFFLOADING, METACOGNITIVE LAZINESS, AND ACADEMIC INTEGRITY

A large and quickly expanding amount of literature identifies the risks that generative AI presents to the development of metacognitive and critical-thinking skills. The most significant single study in this area, carried out by Fan et al., employed a controlled laboratory approach in which Chinese university students did an English-language reading-to-writing task under one of four conditions, one of which allowed unrestricted access to ChatGPT [5]. The students in the ChatGPT group displayed patterns which the authors called "metacognitive laziness": a tendency to accept the content produced by AI with little critical assessment, thereby not only delegating the writing task but also avoiding the metacognitive judgement of whether the output was in fact of good quality. Gerlich extended this finding to a wider sample and found that frequent use of AI tools was associated with decreased engagement in critical thinking, the effect being most evident among younger and less confident users [6]. Additional experimental research by Lee et al. showed that the participants reported undertaking systematically less cognitive effort when carrying out tasks with the help of generative AI, with their self-reported level of critical thinking engagement falling as their confidence in the AI tool increased [27].

The results come together to form what various authors refer to as a "cognitive offloading ladder", which is a conceptual model that separates productive cognitive support (in this case AI taking on low-value, routine sub-tasks so as to free up capacity for higher-order reasoning) from cognitive substitution (where AI carries out the synthesis, self-monitoring and regulation that make up the real intellectual work involved in a task). In such a situation, when learners get a completed literature review or an essay draft from an LLM, they end up with the final product without going through the metacognitive process, that is, the self-monitoring and regulation which would normally promote intellectual development; this has been called an "illusion of competence", because learners then confuse the smoothness of the text produced by the AI with their own understanding.

The cognitive risks are closely linked with concerns regarding academic integrity. A systematic review of 41 studies on generative AI and academic integrity in higher education found consistent evidence that plagiarism-related behaviour had increased after accessible generative AI tools had been introduced, together with ongoing difficulties on the part of both teachers and students in being able to reliably tell the difference between AI-produced and human-written text [28]. Other research has pointed out that the opacity of the algorithms makes this issue worse: since generative AI systems operate as "black boxes", users are generally not aware of how the outputs are generated, which in turn reduces the critical examination that would normally promote metacognitive engagement [29]. The institutional responses described in the literature vary from outright bans to a more structured form of integration, and the reviews all agree that policy by itself is not sufficient to deal with the fundamental pedagogical issue; assessment design and specific instruction in the metacognitive use of AI are necessary additions to any integrity policy [30].

It is important to note that a number of authors warn against interpreting these risks in a deterministic way. The same design principles which have been found to enhance metacognition in studies involving instructional systems and dashboards such as deliberate scaffolding, reflective prompting, and the gradual reduction of support have also been suggested as ways of addressing the offloading risks associated with generative AI, including the use of cognitive-forcing functions that interrupt the automatic acceptance of AI output and Socratic-questioning interfaces which require learners to express their own reasoning before being given AI assistance. The evidence base for these mitigations in higher education settings, however, remains considerably thinner than the evidence base documenting the risks themselves.

11. PERSONALIZED AND ADAPTIVE LEARNING

AI-driven personalised and adaptive learning platforms constitute a further major theme, though one whose connection to metacognition specifically is often implicit rather than explicit. A systematic review of 45 studies on AI and personalised learning in higher education found consistent evidence that adaptive systems which tailor content, pacing, and difficulty to individual learners improve engagement, test scores, and motivation, with course-completion gains particularly notable when adaptive assessment was paired with real-time feedback [31]. A separate global systematic review similarly found that machine learning, natural language processing, and ITS technologies jointly enable dynamic, learner-responsive pathways, though both reviews flagged persistent gaps in theoretical grounding and in evaluation of long-term, transferable metacognitive gains as opposed to short-term performance gains.

The relationship between adaptivity and metacognition is double-edged in ways that echo the offloading literature discussed above. When an adaptive system silently adjusts task difficulty on the learner's behalf, it can reduce the learner's own need to monitor and judge their readiness, the very judgment that constitutes metacognitive monitoring. Reviews of adaptive learning mechanisms note that systems designed to make their adaptivity visible to the learner (for example, by explaining why a task's difficulty changed) tend to preserve more of this monitoring function than systems that adapt invisibly, reinforcing the broader design principle emerging across this review: metacognitive benefit depends on whether AI processes are transparent and reflection-inducing, or opaque and merely convenient.

12. ALIGNMENT WITH THE EDUCATION 4.0 SKILLS AGENDA

The Education 4.0 framework associated with the World Economic Forum's account of skills needed for the Fourth Industrial Revolution provides a useful integrating frame for the literature reviewed above, because it explicitly identifies technology skills, innovation and creativity, interpersonal skills, and global citizenship as core competency domains, alongside personalised, self-paced, and problem-based learning experiences. Recent scholarship connecting large language models specifically to Education 4.0 has found that LLM use can plausibly support the development of critical-thinking competencies central to this agenda, while simultaneously introducing ethical and legal challenges around privacy, bias, and academic integrity that Education 4.0 policy frameworks must now explicitly accommodate. Legal and policy scholarship has likewise argued that governing opaque algorithmic systems in education and the future of work requires explicit attention to justice and accountability [32]. A broader conceptual mapping of "Higher Education 4.0" identifies seven interacting dimensions: competencies, pedagogy, technology-based education, infrastructure, governance, sociocultural and ethical considerations, and research partnerships, underscoring that AI-supported metacognitive learning cannot be treated as a purely technical or purely pedagogical matter, but sits at the intersection of all seven. For pre-service teachers, this alignment is especially consequential: Education 4.0's emphasis on personalised, technology-enabled, learner-centred pedagogy implies that tomorrow's teachers must themselves be metacognitively and technologically literate enough to model and teach these competencies to their own future students a second-order requirement that current teacher-education curricula, per the AI-literacy-framework literature reviewed in Section 7, do not yet consistently address.

13. SYNTHESIS

The conceptual framework and SRL-phase mapping (Figure 1 and Table 1) are presented earlier, in Section 4.

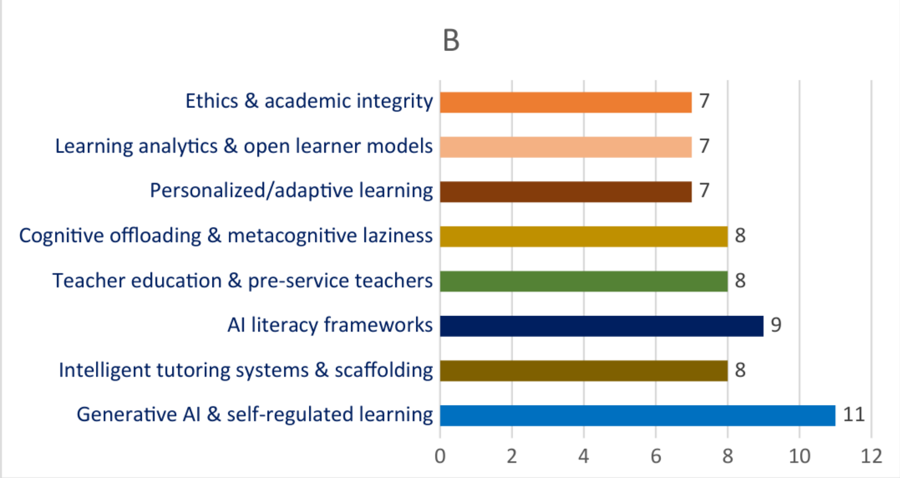

Figure 2. Thematic Distribution of Reviewed Literature (2020-2026)

Table 2 presents a representative (non-exhaustive) summary of key empirical studies discussed above, illustrating the diversity of populations, tools, and outcomes examined across the reviewed literature.

Table 2. Summary of Representative Studies on AI-Supported Metacognitive Learning

Study	Sample/Context	AI Tool Studied	Key Finding Relevant to Metacognition
Azevedo et al. (2022) [7]	University biology learners	MetaTutor (ITS)	Pedagogical agents scaffold planning and monitoring using multichannel learner data
Dever et al. (2023) [8]	Undergraduates (N = 118)	MetaTutor (ITS)	Agent scaffolding interacts dynamically with learner-initiated SRL strategies
Wang et al. (2023) [16]	University ITS users	Computer-based scaffolding	Metacognitive scaffolds improve monitoring accuracy more than content scaffolds
Ng et al. (2024) [17]	Secondary science students (N = 74)	SRLbot (GenAI chatbot)	GenAI chatbot outperforms rule-based bot on SRL and motivation
Xu et al. (2025) [18]	College students (N = 68)	GenAI + metacognitive support	Explicit metacognitive support, not GenAI access alone, drives SRL gains
Fan et al. (2024) [5]	Chinese university students (N = 117)	ChatGPT (writing task)	Unrestricted GenAI access linked to “metacognitive laziness”
Gerlich (2025) [6]	Mixed adult sample	General AI tools	Habitual AI use correlates with reduced critical-thinking engagement
Iqbal et al. (2025) [12]	Pre-service teachers	Generative AI tools	Shared metacognition and cognitive offloading mediate achievement gains
Bae et al. (2024) [25]	Pre-service teachers	Generative AI (general)	Perceived usefulness coexists with concern over reduced creativity/agency
Bittle & El-Gayar (2025) [28]	41 studies (systematic review)	Generative AI (general)	Consistent evidence of integrity risks and AI-text detection difficulty
Merino-Campos (2025) [31]	45 studies (systematic review)	AI personalised learning platforms	Adaptive systems raise engagement, motivation, and completion rates

14. DISCUSSION: AMPLIFIER OR SUBSTITUTE?

When considered as a whole, the literature reviewed here cannot be easily judged in terms of whether AI helps or hinders metacognitive development in higher education, and this review contends that the question is to some extent wrongly formulated. The more accurate conclusion is that AI systems act as amplifiers of the metacognitive tendencies and instructional design that surround their use. In cases where AI tools are based on explicit metacognitive scaffolds, such as pedagogical agents that gradually disappear over time, negotiable open learner models that encourage dialogue rather than passive reception, or generative AI used in conjunction with deliberate reflective prompting, the evidence always shows an improvement in planning, monitoring, and self-evaluation [7, 18]. On the other hand, when AI tools are used without such scaffolding, for example, when students have unrestricted access to generative AI for writing tasks, use invisibly adaptive platforms, or view dashboards that show data without prompting reflection, the evidence is just as consistently geared towards offloading, laziness, and a decline in critical engagement [5, 6].

Another significant tension relates to whether the responsibility for metacognitive regulation lies with the learner or with the system. Some researchers describe AI-assisted metacognition as becoming more 'shared' or 'distributed' between the learner and the system, a way of looking at the issue which has real value when explaining how, for example, an open learner model can enable monitoring even if the learner has not formed that judgement on their own [12, 13]. However, this perspective also has the potential to lead to a gradual shift of regulatory responsibility completely away from the learner, and this is exactly the process that underlies the findings on cognitive offloading and metacognitive laziness. The review thus argues that 'shared metacognition' should not be seen as inherently good or bad, but rather as a design feature; temporary, transparent shared regulation that is aimed at leading to eventual learner independence promotes development, while permanent, opaque or convenience-based shared regulation hinders it. Third, the review surfaces a notable population-specific concern for pre-service teachers. Because this population is simultaneously a student cohort developing its own metacognitive competence and a professional cohort that will soon need to model and teach metacognitive strategies to others, the stakes of AI-related cognitive offloading are arguably higher than for the general undergraduate population. The TAM-based literature (Section 9) demonstrates that pre-service teachers are, on balance, willing adopters of generative AI, but adoption-intention research says little about whether that adoption strengthens or weakens the specific pedagogical-reasoning competencies these future teachers will need. This is a significant and, per the reviewed literature, still underexplored gap.

15. IMPLICATIONS FOR PRACTICE

Several practical implications follow from this synthesis for higher education institutions and teacher-education programmes specifically. First, institutions should prioritise AI tools and platform features that make regulatory processes visible and negotiable dashboards and learner models that invite reflection rather than passive viewing over tools chosen purely for content-generation convenience. Second, AI literacy instruction should be explicitly linked to metacognitive strategy instruction, rather than treated as a separate digital-skills module; the reviewed frameworks (Section 7) converge on the idea that evaluative and ethical AI competencies are themselves metacognitive skills. Third, assessment design should be adapted to require demonstration of the reasoning process, not just the product, given the well-documented risk that generative AI can supply a finished product while the learner bypasses the underlying metacognitive work. Recent work on developmental assessment alternatives in the age of artificial intelligence similarly argues that assessment must evidence authentic learning rather than AI-generatable products [33]. Fourth, teacher-education curricula should treat pre-service teachers' own AI-metacognitive competence as a distinct professional outcome, given the second-order responsibility these students will carry to model such competence for their own future students.

16. LIMITATIONS OF THE REVIEW

As a narrative rather than systematic review, this synthesis does not follow a fully reproducible search-and-screening protocol, and source selection, while guided by relevance and quality, necessarily reflects the reviewer's judgement. The rapidly evolving nature of generative AI research also means that some of the most recent findings cited here derive from preprints or emerging venues that have not yet completed full peer review, and effect sizes reported across studies are not directly comparable given considerable heterogeneity in tools, tasks, and outcome measures. Geographically, the reviewed literature is weighted toward North American, European, and East Asian contexts; empirical evidence from South Asian higher education, including Indian teacher-education contexts, remains comparatively sparse and represents a priority area for future primary research. In particular, this review did not include a PRISMA-style flow diagram of records identified, screened and excluded; it was not performed by two independent reviewers cross-checking inclusion decisions; and it did not involve a single, fixed Boolean search string across a pre-specified set of databases. Thus, the review provides representative, quality-weighted coverage of the field, not a claim of exhaustive coverage.

17. CONCLUSION

This review has gathered a wide range of literature on AI-supported metacognitive learning in higher education. It focuses on intelligent tutoring systems, generative AI chatbots, AI literacy, learning analytics dashboards, open learner models, teacher education, and the related risks of cognitive offloading and academic integrity loss. The main conclusion is that AI's impact on learners' ability to plan, monitor, and assess their thinking isn't determined by the technology itself; it depends on how teaching is designed. Clear, transparent support helps metacognitive development, while vague or overly easy AI assistance can hinder it. This conclusion is especially urgent for pre-service teachers and the broader Education 4.0 skills agenda. The metacognitive habits formed during teacher training are likely to influence the teaching methods these future educators will use in their classrooms. Future research is necessary, particularly long-term, discipline-specific, and culturally relevant studies from underrepresented higher education contexts. This research will help determine how these design principles can foster lasting and transferable metacognitive skills beyond the initial learning task.

REFERENCES

[1] H. Crompton and D. Burke, "Artificial intelligence in higher education: the state of the field," *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 22, Apr. 2023, doi: 10.1186/s41239-023-00392-8.

[2] J. H. Flavell, "Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry.," *Am. Psychol.*, vol. 34, no. 10, pp. 906–911, Oct. 1979, doi: 10.1037/0003-066X.34.10.906.

[3] P. R. Pintrich, "The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing," *Theory Pract.*, vol. 41, no. 4, pp. 219–225, Nov. 2002, doi: 10.1207/s15430421tip4104_3.

[4] B. J. Zimmerman, "Self-Regulated Learning and Academic Achievement: An Overview," *Educ. Psychol.*, vol. 25, no. 1, pp. 3–17, Jan. 1990, doi: 10.1207/s15326985ep2501_2.

[5] Y. Fan *et al.*, "Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance," *Br. J. Educ. Technol.*, vol. 56, no. 2, pp. 489–530, Mar. 2024, doi: 10.1111/bjet.13544.

[6] M. Gerlich, "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking," *Societies*, vol. 15, no. 1, p. 6, Jan. 2025, doi: 10.3390/soc15010006.

[7] R. Azevedo *et al.*, "Lessons Learned and Future Directions of MetaTutor: Leveraging Multichannel Data to Scaffold Self-Regulated Learning With an Intelligent Tutoring System," *Front. Psychol.*, vol. 13, p. 813632, Jun. 2022, doi: 10.3389/fpsyg.2022.813632.

[8] D. A. Dever, N. A. Sonnenfeld, M. D. Wiedbusch, S. G. Schmorrow, M. J. Amon, and R. Azevedo, "A complex systems approach to analyzing pedagogical agents' scaffolding of self-regulated learning within an intelligent tutoring system," *Metacognition Learn.*, vol. 18, no. 3, pp. 659–691, Dec. 2023, doi: 10.1007/s11409-023-09346-x.

[9] G. Schraw and D. Moshman, "Metacognitive theories," *Educ. Psychol. Rev.*, vol. 7, no. 4, pp. 351–371, Dec. 1995, doi: 10.1007/BF02212307.

[10] Winne, Philip H. and Hadwin, Allyson F, "Studying as Self-Regulated Learning," in *Metacognition in Educational Theory and Practice*, Hacker, D. J., Dunlosky, J., and Graesser, A. C., Eds., Mahwah, NJ: Lawrence Erlbaum Associates, 1998, pp. 277–304.

[11] L. Tankelevitch *et al.*, "The Metacognitive Demands and Opportunities of Generative AI," 2023, doi: 10.48550/ARXIV.2312.10893.

[12] 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," *Sci. Rep.*, vol. 15, no. 1, p. 16610, May 2025, doi: 10.1038/s41598-025-01676-x.

[13] M. Tsakeni, S. C. Nwafor, M. Mosia, and F. O. Egara, "Mapping the Scaffolding of Metacognition and Learning by AI Tools in STEM Classrooms: A Bibliometric–Systematic Review Approach (2005–2025)," *J. Intell.*, vol. 13, no. 11, p. 148, Nov. 2025, doi: 10.3390/jintelligence13110148.

[14] M. J. Grant and A. Booth, "A typology of reviews: an analysis of 14 review types and associated methodologies," *Health Inf. Libr. J.*, vol. 26, no. 2, pp. 91–108, Jun. 2009, doi: 10.1111/j.1471-1842.2009.00848.x.

[15] D. A. Dever, M. D. Wiedbusch, S. M. Romero, and R. Azevedo, "Investigating pedagogical agents' scaffolding of self-regulated learning in relation to learners' subgoals," *Br. J. Educ. Technol.*, vol. 55, no. 4, pp. 1290–1308, Jul. 2024, doi: 10.1111/bjet.13432.

[16] F. Wang, X. Zhou, K. Li, A. C. K. Cheung, and M. Tian, "The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English," *Interact. Learn. Environ.*, vol. 33, no. 7, pp. 4633–4652, Aug. 2025, doi: 10.1080/10494820.2025.2470319.

[17] D. T. K. Ng, C. W. Tan, and J. K. L. Leung, "Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study," *Br. J. Educ. Technol.*, vol. 55, no. 4, pp. 1328–1353, Jul. 2024, doi: 10.1111/bjet.13454.

[18] X. Xu, L. Qiao, N. Cheng, H. Liu, and W. Zhao, "Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support," *Br. J. Educ. Technol.*, vol. 56, no. 5, pp. 1842–1863, Sep. 2025, doi: 10.1111/bjet.13599.

[19] L. Sun and L. Zhou, "Does Generative Artificial Intelligence Improve the Academic Achievement of College Students? A Meta-Analysis," *J. Educ. Comput. Res.*, vol. 62, no. 7, pp. 1676–1713, Dec. 2024, doi: 10.1177/07356331241277937.

[20] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, "Conceptualizing AI literacy: An exploratory review," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100041, 2021, doi: 10.1016/j.caeai.2021.100041.

[21] D. T. K. Ng, W. Wu, J. K. L. Leung, T. K. F. Chiu, and S. K. W. Chu, "Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach," *Br. J. Educ. Technol.*, vol. 55, no. 3, pp. 1082–1104, May 2024, doi: 10.1111/bjet.13411.

[22] O. Almatrafi, A. Johri, and H. Lee, "A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023)," *Comput. Educ. Open*, vol. 6, p. 100173, Jun. 2024, doi: 10.1016/j.caeo.2024.100173.

[23] S.-C. Kong, M.-Y. W. Cheung, and O. Tsang, "Developing an artificial intelligence literacy framework: Evaluation of a literacy course for senior secondary students using a project-based learning approach," *Comput. Educ. Artif. Intell.*, vol. 6, p. 100214, Jun. 2024, doi: 10.1016/j.caeai.2024.100214.

[24] I. Jivet, M. Scheffel, M. Specht, and H. Drachsler, "License to evaluate: preparing learning analytics dashboards for educational practice," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, Sydney New South Wales Australia: ACM, Mar. 2018, pp. 31–40. doi: 10.1145/3170358.3170421.

[25] H. Bae, J. Hur, J. Park, G. W. Choi, and J. Moon, "Pre-service Teachers' Dual Perspectives on Generative AI: Benefits, Challenges, and Integrating into Teaching and Learning," *Online Learn.*, vol. 28, no. 3, Sep. 2024, doi: 10.24059/olj.v28i3.4543.

[26] M. Nyaaba *et al.*, "Generative AI as a Learning Buddy and Teaching Assistant: Pre-service Teachers' Uses and Attitudes," 2024, *arXiv*. doi: 10.48550/ARXIV.2407.11983.

[27] H.-P. (Hank) Lee *et al.*, "The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan: ACM, Apr. 2025, pp. 1–22. doi: 10.1145/3706598.3713778.

[28] K. Bittle and O. El-Gayar, "Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda," *Information*, vol. 16, no. 4, p. 296, Apr. 2025, doi: 10.3390/info16040296.

[29] I. M. García-López and L. Trujillo-Liñán, "Ethical and regulatory challenges of Generative AI in education: a systematic review," *Front. Educ.*, vol. 10, p. 1565938, Jun. 2025, doi: 10.3389/feduc.2025.1565938.

[30] N. J. Francis, S. Jones, and D. P. Smith, "Generative AI in Higher Education: Balancing Innovation and Integrity," *Br. J. Biomed. Sci.*, vol. 81, p. 14048, Jan. 2025, doi: 10.3389/bjbs.2024.14048.

[31] C. Merino-Campos, "The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review," *Trends High. Educ.*, vol. 4, no. 2, p. 17, Mar. 2025, doi: 10.3390/higheredu4020017. [32] J. K. Pandey and R. Kumar, "Governing the algorithmic agent: Confronting overt and covert challenges to justice and the future of work," Int. J. Law Manag. Humanit., vol. 8, no. 4, pp. 1974–1984, 2025, doi: 10.10000/IJLMH.1110648.

[33] S. K. Maurya and J. K. Pandey, "Developmental assessment alternatives in business education," in Reimagining Business Education in the Age of Artificial Intelligence, K. Mili, Ed., Hershey, PA: IGI Global, 2026, pp. 271–316, doi: 10.4018/979-8-3373-4377-8.ch009.

