IAA Journal of Education 11 (1):1-4, 2025. www.iaajournals.org
©IAAJOURNALS ISSN: 2636-7270
https://doi.org/10.59298/IAAJE/2025/11114 IAAJB:1200000000

AI as a Metacognitive Mirror: How Reflective Conversations with Generative AI Affect Students' Self-Regulated Learning, Confidence Calibration, and Transfer of Knowledge

Ugwu Jovita Nnenna

Department of Publication and Extension Kampala International University Uganda

ABSTRACT

Generative artificial intelligence is increasingly being explored in education not only as a source of information, but also as a reflective partner that can support students' metacognition. This narrative review examines the idea of AI as a metacognitive mirror, focusing on how reflective conversations with generative AI may influence self-regulated learning, confidence calibration, and transfer of knowledge. The review suggests that AI-guided dialogue can support major phases of self-regulated learning by helping students set goals, monitor understanding, explain reasoning, identify knowledge gaps, and evaluate learning strategies. It may also improve confidence calibration by prompting learners to compare perceived understanding with actual performance, thereby reducing overconfidence and underconfidence. In relation to transfer, reflective AI interactions may encourage students to apply concepts across new contexts through comparison, self-explanation, and principle-based reasoning. At the same time, the review highlights important risks, including over-reliance on AI, superficial engagement, inflated confidence, and reduced independent thinking. Current evidence indicates that the educational value of generative AI depends less on answer generation and more on how it is pedagogically structured to promote active reflection. The review concludes that AI has significant potential as a metacognitive support tool, but stronger empirical studies are needed to determine its long-term effects on learner autonomy, calibration accuracy, and meaningful transfer of knowledge.

Keywords: Generative artificial intelligence; metacognition; self-regulated learning; confidence calibration; knowledge transfer; reflective learning; educational technology; AI in education

INTRODUCTION

The rapid spread of generative artificial intelligence (AI) in education has shifted discussion from whether these tools should be used to how they should be used in ways that genuinely improve learning. Recent reviews and commentaries suggest that systems such as ChatGPT can support idea generation, feedback, personalization, and access to information, but they also raise concerns about inaccuracy, academic integrity, shallow engagement, and student over-reliance on machine-generated responses [1-8]. As universities and schools move from initial caution toward more structured integration, the central educational question is no longer simply whether generative AI can produce answers, but whether it can be designed and used to strengthen students' thinking and learning processes themselves.

One promising way to frame this problem is to view generative AI as a metacognitive mirror rather than merely an answer engine. In this role, AI does not just deliver content; it prompts learners to articulate goals, explain reasoning, monitor progress, evaluate confidence, and revise strategies. This framing is closely aligned with

metacognition and self-regulated learning (SRL) theory, which describes successful learning as an active cycle of planning, monitoring, feedback use, and self-evaluation [4,10-15]. Meta-analytic evidence shows that metacognitive prompts can improve SRL and learning outcomes in digital environments, while research on explainable AI and open learner models indicates that transparency and inspectability are important if students are to question, regulate, and learn from AI-supported interactions rather than passively accept them [10,11,20].

This perspective is especially relevant to two educational outcomes: confidence calibration and transfer of knowledge. Confidence calibration refers to how accurately learners judge what they know and do not know, and it is central to effective self-regulation because study decisions depend on accurate self-monitoring [16-18]. Transfer, by contrast, concerns whether learners can apply what they have learned in new problems or contexts rather than only reproduce information in the original setting [19]. Reflective conversations with generative AI may support both outcomes by asking students to predict performance, justify their confidence, compare their expectations with results, explain why an answer works, and apply the same principle to a new case. Early evidence from educational chatbot research suggests that metacognitive feedback can improve not only learning outcomes but also retention and knowledge transfer, indicating that conversational scaffolds may help learners move beyond answer retrieval toward reflective understanding [9].

At the same time, the benefits of generative AI are not automatic. Because large language models produce fluent, immediate, and often persuasive responses, they can also create an illusion of understanding, reduce productive struggle, and encourage dependence when students outsource monitoring and evaluation to the tool itself [1,3-5,11]. Thus, the educational value of generative AI may depend less on its capacity to generate content and more on whether its use is structured to provoke reflection, uncertainty checking, self-explanation, and strategic revision. Against this background, this review examines generative AI as a metacognitive mirror and synthesizes how reflective human-AI dialogue may influence students' self-regulated learning, confidence calibration, and transfer of knowledge, while also identifying the pedagogical risks and research gaps that should guide future work in this area [1,4,5,9-11,16,19,20]

Thematic synthesis

Generative AI is increasingly being discussed not only as a content generator or tutor, but as a conversational tool that can help learners think about their own thinking. From this perspective, AI functions as a metacognitive mirror**:** it can prompt students to set goals, explain reasoning, monitor uncertainty, compare answers, and revise strategies. Recent reviews agree that generative AI can support personalization, feedback, and interaction in education, but they also warn that its value depends on how it is designed and used pedagogically rather than on the novelty of the tool itself [1-8].

This framing aligns closely with self-regulated learning (SRL), where effective learners plan, monitor, and evaluate their learning in cycles [13-16]. In theory, reflective conversations with AI can scaffold each phase of that cycle. Before studying, AI can ask learners to define goals and predict difficulty. During study, it can ask for self-explanations, identify gaps, and generate comparison prompts. After study, it can support reflection by asking what remains unclear, what strategies worked, and what should change next time. Evidence from metacognitive prompting research in computer-based environments suggests that such prompts improve SRL and learning outcomes, while work on explainable AI in education highlights the importance of transparency if students are to monitor and trust AI support appropriately [11,12].

A second important issue is confidence calibration the degree to which students' confidence matches their actual performance [17-19]. Calibration matters because poorly calibrated learners often stop too early, over-study familiar material, or underestimate what they can do. Classical metacognition research shows that self-evaluation, delayed judgments, and explicit monitoring can improve the accuracy of learners' judgments [17-19]. This is where generative AI may be especially useful: a reflective chatbot can ask students to predict their score before answering, justify their confidence, compare prediction with outcome, and then revise a study plan. In this sense, AI may help students move from passive answer-seeking to active monitoring. However, the same fluency that makes AI engaging can also be dangerous. Reviews of AI use in education warn that over-reliance on dialogue systems may reduce critical thinking and analytical effort, and persuasive AI responses may create a false sense of understanding [4-6].

The third issue is transfer of knowledge**.** Educationally, AI-supported reflection is only valuable if it helps learners apply knowledge beyond the original task. Transfer literature shows that students are more likely to generalize learning when they focus on underlying principles, compare cases, and explain why a solution works rather than merely copying procedures [20]. This is where reflective AI conversations may have a real advantage: generative AI can quickly produce contrasting examples, new scenarios, counterexamples, and prompts that push learners to explain their reasoning across contexts. An emerging experimental study on educational chatbots with metacognitive feedback found benefits for knowledge transfer as well as motivation and retention in science learning [10]. Still, evidence directly connecting LLM-based reflective dialogue to robust far transfer remains limited, so claims about transfer should remain cautious [6,10,20].

Overall, the literature suggests that the most promising educational role for generative AI is not as an answer machine, but as a structured reflective partner. Used well, it can externalize planning, self-questioning, monitoring, and revision—core processes of SRL [13-16]. Used poorly, it can offload thinking, inflate confidence, and weaken independent judgment [4-6]. The present evidence base is still uneven: many studies are reviews, conceptual pieces, or policy analyses, while relatively few classroom experiments directly test reflective AI dialogue against outcomes such as calibration accuracy and transfer [1-8,10-12]. The clearest next step for research is to compare answer-first AI with reflection-first AI**,** using delayed post-tests, calibration measures, and transfer tasks rather than immediate performance alone [5,6,10,17-20].

CONCLUSION

Viewing generative AI as a metacognitive mirror offers a promising shift in educational thinking. Rather than treating AI merely as a tool for producing answers, this perspective emphasizes its role in prompting students to plan, monitor, question, and revise their own learning processes. The literature suggests that reflective conversations with AI can strengthen self-regulated learning, support more accurate confidence judgments, and create opportunities for deeper transfer of knowledge across tasks and contexts. However, these benefits are not automatic. Without careful instructional design, AI may encourage dependency, passive acceptance of responses, and an illusion of understanding. Therefore, the real educational value of generative AI lies in using it to stimulate reflective thinking rather than replace it. Future research should move beyond short-term performance outcomes and examine how reflection-centered AI use shapes learner independence, metacognitive growth, and durable transfer over time.

REFERENCES

1. Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. **Learn Individ Differ.** 2023;103:102274. Available from: **https://doi.org/10.1016/j.lindif.2023.102274**
2. Lo CK. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. **Educ Sci.** 2023;13(4):410. Available from: **https://doi.org/10.3390/educsci13040410**
3. Tlili A, Shehata B, Adarkwah MA, Bozkurt A, Hickey DT, Huang R, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. **Smart Learn Environ.** 2023;10(1):15. Available from: **https://doi.org/10.1186/s40561-023-00237-x**
4. Zhai C, Wibowo S, Li LD. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. **Smart Learn Environ.** 2024;11(1):28. Available from: **https://doi.org/10.1186/s40561-024-00316-7**
5. Tankelevitch L, Kewenig V, Simkute A, Scott AE, Sarkar A, Sellen A, et al. The metacognitive demands and opportunities of generative AI. In: **Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems**; 2024. p. 1-24. Available from: **https://doi.org/10.1145/3613904.3642902**
6. Yan L, Greiff S, Teuber Z, Gašević D. Promises and challenges of generative artificial intelligence for human learning. **Nat Hum Behav.** 2024;8(10):1839-1850. Available from: **https://doi.org/10.1038/s41562-024-02004-5**
7. Wang H, Dang A, Wu Z, Mac S. Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. **Comput Educ Artif Intell.** 2024;7:100326. Available from: **https://doi.org/10.1016/j.caeai.2024.100326**

8. Rejeb A, Rejeb K, Appolloni A, Treiblmaier H, Iranmanesh M. Exploring the impact of ChatGPT on education: A web mining and machine learning approach. **Int J Manag Educ.** 2024;22(1):100932. Available from: https://doi.org/10.1016/j.ijme.2024.100932
9. Romero-Rodríguez JM, Ramírez-Montoya MS, Buenestado-Fernández M, Lara-Lara F. Use of ChatGPT at University as a Tool for Complex Thinking: Students' Perceived Usefulness. **J New Approaches Educ Res.** 2023;12:323-339. Available from: https://doi.org/10.7821/naer.2023.7.1458
10. Yin J, Zhu Y, Goh T-T, Wu W, Hu Y. Using Educational Chatbots with Metacognitive Feedback to Improve Science Learning. **Appl Sci (Basel).** 2024;14(20):9345. Available from: https://doi.org/10.3390/app14209345
11. Guo L. Using metacognitive prompts to enhance self-regulated learning and learning outcomes: A meta-analysis of experimental studies in computer-based learning environments. **J Comput Assist Learn.** 2022;38(3):811-832. Available from: https://doi.org/10.1111/jcal.12650
12. Khosravi H, Buckingham Shum S, Chen G, Conati C, Tsai Y-S, Kay J, et al. Explainable Artificial Intelligence in education. **Comput Educ Artif Intell.** 2022;3:100074. Available from: https://doi.org/10.1016/j.caeai.2022.100074
13. Zimmerman BJ. Becoming a self-regulated learner: An overview. **Theory Into Practice.** 2002;41(2):64-70. Available from: https://doi.org/10.1207/S15430421TIP4102_2
14. Panadero E. A Review of Self-regulated Learning: Six Models and Four Directions for Research. **Front Psychol.** 2017;8:422. Available from: https://doi.org/10.3389/fpsyg.2017.00422
15. Butler DL, Winne PH. Feedback and Self-Regulated Learning: A Theoretical Synthesis. **Rev Educ Res.** 1995;65(3):245-281. Available from: https://doi.org/10.3102/00346543065003245
16. Schraw G, Moshman D. Metacognitive theories. **Educ Psychol Rev.** 1995;7(4):351-371. Available from: https://doi.org/10.1007/BF02212307
17. Stone NJ. Exploring the Relationship between Calibration and Self-Regulated Learning. **Educ Psychol Rev.** 2000;12(4):437-475. Available from: https://doi.org/10.1023/A:1009084430926
18. Thiede KW, Anderson MCM, Therriault D. Accuracy of metacognitive monitoring affects learning of texts. **J Educ Psychol.** 2003;95(1):66-73. Available from: https://doi.org/10.1037/0022-0663.95.1.66
19. Osterhage JL, Usher EL, Douin TA, Bailey WM. Opportunities for self-evaluation increase student calibration in an introductory biology course. **CBE Life Sci Educ.** 2019;18(2):ar16. Available from: https://doi.org/10.1187/cbe.18-10-0202
20. Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. **Psychol Bull.** 2002;128(4):612-637. Available from: https://doi.org/10.1037/0033-2909.128.4.612

CITE as: Ugwu Jovita Nnenna (2025). AI as a Metacognitive Mirror: How Reflective Conversations with Generative AI Affect Students' Self-Regulated Learning, Confidence Calibration, and Transfer of Knowledge. IAA Journal of Education 11 (1):1-4. https://doi.org/10.59298/IAAJE/2025/11114