

Transforming Mathematics Education In The Era Of Generative AI: From Problem Solving To AI-Supported Mathematical Reasoning

Baharuddin

Department of Mathematics, State University of Makassar, Indonesia

Corresponding Author: baharuddin.fmipa@unm.ac.id

Abstract

The development of Generative Artificial Intelligence (GenAI) has brought significant changes to mathematics education, particularly in problem-solving processes and mathematical reasoning. This study aims to analyze the transformation of GenAI utilization from an answer-generating tool into a facilitator of mathematical reasoning. The study employed a Systematic Literature Review (SLR) of relevant studies published between 2023 and 2026. The data were analyzed thematically to identify patterns of AI utilization, the role of AI in mathematical thinking processes, and the implications for mathematics education. The findings indicate that GenAI can be used not only to generate solutions but also to function as a cognitive scaffold by providing guiding questions, alternative strategies, feedback, error detection, and support for reflection. These findings suggest a shift from conventional problem solving toward AI-supported mathematical reasoning, in which learners remain the primary agents of their thinking processes while AI serves as a supportive tool. This transformation necessitates changes in teachers' roles, instructional design, and assessment practices, with greater emphasis on mathematical reasoning, argumentation, verification, and reflection. Therefore, the integration of GenAI should be directed toward enhancing the quality of students' mathematical reasoning rather than replacing their thinking processes.

Keyword: Generative AI, mathematics education, mathematical reasoning, problem solving, cognitive scaffolding.

INTRODUCTION

The development of Generative Artificial Intelligence (GenAI) has brought significant changes to the ways in which humans acquire, process, and generate knowledge, including in mathematics education. The emergence of Large Language Models (LLMs), such as ChatGPT, has

110

IJETER, Volume 04 Issue 04, 2026



Copyright to the author of some rights reserved This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).



transformed the role of technology from a mere computational aid into a dialogic partner capable of explaining concepts, generating examples, providing feedback, outlining solution procedures, and responding adaptively to users' questions. These changes are occurring at an unprecedented pace, requiring educational practices to adapt to the emerging characteristics of increasingly AI supported learning environments. Data from the Digital Education Council Global AI Student Survey 2024, which involved 3,839 students from 16 countries, indicated that 86% of students had used AI for their studies, while 54% reported using it at least weekly. ChatGPT was the most widely used AI tool, with 66% of respondents reporting its use in their studies. Subsequently, the 2025 Higher Education Policy Institute (HEPI) survey of 1,041 full-time undergraduate students in the United Kingdom revealed that 92% of respondents had used AI in some form, and 88% had used GenAI for assignments or assessments. These findings indicate that AI is no longer an experimental technological phenomenon but has become an established component of the learning ecosystem.

In mathematics education, these changes have more complex implications because the objectives of mathematics learning extend beyond obtaining correct answers. They also involve developing the ability to understand problems, select strategies, construct representations, provide explanations, formulate arguments, evaluate procedures, and reflect on the validity of solutions. Traditionally, problem solving has been a central orientation in mathematics education because it positions students as problem solvers who are expected to engage in systematic and independent thinking. However, the emergence of GenAI introduces a new situation in which some problem-solving processes that previously required students' direct engagement can now be performed or simulated by AI. For instance, Wei (2024) demonstrated that ChatGPT-4 and ChatGPT-4o could solve various mathematics items from the National Assessment of Educational Progress (NAEP) and generally outperformed U.S. students in several categories, although challenges remained in geometry, measurement, and more difficult problems. These findings suggest that the central challenge in mathematics education is no longer limited to how students can arrive at solutions. Rather, it increasingly concerns how students can interact critically with AI-generated solutions, verify their validity, identify errors, request justifications, compare strategies, and construct mathematically sound and defensible arguments.

The need for this shift in educational orientation becomes even more pressing when considered in relation to students' mathematical proficiency. The results of PISA 2022 showed that only 18% of Indonesian students achieved at least Level 2 in mathematics, substantially below the OECD average of 69%. Level 2 represents the minimum proficiency level at which





students begin to recognize and use simple mathematical representations in given situations without direct instructions. The OECD also reported that Indonesia's mathematics performance declined in 2022 compared with 2018 and represented one of the lowest levels recorded since Indonesia began participating in PISA. This situation highlights the fundamental need to strengthen students' mathematical thinking abilities. At the same time, the increasingly widespread use of GenAI presents two contrasting possibilities. AI can provide scaffolding, explanations, alternative strategies, and feedback that support students' thinking processes. However, applications focused primarily on obtaining instant answers may reduce opportunities for students to develop independent reasoning. The Digital Education Council survey further indicated that 58% of students felt they lacked sufficient AI knowledge and skills, while more than half expressed concerns about the consequences of excessive dependence on AI. Therefore, the challenge facing mathematics education in the GenAI era is not merely to integrate technology into instruction but to design interaction patterns that ensure AI strengthens, rather than replaces, mathematical thinking.

Several previous studies have begun to reveal the direction of this transformation. Schorcht et al. (2024) investigated the quality of AI supported mathematical problem solving through various prompting techniques. Their findings indicated that certain techniques, particularly Chain of Thought and Ask me Anything, could improve the quality of problem-solving processes, although mathematical content quality did not improve significantly across all prompting techniques. Meanwhile, Dilling & Herrmann (2024) explored the use of LLMs to support prospective mathematics teachers in constructing geometric proofs. Their study indicated that interaction with LLMs could become part of the mathematical reasoning process. Nevertheless, prospective teachers demonstrated limited understanding of how LLMs operate and exhibited several misconceptions about the technology. At the literature review level, systematic reviews of ChatGPT in school mathematics education have also identified benefits, including rapid feedback and personalized support, while highlighting limitations in complex reasoning and contextual understanding, as well as the risk of user dependence. Thus, previous research has provided evidence that GenAI can support problem solving and specific mathematical activities. However, the relationship between AI use and the development of mathematical reasoning still requires deeper conceptual examination.

Although research on AI in mathematics education is expanding rapidly, several limitations remain and provide opportunities for further investigation. First, many studies continue to focus on AI effectiveness in improving learning outcomes, solving mathematical problems, providing feedback, or generating instructional materials. Consequently, the





transformation of mathematical thinking processes has not yet become a primary focus. Second, research on GenAI and mathematics remains fragmented according to technological functions, such as tutoring, problem solving, feedback, assessment, and content generation. Relatively few studies have explicitly explained the paradigm shift from problem solving toward AI supported mathematical reasoning. Third, the available empirical evidence is still relatively recent, with many studies focusing on user perceptions, experiences, or experiments involving a single type of mathematical task. A 2025 systematic review of ChatGPT in mathematics education, which analyzed 20 empirical studies, identified a predominance of qualitative and mixed-methods approaches and highlighted the limited availability of long-term evidence regarding the effects of AI use on mathematics learning. A recent meta-analysis of 22 empirical studies, comprising 46 independent samples and 5,232 participants, found a moderate positive effect of GenAI on mathematics learning outcomes. However, the magnitude of this effect was influenced by the level of GenAI integration, sample size, and type of instructional content. These findings suggest that the literature is moving beyond the question of whether AI helps students learn mathematics toward a more fundamental question: how does AI transform the ways in which students engage in mathematical reasoning?

In response to these research gaps, this study introduces a novel perspective by positioning GenAI not merely as a tool for obtaining answers or solving problems but as a form of cognitive and dialogic support that can facilitate the development of mathematical reasoning. In this study, the concept of AI supported mathematical reasoning refers to processes through which students use AI to develop and test conjectures, request justifications, compare alternative solution strategies, verify the correctness of mathematical procedures, identify errors or misconceptions in AI-generated responses, construct arguments, and reflect on their solutions. From this perspective, the focus of mathematics learning shifts from simply being able to solve problems to being able to explain why a solution is valid, evaluate AI-generated outputs, and make mathematical decisions based on evidence and logical reasoning. Another contribution of this study lies in its integrative mapping of changes in AI functions, student roles, teacher roles, mathematical task characteristics, human AI interaction patterns, and the reasoning competencies identified in the literature. This approach extends previous research, which has often examined the benefits and risks of GenAI separately, by offering a conceptual framework for understanding how AI can function as a scaffold for reasoning rather than as a substitute for mathematical thinking.

Accordingly, the study entitled “Transforming Mathematics Education in the Era of Generative AI: From Problem Solving to AI-Supported Mathematical Reasoning” is important for





developing a more systematic understanding of the ongoing transformation of mathematics education. The urgency of this investigation stems not only from the rapid adoption of GenAI but also from the need to ensure that technological transformation remains aligned with the fundamental goals of mathematics education: fostering logical, critical, argumentative, reflective, and evidence-based thinking. UNESCO has likewise emphasized the importance of a human-centered approach to the use of GenAI in education, together with pedagogical design and validation processes that ensure its safe, ethical, meaningful, and contextually appropriate implementation. Therefore, this literature review seeks to synthesize recent research developments, identify patterns of GenAI use in mathematical reasoning activities, examine its benefits and risks, map the limitations of the available empirical evidence, and formulate a conceptual direction for mathematics education that positions AI as a supporter of reasoning processes. The findings are expected to contribute theoretically to the development of the concept of AI supported mathematical reasoning while providing a foundation for teachers, lecturers, researchers, and curriculum developers to design mathematics learning environments that are not only responsive to technological advances but also firmly oriented toward the quality of human mathematical thinking.

METHOD

This study employed a Systematic Literature Review (SLR) to identify, analyze, and synthesize research on the use of Generative Artificial Intelligence (GenAI) in mathematics education, particularly in relation to problem solving and mathematical reasoning. The review process adapted the PRISMA framework, comprising identification, screening, eligibility assessment, and the selection of studies for analysis. Relevant literature was retrieved through searches of the Google Scholar academic database. The review was limited to publications from 2023 and 2026, following the widespread adoption of ChatGPT and other GenAI technologies.

Data were analyzed using thematic analysis by categorizing the findings into five themes: (1) GenAI as a tool for supporting problem solving; (2) GenAI as a learning scaffold; (3) GenAI support for mathematical reasoning; (4) student AI interaction; and (5) risks and challenges associated with GenAI use, including dependency, inaccurate information, and reduced independent thinking.

Furthermore, the findings were compared to identify the shift in AI's role from merely generating answers (answer generator) to supporting reasoning processes (reasoning scaffold). The synthesis examined patterns, similarities, differences, limitations, and research gaps in previous studies. The resulting synthesis was used to develop a conceptual understanding of AI-



supported mathematical reasoning and to identify directions for mathematics education that leverage GenAI without replacing students' mathematical thinking processes.

RESULTS AND DISCUSSION

RESULT

1. Results and Synthesis of Previous Studies

Based on the literature search, ten articles relevant to the use of Generative Artificial Intelligence (GenAI) in mathematics education were identified, particularly those addressing problem solving, mathematical reasoning, scaffolding, mathematical proof, and student–AI interaction. The studies analyzed were published between 2023 and 2026. The synthesis indicates a shift in research focus from using AI primarily as an answer generating tool toward employing AI to support mathematical thinking and reasoning processes.

Table 1. Synthesis of Previous Studies

No.	Authors/Year	Research Focus	Key Findings	Relevance to the Present Study
1	Son (2023)	The use of ChatGPT in fraction learning and the analysis of mathematical thinking	ChatGPT can assist in analyzing students' responses and supporting mathematics learning; however, AI-generated outputs still require human verification.	Demonstrates the potential of AI to support the analysis of mathematical thinking processes.
2	Dasari et al. (2024)	The role of ChatGPT in mathematics education	Using ChatGPT without teacher guidance does not necessarily lead to improved learning outcomes.	Highlights the importance of collaboration among teachers, students, and AI.
3	Chao & Han (2026)	The use of GenAI to provide mathematical problem-solving hints	GenAI can generate various forms of problem-solving hints through specific prompting techniques.	Illustrates the development of AI from an answer provider into a problem-solving support tool.

4	Schorcht et al. (2024)	AI-supported mathematical problem solving and prompting techniques	The quality of AI responses is influenced by how users formulate their instructions or prompts.	Highlights the importance of AI interaction skills in mathematics learning.
5	Dilling & Herrmann (2024)	The use of LLMs in geometric proof construction	LLMs can support the proof-construction process, although users have limited understanding of how AI works and of its limitations.	Emphasizes the importance of AI literacy and the ability to evaluate mathematical arguments.
6	Yoon et al. (2024)	GenAI in mathematical proof	GenAI can support mathematical proof construction, but users must be able to evaluate the correctness and validity of AI-generated outputs.	Directly relates to the concept of AI-supported mathematical reasoning.
7	Turmuzi et al., (2025)	A systematic review of ChatGPT in mathematics education	ChatGPT provides feedback and learning support but has limitations in complex problem solving and may encourage dependency.	Identifies both the benefits and risks of GenAI use.
8	Liu et al., (2026)	The effects of GenAI on mathematics learning outcomes	GenAI has a positive effect on mathematics learning outcomes, although its effectiveness depends on instructional design and the degree of AI integration.	Demonstrates that AI effectiveness is not automatic but depends on how it is implemented.
9	Malik et al., (2026)	Polya-based prompting for problem-solving scaffolding	Prompts structured around Polya's problem-solving stages can provide systematic support for mathematical problem solving.	Connects established problem-solving theory with the use of GenAI.

10	Davor (2026)	GenAI as a cognitive scaffold for reasoning and geometric proof	GenAI can support connections between visual representations, deductive reasoning, and proof construction.	Reinforces the role of AI as a facilitator of mathematical reasoning.
----	--------------	---	--	---

2. Thematic Synthesis of the Findings

The ten studies reveal several major findings. First, GenAI has evolved from functioning primarily as an answer generator to serving as a tool for supporting problem solving. Research by Son (2023), Schorcht et al. (2024) indicates that AI can do more than provide final answers. It can also generate hints, explain solution procedures, suggest alternative strategies, and provide feedback to learners. These capabilities suggest that GenAI has considerable potential to function as a scaffold in mathematics education.

Second, the effectiveness of GenAI depends substantially on the nature of student–AI interaction. Schorcht et al. (2024) demonstrated that the quality of AI-generated responses is influenced by prompting techniques. This suggests that students need more than mathematical problem-solving skills. They also need to formulate appropriate questions, provide clear instructions, critically assess AI responses, and determine whether the information generated is mathematically valid.

Third, teachers remain essential in GenAI-supported learning environments. Findings attributed to Dasari et al. indicate that using ChatGPT independently does not automatically improve the quality of learning. Teachers continue to play a crucial role in providing guidance, clarifying mathematical concepts, evaluating solutions, and ensuring that AI functions as a learning support tool rather than a substitute for students' thinking processes.

Fourth, recent research suggests a shift from conventional problem solving toward mathematical reasoning. Studies by Dilling & Herrmann (2024), Yoon et al. (2024), and Davor (2026) indicate that GenAI is increasingly being explored in activities involving mathematical proof, argumentation, and reasoning. In these contexts, students are expected not only to obtain answers but also to explain their reasoning, verify the validity of mathematical steps, identify errors, and justify their solutions using sound mathematical arguments.

Fifth, the use of GenAI involves several risks. AI systems may generate incorrect or inconsistent answers or responses that do not adequately address the context of a problem. Furthermore, excessive reliance on AI may encourage dependency, leading students to accept generated answers without verification. Consequently, the ability to critically evaluate AI-generated outputs constitutes an essential competency in mathematics education in the GenAI

era.

3. The Conceptual Transformation of Mathematics Learning

Overall, the findings suggest the following progression in the role of AI in mathematics education.

Conventional Problem Solving

Students understand the problem → select a strategy → execute procedures → obtain a solution.

AI as an Answer Generator

Students submit a problem → AI generates a solution → students receive the answer.

AI as a Learning Scaffold

Students request hints → AI provides solution steps or guiding questions → students work toward solving the problem.

AI-Supported Mathematical Reasoning

Students formulate conjectures → engage in dialogue with AI → compare strategies → verify AI-generated answers → identify errors → provide justifications → evaluate evidence → revise solutions → reflect on the reasoning process.

This progression illustrates a conceptual shift in the role of AI, from producing answers to supporting students' active engagement in mathematical reasoning. In the final stage, AI is not merely a source of solutions but a dialogic and cognitive support mechanism that encourages students to question, verify, justify, and reflect on mathematical ideas.

4. Overall Interpretation of the Findings

The synthesis indicates that GenAI creates opportunities to transform the orientation of mathematics education. The focus of learning should no longer be limited to students' ability to obtain correct answers; it should also encompass their ability to explain mathematical reasoning, evaluate strategies, verify AI-generated outputs, identify errors, construct arguments, and reflect on problem-solving processes.

Accordingly, AI-supported mathematical reasoning can be conceptualized as a further development of AI-supported problem solving. Within this approach, AI is not positioned as a replacement for students' thinking processes but as a dialogic partner and cognitive scaffold that helps learners develop mathematical reasoning in a more critical and reflective manner.

These findings also reveal a research gap: there remains a need for more comprehensive



conceptual studies that integrate GenAI, problem solving, mathematical reasoning, the evaluation of AI-generated outputs, and the teacher's role within a unified framework. Therefore, the present study seeks to address this gap through a synthesis of the literature on the transformation of mathematics education from conventional problem solving toward AI-supported mathematical reasoning.

DISCUSSION

The findings of this study indicate that the transformation of mathematics education in the era of Generative Artificial Intelligence (GenAI) involves more than the adoption of new technologies; it also entails fundamental changes in how students construct knowledge and engage in mathematical activities. The literature suggests that GenAI is gradually shifting from a tool primarily used to generate answers toward a tool capable of providing scaffolding, feedback, guiding questions, and opportunities for dialogue throughout the problem-solving process. For example, Zhou et al. (2025) found that ChatGPT-facilitated scaffolding could improve university students' problem-solving performance, while high-performing students demonstrated more interpretive and evaluative activities associated with metacognitive representations. These findings suggest that the educational value of GenAI lies not primarily in its ability to generate answers quickly but in its potential to stimulate higher-order cognitive activities.

First, constructivism provides a theoretical perspective for explaining this transformation. From a constructivist perspective, mathematical knowledge is not simply transferred from a source to learners; rather, it is constructed through activity, experience, reflection, and interaction with the learning environment. In the context of GenAI, AI can be positioned as a learning environment or dialogic partner that provides information, questions, examples, counterexamples, and alternative strategies. Students must then process this information and connect it with their existing knowledge. Accordingly, a constructivist approach to AI use does not involve simply requesting a final answer but rather using AI-generated responses as resources for developing personal understanding. This perspective is consistent with research on GenAI use in complex mathematical tasks, which indicates that AI can generate responses but may also produce errors. Interaction and prompting can therefore be used to explore, refine, and critically evaluate AI-generated responses.

Second, scaffolding theory provides a strong foundation for explaining how AI can support problem-solving processes. Scaffolding refers to assistance provided when learners encounter tasks they cannot yet complete independently, with that assistance gradually reduced





as their competence develops. In the context of GenAI, such support may take the form of guiding questions, hints, examples, feedback, or requests for students to explain the reasoning behind a particular step. Zhou et al. (2025) demonstrated that ChatGPT-based scaffolding could support problem-solving activities and encourage interpretive and evaluative engagement. Malik et al., (2026) also developed Polya-aligned prompting based on four stages: understanding the problem, devising a plan, carrying out the plan, and looking back. Their findings indicated improved performance among students receiving ChatGPT-based scaffolding, although strategic planning remained a challenge.

Third, instrumental genesis offers an important perspective for understanding why GenAI does not automatically become a learning instrument merely because it is available. Dilling & Herrmann (2024) employed instrumental genesis theory to explain the relationship between learners and LLMs. According to this theory, a technological artifact becomes an instrument through the process of use and the development of activity schemes by the user. In other words, ChatGPT becomes a mathematical instrument only when learners develop ways of using the technology for specific mathematical purposes. In mathematics education, this process may involve learning to formulate productive prompts, request justifications, test alternative solutions, and evaluate AI-generated responses. Dilling and Herrmann found that prospective teachers had limited experience with LLMs, and some demonstrated superficial understanding or misconceptions about how the technology operates. This finding indicates that AI integration requires the development of competencies for using technological instruments rather than merely providing access to the technology.

Fourth, the metacognitive perspective reinforces the argument that the primary goal of AI-supported learning is not to make students increasingly dependent on AI but to help them develop the ability to regulate their own thinking. Metacognition involves the ability to plan, monitor, and evaluate cognitive processes. In mathematical problem solving, students need to determine what they understand about a problem, which strategies they are using, whether those strategies are effective, and whether the resulting answers are reasonable. When GenAI is used appropriately, its responses can become objects of monitoring and evaluation. Students can ask why a particular procedure is appropriate, request alternative solutions, identify inconsistencies, and compare AI-generated answers with their own reasoning. Zhou et al. (2025), in examining differences between high- and low-performing groups, highlighted the importance of interpretive and evaluative activities in AI-supported scaffolding.

Fifth, dialogic learning theory helps explain why interactions between students and GenAI may have pedagogical value. From this perspective, knowledge develops through





questioning, responding, clarification, argumentation, and the negotiation of meaning. GenAI can sustain rapid exchanges, allowing students to ask follow-up questions without waiting for a teacher's response. However, such dialogue becomes pedagogically meaningful only when students do not automatically accept every AI-generated response as correct. Research examining AI responses to challenging mathematical problems indicates that GenAI can make substantial errors, although appropriate prompting may enable it to revise and improve its answers. This situation creates opportunities to use AI errors as objects of mathematical discussion. Accordingly, AI dialogue can be designed to encourage questions such as "Why?", "Is this correct?", "How can this be proved?", and "Is there another strategy?"

Sixth, the perspective of higher-order thinking suggests that the shift from conventional problem solving toward AI-supported mathematical reasoning must be accompanied by changes in the types of activities students undertake. When AI is used merely to request solutions to mathematical problems, students' cognitive engagement may be reduced to receiving information and following procedures. Conversely, when students are required to evaluate AI-generated solutions, compare two strategies, identify errors, provide justifications, or construct their own proofs, they engage in analysis, evaluation, and creation. Zhou et al. (2025) found that university students who performed well when using ChatGPT-based scaffolding were more likely to engage in interpretive and evaluative activities. This finding suggests that AI interaction should be deliberately designed to promote higher-order cognitive processes rather than simply improve the efficiency of obtaining answers.

Seventh, the mathematical reasoning perspective places justification, argumentation, proof, and evaluation at the core of mathematics education. This shift is particularly important because the ability to produce an answer does not necessarily demonstrate mathematical understanding. Dilling & Herrmann (2024) showed that LLMs could be used in geometric proof-construction activities, although such use requires an understanding of the characteristics and limitations of LLMs. Meanwhile, research on formal mathematical reasoning in AI emphasizes the importance of verifiable reasoning and outputs that can be justified formally. In educational contexts, this principle can be translated into the expectation that students should not stop at stating that "the AI says this is the answer" but should instead explain the mathematical reasoning that supports the answer.

Eighth, cognitive apprenticeship theory helps explain the changing role of teachers in GenAI-supported learning environments. In this approach, learning occurs through modeling, scaffolding, coaching, articulation, reflection, and exploration. GenAI may assume some supportive functions, such as providing examples or preliminary feedback. However, teachers





remain essential for designing tasks, establishing cognitive objectives, guiding discussions, and ensuring the quality of students' reasoning. Research on the use of LLMs for mathematics curriculum scaffolding suggests that teacher-led scaffolding involves observing students' needs, formulating strategies, and implementing appropriate support. LLMs can assist by providing adaptations and supplementary materials, but their use must remain within a framework of teacher-led pedagogical decision-making.

Ninth, the perspective of human–AI collaboration suggests that the most productive relationship is neither one of students competing against AI nor one in which AI replaces teachers. Instead, it involves a complementary distribution of responsibilities between humans and technology. GenAI is particularly capable of generating alternative solutions, re-explaining information, and responding rapidly to questions. Humans, by contrast, remain responsible for determining objectives, assessing relevance, verifying correctness, understanding context, and making mathematical decisions. Walkington (2025) emphasized that the emergence of GenAI has important implications for mathematics education because the ability of AI models to respond to complex questions is changing how teachers and students interact with mathematical tasks. A meta-analysis published in 2026 also reported a moderate positive effect of GenAI on mathematics learning outcomes, although the magnitude of this effect varied according to the degree of GenAI integration, sample characteristics, and instructional content. These findings reinforce the argument that AI effectiveness is not an inherent property of the technology alone but also depends on the design of human–AI interactions.

Tenth, these theoretical perspectives can be integrated into an AI-supported mathematical reasoning framework that serves as the conceptual foundation of this study. This framework positions GenAI as a cognitive scaffold, dialogic partner, and mathematical instrument that supports mathematical activity. Students remain the primary agents responsible for understanding problems, developing strategies, generating conjectures, examining AI-generated outputs, providing justifications, evaluating arguments, and reflecting on solutions. Teachers, in turn, function as designers of learning environments, providers of pedagogical scaffolding, and facilitators of reasoning validation. Thus, the transformation does not involve replacing problem solving with AI but rather reorienting problem solving toward reasoning activities supported by AI. Recent research provides an empirical basis for this direction, although a more systematic pedagogical framework is still needed to determine how, when, and for which cognitive purposes GenAI should be used in mathematics education.

Taken together, these ten perspectives lead to one central proposition: GenAI should not be positioned as a mathematical answer-generating machine but as a means of supporting the



processes of understanding, reasoning, verification, argumentation, and reflection. This proposition provides the conceptual foundation for the shift from conventional problem solving toward AI-supported mathematical reasoning.

CONCLUSION

GenAI is driving a transformation in mathematics education, shifting the role of AI from an answer generator to a cognitive scaffold that supports mathematical reasoning. Its use should be directed toward activities involving understanding, evaluating, verifying, justifying, and reflecting on mathematical solutions. Accordingly, AI-supported mathematical reasoning positions students as the primary agents of their thinking processes, AI as a supportive tool, and teachers as facilitators of learning.

Future research should empirically examine the effectiveness of AI-supported mathematical reasoning, develop AI-supported instructional models and mathematical reasoning assessment instruments, and investigate the long-term effects of GenAI use, teacher competencies, AI literacy, and ethical considerations in mathematics education.

REFERENCE

- Chao, Z., & Han, T. (2026). Applications, multidimensional challenges, and innovative pedagogical practices of generative AI in mathematics education: A systematic review 2022–2026. *Frontiers in Education*, 11, 1931535. <https://doi.org/10.3389/feduc.2026.1931535>
- Davor, I. (2026). Generative AI-supported instruction as a cognitive scaffold: Effects on senior high school students' geometric reasoning and proof construction. *Frontiers in Education*, 11, Article 1869961. <https://doi.org/10.3389/feduc.2026.1869961>
- Dasari, D., Hendriyanto, A. H., Sahara, S., Suryadi, D., Muhaimin, L. H., Chao, T., & Fitriana, L. (2024). ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning. *Frontiers in Education*, 8, Article 1295413. <https://doi.org/10.3389/feduc.2023.1295413>
- Digital Education Council. (2024). *Global AI student survey 2024*. <https://www.digitaleducationcouncil.com/resource-library-items/digital-education-council-global-ai-student-survey-2024>
- Dilling, F., & Herrmann, M. (2024). Using large language models to support pre-service teachers' mathematical reasoning—An exploratory study on ChatGPT as an instrument for creating mathematical proofs in geometry. *Frontiers in Artificial Intelligence*, 7, Article 1460337.



- <https://doi.org/10.3389/frai.2024.1460337>
- Liu, B., Zhang, W., & Wang, F. (2026). Can generative artificial intelligence effectively enhance students' mathematics learning outcomes? A meta-analysis of empirical studies from 2023 to 2025. *Education Sciences*, 16(1), Article 140. <https://doi.org/10.3390/educsci16010140>
- Malik, M. A., Lince, R., & Husnaeni. (2026). A Polya-aligned prompting protocol for ChatGPT scaffolding: Evidence from eighth-grade systems-of-linear-equations problem solving. *International Journal of Environment, Engineering and Education*, 8(1). <https://doi.org/10.55151/ijeedu.v8i1.331>
- Organisation for Economic Co-operation and Development. (2023). *PISA 2022 results (Volume I and II): Country note—Indonesia*. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/indonesia_c2e1ae0e-en.html
- Schorcht, S., Buchholtz, N., & Baumanns, L. (2024). Prompt the problem—Investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education*, 9, Article 1386075. <https://doi.org/10.3389/feduc.2024.1386075>
- Son, T. (2023). Exploring the possibility of using ChatGPT in mathematics education: Focusing on student product and pre-service teachers' discourse related to fraction problems. *Education of Primary School Mathematics*, 26(2), 99–113. <https://doi.org/10.7468/jksmec.2023.26.2.99>
- Turmuzi, M., Azmi, S., & Kertiyani, N. M. I. (2025). ChatGPT in school mathematics education: A systematic review of opportunities, challenges, and pedagogical implications. *Teaching and Teacher Education*, 156, Article 105286. <https://doi.org/10.1016/j.tate.2025.105286>
- Walkington, C. (2025). The implications of generative artificial intelligence for mathematics education. *School Science and Mathematics*, 126(4), 356–365. <https://doi.org/10.1111/ssm.18356>
- Wei, X. (2024). Evaluating ChatGPT-4 and ChatGPT-4o: Performance insights from NAEP mathematics problem solving. *Frontiers in Education*, 9, Article 1452570. <https://doi.org/10.3389/feduc.2024.1452570>
- Yoon, H., Hwang, J., Lee, K., Roh, K. H., et al. (2024). Students' use of generative artificial intelligence for proving mathematical statements. *ZDM—Mathematics Education*, 56, 1531–1551. <https://doi.org/10.1007/s11858-024-01629-0>
- Zhou, R., He, X., Fan, Q., Li, Y., Li, Y., Xiao, X., & Fang, J. (2025). Exploring ChatGPT-facilitated scaffolding in undergraduates' mathematical problem solving. *Journal of Computer*



