International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 3, pp. 270-289, March 2026
https://doi.org/10.26803/ijlter.25.3.12
Received Dec 13, 2025; Revised Feb 18, 2026; Accepted Feb 19, 2026

AI in Students' Mathematics Learning: A PRISMA-Based Systematic Review of Challenges and Solutions (2021-2025)

Yaqian Song, Mohamed Yusoff Mohd Nor*[*] and Bity Salwana Alias
Universiti Kebangsaan Malaysia,
Bangi, Malaysia

Abstract. Digital technology has fundamentally transformed the landscape of education. Recent advances in artificial intelligence (AI) are redefining educational practices and creating new opportunities for mathematics learning. However, despite growing interest, AI usage in mathematics learning remains in its early stages. Existing studies demonstrate the limited integration of AI tools into mathematics-specific pedagogy and insufficient focus on students' mathematical problem-solving abilities, while there remains a need for the challenges associated with AI implementation to be systematically categorized. To address these gaps, this study conducted a systematic literature review (SLR) following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. Twelve eligible quantitative studies published between 2021 and 2025 were retrieved from Web of Science (WoS) and Scopus databases using rigorous screening procedures and predefined inclusion criteria. Subsequently, these eligible studies underwent in-depth analysis to synthesize research characteristics, AI technologies implemented, challenges encountered in AI-supported mathematics learning, and corresponding solutions. The findings indicate that China contributed the largest number of publications. The reviewed studies predominantly employed quasi-experimental designs with short intervention durations and small sample sizes, with most research being conducted in primary education settings. Three primary types of AI technologies were identified: adaptive learning systems, intelligent tutoring systems, and chatbots. This review highlighted three major categories of challenges and solutions: research-related, technical and design-related, and educational-pedagogical. Finally, this study offers evidence-based insights to support researchers, developers, educators, and policymakers in aligning AI technologies with the cognitive and problem-solving demands of mathematics learning, informing more focused instructional design.

[*]Corresponding author: Mohamed Yusoff Mohd Nor; *yusoff1963@ukm.edu.my*

Keywords: AI; mathematics learning; challenges; solutions; PRISMA

1. Introduction
As society transitions to the digital era, digital transformation is reshaping education by redefining the way in which knowledge is delivered, accessed, and experienced, fostering more inclusive, engaging, and adaptive learning environments (Aditya & Suranto, 2024; Zou et al., 2025). Emerging technologies such as AI, virtual and augmented reality, cloud-based learning management systems, and learning analytics are reforming instructional practices. Together, they are shifting traditional teacher-centered approaches toward more dynamic, interactive, and learner-centered learning experiences that enhance understanding, flexibility, and learner autonomy (Alam et al., 2025; Timotheou et al., 2023; Wang et al., 2024). Within this landscape, AI has emerged as an especially powerful catalyst for innovation, extending beyond conventional digital tools to drive systemic improvements in teaching, learning, and educational design (Merino-Campos, 2025).

Characterized by its data-driven processes, pattern recognition, and adaptive feedback, AI has gained increasing prominence as it continues to permeate daily life (Guan et al., 2020). It now serves as a key driver of economic and social progress, supporting automation, technological advancements, and new productivity structures (Filippucci et al., 2024). Particularly in the realm of education, however, AI's rapid expansion has raised ethical and academic concerns. Extensive data collection introduces risks such as privacy breaches, data insecurity, and algorithmic bias (Golda et al., 2024; Sajja et al., 2025), while excessive reliance on AI-driven systems may erode students' higher-order thinking skills, including analytical reasoning, critical thinking and creativity (Adiyono et al., 2025; Rahmatika et al., 2025). Such concerns underscore the need for responsible and balanced AI integration in educational practice, such as pedagogical strategies integration.

The adoption of AI in education has been widely recognized for its potential to enhance pedagogical innovation and foster higher levels of student engagement and motivation (Khazanchi et al., 2025; Liu et al., 2025). AI applications such as adaptive learning systems, intelligent tutoring systems, and AI-powered learning analytics further contribute to cognitive and emotional learning by tailoring instructional materials to learners' needs, preferences, and prior knowledge (Luo et al., 2025; Vistorte et al., 2024).

Despite these benefits, insufficient AI literacy and inadequate professional training continue to hinder teachers' ability to effectively integrate, evaluate, and ethically apply AI tools in classroom settings (Daher, 2025; Zhao et al., 2022). Consequently, AI is often used as a technological supplement rather than being pedagogically embedded within subject-oriented instructional design. Notably, research on subject-focused AI integration, particularly in mathematics learning, remains limited (Mohamed et al., 2022; Opesemowo & Adewuyi, 2024), with few SLR articles to date addressing this domain.

Mathematics learning has long been regarded as a fundamental pillar of students' cognitive development and academic success. As an essential discipline underpinning science, technology, engineering, and mathematics (STEM), it plays a central role in shaping students' analytical and problem-solving abilities (Goos et al., 2023; Just & Siller, 2022). With the increasing integration of AI into STEM education (Xu & Ouyang, 2022), new opportunities emerge to reshape mathematics education and enhance students' comprehensive mathematical literacy. Specifically, the effective use of AI can strengthen students' higher-order thinking skills, including analytical processing, evaluative judgment, and creative problem-solving (Alvarez, 2024; Wang & Fan, 2025).

Despite the growing importance of AI in mathematics learning, notable gaps remain in the existing research. First, AI applications in mathematics learning are still in their infancy. Pai et al. (2021) noted that dialogue-based intelligent tutoring systems remain underexplored in Chinese-language contexts. Existing studies have focused mainly on generic AI applications, while paying limited attention to AI tools for problem-solving (Alvarez, 2024; Awang et al., 2025; Nguyen & Pham, 2025). Second, the pedagogical integration of AI in mathematics learning remains insufficiently aligned with subject-specific needs. Many applications lack adequate disciplinary responsiveness and consequently struggle to effectively support key tasks such as mathematical word problem solving (Liu et al., 2025). Furthermore, although AI integration in mathematics learning presents multiple challenges, few systematic literature reviews have yet attempted to categorize these challenges or synthesize corresponding solutions.

Therefore, in order to address these gaps, it is necessary to systematically synthesize existing studies to map the current research landscape, examine the types of AI technologies employed, and identify the practical challenges and corresponding solutions associated with AI implementation in mathematics learning. Such a synthesis can serve to deepen understanding of AI's educational value and human–AI interaction dynamics. Guided by this rationale, the present review was structured around the following research questions:
1）What are the characteristics of existing research on AI in students' mathematics learning?
2）What types of AI technologies have been implemented to support mathematics learning?
3）What are the main challenges encountered in integrating AI into students' mathematics learning?
4）What solutions have been proposed or implemented to address these challenges?

2. Methodology

To ensure methodological rigor, transparency, and replicability throughout the review process, this study adhered to PRISMA guidelines (Moher et al., 2010; Page et al., 2021). By standardizing the procedures for identification, screening, and inclusion, the PRISMA framework enhances the reliability of systematic reviews and helps reduce potential bias (Shamseer et al., 2015). The study selection process was based on the PRISMA 2020 Statement developed by Page et al. (2021), which includes a rigorous checklist and flow diagram. The overall screening and selection procedures are presented in Figure 1.

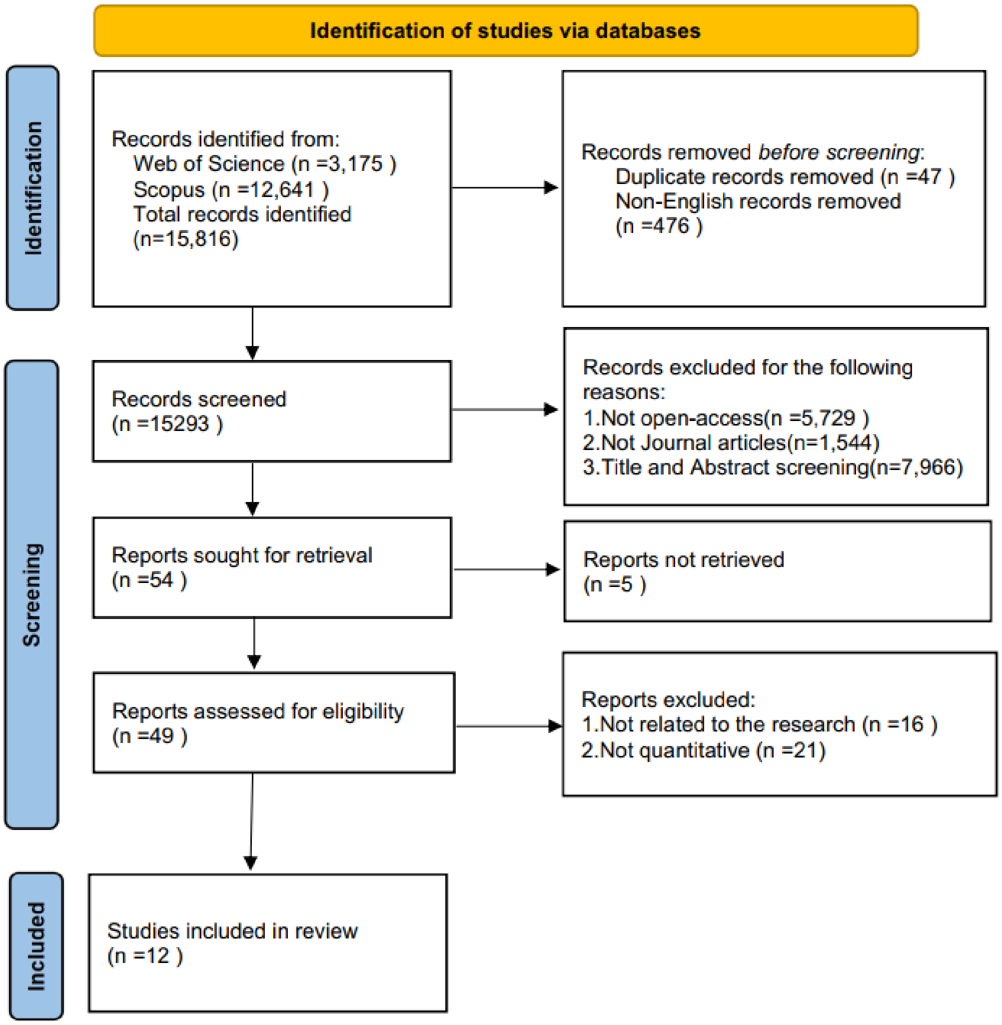

Figure 1: Data Collection Process Based on the PRISMA Guidelines

2.1 Data search sources

In this study, WoS and Scopus were selected because of their extensive coverage, rigorous quality standards, and high academic prestige (Gasparyan et al., 2013). Although the search was restricted to two databases, both provide broad access to high-quality articles in the field of social sciences (Munárriz & Rincón, 2025).

During the initial retrieval, the WoS search yielded an exceptionally large volume of records; therefore, the search was restricted to the Education/Educational Research category to ensure disciplinary relevance. In WoS, the search was conducted within the TITLE field, whereas in Scopus, the search was performed within the TITLE-ABS-KEY field. Two main search strings were applied, and the complete search formula is presented in Table 1.

Table 1: Database search field and search strings

Database	Search field	Search string
WoS	TITLE	("AI" OR "artificial intelligence" OR "intelligent tutoring system" OR "personalized learning") AND ("mathematics learning" OR "math learning" OR "mathematics education" OR "math education" OR "mathematics achievement" OR "math achievement" OR "problem solving")
Scopus	TITLE-ABS-KEY	("AI" OR "artificial intelligence" OR "intelligent tutoring system" OR "personalized learning") AND ("mathematics learning" OR "math learning" OR "mathematics education" OR "math education" OR "mathematics achievement" OR "math achievement" OR "problem solving")

2.2 Eligibility and exclusion criteria

In systematic reviews, it is essential to define clear inclusion and exclusion criteria to ensure a structured and unbiased selection of studies (Granados-Duque & García-Perdomo, 2021). Accordingly, this study established specific criteria to guide the screening process; these are summarized in Table 2.

Table 2: The inclusion and exclusion criteria

Criteria	Eligibility	Exclusion
Timeline	2021-2025	Prior to 2021
Scope	Studies related to AI in mathematics learning	Studies not related to AI in mathematics learning
Language	English	Non-English
Research method	Quantitative	Non-quantitative
Literature type	Journal articles	Book series, books, chapters in book, conferences proceedings and editorials
Article access	Full text available	No full-text access

The first criterion concerned the timeline: a five-year period (2021–2025) was selected to capture the most recent research trends, given the rapid evolution of AI technologies and their expanding applications in education. Limiting the review to English-language publications ensured consistency in interpretation and evaluation, thereby enhancing the overall reliability of the review (Awang et al., 2025).

Only quantitative studies were included, as these provide objective and comparable results that support the synthesis of reliable findings in a systematic review (Gulz et al., 2020; Mills, 2021). Research on AI-assisted mathematics learning commonly evaluates learning outcomes using measurable indicators such as test scores and standardized questionnaires (Huang et al., 2021; Mohamed et al., 2022). In order to maintain scholarly standards, only journal articles were included, while book series, books, book chapters, conference proceedings, and editorials were excluded. Additionally, only studies with full-text availability were retained to ensure the analysis was based on complete and verifiable information.

Initially, 15,816 records were identified across both databases, including 3,175 from WoS and 12,641 from Scopus. After removing 47 duplicates and 476 non-English publications, 15,293 records remained for screening. Subsequently, 5,729 non–open-access papers and 1,544 non-journal articles were excluded. An additional 7,966 studies were removed during title and abstract screening due to being outside the scope of the review, leaving 54 reports for full-text retrieval. The full texts of five articles were inaccessible and were therefore excluded, resulting in 49 full-text articles being assessed for eligibility. Of these, 16 studies were excluded for lack of relevance, and 21 were removed because they employed non-quantitative methods. Ultimately, 12 quantitative studies met all inclusion criteria and were included in the final review.

2.3 Inclusion

A data extraction table was developed to summarize the key characteristics of the selected studies, as shown in Table 3.

Table 3: Key characteristics of the selected studies

Author（s）	Year	Country or Region	Study Design	Educational Level	Sample Size	Intervention Duration
Pai et al.	2021	Taiwan	Quasi-experiment	Primary school	134	2 weeks
Thai et al.	2022	The United States	Quasi-experiment	Early childhood	453	12-14 weeks
del Olmo-Muñoz et al.	2022	Spain	Quasi-experiment	Primary school	110	8 weeks
Shih et al.	2023	Taiwan	Quasi-experiment	Primary school	66	2 hours
Huang et al.	2024a	China	Quasi-experiment	Primary school	69	4weeks
Alkhasawneh	2025	Jordan	Quasi-experiment	Secondary school	76	Not specified
Khazanchi et al.	2025	The United States	Quasi-experiment	Secondary school	78	5 weeks
Li and Lyu	2025	The United States	Quasi-experiment	Higher education	151	45 minutes
Lin et al.	2025	Taiwan	Quasi-experiment	Primary school	4442	8 months
Lin and Jiang	2025	China	Questionnaire	Higher education	331	3 weeks
Liu et al.	2025	China	Quasi-experiment	Primary school	104	40 minutes
Qiu and Ishak	2025	China	Questionnaire	Primary school	387	8 weeks

3. Results and Discussion

To provide a comprehensive understanding of the findings, this section synthesized the included studies across the following four key dimensions: research characteristics; AI technologies used in mathematics learning; challenges associated with AI integration; and the solutions proposed or implemented to address these challenges.

3.1 The characteristics of selected studies

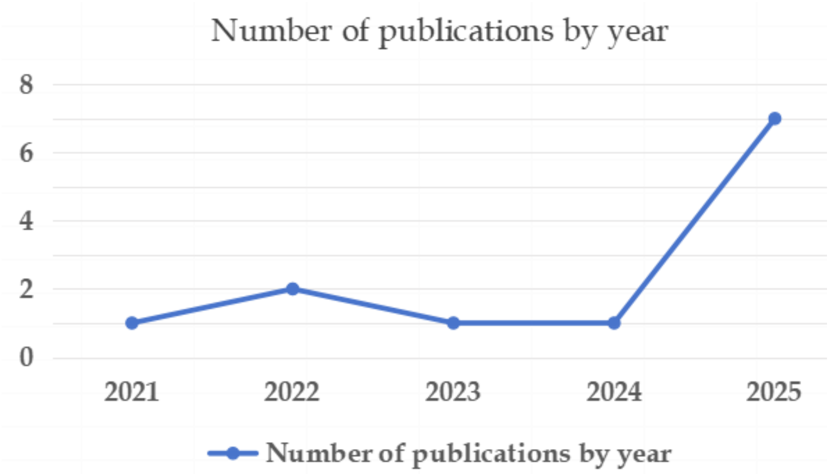

Figure 2: Number of publications by year

Figure 2 shows the annual distribution of the included studies from 2021 to 2025. Research output remained limited between 2021 and 2024, with only one publication in 2021, 2023, and 2024, and a modest increase in 2022 (n = 2). A substantial rise occurred in 2025, with seven studies published. This upward trend indicates that empirical research on AI in students' mathematics learning remains in an early developmental stage but is gaining momentum as AI, particularly generative AI, becomes more accessible and increasingly embedded within educational practice and policy (Nedungadi et al., 2024; Walkington, 2025).

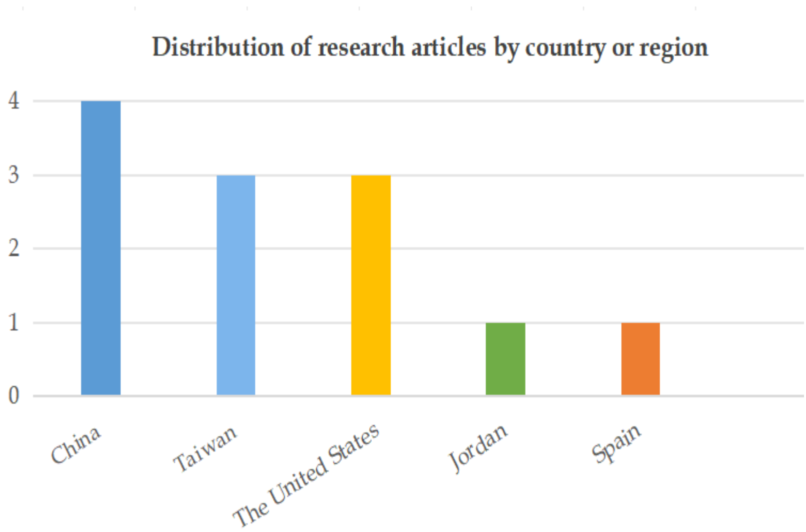

Figure 3: Distribution of research articles by country or region

Figure 3 illustrates the geographical distribution of the reviewed studies. As can be seen, China produced the highest number of publications (n = 4), followed by Taiwan and the United States, each contributing three studies (n = 3). Jordan and Spain each published one study (n = 1). China's leading output may be attributed to sustained governmental investment in digital transformation and national policy initiatives, such as the Smart Education strategy (Ding & Wu, 2024; Huang et al., 2024b).

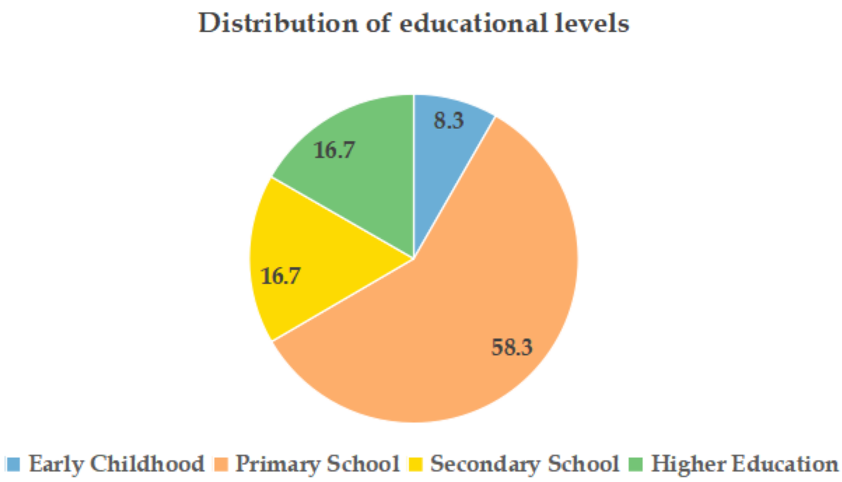

Figure 4: Distribution of educational levels

The reviewed studies spanned multiple educational levels, with primary education receiving the greatest attention (58.3%, n = 7), followed by secondary education (16.7%, n = 2) and higher education (16.7%, n = 2). As depicted in Figure 4, only one study focused on early childhood education (8.3%, n = 1). The findings indicate the adaptability of AI tools in catering to learners' diverse developmental needs. In K–12 education, AI was primarily used to strengthen foundational skills, whereas in higher education, it supported more advanced cognitive processes. The dominance of primary education studies may reflect the structured nature of early mathematics curricula, which aligns well with AI–driven learning models and enables more feasible classroom implementation.

Table 4: Types of quantitative research designs

Type of quantitative studies	**Number**	**References**
Quasi-experiment	10	Pai et al. (2021); del Olmo-Muñoz et al. (2022); Thai et al. (2022); Shih et al. (2023); Huang et al. (2024a); Khazanchi et al. (2025); Alkhasawneh (2025); Liu et al. (2025); Lin et al. (2025); Li and Lyu (2025)
Questionnaire	2	Lin and Jiang (2025); Qiu and Ishak (2025)

Regarding the quantitative methodologies employed in the reviewed studies, Table 4 demonstrates that quasi-experimental designs dominated the current evidence base. Specifically, ten studies adopted quasi-experimental designs, whereas only two relied on questionnaire-based designs (Lin & Jiang, 2025; Qiu & Ishak, 2025). This methodological predominance suggests that quasi-experimental research is widely viewed as offering a stronger basis for causal inference and greater empirical validity compared to other non-experimental quantitative methods (Kim & Steiner, 2016; Steiner & Kim, 2024).

Table 5: Classification of studies by sample size

Sample size	Number	Percentage (%)
<100	4	33.3
100-300	4	33.3
300-500	3	25.0
≥500	1	8.3

Table 5 summarizes the sample sizes of the reviewed studies. Four studies involved fewer than 100 participants and another four recruited between 100–300 students, indicating that 66.7% of the studies relied on sample sizes below 300. Three studies involved 300–500 participants, while only one study employed a large-scale sample with 4,442 students. This distribution demonstrates that research on AI-assisted mathematics learning remains largely small-scale; this is consistent with the findings of Garzón et al. (2025), who highlighted the scarcity of large-sample empirical evidence in AI education research.

The predominance of small sample sizes in the reviewed studies may be attributed to the high implementation costs and technical demands associated with AI-based instructional interventions, as well as the exploratory stage of AI application in mathematics learning, whereby research remains primarily limited to pilot or evaluation studies. Consequently, this pattern may constrain the external validity of the findings.

Table 6: Duration of study interventions

Subcategory	Number	Percentage (%)
Less than 1 week	3	25.0
Between 1 and 4 weeks	2	16.7
Between 4 and 12 weeks	4	33.3
More than 12 weeks	2	16.7
Not specified	1	8.3

Table 6 summarizes the duration of the study interventions. Three studies (25.0%) lasted less than one week, two (16.7%) spanned one to four weeks, and four (33.3%) lasted four to twelve weeks. Only two studies (16.7%) extended beyond twelve weeks, and one study (8.3%) did not specify its duration. Overall, 75% of the interventions lasted no longer than twelve weeks, indicating that AI-assisted mathematics learning research is predominantly based on short-term studies with limited longitudinal validation. Thus, most of the reviewed studies employed short-term interventions, which may be attributable to the fact that extended implementations are constrained by curricular schedules, methodological control requirements, and practical limitations related to time, resources, and feasibility.

3.2 Types of AI technologies implemented in mathematics learning

Table 7: Classification of AI technologies implemented in mathematics learning

Category	Type of AI Tools	AI Tools	References
Adaptive Learning Systems	Educational AI	My Math Academy	Thai et al. (2022)
		Interactive e-book editing platform	Huang et al. (2024a)
		Nerd AI	Alkhasawneh (2025)
		Edmentum Exact Path	Khazanchi et al. (2025)
		AI-assisted Adaptive Learning System	Qiu and Ishak (2025)
		TALP	Lin et al. (2025)
Intelligent Tutoring Systems	Educational AI	Mathematical ITS	Pai et al. (2021)
		HINTS	del Olmo-Muñoz et al. (2022)
		Math ITS	Shih et al. (2023)
Chatbots	Generative AI	GAI	Lin and Jiang (2025)
		ChatGPT-MPS	Liu et al. (2025)
		ConvAI	Li and Lyu. (2025)

Table 7 summarizes the AI technologies integrated into mathematics learning, which are classified into three categories: adaptive learning systems (n = 6), intelligent tutoring systems (n = 3), and chatbots (n = 3). To clarify their pedagogical purposes, these tools were further grouped by functionality. Most of these tools (n = 9) were categorized as Educational AI, and three studies employed Generative AI.

Adaptive learning systems emerged as the most widely adopted technology. Through AI-driven diagnostics, personalized learning pathways, and real-time feedback, they provide differentiated support that accommodates diverse learner needs. Furthermore, empirical evidence indicates that individualized adaptive learning systems can produce significantly greater gains in mathematics performance compared with traditional instruction (Son, 2024; Wang et al., 2023).

By contrast, intelligent tutoring systems and chatbots appeared less frequently but highlighted two complementary trajectories in AI-supported mathematics learning (Panqueban & Huincahue, 2024). Intelligent tutoring systems integrated detailed domain models and instructional scaffolding to deliver targeted hints, diagnose errors, and provide formative feedback on specific mathematical topics (del Olmo-Muñoz et al., 2022; Pai et al., 2021; Shih et al., 2023). Additionally, meta-analytic findings showed that intelligent tutoring systems produce small yet positive effects on K–12 students' mathematics learning (Steenbergen-Hu & Cooper, 2013).

Chatbots were found to facilitate conversational interaction, enabling students to ask questions in natural language, access alternative explanations, and articulate their reasoning. Recent studies in mathematics education have reported improvements in engagement, personalized support, and problem-solving when AI-driven conversational agents are integrated into instruction (Almarashdi et al., 2024; Pala et al., 2025). Compared with adaptive learning systems and intelligent tutoring systems, chatbots have demonstrated stronger capability in supporting complex mathematical problem-solving (Chau et al., 2025).

Within the reviewed literature, chatbot tools were used mainly in higher-education studies, suggesting that learners in these contexts may benefit more from AI systems that support complex, dialogue-based interactions. In contrast, adaptive learning systems were predominantly implemented in early childhood, primary, and secondary settings, and all intelligent tutoring system applications focused on primary students. This pattern suggests that the structured guidance and adaptive instructional support provided by these systems align well with the developmental and learning needs of younger learners.

3.3 The main challenges encountered in integrating AI into students' mathematics learning

Table 8: Challenges in integrating AI into students' mathematics learning

Category	Refined Challenges	References
Research-related	Limited empirical evidence and restricted research contexts	Alkhasawneh (2025); Khazanchi et al. (2025); Lin and Jiang. (2025)
	Methodological limitations	Lin et al. (2025)
	Fragmented research focus	del Olmo-Muñoz et al. (2022); Lin and Jiang. (2025); Lin et al. (2025)
Technical and design-related	Insufficient adaptivity, personalization, and cognitive responsiveness	Shih et al. (2023); Qiu and Ishak (2025)
	Early-stage development of AI-supported learning systems	Pai et al. (2021); Qiu and Ishak (2025)
	Inadequate support for complex mathematical problem solving	Liu et al. (2025); Shih et al. (2023)
Educational-pedagogical	Insufficient teacher and instructional support	Huang et al. (2024a); Qiu and Ishak (2025)
	Low student engagement and motivation	Khazanchi et al. (2025); Qiu and Ishak (2025)
	Lack of engaging, evidence-based learning resources	Thai et al. (2022)
	Unequal access and usability barriers	Qiu and Ishak (2025); Shih et al. (2023)

Outlined in Table 8 are the three major categories of challenges that were found to constrain the integration of AI in students' mathematics learning: research-related, technical and design-related, and educational-pedagogical.

The first category focused on research-related challenges. Empirical evidence on AI-supported mathematics learning remains limited, as many studies to date have relied on small samples, short intervention periods, or single-site designs, thereby restricting the generalizability of their findings (Alkhasawneh, 2025; Khazanchi et al., 2025; Lin & Jiang, 2025). Methodological issues such as non-randomized research designs and heavy reliance on self-reported instruments further weaken the robustness of the currently available evidence (Lin et al., 2025). Furthermore, research topics also appear fragmented, with core mechanisms including learning flexibility, motivation, and self-regulated learning often remaining underexamined (del Olmo-Muñoz et al., 2022; Lin et al., 2025; Lin & Jiang, 2025).

The second category concerned technical and design-related challenges. The literature indicates that the effectiveness of AI tools remains constrained by limited adaptivity, insufficient personalization, and weak cognitive responsiveness (Qiu & Ishak, 2025; Shih et al., 2023). Although AI-driven systems are increasingly used in mathematics learning, most remain in an early stage of development (Pai et al., 2021; Qiu & Ishak, 2025). Liu et al. (2025) and Shih et al. (2023) further observed that existing AI technologies struggle to support complex, cognitively demanding tasks such as mathematical word-problem solving, which requires deeper conceptual understanding and analytical reasoning.

The third category related to educational-pedagogical challenges, which also influence the effectiveness of AI in mathematics learning. The reviewed studies highlight insufficient teacher preparation and limited instructional support as major obstacles to meaningful AI integration (Huang et al., 2024a; Qiu & Ishak, 2025). Learner-related factors such as low engagement and weak motivation further limited the effectiveness of AI-assisted learning activities (Khazanchi et al., 2025; Qiu & Ishak, 2025). In addition, Thai et al. (2022) observed that many AI-supported platforms still lack engaging, evidence-based learning resources, thereby constraining opportunities for deeper mathematical learning. Persistent access and usability barriers further create and perpetuate inequities that restrict the pedagogical benefits of AI tools (Qiu & Ishak, 2025; Shih et al., 2023).

In summary, these challenges reflect deeper structural obstacles to the integration of AI into mathematics learning. First, a fundamental misalignment was identified between the cognitive demands of mathematics and the current capabilities of AI technologies. Mathematics requires abstraction, logical reasoning, and multi-step problem-solving (Ni et al., 2018), yet many AI systems still lack the pedagogical sensitivity and adaptive depth needed to support such complex cognitive processes. Second, many educators lack the specialized competencies and pedagogical frameworks required to effectively integrate AI into instruction, making it difficult to align AI use with mathematics-specific learning goals (Ng et al., 2024; Zou et al., 2025). Third, policy-level limitations, such as insufficient strategic planning, uneven infrastructural development, and inequitable resource distribution, create additional barriers to AI integration, particularly for under-resourced institutions (Bawa & Bawa, 2025; Waqar et al., 2024).

3.4 Solutions proposed or implemented for integrating AI into students' mathematics learning

Table 9: Solutions for integrating AI into students' mathematics learning

Category of Identified Challenges	Solutions Refined	References
Research-related	Rigorous longitudinal validation	Thai et al. (2022); Khazanchi et al. (2025)
	Motivational and mediating mechanisms	Lin and Jiang (2025); Khazanchi et al. (2025); Lin et al.(2025)
	Cross-context expansion	Pai et al. (2021); Thai et al. (2022); Lin et al.(2025); del Olmo-Muñoz et al. (2022) Liu et al. (2025);
Technical and design-related	Adaptivity, personalization and intelligent feedback	Alkhasawneh (2025); Liu et al. (2025); Shih et al. (2023); del Olmo-Muñoz et al. (2022)
	Development of dialogue-based AI	Pai et al. (2021)
	Cross-platform accessibility	Liu et al. (2025); Alkhasawneh (2025)
Educational and pedagogical	Strengthened teacher participation and diagnostic instruction	Shih et al. (2023)
	Support for self-efficacy, and social interaction	Qiu and Ishak (2025)
	Pedagogical strategies integration	Shih et al. (2023); Li and Lyu (2025); Alkhasawneh (2025);
	Data-driven feedback	Huang et al. (2024a)

To address these challenges, the reviewed studies converged on a set of improvement strategies spanning research-related, technical and design-related, and educational-pedagogical, as summarized in Table 9.

At the research-related level, the reviewed studies call for rigorous longitudinal validation to assess the sustained effects of AI interventions (Khazanchi et al., 2025; Thai et al., 2022). A further recommendation is the incorporation of motivational and mediating mechanisms to deepen understanding of the way in which AI shapes mathematics learning processes and outcomes (Khazanchi et al., 2025; Lin et al., 2025; Lin & Jiang, 2025). Lastly, the studies call for the expansion of research across diverse educational levels, cultural contexts, and institutional settings in order to enhance the generalizability of findings (del Olmo-Muñoz et al., 2022; Lin et al., 2025; Liu et al., 2025; Pai et al., 2021; Thai et al., 2022).

At the technical and design level, the reviewed studies underscore the need for enhancing system adaptivity, personalization, and intelligent feedback to provide more responsive and individualized learning support (Alkhasawneh, 2025; del Olmo-Muñoz et al., 2022; Liu et al., 2025; Shih et al., 2023). In addition, Pai et al. (2021) highlight the potential for developing dialogue-based AI that engages learners in interactive problem-solving by enabling them to articulate, test, and refine their mathematical reasoning through conversational exchanges with computer agents. Furthermore, improving mobile and cross-platform accessibility will help to reduce digital inequities and expand opportunities for students to engage with AI-supported mathematics learning across varied contexts (Alkhasawneh, 2025; Liu et al., 2025).

At the educational and pedagogical level, the reviewed studies emphasize that AI should complement rather than replace teachers, identifying that strengthened teacher participation and diagnostic instruction enhance learning personalization and ensure appropriate guidance during AI use (Shih et al., 2023). In addition, Qiu and Ishak (2025) note that AI tools should be designed to strengthen students' self-efficacy and social support, as the combined influence of these factors is critical for fostering deeper mathematical learning.

Several studies further highlighted the importance of embedding evidence-based pedagogical strategies into AI systems to ensure a strong alignment between learners' cognitive needs and the instructional goals in mathematics learning contexts (Alkhasawneh, 2025; Li & Lyu, 2025; Shih et al., 2023). Complementing these efforts, data-driven feedback has been found to strengthen students' self-regulated learning and enhance academic achievement in mathematical word problem solving and cognitive strategies (Huang et al., 2024a).

Although the reviewed studies provide a valuable foundation by demonstrating the effectiveness of AI tools in improving student engagement and mathematics learning performance, the evidence with regard to long-term learning outcomes and scalable classroom integration remains limited. Therefore, further research is urgently needed to support the more sustainable integration of AI into mathematics learning. In particular, greater alignment with STEM education frameworks is essential to enhance students' higher-order thinking skills.

AI should be leveraged to support inquiry-based learning and authentic problem-solving, thereby aligning AI-enhanced mathematics learning with broader sustainable educational development goals. Furthermore, it is essential to expand AI-related research in higher education, where the advanced cognitive demands of university-level mathematics can guide the design of more sophisticated AI tools and pedagogical models. Moreover, enhancing AI literacy among both teachers and students is critical, as robust competencies across stakeholders are necessary for effective AI integration into pedagogical practices in real classrooms in order to ensure improved learning outcomes.

4. Conclusion

This review systematically examined the integration of AI into students' mathematics learning using the PRISMA framework, drawing on 12 quantitative studies published between 2021 and 2025. The findings indicate that this field remains in an early stage of scholarly inquiry, underscoring the need for more extensive and rigorous research. To date, the existing evidence base remains fragmented, characterized by small-scale exploratory designs, pedagogically generic AI tools, and the limited incorporation of discipline-specific instructional approaches.

By synthesizing the findings across the four key dimensions of research characteristics, AI technologies, challenges, and solutions, the review provides a clear and cohesive overview of the current developments as well as persisting gaps. Not only do these findings inform the strategic use of AI to enhance mathematics learning but they also extend the application of learning theories in AI-supported contexts and provide practical implications for policymakers in guiding future research, instructional innovation, and policy development.

Despite these important contributions, several limitations of the current study should be acknowledged. First, the limited number of empirical studies constrains the strength of the evidence base; future research should therefore include large-scale investigations. Second, limiting the review to the WoS and Scopus databases may have reduced coverage; thus, subsequent work should include additional databases to enhance comprehensiveness. Third, focusing solely on quantitative studies limits explanatory depth; therefore, further research should incorporate qualitative or mixed methods approaches to enrich interpretation.

In addition, the inclusion of only English-language articles may lead to language bias. To mitigate this, future work should expand multilingual coverage to improve global representation. Finally, most of the included studies were conducted in a small number of countries and relied on short-term interventions, limiting the generalizability of the findings and the understanding of long-term learning effects. Consequently, future work should adopt cross-national and longitudinal designs.

5. Acknowledgements

The authors would like to acknowledge using ChatGPT during the writing process only for language polishing and grammatical revision. The tool did not contribute to the conceptual development or the main content of the article. The manuscript represents the authors' original work, and the authors take full responsibility for the accuracy and originality of the content.

6. References

Aditya, R. Q., & Suranto, S. (2024). The role of educational transformation in the digital era in improving student quality. *Al Qalam: Jurnal Ilmiah Keagamaan Dan Kemasyarakatan*, *18(3)*, 1756-1776. https://doi.org/10.35931/aq.v18i3.3301

Adiyono, A., Suwartono, T., Nurhayati, S., Dalimarta, F. F., & Wijayanti, O. (2025). Impact of artificial intelligence on student reliance for exam answers: A case study in

IRCT Indonesia. *International Journal of Learning, Teaching and Educational Research*, *24*(3), 455–479. https://doi.org/10.26803/ijlter.24.3.22

Alam, T. M., Stoica, G. A., Sharma, K., & Özgöbek, Ö. (2025). Digital technologies in the classrooms in the last decade (2014–2023): A bibliometric analysis. *Frontiers in Education*, *10*, 1533588. https://doi.org/10.3389/feduc.2025.1533588

Alkhasawneh, S. (2025). AI-driven personalized mathematics learning through interactive mobile platforms: Effects on achievement and motivation. *International Journal of Interactive Mobile Technologies*, *19*(13), 33–54. https://doi.org/10.3991/ijim.v19i13.54947

Almarashdi, H. S., Jarrah, A. M., Abu Khurma, O., & Gningue, S. M. (2024). Unveiling the potential: A systematic review of ChatGPT in transforming mathematics teaching and learning. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(12), em2555. https://doi.org/10.29333/ejmste/15739

Alvarez, J. I. (2024). Evaluating the impact of AI–powered tutors MathGPT and Flexi 2.0 in enhancing calculus learning. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi, 8*(2), 495–508. https://doi.org/10.22437/jiituj.v8i2.34809

Awang, L. A., Yusop, F. D., & Danaee, M. (2025). Current practices and future direction of artificial intelligence in mathematics education: A systematic review. *International Electronic Journal of Mathematics Education*, *20*(2), em0823. https://doi.org/10.29333/iejme/16006

Bawa, I., & Bawa, S. (2025). Bridging EdTech gaps: Examining learning equity in low-income educational settings. *International Journal of Educational Development*, *118*, 103398. https://doi.org/10.1016/j.ijedudev.2025.103398

Chau, D. B., Luong, V. T., Long, T. T., & Linh, N. T. T. (2025). Personalized mathematics teaching with the support of AI chatbots to improve mathematical problem-solving competence for high school students in Vietnam. *European Journal of Educational Research*, *14*(1), 323–333. https://doi.org/10.12973/eu-jer.14.1.323

Daher, R. (2025). Integrating AI literacy into teacher education: A critical perspective paper. *Discover Artificial Intelligence*, *5*(1), 217. https://doi.org/10.1007/s44163-025-00475-7

del Olmo-Muñoz, J., González-Calero, J. A., Diago, P. D., Arnau, D., & Arevalillo-Herráez, M. (2022). Using intra-task flexibility on an intelligent tutoring system to promote arithmetic problem-solving proficiency. *British Journal of Educational Technology*, *53*(6), 1976–1992. https://doi.org/10.1111/bjet.13228

Ding, L., & Wu, S. (2024). Digital transformation of education in China: A review against the backdrop of the 2024 world digital education conference. *Science Insights Education Frontiers*, *20*(2), 3283–3299. https://doi.org/10.15354/sief.24.re340

Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A., & Nicoletti, G. (2024). *The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges*. (OECD Artificial Intelligence Papers No. 15). OECD Publishing. https://doi.org/10.1787/8d900037-en

Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, *9*(8), 84. https://doi.org/10.3390/mti9080084

Gasparyan, A. Y., Ayvazyan, L., & Kitas, G. D. (2013). Multidisciplinary bibliographic databases. *Journal of Korean Medical Science*, *28*(9), 1270. https://doi.org/10.3346/jkms.2013.28.9.1270

Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (2024). Privacy and security concerns in generative AI: A comprehensive survey. *IEEE Access*, *12*, 48126–48144. https://doi.org/10.1109/ACCESS.2024.3381611

Goos, M., Carreira, S., & Namukasa, I. K. (2023). Mathematics and interdisciplinary STEM education: Recent developments and future directions. *ZDM – Mathematics Education*, *55*(7), 1199–1217. https://doi.org/10.1007/s11858-023-01533-z

Granados-Duque, V., & García-Perdomo, H. A. (2021). Systematic review and meta-analysis: Which pitfalls to avoid during this process. *International Braz J Urol*, *47*(5), 1037–1041. https://doi.org/10.1590/s1677-5538.ibju.2020.0746

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, *4*(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001

Gulz, A., Londos, L., & Haake, M. (2020). Preschoolers' understanding of a teachable agent-based game in early mathematics as reflected in their gaze behaviors – an experimental study. *International Journal of Artificial Intelligence in Education*, *30*(1), 38–73. https://doi.org/10.1007/s40593-020-00193-4

Huang, J., Cai, Y., Lv, Z., Huang, Y., & Zheng, X.-L. (2024a). Toward self-regulated learning: Effects of different types of data-driven feedback on pupils' mathematics word problem-solving performance. *Frontiers in Psychology*, *15*, 1356852. https://doi.org/10.3389/fpsyg.2024.1356852

Huang, J., Saleh, S., & Liu, Y. (2021). A review on Artificial intelligence in Education. *Academic Journal of Interdisciplinary Studies*, *10*(3), 206. https://doi.org/10.36941/ajis-2021-0077

Huang, R., Liu, D., Kanwar, A. S., Zhan, T., Yang, J., Zhuang, R., Liu, M., Li, Z., & Adarkwah, M. A. (2024b). Global understanding of smart education in the context of digital transformation. *Open Praxis*, 16(4), 663–676. https://doi.org/10.55982/openpraxis.16.4.761

Just, J., & Siller, H.-S. (2022). The role of mathematics in STEM secondary classrooms: A systematic literature review. *Education Sciences*, *12*(9), 629–647. https://doi.org/10.3390/educsci12090629

Khazanchi, R., Di Mitri, D., & Drachsler, H. (2025). The effect of AI-based systems on mathematics achievement in rural context: A quantitative study. *Journal of Computer Assisted Learning*, *41*(1), e13098. https://doi.org/10.1111/jcal.13098

Kim, Y., & Steiner, P. (2016). Quasi-experimental designs for causal inference. *Educational Psychologist*, *51*(3–4), 395–405. https://doi.org/10.1080/00461520.2016.1207177

Li, C., & Lyu, B. (2025). Investigating the motivational and knowledge affordances of conversational AI using induction, concretization and exemplification in math learning. *British Journal of Educational Technology*, *56*(5), 1814–1841. https://doi.org/10.1111/bjet.13612

Lin, C.-H., Kuo, B.-C., & Chang, F. T. Y. (2025). The impact of Taiwan Adaptive Learning Platform (TALP) on self-regulated learning and mathematics achievement. *Educational Psychology*, 1–22. https://doi.org/10.1080/01443410.2025.2561028

Lin, W., & Jiang, P. (2025). Factors influencing college students' generative artificial intelligence usage behavior in mathematics learning: A case from China. *Behavioral Sciences*, *15*(3), 295–322. https://doi.org/10.3390/bs15030295

Liu, J., Sun, D., Sun, J., Wang, J., & Yu, P. L. H. (2025). Designing a generative AI enabled learning environment for mathematics word problem solving in primary schools: Learning performance, attitudes and interaction. *Computers and Education: Artificial Intelligence*, *9*, 100438. https://doi.org/10.1016/j.caeai.2025.100438

Luo, J., Zheng, C., Yin, J., & Teo, H. H. (2025). Design and assessment of AI-based learning tools in higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, *22*(1), 42. https://doi.org/10.1186/s41239-025-00540-2

Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. *Trends in Higher Education*, *4*(2), 17. https://doi.org/10.3390/higheredu4020017

Mills, N. J. (2021). ALEKS constructs as predictors of high school mathematics achievement for struggling students. *Heliyon*, *7*(6), e07345. https://doi.org/10.1016/j.heliyon.2021.e07345

Mohamed, M. Z. B., Hidayat, R., Suhaizi, N. N. B., Sabri, N. B. M., Mahmud, M. K. H. B., & Baharuddin, S. N. B. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education*, *17*(3), em0694. https://doi.org/10.29333/iejme/12132

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Munárriz, A., & Rincón, Y. R. (2025). Mind the gap: Linking quantitative skill deficits and academic success in economics and business degrees — a systematic literature review. *The International Journal of Management Education*, *23*(3), 101277. https://doi.org/10.1016/j.ijme.2025.101277

Nedungadi, P., Tang, K.-Y., & Raman, R. (2024). The transformative power of generative artificial intelligence for achieving the sustainable development goal of quality education. *Sustainability*, *16*(22), 9779. https://doi.org/10.3390/su16229779

Ng, D. T. K., Su, J., Leung, J. K. L., & Chu, S. K. W. (2024). Artificial intelligence (AI) literacy education in secondary schools: A review. *Interactive Learning Environments*, *32*(10), 6204–6224. https://doi.org/10.1080/10494820.2023.2255228

Nguyen, D. T., & Pham, Q. V. (2025). The evolving landscape of AI integration in mathematics education: A systematic review of trends (2015-2025). *Eurasia Journal of Mathematics, Science and Technology Education*, 21(10), em2714. https://doi.org/10.29333/ejmste/17078

Ni, Y., Zhou, D.-H. R., Cai, J., Li, X., Li, Q., & Sun, I. X. (2018). Improving cognitive and affective learning outcomes of students through mathematics instructional tasks of high cognitive demand. *The Journal of Educational Research*, *111*(6), 704–719. https://doi.org/10.1080/00220671.2017.1402748

Opesemowo, O. A. G., & Adewuyi, H. O. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(7), em2478. https://doi.org/10.29333/ejmste/14762

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Pai, K.-C., Kuo, B.-C., Liao, C.-H., & Liu, Y.-M. (2021). An application of Chinese dialogue-based intelligent tutoring system in remedial instruction for mathematics learning. *Educational Psychology*, *41*(2), 137–152. https://doi.org/10.1080/01443410.2020.1731427

Pala, C. A., Lagrimas, D., & Nobis Jr., M. L. (2025). Exploration of chatbots in mathematics education for innovative learning process. *Asian Journal of Advanced Research and Reports*, *19*(5), 178–194. https://doi.org/10.9734/ajarr/2025/v19i51010

Panqueban, D., & Huincahue, J. (2024). Inteligencia artificial en educación matemática: Una revisión sistemática. *Uniciencia*, *38*(1), 1–17. https://doi.org/10.15359/ru.38-1.20

Qiu, Y., & Ishak, N. A. (2025). AI-assisting technology and social support in enhancing deep learning and self-efficacy among primary school students in mathematics in China. *International Journal of Learning, Teaching and Educational Research*, *24*(2), 21–37. https://doi.org/10.26803/ijlter.24.2.2

Rahmatika, W., Mursalin, M., Saputra, E., & Fonna, M. (2025). The influence of the use of AI in mathematics learning at the high school level: A systematic literature review. *Electronic Journal of Education, Social Economics and Technology*, *6*(2), 1068. https://doi.org/10.33122/ejeset.v6i2.1068

Sajja, R., Sermet, Y., & Demir, I. (2025). End-to-end deployment of the educational AI hub for personalized learning and engagement: A case study on environmental science education. *IEEE Access*, *13*, 55169–55186. https://doi.org/10.1109/ACCESS.2025.3554222

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & the PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647. https://doi.org/10.1136/bmj.g7647

Shih, S.-C., Chang, C.-C., Kuo, B.-C., & Huang, Y.-H. (2023). Mathematics intelligent tutoring system for learning multiplication and division of fractions based on diagnostic teaching. *Education and Information Technologies*, *28*(7), 9189–9210. https://doi.org/10.1007/s10639-022-11553-z

Son, T. (2024). Intelligent tutoring systems in mathematics education: A systematic literature review using the substitution, augmentation, modification, redefinition model. *Computers*, *13*(10), 270. https://doi.org/10.3390/computers13100270

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, *105*(4), 970–987. https://doi.org/10.1037/a0032447

Steiner, P. M., & Kim, Y. (2024). Causal research designs and analysis in education. *Asia Pacific Education Review*, *25*(3), 555–556. https://doi.org/10.1007/s12564-024-09988-9

Thai, K.-P., Bang, H. J., & Li, L. (2022). Accelerating early math learning with research-based personalized learning games: A cluster randomized controlled trial. *Journal of Research on Educational Effectiveness*, *15*(1), 28–51. https://doi.org/10.1080/19345747.2021.1969710

Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino, S. V., Giannoutsou, N., Cachia, R., Monés, A. M., & Ioannou, A. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. *Education and Information Technologies*, *28*(6), 6695–6726. https://doi.org/10.1007/s10639-022-11431-8

Vistorte, A. O. R., Deroncele-Acosta, A., Ayala, J. L. M., Barrasa, A., López-Granero, C., & Martí-González, M. (2024). Integrating artificial intelligence to assess emotions in learning environments: A systematic literature review. *Frontiers in Psychology*, *15*, 1387089. https://doi.org/10.3389/fpsyg.2024.1387089

Walkington, C. (2025). The implications of generative artificial intelligence for mathematics education. *School Science and Mathematics*, ssm.18356. https://doi.org/10.1111/ssm.18356

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*, *12*(1), 621. https://doi.org/10.1057/s41599-025-04787-y

Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L., & Feng, M. (2023). When adaptive learning is effective learning: Comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, *31*(2), 793–803. https://doi.org/10.1080/10494820.2020.1808794

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Waqar, Y., Rashid, S., Anis, F., & Muhammad, Y. (2024). Digital divide & inclusive education: Examining how unequal access to technology affects educational inclusivity in urban versus rural Pakistan. *Journal of Social & Organizational Matters*, *3*(3), 1–13. https://doi.org/10.56976/jsom.v3i3.97

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, *9*(1), 59. https://doi.org/10.1186/s40594-022-00377-5

Zhao, L., Wu, X., & Luo, H. (2022). Developing AI literacy for primary and middle school teachers in China: Based on a structural equation modeling analysis. *Sustainability*, *14*(21), 14549. https://doi.org/10.3390/su142114549

Zou, Y., Kuek, F., Feng, W., & Cheng, X. (2025). Digital learning in the 21st century: Trends, challenges, and innovations in technology integration. *Frontiers in Education*, *10*, 1562391. https://doi.org/10.3389/feduc.2025.1562391