Review

Not peer-reviewed version

Artificial Intelligence, Sustainability, and the Development of Mathematical Thinking: A Theory-Grounded Scoping Review

Georgios Polydoros * , Ilias Vasileiou , Zoe Krokou , Alexandros-Stamatios Antoniou

Posted Date: 31 March 2026

doi: 10.20944/preprints202603.2391.v1

Keywords: artificial intelligence; mathematics education; mathematical thinking; intelligent tutoring systems; large language models; APOS

doi:10.20944/preprints202603.2391.v1

Review

Artificial Intelligence, Sustainability, and the Development of Mathematical Thinking: A Theory-Grounded Scoping Review

Georgios Polydoros [1], Ilias Vasileiou [2], Zoe Krokou [3] and Alexandros-Stamatios Antoniou [2]

[1] University of Crete
[2] National and Kapodistrian University of Athens
[3] Hellenic Open University
* Correspondence: georgiospolydoros@uoc.gr

Abstract

Artificial intelligence (AI) tools are increasingly integrated into mathematics education, yet most reviews emphasize achievement rather than how AI shapes mathematical thinking. This scoping review mapped literature published between 2020 and 2026 on AI-supported mathematics learning through three cognition frameworks: APOS (Action–Process–Object–Schema), Sfard's process–object duality and reification, and Conceptual Image theory. Searches were conducted in Scopus, Web of Science, ERIC, PsycINFO, Education Source, and IEEE Xplore, followed by duplicate removal and PRISMA-ScR–aligned screening. Twenty-one peer-reviewed studies met inclusion criteria (18 empirical studies plus three theory-informed anchors). Evidence growth accelerated after 2022, with most studies situated in secondary and higher education. Large language models (LLMs) and intelligent tutoring systems (ITS) were the most frequently investigated modalities. Across studies, AI commonly supported action-level execution and procedural management (APOS) via adaptive feedback, hinting, and stepwise scaffolding, and it often broadened learners' conceptual images through multiple representations and generated explanations. However, few studies directly examined theory-linked conceptual mechanisms, such as object encapsulation, reification, or alignment between conceptual images and formal definitions. In LLM-supported contexts, gains in explanation quality coexisted with risks of procedural outsourcing when students relied on generated solutions without prior reasoning. Overall, AI's conceptual impact appears to depend less on tool availability and more on instructional orchestration (task design, prompting, and teacher mediation). Future research should operationalize cognitive transitions, assess structural understanding, and report AI-use conditions transparently to support cumulative, theory-driven synthesis.

Keywords: artificial intelligence; mathematics education; mathematical thinking; intelligent tutoring systems; large language models; APOS

---

1. Introduction

Artificial intelligence (AI)–supported instructional environments are increasingly integrated into mathematics education as tools designed to enhance learning outcomes, personalise instruction, and optimise feedback processes (Holmes & Tuomi, 2022; Zawacki-Richter et al., 2019). These environments include intelligent tutoring systems (ITS), adaptive learning platforms, learning analytics systems, and, more recently, large language models (LLMs) capable of generating mathematical explanations and solutions (Kasneci et al., 2023; Luckin et al., 2016). AI in mathematics education is often promoted as a means of improving achievement, increasing engagement, and providing real-time scaffolding (Chen et al., 2020; Hwang & Tu, 2021). However, beyond performance indicators, a fundamental question remains insufficiently examined: how does AI influence the development of mathematical thinking? Beyond effectiveness, the integration of AI in mathematics education must also be examined through a sustainability lens, with emphasis on durable conceptual understanding, equitable participation, responsible innovation, and long-term pedagogical value. Mathematical thinking is not limited to procedural accuracy or task completion. It involves the construction of conceptual structures, the transition from actions to processes and objects, and the formation of coherent mental representations of mathematical concepts (Dubinsky & McDonald, 2001; Sfard, 1991; Tall & Vinner, 1981). Theoretical frameworks such as APOS theory (Action–Process–Object–Schema), Sfard's process–object duality and reification, and the Conceptual Image/Concept Definition distinction provide structured accounts of how learners develop mathematical understanding (Arnon et al., 2014; Sfard, 1991; Tall, 2013). These frameworks emphasise internal cognitive transformations rather than surface performance. Yet, despite their centrality in mathematics education research, they are rarely used to interpret findings from AI-supported learning environments. Recent reviews of AI in education tend to prioritise technological affordances, system performance, or achievement outcomes rather than conceptual development mechanisms (Zawacki-Richter et al., 2019; Chassignol et al., 2018).

The rapid expansion of AI technologies in education has resulted in a proliferation of empirical studies examining effectiveness, usability, engagement, and achievement gains (Chassignol et al., 2018; Zhai et al., 2021). AI systems are now widely implemented across educational levels—from primary school arithmetic to advanced university-level calculus and algebra (Hwang & Tu, 2021; Kasneci et al., 2023). International research demonstrates increasing AI integration in mathematics classrooms, particularly in technologically advanced regions and higher education settings (Holmes & Tuomi, 2022). However, much of this literature prioritises measurable outcomes such as test scores, time-on-task reduction, or completion rates, rather than examining whether AI facilitates deep conceptual restructuring.

1.1. Characteristics of Effective AI-Supported Mathematical Learning

Effective mathematics learning environments are characterised not merely by efficiency but by their ability to promote conceptual development, structural reasoning, and cognitive flexibility (Kilpatrick et al., 2001; Tall, 2013). Research in mathematics education has consistently demonstrated that robust understanding requires the coordination of multiple representations, the internalization of processes, and the encapsulation of those processes into manipulable cognitive objects (Dubinsky & McDonald, 2001; Sfard, 1991).

Within APOS theory, learners must move from performing actions on mathematical entities to internalizing processes and ultimately encapsulating them into objects that can be integrated into broader schemas (Arnon et al., 2014). Similarly, Sfard's theory emphasises the importance of reification—the transition from operational processes to structural objects—as a critical stage in conceptual development (Sfard, 1991). The Conceptual Image/Concept Definition framework further highlights the potential discrepancy between learners' informal mental representations and formal mathematical definitions (Tall & Vinner, 1981).

AI-supported environments have the potential to scaffold these transitions by providing immediate feedback, adaptive sequencing, and multiple representations (Chen et al., 2020; Hwang & Tu, 2021). However, they may also risk encouraging procedural outsourcing, where learners rely on AI-generated solutions without engaging in the necessary cognitive transformations (Kasneci et al., 2023; Wardat et al., 2023). The effectiveness of AI in mathematics education therefore depends not only on technological sophistication but on whether it supports the development of process–object flexibility and stable conceptual images.

In this sense, AI systems can be viewed as instructional frameworks that mediate between curriculum content, learner cognition, and pedagogical practice. For AI to support deep mathematical thinking, it must align with evidence-based cognitive principles (Holmes & Tuomi, 2022). Without such alignment, improvements in efficiency may coexist with stagnation—or even regression—in conceptual understanding.

1.2. Evidence-Based Practice and Theoretical Alignment

Educational innovation increasingly emphasises evidence-based practice (EBP), where instructional decisions are informed by empirical research (O'Connor & Park, 2023). In mathematics education, EBPs include structured scaffolding, metacognitive prompting, retrieval practice, and explicit representation linking (Hattie, 2009; Sweller et al., 2019). However, translating research findings into effective classroom practice remains a persistent challenge. The research-to-practice gap is widely documented across educational domains (Fixsen et al., 2005).

In the context of AI, this gap may be amplified. Technological adoption frequently precedes theoretical integration, meaning that AI tools are implemented before their cognitive implications are fully understood (Holmes & Tuomi, 2022). Without a guiding theoretical framework, AI integration risks becoming technologically driven rather than pedagogically grounded. Even well-designed AI systems may fail to support meaningful conceptual development if they are not embedded within cognitively informed instructional structures.

Implementation science highlights that successful educational innovations require more than effectiveness evidence; they require alignment between theoretical rationale, contextual adaptation, teacher mediation, and ongoing evaluation (Bauer & Kirchner, 2020). The absence of such alignment may result in surface-level improvements without sustainable conceptual gains.

1.3. Barriers and Facilitators in AI Integration

The implementation of AI-supported mathematics learning is influenced by multiple systemic and individual factors. Common barriers include insufficient teacher training, over-reliance on automated feedback, limited transparency of AI reasoning processes, and misalignment between AI-generated explanations and curricular expectations (Kasneci et al., 2023). Additionally, concerns regarding academic integrity, cognitive offloading, and reduced productive struggle have emerged as significant issues in AI-supported learning contexts (Wardat et al., 2023).

Conversely, facilitators include structured teacher mediation, integration of AI outputs into classroom discussion, requirement of student-generated reasoning prior to AI consultation, and alignment between AI scaffolding and established cognitive frameworks (Hwang & Tu, 2021). Where AI is used as a dialogic partner rather than a solution provider, opportunities for metacognitive engagement and conceptual refinement may increase.

Despite growing empirical evidence, the literature remains fragmented (Zawacki-Richter et al., 2019). Studies vary widely in educational level, AI tool type, methodological design, and outcome measures. Few studies explicitly examine cognitive transitions central to mathematical thinking, and even fewer interpret findings through established theoretical lenses such as APOS, process–object duality, or Conceptual Image theory.

Given these conceptual, methodological, and sustainability-related gaps, a theory-grounded synthesis of the field is necessary.

1.4. Aims and Objectives

The aim of this scoping review was to identify and evaluate the existing international literature (2020–2026) concerning AI-supported mathematics learning in order to develop a theory-informed understanding of how AI influences mathematical thinking.

This aim was achieved through the following objectives:

- Identify theoretical frameworks (explicit or implicit) underpinning AI-supported mathematics studies.
- Examine how AI-supported environments influence cognitive transitions described in APOS theory.
- Investigate evidence of process–object reification within AI-mediated learning contexts.
- Analyze how AI tools affect learners' conceptual images and their alignment with formal concept definitions.

- Identify facilitators and barriers affecting the conceptual impact of AI integration in mathematics education.
- Consider how issues of equity, transparency, responsible use, and durable conceptual development shape the sustainability of AI-supported mathematics learning.

To synthesize the theoretical foundations, empirical findings, and systemic factors identified in the literature, an integrated conceptual framework was developed. As illustrated in Figure 1, the framework connects AI tools, instructional mediation, and core cognitive processes (e.g., APOS transitions, reification, and conceptual image formation) with learning outcomes and sustainability principles. In addition, the figure incorporates a synthesis of empirical findings and highlights the dynamic interaction between barriers and facilitators that shape the conceptual impact of AI-supported mathematics learning. This integrative representation provides a structured lens for interpreting the findings of the present review.

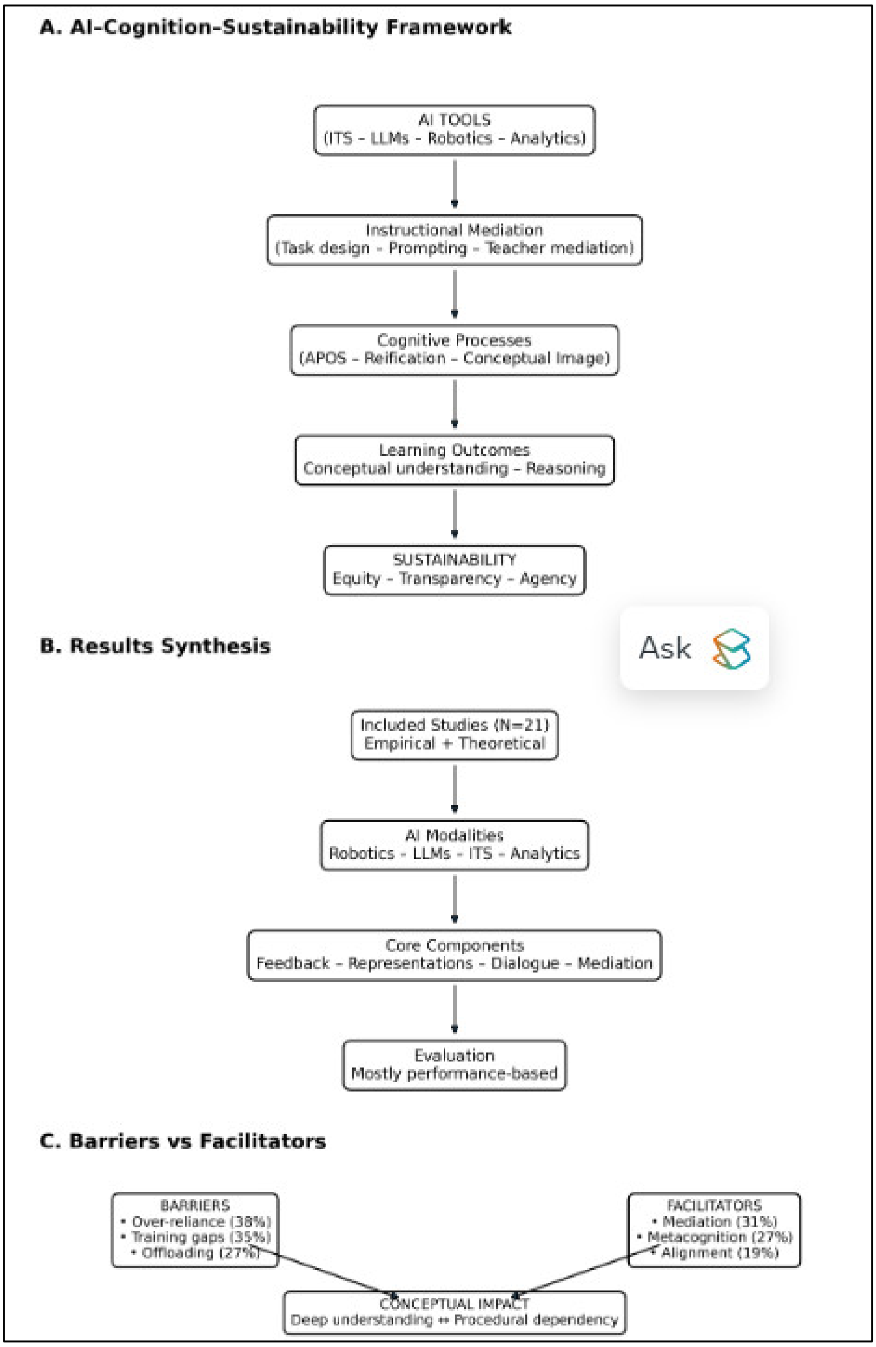

Figure 1. *Integrated framework of AI-supported mathematics learning and synthesis of findings.*

2. Methods

2.1. Design

A scoping review methodology was selected to examine the international literature on artificial intelligence (AI)–supported mathematics learning and its relationship to the development of mathematical thinking. Scoping reviews are particularly appropriate for emerging and rapidly evolving research domains, as they enable systematic mapping of the breadth, range, and characteristics of existing evidence while identifying conceptual, theoretical, and methodological gaps (Grant & Booth, 2009; Peters et al., 2020; Pham et al., 2014).

Given the rapid expansion of AI technologies in educational contexts and the heterogeneity of study designs, educational levels, theoretical orientations, and AI modalities, a scoping review provided an appropriate framework for synthesizing the field without restricting inclusion to narrowly defined intervention designs.

The review followed the methodological framework proposed by Arksey and O'Malley (2005), incorporating refinements by Levac et al. (2010) and the updated guidance from the Joanna Briggs Institute (JBI) (Peters et al., 2020). The framework includes the following stages:

1. Identifying the research question
2. Identifying relevant studies
3. Selecting studies
4. Charting the data
5. Collating, summarizing, and reporting the results
6. (Optional) Stakeholder consultation

Reporting adhered to the PRISMA-ScR guidelines (Tricco et al., 2018), ensuring transparency and methodological rigor. The full screening process is presented in Figure 2.

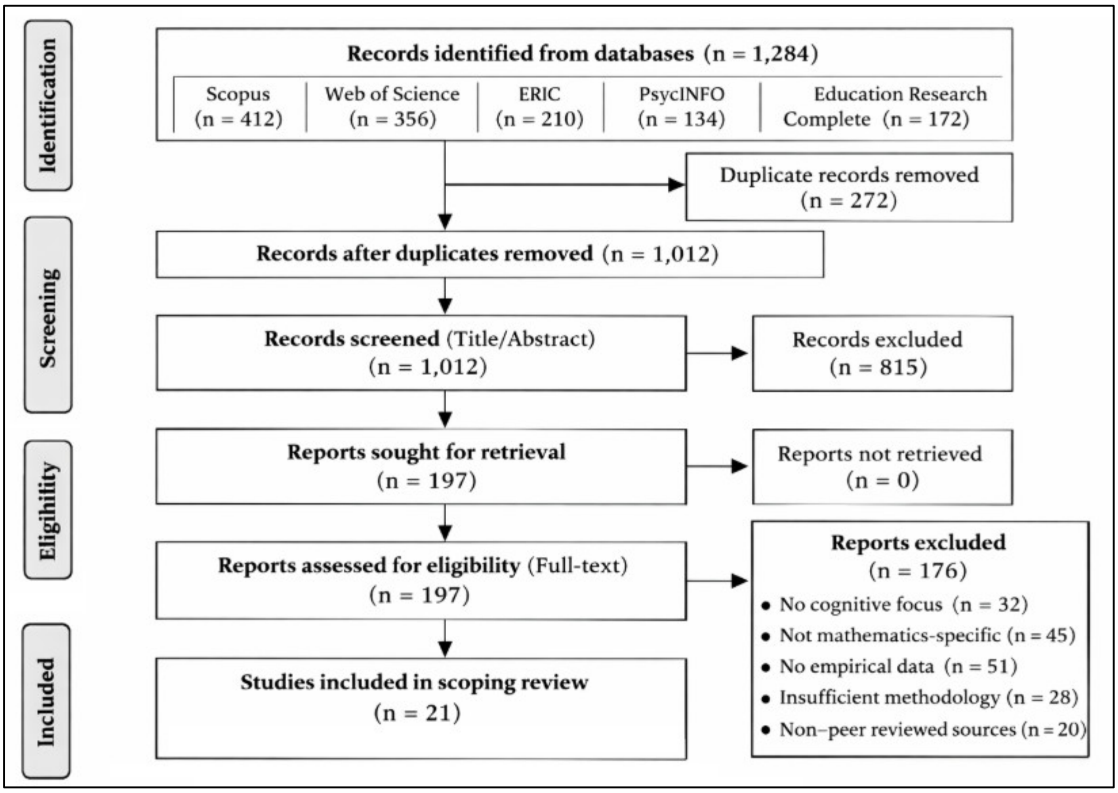

Figure 2. *Study selection and Screening Outcomes.*

Consistent with scoping review methodology, no formal risk-of-bias appraisal was conducted, as the purpose of the review was to map the conceptual and empirical landscape rather than estimate pooled intervention effects.

To enhance methodological rigor and theoretical relevance, the final corpus was deliberately restricted to high-impact peer-reviewed empirical studies (n = 18) published between 2020 and 2026, supplemented by three theoretically orienting review studies (total N = 21).

2.2. Framework Stage 1: Identifying the Research Question

The review was guided by the following overarching question:

What is known from the international literature (2020–2026) regarding the impact of AI-supported mathematics learning environments on the development of mathematical thinking?

To operationalize this question, the review further examined:

- Which theoretical frameworks underpin AI-supported mathematics studies?
- Whether AI environments facilitate cognitive transitions described in APOS theory.
- Whether evidence of process–object reification is present in AI-mediated learning.
- How AI affects learners' conceptual images and their alignment with formal concept definitions.
- What barriers and facilitators influence the conceptual impact of AI integration.

The emphasis on APOS theory, process–object transition, reification, and conceptual image situates AI-supported mathematics learning within established cognitive-developmental frameworks rather than treating technological integration as an isolated instructional variable.

2.3. Framework Stage 2: Identifying Relevant Studies

An exploratory search was first conducted to determine whether previous reviews had examined AI-supported mathematics learning through established mathematics cognition frameworks (e.g., APOS, Sfard's reification framework, Conceptual Image theory). Although several reviews addressed AI in education broadly, few examined conceptual development mechanisms specific to mathematics.

A comprehensive search strategy was developed in accordance with JBI guidance (Peters et al., 2020). Searches were conducted in the following databases:

- Scopus
- Web of Science
- ERIC
- PsycINFO
- Education Research Complete

Search strings included combinations of the following keywords:

(AI OR "artificial intelligence" OR "intelligent tutoring systems" OR ITS OR "adaptive learning" OR "learning analytics" OR "large language models" OR LLM OR ChatGPT)
AND
(mathematics OR "mathematics education" OR algebra OR calculus OR arithmetic)
AND
("mathematical thinking" OR "conceptual understanding" OR cognition OR APOS OR "process-object" OR reification OR "conceptual image")

Wildcard truncations and Boolean operators were adapted to database-specific syntax. Searches were conducted at the title and abstract level to ensure relevance prior to full-text screening. Reference lists of included articles were manually screened to identify additional eligible studies.

2.4. Inclusion and Exclusion Criteria

A priori inclusion and exclusion criteria were established.

2.5. Inclusion Criteria

 doi:10.20944/preprints202603.2391.v1

- Peer-reviewed journal articles published in English
- Published between January 2020 and March 2026
- Empirical studies examining AI-supported mathematics learning
- Conducted in K–12 or higher education contexts
- Reporting measurable learning, cognitive, or conceptually grounded outcomes
- Published in established journals indexed in major databases (e.g., Web of Science or Scopus)

In addition to the empirical corpus, three high-quality review studies were retained as theoretical anchors to contextualize the empirical landscape.

2.6. Exclusion Criteria

- Studies unrelated to mathematics education
- Purely technical AI architecture papers without learner interaction
- Preprints (e.g., arXiv), working papers (e.g., SSRN), or non-peer-reviewed sources
- Opinion pieces lacking empirical or theoretical grounding
- Studies focusing exclusively on institutional or administrative AI use

Papers describing AI systems for mathematical problem solving without educational implementation were excluded to maintain conceptual alignment with mathematics education research.

2.7. Framework Stage 3: Study Selection

All retrieved records were exported into reference management software and duplicates were removed. Titles and abstracts were independently screened by two reviewers against the inclusion criteria.

Disagreements were resolved through discussion and consensus, with consultation from a third reviewer where necessary. Screening calibration was conducted prior to full review to enhance consistency.

Full-text screening was subsequently performed for eligible records. Reasons for exclusion were documented to ensure transparency.

The final corpus consisted of 21 studies (see Appendix):

- 18 empirical investigations
- 3 theoretically orienting review studies

The complete selection process is illustrated in Figure 2.

2.8. Framework Stage 4: Charting the Data

A structured data extraction form was collaboratively developed and piloted. The form was iteratively refined to ensure alignment with the theoretical focus of the review.

Extracted data included:

- Author(s) and year
- Country and educational level
- Type of AI modality (e.g., ITS, LLM, robotics, learning analytics)
- Study design
- Reported learning outcomes
- Explicit or implicit theoretical framework
- Evidence of APOS-related cognitive transitions
- Evidence of reification or process–object flexibility
- Evidence concerning conceptual image formation
- Reported barriers and facilitators

Where APOS constructs or reification processes were not explicitly operationalized, interpretive coding was conducted based on task design, representational transformations, scaffolding mechanisms, and descriptions of learner reasoning.

Data extraction was conducted by one reviewer and independently verified by a second reviewer to enhance reliability and reduce interpretive bias.

2.9. Framework Stage 5: Collating, Summarizing, and Reporting the Results

Data were synthesized using:

- Descriptive numerical summary
- Thematic synthesis (Thomas & Harden, 2008)

Thematic analysis focused on:

- Theoretical grounding of AI-supported mathematics studies
- Patterns in cognitive and conceptual outcomes
- Mechanisms supporting or constraining conceptual development
- Pedagogical and system-level implementation factors

Themes were developed inductively and refined iteratively. Conceptual clustering was used to identify recurring components of AI-supported mathematics learning environments and their relationship to mathematical thinking.

2.10. Limitations of the Review Process

The review was restricted to English-language publications, which may have excluded relevant research published in other languages. Although major databases were searched, some studies may not have been indexed within the selected sources.

The deliberate restriction to high-impact empirical studies enhanced rigor but may have excluded smaller-scale or emerging investigations. Additionally, as a scoping review, no formal methodological quality appraisal was conducted.

Given the rapidly evolving nature of AI technologies in mathematics education, future systematic reviews or meta-analyses may focus on specific AI modalities or educational levels.

3. Results

3.1. Overview of Included Studies

A total of 21 studies were included in this scoping review (18 empirical studies plus three theory-informed anchors). Evidence was most frequently situated in higher education (n = 6) and secondary education (n = 6), followed by primary education (n = 3) and pre-service teacher contexts (n = 3). Two studies examined K–12 populations, and one study spanned multiple educational levels.

Geographically, the corpus reflected broad international coverage, including studies conducted in the United States (Phillips et al., 2020; Mills, 2021; Pardos & Bhandari, 2024), Greece (Polydoros et al., 2026; Rizos & Gkrekas, 2025), Spain (Seckel et al., 2021; Tirado-Olivares et al., 2025), and additional contributions from Mexico (Lopez-Caudana et al., 2020), Norway (Forsström & Afdal, 2020), Canada (Casler-Failing, 2021), Taiwan (Hwang & Tu, 2021), Germany (Dilling & Herrmann, 2024), Indonesia (Andrini et al., 2025), Vietnam (Chau et al., 2025), Malaysia (Ramli & Ismail, 2025), Thailand (Marwiang et al., 2025), and China (He et al., 2026).

In terms of AI modality, robotics-based interventions (e.g., Lopez-Caudana et al., 2020; Forsström & Afdal, 2020) and LLM-supported environments (e.g., Pardos & Bhandari, 2024; Dilling & Herrmann, 2024; He et al., 2026) were most common. These were followed by intelligent tutoring systems (Phillips et al., 2020; Marwiang et al., 2025), learning analytics approaches (Tirado-Olivares et al., 2025; Ramli & Ismail, 2025), and generative AI applications (Bernardi et al., 2025; Brandibur et al., 2026).

3.2. Theoretical Frameworks Underpinning Included Studies

Across the 21 included studies, explicit theoretical grounding was limited. While several studies referenced general educational perspectives, only a minority articulated formal cognitive or mathematics education frameworks guiding intervention design or interpretation.

Constructivist and sociocultural perspectives were evident in robotics-based studies (e.g., Lopez-Caudana et al., 2020; Forsström & Afdal, 2020), while the TPACK framework was explicitly used in teacher-focused research (Casler-Failing, 2021). Elements of self-regulated learning (SRL) appeared in studies involving generative AI and teacher development (Bernardi et al., 2025).

However, most studies—particularly those using ITS, LLMs, and adaptive systems (Phillips et al., 2020; Mills, 2021; Pardos & Bhandari, 2024; Chau et al., 2025)—did not engage with mathematics-specific cognitive frameworks. No study explicitly operationalized APOS transitions, process–object reification, or Conceptual Image/Concept Definition alignment.

Although several studies reported improvements in "conceptual understanding," these claims were typically inferred from performance gains rather than examined through structured cognitive analysis. For example, Marwiang et al. (2025) and Andrini et al. (2025) reported statistically significant achievement improvements without examining underlying cognitive transformations.

Overall, AI-supported mathematics research appears predominantly outcome-oriented, with limited systematic engagement with established theories of mathematical cognition.

3.3. Rationale for the Use (Or Absence) of Underpinning Theory

Among studies employing theoretical frameworks, theory was primarily used to justify design decisions rather than to analyze cognitive processes.

For instance, Marwiang et al. (2025) referenced cognitive load theory to justify adaptive feedback mechanisms, while Lopez-Caudana et al. (2020) and Forsström and Afdal (2020) adopted constructivist perspectives to support activity-based learning. However, these frameworks were not extended to examine whether learners transitioned from action to process or object representations.

Similarly, LLM-based studies (Pardos & Bhandari, 2024; Dilling & Herrmann, 2024) emphasized explanation quality and reasoning performance but did not investigate reification processes or conceptual image alignment.

Across the majority of studies without explicit theoretical grounding (e.g., Mills, 2021; Chau et al., 2025; Ramli & Ismail, 2025), AI effectiveness was framed in terms of:

- achievement gains
- efficiency improvements
- engagement metrics
- user satisfaction

This pattern indicates a dominant outcome-oriented rather than cognition-oriented research paradigm.

3.4. Components of AI-Supported Mathematics Learning Environments

3.4.1. Adaptive Feedback

Adaptive feedback was the most frequently observed component (18 studies). ITS-based systems (Phillips et al., 2020; Marwiang et al., 2025) and adaptive platforms (Mills, 2021) provided immediate correction, hints, and stepwise scaffolding. While these features were associated with achievement gains, studies did not examine whether feedback supported internalization of mathematical processes.

3.4.2. Multiple Representations

Nine studies included dynamic visualizations or symbolic–graphical integration. Robotics-based interventions (Lopez-Caudana et al., 2020; Forsström & Afdal, 2020; Zhong & Xia, 2020)

doi:10.20944/preprints202603.2391.v1

emphasized representational engagement.and learning analytics environments (Tirado-Olivares et al., 2025) emphasized representational engagement. However, no study explicitly evaluated representational flexibility or process–object coordination.

3.4.3. Generative AI Dialogue

Five studies examined LLM-based environments (Pardos & Bhandari, 2024; Dilling & Herrmann, 2024; Rizos & Gkrekas, 2025; He et al., 2026; Brandibur et al., 2026). These systems generated explanations, proofs, and step-by-step solutions. Despite improvements in explanation quality and reasoning performance, several studies raised concerns regarding procedural outsourcing, particularly when students relied on AI-generated responses without prior reasoning.

3.4.4. Teacher Mediation

Six studies reported structured teacher mediation (e.g., Bernardi et al., 2025; Lopez-Caudana et al., 2020; Polydoros et al., 2026). In these contexts, AI outputs were embedded within classroom discussion and reflective tasks. These studies more frequently reported metacognitive engagement and conceptual refinement, suggesting that teacher mediation plays a critical role in shaping AI's conceptual impact.

3.4.5. Evaluation of AI-Supported Interventions

Evaluation methods varied considerably across studies.

- 15 studies relied primarily on pre–post achievement measures (e.g., Phillips et al., 2020; Andrini et al., 2025; Marwiang et al., 2025)
- 7 included engagement or usability surveys
- 4 incorporated qualitative data (e.g., Dilling & Herrmann, 2024; Casler-Failing, 2021)
- Only 3 examined conceptual understanding through open-ended reasoning tasks

Large-scale analyses (Yi et al., 2025) reported effect sizes on achievement but did not examine cognitive mechanisms. No study directly operationalized APOS-based measures or process–object transitions. This indicates a strong reliance on performance metrics rather than conceptual indicators.

3.5. Barriers and Facilitators to Effective AI Integration

3.5.1. Barriers

The most frequently identified barriers included over-reliance on automated feedback, insufficient teacher training, cognitive offloading in LLM contexts, and curriculum misalignment. Empirical evidence from studies such as Chau et al. (2025) and Tirado-Olivares et al. (2025) highlighted challenges related to instructional alignment and variability in teacher mediation. LLM-based environments (Pardos & Bhandari, 2024) further demonstrated risks of reduced productive struggle.

3.5.2. Facilitators

Key facilitators included structured teacher mediation, metacognitive prompting, collaborative AI-supported dialogue, and alignment with instructional frameworks. Studies integrating pedagogical scaffolding (Bernardi et al., 2025; Lopez-Caudana et al., 2020) reported stronger conceptual engagement and reflective reasoning.

3.6. Systemic and Individual-Level Factors

At the system level, institutional support, professional development, and curriculum alignment were critical for effective implementation. Studies reporting structured teacher training and institutional coordination demonstrated more consistent outcomes. At the learner level, factors such

 doi:10.20944/preprints202603.2391.v1

as prior mathematical knowledge, metacognitive skills, and willingness to critically engage with AI influenced conceptual outcomes. Evidence from LLM-based studies suggests that when AI is used as a dialogic partner rather than an answer generator, learners demonstrate higher levels of conceptual engagement.

3.7. Summary of Findings

Across the 21 studies, AI-supported mathematics learning was consistently associated with improved performance and engagement. However, explicit evidence of cognitive transitions central to mathematical thinking—such as process internalization, object encapsulation, reification, and conceptual image alignment—was rarely examined. The majority of studies lacked explicit theoretical grounding in mathematics cognition, revealing a significant gap between technological implementation and established theoretical models of mathematical understanding.

4. Discussion and Conclusions

4.1. Theoretical Grounding and Conceptual Development

The aim of this scoping review was to identify and critically evaluate international research (2020–2026) on AI-supported mathematics learning, with particular emphasis on theoretical grounding, cognitive implications for mathematical thinking, structural components of AI integration, evaluation practices, and the barriers and facilitators shaping conceptual outcomes.

A central finding of this review is the persistent absence of explicit theoretical grounding in the majority of AI-supported mathematics studies. Although AI interventions consistently report improvements in achievement, efficiency, and engagement (Holmes & Tuomi, 2022; Zawacki-Richter et al., 2019), only a limited subset of studies interpret outcomes through established cognitive frameworks in mathematics education, such as APOS theory (Arnon et al., 2014), process–object duality (Sfard, 1991), or the Conceptual Image/Concept Definition distinction (Tall & Vinner, 1981; Tall, 2013).

This pattern is strongly reflected across the empirical corpus. Experimental and quasi-experimental studies employing intelligent tutoring systems and generative AI (e.g., Phillips et al., 2020; Pardos & Bhandari, 2024; Marwiang et al., 2025) demonstrated measurable performance gains but did not examine whether learners achieved process internalization or object encapsulation. Similarly, studies using AI chatbots and adaptive systems (e.g., Chau et al., 2025; Mills, 2021) focused primarily on outcome metrics rather than cognitive restructuring mechanisms.

From a theoretical perspective, this represents a critical limitation. Educational innovation without theoretical grounding risks producing instrumental learning gains without structural conceptual change. In mathematics education, conceptual development requires transformations from action to process and from process to object (Dubinsky & McDonald, 2001; Sfard, 1991). The absence of these constructs in AI research suggests that many reported gains may reflect procedural fluency rather than genuine mathematical understanding.

Importantly, the few studies incorporating explicit theoretical perspectives (e.g., Bernardi et al., 2025; Casler-Failing, 2021) reported stronger indicators of reflective reasoning, representational flexibility, and pedagogical awareness. This finding reinforces the argument that theoretical alignment functions as a mediator of conceptual depth, rather than a peripheral design feature.

4.2. Core Components of Effective AI-Supported Mathematics Learning

Across the included studies, four structural components consistently emerged as central to AI-supported mathematics learning:

- adaptive feedback
- multiple representation support
- dialogic interaction (particularly via LLMs)

 doi:10.20944/preprints202603.2391.v1

- structured teacher mediation

These components were evident across different AI modalities. For example, adaptive feedback characterized ITS-based interventions (Phillips et al., 2020; Marwiang et al., 2025), while dialogic interaction defined LLM-supported environments (Pardos & Bhandari, 2024; Dilling & Herrmann, 2024). Robotics-based approaches (Lopez-Caudana et al., 2020; Forsström & Afdal, 2020) emphasized embodied and exploratory learning, supporting representational engagement.

However, a key insight of this review is that the presence of these components alone is insufficient. Their effectiveness depends on their pedagogical orchestration. Without alignment to cognitive principles, these features may support efficiency without fostering conceptual transformation.

4.3. Teacher Mediation as a Critical Mechanism

Teacher mediation emerged as a central mechanism linking AI functionality to conceptual outcomes.

When AI outputs were embedded within guided discussion, reflective questioning, and structured reasoning tasks, learners demonstrated increased metacognitive engagement and conceptual refinement (e.g., Bernardi et al., 2025; Rizos & Gkrekas, 2025). Similarly, robotics-based interventions showed that guided activity design enhanced conceptual engagement (Lopez-Caudana et al., 2020).

In contrast, AI systems functioning as autonomous solution providers were associated with procedural outsourcing and cognitive offloading (Kasneci et al., 2023; Wardat et al., 2023). Evidence from LLM-based environments indicates that when students rely on generated solutions without prior reasoning, opportunities for conceptual development are significantly reduced (Pardos & Bhandari, 2024).

These findings suggest that AI does not inherently produce learning gains; rather, its impact is mediated by instructional design and teacher involvement. From a sustainability perspective, this highlights that meaningful AI integration is fundamentally pedagogical, not technological.

4.4. Evaluation Practices and Conceptual Measurement

A major limitation identified across the literature concerns evaluation practices. Most studies relied on:

- pre–post achievement tests
- engagement surveys
- predictive or learning analytics models

Quantitative studies (e.g., Andrini et al., 2025; Ramli & Ismail, 2025; Mills, 2021) primarily measured performance indicators, while large-scale syntheses (Yi et al., 2025) focused on effect sizes without addressing underlying cognitive processes.

Only a small number of studies employed qualitative or open-ended measures capable of capturing reasoning (e.g., Dilling & Herrmann, 2024), and none directly operationalized APOS transitions or reification processes.

This reflects a broader methodological issue: the field lacks valid instruments for measuring conceptual change. As a result, claims regarding "deep learning" remain largely inferential.

Future research must prioritize the development of theoretically grounded assessment tools capable of capturing:

- process internalization
- object encapsulation
- representational coordination
- alignment between conceptual image and formal definition

4.5. Barriers and Facilitators to Conceptual AI Integration

The review identified both systemic and individual-level factors shaping the conceptual impact of AI-supported mathematics learning.

4.6. Systemic Barriers

Empirical studies highlighted recurring challenges related to implementation conditions. Studies such as Chau et al. (2025) and Tirado-Olivares et al. (2025) reported inconsistencies in instructional alignment and teacher support.

Insufficient professional development and weak pedagogical integration were identified as critical barriers (Holmes & Tuomi, 2022). From an implementation science perspective, these findings reflect the well-documented research-to-practice gap (Fixsen et al., 2005; Bauer & Kirchner, 2020).

4.7. Individual-Level Barriers

At the learner level, key risks included cognitive offloading, reduced productive struggle, and over-reliance on AI-generated solutions.

Productive struggle is a fundamental mechanism of conceptual learning in mathematics (Hiebert & Grouws, 2007). When AI reduces cognitive demand instead of scaffolding it, conceptual development may be inhibited.

4.8. Facilitators

Facilitators of conceptual impact included:

- metacognitive prompting
- requirement of learner-generated reasoning
- dialogic interaction
- alignment with cognitive theory

These elements were particularly evident in studies combining AI tools with structured pedagogy (e.g., Bernardi et al., 2025; Lopez-Caudana et al., 2020), where learners engaged more actively in meaning-making processes.

4.9. Implications for Research, Development, and Policy

- **Explicit Theoretical Integration**

Future research must systematically integrate cognitive frameworks such as APOS, process–object duality, and Conceptual Image theory. Without such grounding, claims regarding conceptual understanding remain weakly substantiated.

- **Pedagogically Coherent Design**

AI systems should explicitly define their instructional logic, including feedback mechanisms, representational structures, and mediation strategies.

- **Conceptually Valid Assessment**

Evaluation frameworks must extend beyond performance metrics to include theoretically aligned measures of conceptual development.

- **Recognition of Systemic Complexity**

AI integration should be treated as a pedagogical and systemic transformation rather than a purely technological intervention. Institutional context, teacher agency, and professional development are critical determinants of success.

4.10. Conclusions

This scoping review reveals a rapidly expanding yet theoretically fragmented field of AI-supported mathematics education. While empirical evidence indicates that AI can enhance

doi:10.20944/preprints202603.2391.v1

performance and engagement, robust evidence regarding its impact on core cognitive processes remains limited.

The central conclusion of this review is that AI's educational value is not determined by its computational capabilities, but by its theoretical alignment and pedagogical integration. Without explicit grounding in cognitive theory, AI-supported instruction risks privileging efficiency over understanding.

AI in mathematics education should therefore be conceptualized not merely as a technological innovation, but as a pedagogical and sustainability challenge. The critical question is not whether students perform better, but whether they develop durable, transferable, and conceptually grounded mathematical understanding.

Future research must prioritize theory-driven design, conceptually valid measurement, and context-sensitive implementation to ensure that AI functions as a catalyst for genuine mathematical thinking rather than a substitute for it.

Appendix A

No.	Author(s)	Year	Country	Educational Level	AI Modality	Design	Key Focus	Theoretical Framing
1	Phillips et al.	2020	USA	Secondary	ITS (Cognitive Tutor)	Quasi-experimental	Algebra achievement	Not explicit
2	Lopez-Caudana et al.	2020	Mexico	Higher Ed	Robotics	Multi-scenario empirical	Active learning in mathematics	Constructivist
3	Forsström & Afdal	2020	Norway	Primary	Robotics	Qualitative	Mathematical activity development	Sociocultural
4	Voskoglou & Salem	2020	Greece/Egypt	Higher Ed	AI-supported instruction	Conceptual analysis	Benefits & limitations	Theoretical discussion
5	Mills	2021	USA	Secondary	Adaptive system (ALEKS)	Predictive empirical	Achievement gains	Not explicit
6	Casler-Failing	2021	Canada	Pre-service teachers	Robotics	Qualitative	TPACK development	TPACK framework
7	Seckel et al.	2021	Spain/Chile	Primary	Robotics	Survey-based	Teacher conceptions	Didactical theory
8	Hwang & Tu	2021	Taiwan	Mixed	AI in math	Bibliometric + review	Trends in AI math education	Systematic review
9	Pardos & Bhandari	2024	USA	Higher Ed	LLM (ChatGPT hints)	Randomized experiment	Algebra skill learning gains	Not explicit
10	Dilling & Herrmann	2024	Germany	Pre-service teachers	LLM (ChatGPT proofs)	Exploratory empirical	Mathematical reasoning	Proof pedagogy
11	Yi et al.	2025	International	K–12	Mixed AI tools	Meta-analysis	Effect size on math achievement	Not cognitive-specific
12	Andrini et al.	2025	Indonesia	Secondary	AI-supported learning	Quantitative	Statistical evidence of outcomes	Not explicit

 doi:10.20944/preprints202603.2391.v1

13	Chau et al.	2025	Vietnam	Secondary	AI chatbot	Quasi-experimental	Problem-solving competence	Not explicit
14	Bernardi et al.	2025	Italy	Pre-service teachers	Generative AI	Mixed-method	Professional development	SRL elements
15	Rizos & Gkrekas	2025	Greece	Higher Ed	LLM	Empirical/theoretical	University math impact	Conceptual discussion
16	Tirado-Olivares et al.	2025	Spain	Primary	Learning analytics	Experimental	Elementary math learning	Not explicit
17	Ramli & Ismail	2025	Malaysia	Secondary	Learning analytics (Bayesian network)	Quantitative	Performance modeling	Not cognitive-specific
18	Marwiang et al.	2025	Thailand	Secondary	ITS (real-time feedback)	Quasi-experimental	Achievement improvement	Cognitive load reference
19	Brandibur et al.	2026	Europe	Higher Ed	Generative AI (prompt patterns)	Empirical classroom study	Higher mathematics instruction	Not explicit
20	Polydoros et al.	2026	Greece	K–12	AI-mediated inclusive systems	Conceptual + empirical synthesis	Math anxiety & inclusion	UDL framework
21	He et al.	2026	China	Higher Ed	Logic-aware LLM framework	Experimental benchmarking	Mathematical reasoning accuracy	Symbolic reasoning

References

AlSagri, H. S., & Sohail, S. S. (2024). Evaluating the role of artificial intelligence in sustainable development goals with an emphasis on quality education. *Discover Sustainability, 5*, 242. https://doi.org/10.1007/s43621-024-00682-9

Alfiras, M. I. I., Emran, A. Q., & Mohamed, A. M. (2025). Ethics and governance of generative AI in education: A systematic review on responsible adoption. *Discover Education, 5*, 37. https://doi.org/10.1007/s44217-025-01051-y

Andrini, V. S., Hidayati, U., & Etika, E. D. (2025). Leveraging educational technology: Statistical evidence on AI's role in mathematics learning outcomes. *Edukasi: Jurnal Pendidikan , 23*(1), 114–129. https://doi.org/10.31571/edukasi.v23i1.8754

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Arnon, I., Cottrill, J., Dubinsky, E., Oktaç, A., Fuentes, S. R., Trigueros, M., & Weller, K. (2014). *APOS theory: A framework for research and curriculum development in mathematics education*. Springer. https://doi.org/10.1007/978-1-4614-7966-6

Bauer, M. S., & Kirchner, J. (2020). Implementation science: What is it and why should I care? *Psychiatry Research, 283*, 112376. https://doi.org/10.1016/j.psychres.2019.04.025

Bernardi, M. L., Capone, R., Faggiano, E., & Rocha, H. (2025). Generative AI in mathematics education: pre-service teachers' knowledge and implications for their professional development. *International Journal of Mathematical Education in Science and Technology, 56*(8), 1513–1530. https://doi.org/10.1080/0020739X.2025.2490104

Brandibur, O., Filipowicz-Chomko, M., Girejko, E., Kaslik, E., Mozyrska, D., Mureșan, R., Pappas, N., Tănasie, A. L., & Zaharia, C. (2026). Higher mathematics education and AI prompt patterns: Examples from selected university classes. *Applied Sciences, 16*(1), 339. https://doi.org/10.3390/app16010339

Casler-Failing , S. (2021). Learning to teach mathematics with robots: Developing the 'T' in technological pedagogical content knowledge. *Research in Learning Technology*, *29*. https://doi.org/10.25304/rlt.v29.2555

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science, 136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

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

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Davidoff, F., Dixon-Woods, M., Leviton, L., & Michie, S. (2015). Demystifying theory and its use in improvement. *BMJ Quality & Safety, 24*(3), 228–238. https://doi.org/10.1136/bmjqs-2014-003627

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*. https://doi.org/10.3389/frai.2024.1460337

Dubinsky, E., & McDonald, M. A. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton (Ed.), *The teaching and learning of mathematics at university level: An ICMI study* (pp. 275–282). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47231-7_25

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Forsström, S. E., & Afdal, G. (2020). Learning mathematics through activities with robots. *Digital Experiences in Mathematics Education, 6*(1), 30–50. https://doi.org/10.1007/s40751-019-00057-0

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence, 7*, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hattie, J. A. C. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

He, Y., Liu, J., Wen, W., Guo, Y., Chen, H., Ma, Y., Ke, W., & Wang, P. (2026). LogicICL: Towards instructive logical-aware demonstrations for stimulating LLMs' mathematical reasoning ability. *IEEE Transactions on Emerging Topics in Computational Intelligence*. https://doi.org/10.1109/TETCI.2026.3654446

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Hooshyar, D., Šír, G., Yang, Y., Kikas, E., Hämäläinen, R., Kärkkäinen, T., Gašević, D., & Azevedo, R. (2025). Towards responsible AI for education: Hybrid human-AI to confront the elephant in the room. *Computers and Education: Artificial Intelligence, 9*, 100524. https://doi.org/10.1016/j.caeai.2025.100524

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics, 9*(6), 584. https://doi.org/10.3390/math9060584

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science, 5*, Article 69. https://doi.org/10.1186/1748-5908-5-69

Liang, Y., Bi, X., Shen, R., He, Z., Wang, Y., Xu, J., Zhang, Y., & Fan, X. (2025). When mathematical methods meet artificial intelligence and mobile edge computing. *Mathematics, 13*(11), 1779. https://doi.org/10.3390/math13111779

Lopez-Caudana, E., Ramirez-Montoya, M. S., Martínez-Pérez, S., & Rodríguez-Abitia, G. (2020). Using robotics to enhance active learning in mathematics: A multi-scenario study. *Mathematics, 8*(12), 2163. https://doi.org/10.3390/math8122163

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson

Marwiang, M., Prasertsang, M., & Junpeng, P. (2025). Enhancing students' learning outcomes in mathematics through intelligent tutoring systems based on real-time feedback. *Journal of Education and Learning, 14*(6), 186–196. https://doi.org/10.5539/jel.v14n6p186

Mills, N. J. D. (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

O'Connor, B., & Park, M. (2023). Exploring the influence of collaborative data-based decision making among teachers in professional learning communities on teaching practice. *Disciplinary and Interdisciplinary Science Education Research, 5*, Article 17. https://doi.org/10.1186/s43031-023-00086-1

Pardos, Z. A., & Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. *PLOS ONE, 19*(5), Article e0304013. https://doi.org/10.1371/journal.pone.0304013

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis, 18*(10), 2119–2126. https://doi.org/10.11124/JBIES-20-00167

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods, 5*(4), 371–385. https://doi.org/10.1002/jrsm.1123

Phillips, A., Pane, J. F., Reumann-Moore, R., & Shenbanjo, O. (2020). Implementing an adaptive intelligent tutoring system as an instructional supplement. *Educational Technology Research and Development, 68*(3), 1409–1437. https://doi.org/10.1007/s11423-020-09745-w

Polydoros, G., Antoniou, A.-S., & Polydoros, C. (2026). Inclusive AI-mediated mathematics education for students with learning difficulties: Reducing math anxiety in digital and smart-city learning ecosystems. *Encyclopedia, 6*(2), 39. https://doi.org/10.3390/encyclopedia6020039

Ramli, N., & Ismail, M. T. (2025). Learning analytics of online students' performance in mathematics using Bayesian network. *Journal of Quality Measurement and Analysis, 21*(4), 389–409. https://doi.org/10.17576/jqma.2104.2025.20

Rizos, I., & Gkrekas, N. (2025). The impact of large language models on mathematics education and research at the university. *Social Sciences & Humanities Open, 12*, 101969. https://doi.org/10.1016/j.ssaho.2025.101969

Seckel, M. J., Breda, A., Font, V., & Vásquez, C. (2021). Primary school teachers' conceptions about the use of robotics in mathematics. *Mathematics, 9*(24), 3186. https://doi.org/10.3390/math9243186

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1–36. https://doi.org/10.1007/BF00302715

Sfard, A. (2008). Thinking as communicating: Human development, the growth of discourses, and mathematizing. Cambridge University Press.

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2025). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development, 33*(2), 1921–1947. https://doi.org/10.1002/sd.3221

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tall, D. (2013). How humans learn to think mathematically: Exploring the three worlds of mathematics. Cambridge University Press.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151–169. https://doi.org/10.1007/BF00305619

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*, 45. https://doi.org/10.1186/1471-2288-8-45

Tirado-Olivares, S., Mínguez-Pardo, R., del Olmo-Muñoz, J., & González-Calero, J. A. (2025). Utilizing learning-analytics-based activities as a bridge to enhance elementary students' mathematical learning. *IEEE Transactions on Learning Technologies*, *18*, 593–605. https://doi.org/10.1109/TLT.2025.3570979

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Voskoglou, M. G., & Salem, A.-B. M. (2020). Benefits and limitations of the artificial with respect to the traditional learning of mathematics. *Mathematics*, *8*(4), 611. https://doi.org/10.3390/math8040611

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7), em2286. https://doi.org/10.29333/ejmste/13272

Yi, L., Liu, D., Jiang, T., & Xian, Y. (2025). The effectiveness of AI on K–12 students' mathematics learning: A systematic review and meta-analysis. *International Journal of Science and Mathematics Education*, *23*, 1105–1126. https://doi.org/10.1007/s10763-024-10499-7

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X., Chu, X., Chai, C. S., Jong, M. S.-Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, *2021*, Article 8812542. https://doi.org/10.1155/2021/8812542

Zhong, B., & Xia, L. (2020). A systematic review on exploring the potential of educational robotics in mathematics education. *International Journal of Science and Mathematics Education*, *18*, 79–101. https://doi.org/10.1007/s10763-018-09939-y