

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2026
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2026
*Proceedings of the Fifty-Fifth Spring Conference
of the Union of Bulgarian Mathematicians
Tryavna, Bulgaria, April 5–9, 2026*

THE EXPANDING SCOPE OF ARTIFICIAL INTELLIGENCE
INTEGRATION IN MATHEMATICS EDUCATION:
A COMPREHENSIVE REVIEW

Tsvetelina Stefanova¹, Slavi Georgiev^{2,3}

¹Department of Computer Systems and Technologies, University of Ruse

²Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

³Department of Applied Mathematics and Statistics, University of Ruse

e-mails: ¹tsstefanova@uni-ruse.bg, ^{2,3}sggeorgiev@uni-ruse.bg

As of early 2026, Artificial Intelligence (AI) has transitioned from an experimental auxiliary tool to a foundational pillar of modern mathematical pedagogy. This paper evaluates the rapid evolution of AI-driven mathematics tools over the academic year 2025–2026, emphasizing the shift toward adaptive, Socratic tutoring systems. We analyze key platforms and tools such as Khanmigo, Mathos AI, and Tutor CoPilot, assessing their impact on student mastery and teacher efficiency. Empirical studies from recent years suggest that AI-assisted human tutoring can lead to measurable improvements in student topic mastery. However, the integration of AI in mathematics education faces ongoing challenges regarding cognitive over-reliance and data privacy. This review synthesizes contemporary research to provide a framework for balanced AI integration that prioritizes conceptual understanding over automated problem-solving.

Keywords: artificial intelligence, mathematics education, intelligent tutoring systems, adaptive learning systems, Socratic AI tutoring, human–AI hybrid instruction, conceptual understanding

РАЗШИРЯВАЩИЯТ СЕ ОБХВАТ НА ИНТЕГРИРАНЕТО
НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В ОБУЧЕНИЕТО ПО
МАТЕМАТИКА: ВСЕОБХВАТЕН ОБЗОР

Цветелина Стефанова¹, Слави Георгиев^{2,3}

¹Катедра „Компютърни системи и технологии“, Русенски университет „Ангел Кънчев“

²Институт по математика и информатика, Българска академия на науките

³Катедра „Приложна математика и статистика“, Русенски университет „Ангел Кънчев“

e-mails: ¹tsstefanova@uni-ruse.bg, ^{2,3}sggeorgiev@uni-ruse.bg

<https://doi.org/10.55630/mem.2026.55.242-254>

2020 Mathematics Subject Classification: 97D40, 97P80, 97U50, 97U70.

Към началото на 2026 г. изкуственият интелект (ИИ) преминава от експериментален помощен инструмент към основна опора на съвременната математическа педагогика. Настоящата статия оценява бързата еволюция на ИИ-базираните платформи и инструменти за обучение по математика през академичната 2025–2026 година, като акцентира върху прехода към адаптивни, сократически системи за подпомагане на ученето. Анализирани са ключови платформи като Khanmigo, Mathos AI и Tutor CoPilot, като се оценява влиянието им върху усвояването на учебното съдържание от учениците и ефективността на учителите. Емпирични проучвания от последните години показват, че обучението с помощта на изкуствен интелект може да доведе до измерими подобрения в овладяването на учебните теми от учениците. Въпреки това приложението на ИИ в обучението по математика продължава да се сблъсква със съществени предизвикателства, свързани с предоверяване на ИИ и защита на личните данни. Настоящият обзор синтезира актуални изследвания и предлага рамка за балансирана интеграция на ИИ, която поставя приоритет върху концептуалното разбиране, а не върху автоматизираното решаване на задачи.

Ключови думи: изкуствен интелект, обучение по математика, интелигентни обучаващи системи, адаптивни обучаващи системи, сократическо ИИ-обучение, хибридно обучение човек–ИИ, концептуално разбиране

1. Introduction

The period 2025–2026 represents a watershed moment in the digital transformation of education. Global investment in the EdTech AI sector reached an estimated \$7.6 billion in 2025, with projections suggesting a rise to over \$30 billion by 2029, reflecting both the growing maturity of AI solutions and their mainstream adoption in schools and universities [32]. In mathematics education, AI integration serves two primary functions: enhancing student learning through personalized, adaptive exercises and feedback, and improving teacher efficiency by automating time-consuming administrative burdens such as grading, progress monitoring, and lesson planning. These capabilities increasingly position AI as not merely a supplementary tool, but as an integral component of the instructional ecosystem. Current leadership sentiment reflects this shift, with 69% of K–12 leaders viewing AI's impact on mathematics instruction as inherently positive, citing its potential to support differentiation, remediation, and enrichment at scale [3, 20].

2. Literature Review: Trends in AI-Mediated Mathematics Education (2025–2026)

Recent research indicates that AI-mediated mathematics education has entered a phase of consolidation, moving from experimental use toward systematic integration into instructional practice. Whereas earlier studies largely framed AI as a computational or assessment aid, literature from 2025–2026 increasingly conceptualizes AI as a pedagogical mediator embedded within human-guided learning environments [28, 36, 40]. This shift is evident across systematic reviews, bibliometric analyses, and empirical studies.

2.1. From Automation to Pedagogical Mediation

Systematic reviews document a clear move away from solver-centric automation toward adaptive tutoring, formative feedback, and instructional scaffolding. The longitu-

dinal review (2015–2025) reports a decline in purely answer-generating systems and a rise in tools designed to support learning processes and conceptual understanding [28]. It is similarly noted in [36] that recent studies increasingly evaluate AI tools based on pedagogical impact rather than computational efficiency. These trends extend earlier concerns regarding over-reliance on automated solvers, as highlighted in our previous research [35], and place greater emphasis on design features such as delayed answer release, adaptive hints, and reflective prompts.

2.2. Learning Outcomes and Mastery

Empirical studies suggest that AI-supported mathematics instruction can lead to measurable learning gains when pedagogically aligned. The review of AI-powered assessments [25] reports improvements in student performance, engagement, and feedback quality, particularly in algebra and calculus. Research on intelligent tutoring systems (ITS) such as DreamBox, ALEKS, and Mathspace shows that adaptive readiness models can enhance mastery by sequencing content and enforcing prerequisite knowledge [37]. Several platforms now hold “strong” evidence ratings under the Every Student Succeeds Act (ESSA) framework.

Human–AI hybrid tutoring models receive particular attention. [39] demonstrates that AI-augmented human tutoring significantly increases the likelihood of topic mastery, with especially strong effects for less-experienced tutors, indicating that AI can amplify pedagogical expertise rather than replace it.

2.3. Socratic and Dialogue-Based Approaches

A prominent trend in the 2025–2026 literature is the rise of Socratic and dialogue-based tutoring systems. Bibliometric analyses identify dialogic scaffolding and non-directive tutoring as emerging focal concepts in mathematics education research [40]. Enabled largely by Large Language Models (LLMs), these systems aim to promote metacognitive regulation by guiding learners through questions and hints instead of providing direct solutions. However, studies caution that generative dialogue alone does not guarantee pedagogical soundness, noting risks of fluent but mathematically incorrect explanations when models are insufficiently constrained [5]. Consequently, teacher-in-the-loop designs are increasingly emphasized as a necessary safeguard [22].

2.4. Teacher Support, Risks, and Equity

Recent literature also frames AI as a tool for teacher support. Survey evidence indicates that educators using AI-based tools for lesson planning, assessment, and feedback report substantial reductions in administrative workload [3], potentially enabling greater focus on student reasoning and targeted intervention [34]. At the same time, studies consistently highlight risks of cognitive over-reliance. Excessive dependence on solver tools has been linked to weakened independent problem-solving and retention, particularly in scan-to-solve workflows [6, 38].

Equity and governance concerns remain unresolved. Uneven access to advanced AI tutoring systems may exacerbate existing educational inequalities [7], while policy-oriented research underscores growing challenges related to data privacy, procurement transparency, and accountability as AI becomes embedded in assessment and learning analytics [24].

2.5. Summary

Overall, the 2025–2026 literature portrays AI-mediated mathematics education as shifting toward integrated, dialogic, and human–AI hybrid models. Evidence suggests that educational impact depends less on raw problem-solving capability and more on pedagogical alignment – specifically, the ability of AI systems to support explanation, verification, and conceptual understanding under human oversight. The focus shifts from the question “Should we use AI?” to “How do we integrate it ethically and pedagogically?”. The main challenge remains the balance between automated assistance and preserving students’ ability to think critically without external help. These findings motivate the comparative analysis of tools and use cases presented in the following sections.

3. Comparative Analysis of Contemporary AI Tools Used in Mathematics Education

The 2025–2026 AI-in-education landscape is characterized by the convergence of large language models (LLMs) and natural language processing (NLP) with domain-specific symbolic mathematics engines. Rather than treating AI as a monolithic category, current tools differ substantially in pedagogical intent, degree of automation, and the role assigned to teachers and learners.

Table 1 provides a functional comparison of major AI tool categories currently used in mathematics education, organized by pedagogical designation rather than underlying technology alone. For each category, the table summarizes representative platforms, key innovations introduced during 2025–2026, and the strongest forms of research-backed educational impact reported in recent studies. This structure reflects the paper’s focus on how design choices – such as answer withholding, adaptive sequencing, or human-in-the-loop support – shape learning outcomes.

Table 1: Pedagogical Design, Innovation, and Impact of Contemporary AI Mathematics Tools

<i>Core Designation</i>	<i>Sample Tools</i>	<i>2025–2026 Innovation</i>	<i>Research-Backed Impact</i>
Socratic Tutor	Khanmigo, Google Socratic, Astra AI, Theta-wise, Rori	Teacher-in-the-Loop Safeguard: Tools now feature dashboards that allow teachers to toggle “reveal answer” privileges and view chat logs to identify misconceptions without manual grading.	Active Recall Efficacy: By refusing to provide direct answers, these tools enforce active recall. Cognitive science research confirms this method increases long-term retention by up to 50% compared to passive review.

Adaptive Solver	Wolfram Alpha, Mathos AI, Microsoft Math Solver, Photomath, Symbolab	Domain-Specific Fine-Tuning: Unlike general chatbots, these solvers are now fine-tuned specifically on mathematical notation and step-by-step logic, aiming to eliminate calculation hallucinations.	Procedural Fluency: Studies on tools like Microsoft Math Solver show that when combined with “similar problem” generation, they significantly improve students’ ability to solve multi-step calculus and algebra problems.
Intelligent Tutoring System (ITS) with Real-time Assessment	ALEKS, DreamBox Learning, Mathspace, Knewton Alta, Eedi	Predictive Readiness Models: Systems like Mathspace and ALEKS now use AI to predict a student’s knowledge state, blocking advanced topics until prerequisite skills are statistically mastered.	ESSA Strong Evidence: ITS platforms like DreamBox hold the highest strong rating under the Every Student Succeeds Act (ESSA) [14], proving students achieve more growth in math proficiency than peers.
Visual Scaffolding	GeoGebra, Desmos, Solvely, Graspable Math, Polypad	Dynamic 3D & AR Visualization: Recent updates (e. g., Desmos 3D) allow students to manipulate variables in real-time to visualize abstract concepts like multivariable calculus and geometric transformations.	Conceptual Grounding: Research indicates that dynamic graphing tools help ground abstract symbols in concrete visuals, measurably improving student performance on conceptual questions regarding functions.
Human-AI Hybrid	Tutor CoPilot, Thinkster Math, ASSISTments, MATHia, Brainly, TutorMe	Real-Time Pedagogical Prompting: AI analyzes the live session and whispers suggestions to the human tutor (e. g., “ <i>Ask them to explain the next step</i> ”), turning novice tutors into experts.	Mastery Gains: A Stanford University RCT (2024-25) on Tutor CoPilot proved that students were 4% more likely to master a topic when their human tutor used AI prompts [39].

General AI Model	ChatGPT, Gemini, DeepSeek, Claude, Microsoft Copilot	Chain-of-Thought (CoT) Processing: These models generate a hidden internal monologue to plan, critique, and verify their logic before outputting a final answer.	Benchmark Breakthroughs: CoT reasoning has raised AI performance on difficult math competitions (like the AIME) from ~11.67% (GPT-4o) to ~81.67% (o1), significantly reducing logic errors [5].
------------------	---	--	--

The table highlights a clear shift toward Socratic and adaptive tutoring systems, in which AI is designed to prompt reasoning, enforce active recall, and support conceptual understanding rather than deliver immediate solutions. At the same time, solver-oriented systems remain prominent, particularly for procedural fluency, benefiting from recent advances in domain-specific fine-tuning that reduce mathematical hallucinations. Intelligent Tutoring Systems (ITS) occupy a middle ground, combining automated assessment with mastery-based progression models that have accumulated strong empirical validation.

Notably, human–AI hybrid tools emerge as the category with the most consistent evidence of learning gains, aligning closely with the abstract’s emphasis on AI-assisted human tutoring. By augmenting – not replacing – human instructors through real-time pedagogical prompts, these systems demonstrate measurable improvements in topic mastery while mitigating risks of cognitive over-reliance. General-purpose AI models, while increasingly capable due to chain-of-thought reasoning, remain pedagogically ambiguous and require careful instructional framing to avoid reverting to answer-centric use.

Table 1 makes explicit that contemporary mathematics AI tools are converging on natural-language interaction while diverging sharply in pedagogical posture – from dialogue-driven scaffolding to solution-centric automation to human-in-the-loop augmentation. The distribution of research-backed outcomes suggests that learning gains are most robust when AI systems are embedded within pedagogically constrained or hybrid instructional designs, rather than used as fully autonomous solvers.

3.1. Advancements in Socratic and Generative Assistants

The most significant technical leap in 2025 was the reduction of hallucinations in general-purpose models through the adoption of reinforcement learning and chain-of-thought architectures. **DeepSeek**, with its R1 model released in early 2025, exemplified this shift by introducing a thinking phase, where the AI plans multi-step solutions before generating output. This architecture raised performance on high-stakes benchmarks like the AIME to levels previously attainable only by closed-source proprietary models [11]. Similarly, **Google’s Gemini** ecosystem evolved with the release of the Deep Think capability in its 2.5 Pro model, which iteratively explores hypotheses for complex logic puzzles. Google further refined this for education with “Gems” – customized AI personas that educators can fine-tune to align with specific curricular standards, effectively creating bespoke teaching assistants for distinctive classroom needs [19]. Anthropic’s **Claude** continued this trend with its 3.5 and 4.0 model updates in mid-2025, introducing “Artifacts” – a feature that renders interactive code and quizzes in a dedicated window, transforming static text interactions into dynamic, visual learning experiences [2]. In

the realm of dedicated educational assistants, **Khanmigo** expanded its utility beyond student tutoring. As of April 2025, it features a comprehensive “Teacher Tools” suite integrated directly into Learning Management Systems (LMS) like Canvas. This update allows educators to automate the creation of rubrics, exit tickets, and lesson plans without leaving their gradebook environment, significantly reducing administrative overhead [33].

3.2. Innovations in Symbolic, Visual, and Solver Tools

Tools rooted in symbolic computation have integrated Large Language Model (LLM) capabilities to enhance accessibility without sacrificing accuracy. Mathos AI (formerly MathGPTPro) released a “PDF Homework Helper” in 2025 that goes beyond simple photo scanning. Students can now upload entire document files, where the AI overlays annotations and step-by-step guidance directly onto the page. This system utilizes a specialized model that claims 20% higher accuracy than standard GPT-4 implementations for calculus tasks [23]. Wolfram Alpha has solidified its role as the grounding engine for the AI ecosystem. Through a robust 2025 plugin architecture, it now integrates seamlessly with ChatGPT and other LLMs, allowing them to verify their computational outputs against Wolfram’s structured data libraries, effectively solving the arithmetic hallucination problem [30].

Visual graphing tools have also seen substantial updates to support higher education. **Desmos Studio** updated its 3D Calculator in early 2026 to include enhanced visualization controls – such as disabling lighting for clearer surface viewing – and expanded regression capabilities (cubic, quartic, and sinusoidal) that allow students to model complex, real-world data sets [12]. **GeoGebra** similarly launched “Math Practice”, an adaptive mode that provides step-by-step feedback on algebraic problems, encouraging students to explore multiple solution paths rather than adhering to a single rigid method [18]. **Photomath** refined its visual scaffolding approach by launching “Animated Steps”, which break down static scan results into dynamic progressions to help learners visualize the procedural flow of a solution [31]. **Solvely** and **Microsoft Math Solver** have also integrated similar scan-to-solve upgrades, with Microsoft’s **Math Progress** in Teams now using AI to identify class-wide error trends and automatically generate targeted practice assignments [26].

3.3. Evolution of Adaptive Learning Platforms

Comprehensive learning platforms have focused heavily on accessibility and verified content reliability. **DreamBox Learning** (now part of Discovery Education) released significant updates in January 2025, including new “Racecourse” lessons for fraction fluency and enhanced accessibility features that allow students with low vision to navigate complex manipulatives using keyboard commands [9]. **Brainly** addressed the reliability crisis with its 2025 “AI Learning Companion”, which uses scan-to-solve technology to provide explanations verified against a database of textbooks, distinguishing itself from raw generative text by citing specific academic sources [8]. **ALEKS** and **Mathspace** continued to refine their adaptive engines, with Mathspace’s 2025 updates focusing on its Mixture-of-Experts AI, which now provides voice-based coaching to support students with reading difficulties. Finally, human-hybrid systems like **Tutor CoPilot** and **Astra AI** have moved toward a teacher-in-the-loop model, where the AI provides real-time sug-

gestions to human tutors or parents, effectively augmenting rather than replacing human instruction [39].

3.4. Evidence of Student Usage

Available surveys and institution-level studies indicate that GenAI use for learning is now mainstream across secondary and higher education, though prevalence varies by sampling and definitions. In the United States, high-school students' self-reported use of GenAI for schoolwork rose from 79% (Jan 2025) to 84% (May 2025), with ChatGPT the dominant tool (69% reporting use for assignments/homework) [1]. Meanwhile, among U.S. teens overall, 64% report using AI chatbots; usage concentrates in a few mainstream systems – ChatGPT (59%), Gemini (23%), and Meta AI (20%) – with others trailing [15].

In higher education, the same concentration pattern appears. An institutional study reports ChatGPT as the primary tool (89.3% of AI users), with Gemini (13.5%) and Copilot (7.7%) far behind and other platforms each <5% [10]. A separate survey similarly shows students assembling tool portfolios, but with strong skew: ChatGPT is cited most (606 mentions), followed by Copilot (67) and Gemini (51), with a longer tail including Grammarly, Claude, and math-oriented apps such as Gauth [27].

For mathematics, students commonly combine general chatbots with specialized solvers. A 2025 study of engineering students reports frequent use of ChatGPT (62.7%) alongside high usage of Mathway (51.7%) and Photomath (48%), reflecting a pragmatic blend of explanation-seeking and scan/stepwise solution workflows [16]. Finally, tool choice can be region- and access-dependent: qualitative evidence from pre-sessional international students (UK) documents multi-tool habits shaped by VPN and language constraints, including reliance on region-specific systems such as DeepSeek alongside global tools and routine AI-to-AI cross-checking [4].

4. Primary Use Cases in Education

AI-enabled mathematics platforms are increasingly used to build *personalized learning pathways*, where learner data are continuously collected and analyzed to adapt the sequence, difficulty, and pacing of practice. In adaptive learning platforms, this typically involves model-based estimation of student knowledge and readiness, allowing systems to recommend targeted exercises and feedback aligned with the learner's evolving profile [35, 37].

Closely related is *misconception remediation*: rather than treating errors as generic “wrong answers”, diagnostic approaches aim to infer which misconception is most likely and to trigger immediate, targeted interventions. Tools such as Eedi explicitly operationalize this idea through diagnostic questioning and misconception mapping to support rapid, instructionally meaningful feedback [17].

A second dominant use case is *multimodal interaction*, reflecting the broader movement toward systems that combine text with image and speech. In practice, this enables students to scan printed or handwritten mathematics and receive step-by-step guidance (a workflow popularized by tools such as Photomath), while emerging multimodal foundation models expand the feasibility of voice-based and visually grounded tutoring scenarios [25].

Finally, AI is increasingly framed as a *teacher-empowerment layer*, accelerating routine tasks such as generating practice sets, drafting lesson materials, and producing feedback at scale. Survey evidence from 2025 suggests that teachers who use AI tools at least weekly estimate time savings of roughly six hours per week on average – highlighting why administrative relief is often cited as a primary driver of adoption [3].

5. Challenges and Ethical Considerations

There are many challenges arising with the introduction of AI mathematical instruments in education [29]. While the potential of AI-supported mathematics instruction is widely acknowledged, the 2025–2026 literature is equally explicit about the risk of *cognitive disengagement*. When students treat AI as an “answer vending machine”, learning can shift from sense-making to output optimization, weakening productive struggle and durable problem-solving habits. This is particularly acute in mathematics because low-friction workflows (scan → answer) invite students to bypass verification and explanation; in Bulgaria, practitioners and education media have repeatedly noted extensive use of Photomath for homework and even tests, often with limited critical assessment – an archetypal “vending machine” pattern that prioritizes completion over comprehension [38]. The same shortcut dynamic is increasingly reinforced by general AI chatbots (especially ChatGPT), which are now widely used for schoolwork and for clarifying/double-checking solutions in policy and survey reports [1, 21]. Large surveys of educators therefore report substantial concern that AI can erode critical-thinking and research skills, underscoring the need for classroom norms that require justification, reflection, and error-checking rather than solution acquisition [6].

A second constraint is *reliability*, which matters because students increasingly consult general-purpose LLMs as if they were authoritative math engines. Even as frontier systems have achieved striking results on demanding tasks (including widely reported near “gold medal” performance on Olympiad-level subsets), outputs can still contain subtle reasoning gaps, incorrect intermediate steps, or explanations misaligned with a learner’s grade level. Recent “live” evaluations such as MathArena highlight this asymmetry: many leading models score very highly on final-answer competitions, yet are less dependable on proof-style/justification-heavy tasks, with proof-based Olympiad settings still leaving substantial room for improvement [5]. As a result, a student may receive the correct final answer while internalizing an invalid method or a fluent but mathematically unsound explanation. Consequently, “Socratic” scaffolding and teacher oversight are design requirements, not add-ons: the key question is whether systems can *consistently* support valid, student-appropriate reasoning, surface assumptions, and promote verification. This aligns with evaluation-focused work emphasizing representative datasets and meaningful metrics, since superficial accuracy can mask pedagogically harmful failure modes (e. g., plausible-but-wrong steps and curriculum misalignment). In practice, the photo-based convenience and step-by-step outputs common in student-facing math tools can be beneficial, but without structured norms for checking and explaining, they may unintentionally amplify uncritical acceptance [13].

Finally, the expansion of AI in classrooms intensifies *equity and privacy* concerns. Access to high-quality AI tutoring and teacher-support systems may widen existing gaps if advanced tools are concentrated in well-funded schools, contributing to a broader AI literacy divide [7]. In parallel, schools face growing scrutiny around student-data gover-

nance, procurement transparency, and the downstream risks of breaches or surveillance-oriented deployments – issues that become more acute as AI tools integrate deeper into assessment, learning analytics, and identity-related systems [24].

Conclusions

Evidence from 2025–2026 indicates that AI is becoming an increasingly entrenched component of mathematics education – less a short-lived trend than a structural shift in how learners practice, receive feedback, and access support. The most effective deployments to date align with a human–AI hybrid approach, in which technology amplifies teacher and tutor expertise – through targeted scaffolding, formative feedback, and workflow automation – rather than attempting to substitute for instructional judgment. In parallel, the rapid rise of model capability, including demonstrations of near-gold-medal performance on elite mathematical benchmarks, underscores a crucial pedagogical inflection point: the central question is no longer whether AI can solve difficult problems, but whether it can reliably help students learn to solve them through age-appropriate reasoning, explanation, and metacognitive support.

Solver-style apps (camera-based and symbolic engines) make mathematics fast and frictionless via scan-to-solution, stepwise transformations, visualization, and instant checking – useful for practice, but also prone to turning learning into “outsourced thinking”. At the same time, benchmarking shows rising accuracy on answer-checkable tasks (high scores on final-answer competitions) while proof-quality reasoning and fully reliable explanations remain less stable. With student adoption now high (e. g., 84% of US high-schoolers report using GenAI for schoolwork, and teens’ chatbot use is led by ChatGPT and Gemini), teachers and tutors should integrate AI as a *coach* and *checker* – for hints, what-if prompts, explanation requests, and misconception diagnosis – while reserving full solver mode for moments that require student justification, comparison of methods, and reflection (e. g., a written reasoning chain or error analysis before consulting full solutions).

Future research should prioritize (i) causal classroom evidence beyond short interventions, including longitudinal studies that test retention, transfer, and mathematical identity; (ii) measurement frameworks that distinguish conceptual growth from answer productivity (e. g., explanation quality, error diagnosis, and strategy diversity); and (iii) implementation science for equitable and privacy-preserving adoption, covering procurement, data governance, and teacher professional development. A particularly promising direction is the design and evaluation of “Socratic-by-default” systems that can adapt their questioning strategies to student level while remaining robust to hallucinations and misuse – shifting AI evaluation from raw problem-solving accuracy to instructional reliability and learning impact.

Acknowledgements

This research is supported by the Scientific Research Fund of “Angel Kanchev” University of Ruse, Bulgaria.

References

- [1] ADAIR, A., ET AL. U.S. High School Students' Use of Generative Artificial Intelligence. New Evidence from High School Students, Parents, and Educators. College Board Research, 2025, https://research.collegeboard.org/media/pdf/AI%20Research%20Brief%201_vf_0.pdf, last visited 04.02.2026
- [2] Anthropic. Claude 3.5 Sonnet, 2024. <https://www.anthropic.com/news/claude-3-5-sonnet>, last visited 04.02.2026
- [3] ASH, A.M. Three in 10 Teachers Use AI Weekly, Saving Six Weeks a Year, 2025. <https://news.gallup.com/poll/691967/three-teachers-weekly-saving-six-weeks-year.aspx>, last visited 05.01.2026
- [4] ATTEWELL, S. Student Perceptions 2025: Insights from a Group of Pre-sessional International Students at the University of Manchester. AI in universities and colleges, 2025. <https://nationalcentreforai.jiscinvolve.org/wp/2025/09/01/student-perceptions-2025-insights-from-a-group-of-pre-sessional-international-students-at-the-university-of-manchester/>, last visited 04.02.2026
- [5] BALUNOVIĆ, M., ET AL. MathArena: Evaluating LLMs on Uncontaminated Math Competitions. arXiv:2505.23281v3 [cs.AI], 2026. <https://doi.org/10.48550/arXiv.2505.23281>
- [6] BENEDEK, M., B.R. SZIKLAI. Impact of AI tools on learning outcomes: decreasing knowledge and over-reliance. arXiv:2510.16019v1 [cs.CY], 2025. <https://doi.org/10.48550/arXiv.2510.16019>
- [7] BOOTH, R. Generation AI: fears of 'social divide' unless all children learn computing skills, 2026, <https://www.theguardian.com/education/2026/jan/05/generation-on-ai-fears-of-social-divide-unless-all-children-learn-computing-skills>, last visited 05.01.2026
- [8] Brainly Staff. Report: How Students Are Embracing AI This Back-to-School Season (2025), Insights: The Brainly Blog, 2025. <https://brainly.com/insights/report-how-students-are-embracing-ai-this-back-to-school-season-2025>, last visited 04.02.2026
- [9] CHARLOTTE, N.C. New Enhancements to DreamBox Math by Discovery Education Improves Accessibility for Elementary Students. Discovery Education, 2025. <https://www.discoveryeducation.com/details/new-enhancements-to-dreambox-math-by-discovery-education-improves-accessibility-for-elementary-students/>, last visited 04.02.2026
- [10] CONTRACTOR, Z., G. REYES. Generative AI in Higher Education: Evidence from an Elite College. arXiv:2508.00717 [econ.GN], 2025. <https://doi.org/10.48550/arXiv.2508.00717>
- [11] DeepSeek. DeepSeek-R1 Release: Reasoning Models and Benchmarks. DeepSeek API Docs, 2025. <https://api-docs.deepseek.com/news/news250120>, last visited 04.02.2026
- [12] Desmos Studio. What's New at Desmos Studio. Desmos Help Center, 2025. https://help.desmos.com/hc/en-us/articles/4405017454477--What-s-New-at-Desmos-Studio#h_01KC4GCCYWN5KTS2GYYJWWC2QY, last visited 04.02.2026
- [13] DYUZHOV, A. Top 5 Best AI Tools for Math in 2025. AI Hub, Overchart AI, 2025.

- <https://overchat.ai/ai-hub/best-ai-tools-for-math-in-2025>, last visited 04.02.2026
- [14] Evidence for Essa. DreamBox Math, DreamBox Learning, 2025. <https://www.evidenceforessa.org/program/dreambox-learning/>, last visited 04.02.2026
- [15] FAVEIRO, M., O. SIDOTI. Teens, Social Media and AI Chatbots 2025. Pew Research Center, 2025. <https://www.pewresearch.org/internet/2025/12/09/teens-social-media-and-ai-chatbots-2025/>, last visited 04.02.2026
- [16] GARCIA, K.F., ET AL. Engineering students' perceptions and actual use of AI-based math tools for solving mathematical problems. *Acta Psychologica*, 256, 105004, 2025. <https://doi.org/10.1016/j.actpsy.2025.105004>
- [17] GARCIA, M.B., ET AL. Modernizing mathematics education with artificial intelligence: a narrative review of AI-powered tools, thematic trends, and instructional applications, SSRN, 2025, <https://dx.doi.org/10.2139/ssrn.5621770>
- [18] GeoGebra. GeoGebra Math Practice, 2026. <https://www.geogebra.org/mathpractice/en>, last visited 04.02.2026
- [19] Google Support. Use Deep Think in Gemini Apps. Gemini Apps Help, 2025. <https://support.google.com/gemini/answer/16345172>, last visited 04.02.2026
- [20] GQOLI, N., J. BAIDOO. The integration of artificial intelligence in mathematics education: towards developing a conceptual framework for understanding student perspectives in higher education. *E-Journal of Humanities, Arts and Social Sciences* Vol.6 No.10, pp 2660-2674, 2025. <https://doi.org/10.38159/ehass.202561031>
- [21] International Bank for Reconstruction and Development, Artificial Intelligence In Bulgarian School Education. *Digital Innovations in Education*, Brief №6, The World Bank, 2024. <https://documents1.worldbank.org/curated/en/099609411142514872/pdf/IDU-2c378d78-1beb-40e2-889c-a389b89cd321.pdf>, last visited 04.02.2026
- [22] Khanmigo ECHO: Transforming classroom education by training teachers to use AI. The University of New Mexico, 2025, <https://projectecho.unm.edu/story/khanmigo-echo-transforming-classroom-education-by-training-teachers-to-use-ai/>, last visited 05.01.2026
- [23] Mathos AI. Mathos AI Vs. MathGPT.AI: Which AI Math Tutor Is Right for You?, 2025. <https://www.mathgptpro.com/blog/mathos-ai-mathgpt-ai-math-tutor-right>, last visited 04.02.2026
- [24] MEROD, A. Risks from AI use are growing alongside its popularity in schools, 2025, <https://www.k12dive.com/news/risks-from-ai-use-are-growing-alongside-its-popularity-in-schools/802387/>, last visited 05.01.2026
- [25] MEYLANI, R. AI-powered assessments in mathematics education: a systematic review of contemporary research literature. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi* 66:3642–3674, 2025. <https://doi.org/10.53444/deubefd.1527781>
- [26] Microsoft Support. Getting Started with Math Progress, 2025. . <https://support.microsoft.com/en-us/topic/getting-started-with-math-progress-a3e786d8-afcd-43e7-bc45-ef65d6c72bc9#id0edd=educators>, last visited 04.02.2026
- [27] MOHAMMAD, N.M., ET AL. How college students use ChatGPT. *Pedagogical Research*, 10(4), em0250, 2025. <https://doi.org/10.29333/pr/17428>
- [28] NGUYEN, D.T., Q.V. PHAM. The evolving landscape of AI integration in math-

- ematics education: A systematic review of trends (2015-2025). *EURASIA Journal of Mathematics, Science and Technology Education*, 21(10), em2714, 2025. <https://doi.org/10.29333/ejmste/17078>
- [29] NIKIFOROVA-ILIEVA, K., T. GEORGIEV. Understanding the role of AI chatbots in education: a cross-context analysis of awareness, trust, and implementation. 2025 10th International Conference on Energy Efficiency and Agricultural Engineering, Stara Zagora, Bulgaria, 2025. <https://doi.org/10.1109/EEAE65901.2025.11273728>
- [30] PALACIO, A.B. *Unleashing the Power of Wolfram Alpha API with Python and ChatGPT.*, Packt, 2023. <https://www.packtpub.com/en-us/learning/how-to-tutorials/unleashing-the-power-of-wolfram-alpha-api-with-python-and-chatgpt>, last visited 04.02.2026
- [31] Photomath. Math from all angles: Photomath for different learning styles, Build Your Math Mind, 2025. <https://photomath.com/articles/math-from-all-angles-photomath-for-different-learning-styles/>, last visited 04.02.2026
- [32] Research and Markets. AI in Education Market Report 2025, 2025. <https://www.researchandmarkets.com/reports/5896034/ai-in-education-market-report?srsltid=Afmb0oq9jGNebo6m0S3v6D085N6mokSVr0YfYY6wiaIicT97QCHuZqRz>, last visited 04.02.2026
- [33] Rutgers Canvas. Feature Release: Khanmigo Teacher Tools, 2025. <https://canvas.rutgers.edu/2025/04/24/feature-release-khanmigo-teacher-tools/>, last visited 04.02.2026
- [34] SEVILLA, A., CUEVAS-RUIZ, P., RELLO, L. ET AL. Artificial Intelligence in Education: Computer-Assisted Learning and AI-guided Tutors. *Ital Econ J* (2025). <https://doi.org/10.1007/s40797-025-00354-1>
- [35] STEFANOVA, T., S. GEORGIEV. Possibilities for using AI in mathematics education. *Mathematics and Education in Mathematics*, 53, 117-125, 2024. <https://doi.org/10.55630/mem.2024.53.117-125>
- [36] SUTOMO, W.A.B., T. TURMUDI. Integration of artificial intelligence in mathematics learning: systematic literature review. *Supremum Journal of Mathematics Education* 9(1):29-41, 2025. <https://doi.org/10.35706/sjme.v9i1.100>
- [37] TAN, L.Y., ET AL. Artificial intelligence-enabled adaptive learning platforms: A review. *Computers and Education: Artificial Intelligence* 9, 100429, 2025. <https://doi.org/10.1016/j.caeai.2025.100429>
- [38] VASILEVA, G. Mobile app solves math problems in seconds, marica.bg, 2017. <https://uchiteli.bg/news/mobilno-prilozhenie-reshava-zadachi-po-matematika-a-za-sekundi/3579>, last visited 04.02.2026
- [39] WANG, R.E., ET AL. Tutor CoPilot: a human-AI approach for scaling real-time expertise. *arXiv:2410.03017v2 [cs.CL]*, 2025. <https://doi.org/10.48550/arXiv.2410.03017>
- [40] YILDIZ, S.G., S.G. KÖRPEOĞLU. Trends and insights of AI in mathematics education: A bibliometric analysis. *International Electronic Journal of Mathematics Education* 20(3), em0837, 2025. <https://doi.org/10.29333/iejme/16401>