

Effect of Artificial Intelligence on Mathematics Performance: A Meta-Analysis

Vivian Maanu¹, Francis Ohene Boateng², Ernest Larbi³

¹Akrokerri College of Education, Akrokerri, Adansi, Ghana

^{2,3}Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development, Ghana

ARTICLE INFO	ABSTRACT
Published Online: 26 September 2025 Corresponding Author: Vivian Maanu	Background: Artificial intelligence (AI) has emerged as a transformative tool in education, offering personalized feedback, adaptive learning, and data-driven instructional strategies. While individual studies have explored its role in mathematics, findings remain fragmented. This study employs a meta-analysis to systematically evaluate the overall effect of AI applications on students' mathematics performance. Methods: A comprehensive search was conducted across five databases (Google Scholar, Scopus, ERIC, Web of Science, and SCISPACE) for peer-reviewed quantitative studies published between 2017 and 2025. A total of 1,042 articles were identified, and after applying strict inclusion criteria guided by the PRISMA framework, 7 studies were retained. Data extraction followed a standardized protocol, and methodological quality was appraised using the Joanna Briggs Institute Critical Appraisal Checklist. Statistical analyses were performed with JASP, including effect size estimation, heterogeneity testing, and publication bias diagnostics. Results: The pooled effect size was $g = 0.603$ ($p < 0.001$, 95% CI = 0.393–0.813), indicating a moderate-to-large positive impact of AI on mathematics performance. Heterogeneity was substantial ($Q(6) = 52.41$, $I^2 = 90.08\%$), suggesting variability across contexts and intervention types. However, fail-safe N analysis ($N = 520.124$) confirmed robustness against publication bias. Conclusion: Findings demonstrate that AI significantly enhances mathematics learning, supporting constructivist and socio-cultural theories of education where AI scaffolds learner self-efficacy and engagement. Policymakers and educators are encouraged to strategically integrate AI into curricula and teacher training to maximize its transformative potential in mathematics education.
KEYWORDS: Artificial Intelligence; Mathematics Performance; Meta-Analysis	

INTRODUCTION

Mathematics serves as a cornerstone for STEM education, providing students with the logical reasoning and quantitative skills necessary in science, technology, engineering, and mathematics disciplines (Eshaq, 2024). In recent years, artificial intelligence (AI) has entered the educational arena in various forms, such as adaptive learning platforms, intelligent tutoring systems, and AI-driven problem-solving tools, offering personalized pacing, automated feedback, and differentiated practice targeting each learner's strengths and weaknesses. For example, AI-powered adaptive learning research shows that engagement metrics and post-assessment

scores increase significantly when using these tools, likely because they adjust difficulty and guidance to each student's performance level. Despite this promise, empirical evidence remains mixed. Some studies report moderate to substantial improvements, for instance, an AI tutor in Ghana raised mathematics growth significantly (effect size ≈ 0.37) over eight months of twice-weekly sessions via WhatsApp (Henkel et al., 2024); yet other meta-analytical work focusing on elementary school mathematics shows smaller effects, around $d = 0.35$, under random-effects models and varying substantially by learning topic and grade level (Ran et al., 2022). Further complexity

emerges from qualitative differences: different AI types (intelligent tutoring vs. adaptive platforms), intervention durations, and student populations all seem to moderate the effectiveness, making direct comparisons challenging.

Given the inconsistent findings and the absence of a fully up-to-date synthesis covering both primary, secondary, and LMIC (low- and middle-income country) settings, there remains a pressing need for a comprehensive meta-analysis that quantifies AI's overall effect on mathematics performance. Such a study would collate data across diverse contexts, control for moderators like intervention type, duration, and grade level, and provide more reliable effect estimates with subgroup analyses.

RESEARCH OBJECTIVE

1. To assess the effect size of artificial intelligence on mathematics performance.

METHODOLOGY

Design

In this study, a meta-analysis was adopted as the primary research approach to provide a systematic and comprehensive synthesis of the body of literature on the relationship between artificial intelligence and mathematics performance. By utilizing this method, the study was able to gather, integrate, and statistically analyze findings from multiple independent research works, thereby offering a more objective and reliable assessment of the overall effect of artificial intelligence tools and applications on students' achievement in mathematics. The meta-analysis not only enabled the identification of consistent patterns across diverse studies but also helped in quantifying the magnitude of the influence, highlighting whether the application of AI-based interventions has a significant positive, negative, or negligible effect on mathematics learning outcomes. Through this process, the study sought to overcome the limitations of individual studies, such as small sample sizes or contextual biases, and instead present a generalized and evidence-based conclusion regarding the role of artificial intelligence in enhancing mathematics performance.

Eligibility Criteria and Study Selection

In order to guarantee both methodological rigor and scholarly relevance, this meta-analysis applied well-defined and stringent eligibility criteria to determine which studies were suitable for inclusion. Specifically, only peer-reviewed journal articles published between 2017 and 2025 were considered, thereby ensuring that the findings are grounded in contemporary research and meet established standards of academic credibility. To maintain consistency in the type of evidence analyzed, eligible studies were required to employ quantitative research methods and to report clear statistical indicators such as correlation coefficients (r) or standardized effect sizes (β), which provide measurable evidence of relationships and effects. The focus was placed on school and university students as the target population, since these

developmental stages are critical for understanding the role of artificial intelligence in shaping mathematics performance and learning outcomes. Furthermore, only studies published in English were included, and each had to utilize at least one psychometrically validated instrument that measured artificial intelligence applications alongside mathematics performance to ensure reliability and comparability across studies. Conversely, non-quantitative studies, articles not subjected to peer review, and research involving non-student populations were excluded to safeguard the interpretability, validity, and comparability of the synthesized results.

Population of the Study

The target population for this study was defined as published empirical research articles that specifically investigated the impact of artificial intelligence (AI) on mathematics performance. The selection focused on studies published within the period 2017 to 2025, ensuring that only recent and relevant contributions reflecting advancements in AI applications to mathematics education were considered. To comprehensively address the research objectives and maintain consistency with the overall study design, a rigorous search was conducted across multiple reputable academic databases. This extensive search process yielded a total of 1,042 articles, representing a broad and diverse pool of empirical evidence that provided the foundation for subsequent analysis and synthesis.

Sample size and Technique

Following an extensive review process involving 828 published articles, only seven (7) studies met the inclusion requirements, underscoring the rigorous and highly selective nature of the screening procedure. The decision to retain this smaller pool of studies was guided by the need to ensure both methodological rigor and direct relevance to the research focus, particularly the connection between artificial intelligence and mathematics performance. To accomplish this, a purposive sampling strategy was adopted, whereby studies were deliberately chosen according to well-defined criteria, including the quality and robustness of their research design, the clarity of their focus on artificial intelligence applications in mathematics, and the availability of adequate data for extraction and comparison. This meticulous and targeted selection process not only filtered out less reliable or tangential works but also guaranteed that the final set of studies would provide coherent, meaningful, and comparable evidence, thereby strengthening the validity and reliability of the meta-analysis.

Search Strategy

A thorough and systematic literature search was undertaken across five widely recognized multidisciplinary databases that are highly regarded for their coverage of educational and psychological research, namely Google Scholar, Scopus, ERIC, Web of Science, and SCISPACE. The search process was structured around the use of Boolean logic, incorporating carefully selected keywords such as “artificial intelligence”

and “mathematics performance.” To strengthen the search strategy, the terms were refined and adjusted iteratively, ensuring that the final set of studies reflected both breadth and depth of coverage within the research domain. In addition to direct database searching, backward citation tracking (examining reference lists of included studies) and forward citation tracking (identifying more recent works that cited the

included studies) were employed to capture studies that might have otherwise been missed. This dual strategy enhanced both the sensitivity and specificity of the search, thereby ensuring that the review incorporated not only well-established contributions but also newly emerging research situated at the intersection of artificial intelligence and mathematics performance.

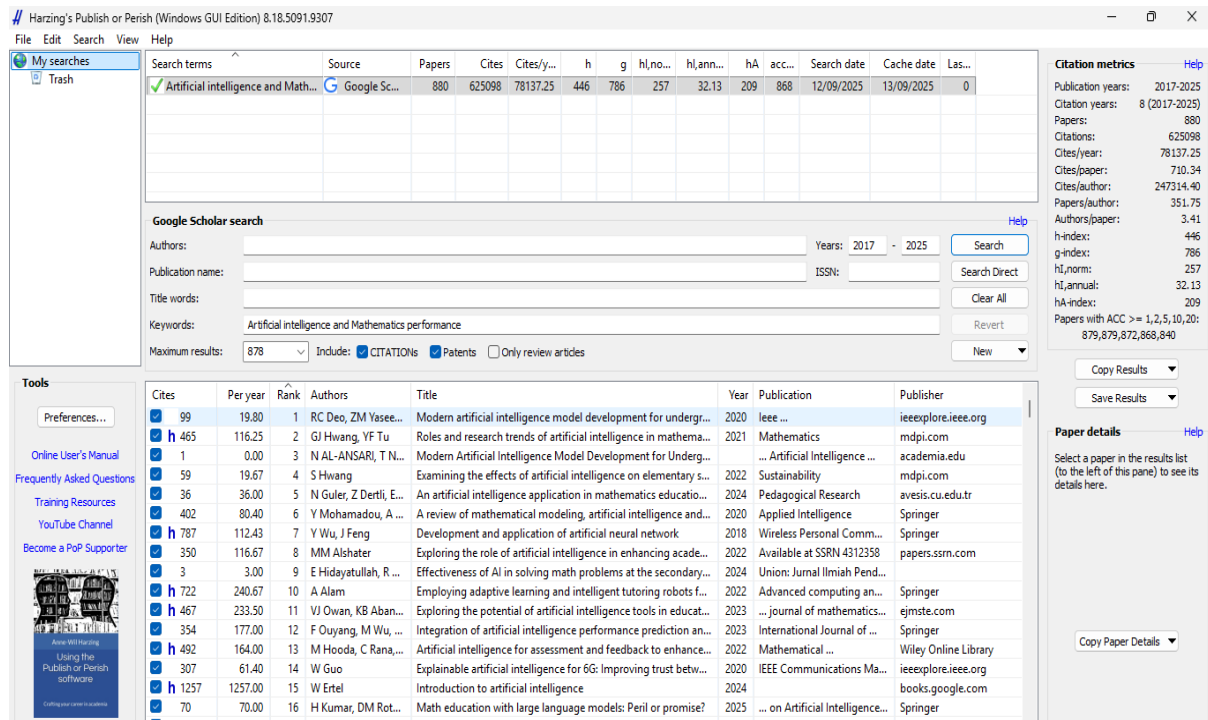


Figure 1. SCISPACE Interface

Study Selection and Data Extraction

The process of selecting studies for inclusion in this meta-analysis was carefully guided by the Preferred Reporting Items for Meta-Analyses (PRISMA) framework, which provided a structured and transparent procedure for identifying, screening, and evaluating relevant research. The initial search results obtained from multiple databases were systematically exported, after which duplicate records were carefully removed with the aid of reference management software to avoid redundancy and bias. The remaining records were subjected to a two-stage screening process: first, titles and abstracts were examined to exclude studies that clearly did not meet the inclusion criteria, and second, the full texts of potentially relevant studies were assessed in detail against the established eligibility requirements. Once the

eligible studies were confirmed, data extraction was conducted using a standardized template designed to maintain accuracy and consistency across studies. The template captured essential information such as the author(s), year of publication, country of study, sample size, and the level of education targeted (elementary, secondary, or tertiary), as well as methodological details including effect size values (β), and the specific measurement instruments used to assess artificial intelligence applications and mathematics performance. This rigorous and systematic procedure ensured that the final dataset was comprehensive, reliable, and suitable for synthesizing effect sizes and study characteristics, thereby providing a strong foundation for the subsequent meta-analytical computations.

“Effect of Artificial Intelligence on Mathematics Performance: A Meta-Analysis”

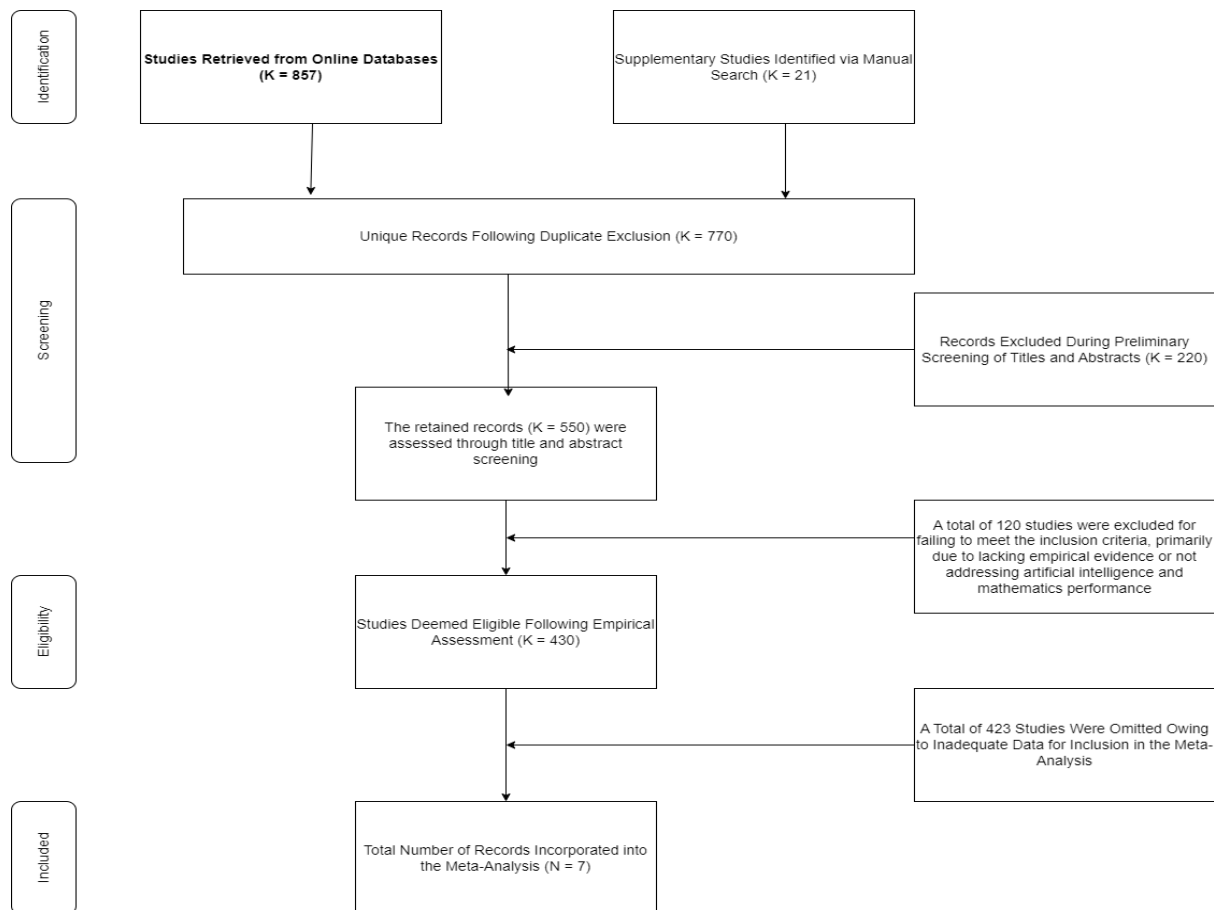


Figure 2. PRISMA-Based Study Inclusion Flowchart

Figure 2 illustrates a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart, which illustrates the process of identifying, screening, and selecting studies for inclusion in a meta-analysis. The process begins with the identification phase, where 857 studies were retrieved from online databases, and an additional 21 supplementary studies were identified through manual search, leading to an initial pool of 878 studies. After removing duplicates, 770 unique records remained. During the screening phase, 220 records were excluded after preliminary title and abstract review, leaving 550 studies for further evaluation. In the eligibility phase, these 550 records were assessed more thoroughly, and 120 were excluded for not meeting the inclusion criteria, mainly due to a lack of empirical evidence or not addressing the relationship between artificial intelligence and mathematics performance. This left 430 studies deemed eligible. However, during the final assessment for inclusion in the meta-analysis, 423 of these were excluded because they contained inadequate data. Ultimately, only 7 studies were included in the final meta-analysis, forming the evidence base for the study's findings. This systematic process ensures that only the most relevant and methodologically sound studies contribute to the analysis.

Coding and Quality Assessment

The coding process for the selected studies was first carried out using Microsoft Excel, where initial data entry, organization, and classification of relevant variables were

performed. After this preliminary stage, the dataset was exported into JASP (version 0.19.3.0), an advanced statistical software, to enable more rigorous quantitative synthesis and analysis. To ensure the credibility and methodological soundness of the included studies, the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Analytical Cross-Sectional Studies was systematically applied. This tool provided a structured framework for evaluating critical aspects such as the appropriateness of the sample, the reliability of measurement instruments, the internal validity of study designs, and the potential risk of bias in reported findings. Based on this appraisal, only studies that achieved a rating of moderate to high quality were retained for the meta-analysis. This careful screening process ensured that the synthesis was built upon strong empirical evidence, thereby minimizing methodological weaknesses and enhancing the reliability, validity, and overall robustness of the conclusions drawn.

Publication Bias Check

One of the most critical methodological considerations in conducting meta-analytical research is the possibility of publication bias, as it has the potential to distort findings and result in an overestimation of the actual effect size. To address this concern, funnel plots were employed, offering a visual means of detecting asymmetry that may signal the presence of bias, particularly when smaller or non-significant studies remain unpublished or when selective reporting is evident. Beyond the visual inspection, a fail-safe N analysis was also

carried out, which serves to estimate the number of missing or unpublished studies required to nullify the observed results. By imputing these hypothetical studies, this technique generates an adjusted pooled effect size, thereby offering a more balanced representation of the true relationship under investigation. The combination of these diagnostic approaches strengthens both the credibility and the transparency of the findings, ensuring that the meta-analysis provides a more reliable contribution to evidence-based practices within the field of education.

Ethical Consideration

This meta-analysis relied solely on data already published in scholarly sources, eliminating the need for direct involvement with human participants and thereby ensuring compliance with established ethical research standards. By using secondary data drawn from peer-reviewed journals and reputable academic publications, the study upholds a strong level of academic integrity, while every source is carefully documented to allow for transparency and traceability. The deliberate and meticulous citation of all materials not only reinforces the reliability of the research process but also respects intellectual property rights, reducing the risk of

plagiarism or misrepresentation of findings. This approach highlights the study’s commitment to ethical responsibility, scholarly rigor, and credibility in contributing to the academic community.

RESEARCH RESULT

Heterogeneity Test

The results of the heterogeneity test provide important insights into the variability of study outcomes regarding the effect of artificial intelligence (AI) on mathematics performance. The Q statistic of 52.410 with 6 degrees of freedom was found to be significant at $p < 0.001$, indicating that the observed variability across studies is not attributable to random error alone. In other words, the studies included in the analysis do not converge toward a single, common effect size; rather, they display significant divergence in how AI influences students’ performance in mathematics. This suggests that contextual or methodological factors may be shaping the extent to which AI-based interventions are effective.

Table 1. Heterogeneity Test Data Summary for the effect of artificial intelligence on mathematics performance.

<i>Q</i>	<i>df</i>	<i>p-value</i>	<i>I²</i>
52.410	6	< 0.001	90.080

The magnitude of this heterogeneity is further confirmed by the I^2 statistic, which reached 90.080%. In meta-analysis, an I^2 value above 75% is generally considered indicative of substantial heterogeneity. Thus, the present findings show that almost all of the variability in effect sizes, approximately 90%, arises from real differences among the included studies, rather than chance. This high level of inconsistency suggests that the impact of AI on mathematics achievement is likely contingent upon diverse moderating variables, such as the type of AI application (e.g., intelligent tutoring systems, adaptive learning platforms, or automated feedback tools), the educational level of participants, or the duration and intensity of the intervention.

The combined effect size estimation for artificial intelligence on mathematics performance.

The findings presented in Table 2 demonstrate that the integration of artificial intelligence (AI) into mathematics

education exerts a substantial influence on students’ performance outcomes. The overall effect size (Hedges’ $g = 0.603$) falls within the moderate-to-large range, underscoring the pedagogical value of AI-based tools and approaches in enhancing mathematical learning. This magnitude of effect suggests that AI does more than provide supplementary support; it can significantly reshape instructional delivery by offering adaptive feedback, personalized learning pathways, and data-driven insights that are often challenging for traditional methods to achieve. The statistical significance ($p < 0.001$) reinforces the robustness of these findings, highlighting AI’s capacity to address long-standing challenges in mathematics learning, such as conceptual misunderstandings, lack of engagement, and difficulties with problem-solving.

Table 2. Presents the results of the combined effect size estimation for artificial intelligence on mathematics performance.

				95% Confidence Interval	
	<i>Effect Size (g)</i>	<i>Std. Error</i>	<i>p-value</i>	<i>Lower Limit</i>	<i>Upper Limit</i>
Overall	0.603	0.086	< 0.001	0.393	0.813

The 95% confidence interval (0.393–0.813) provides further evidence of the consistency and reliability of AI’s effect. Even at the lower bound, the impact remains moderate, confirming

that AI interventions yield tangible benefits across diverse educational contexts. At the upper bound, the large effect points to the transformative potential of AI when

implemented effectively, particularly in environments where it complements teacher instruction and fosters student-centered learning. From a theoretical perspective, these findings align with constructivist and socio-cultural learning theories, where AI functions as a scaffold that enhances learners' agency and self-efficacy in mathematics. Practically, the results suggest that policymakers and educators should prioritize the strategic adoption of AI in curricula, teacher professional development, and classroom practices, as its consistent positive influence can contribute to closing achievement gaps and promoting deeper mathematical understanding.

Figure 3 presents a forest plot, a common visualization in meta-analysis, showing the effect sizes of multiple studies

and a pooled estimate. Each row corresponds to a study, with the black square representing the study's effect size and the horizontal line representing its confidence interval (CI). Larger squares indicate greater weight in the analysis, often based on sample size or precision. For example, Mills (2021) has a relatively narrow CI, suggesting high precision, while Karaman & Aras (2024) show a wider CI, indicating greater uncertainty. The values on the right show the effect size with its 95% CI, while the vertical dashed line at 1.0 serves as a reference point (no effect). All studies included have effect sizes less than 1.0, but their magnitudes vary, suggesting some degree of heterogeneity.

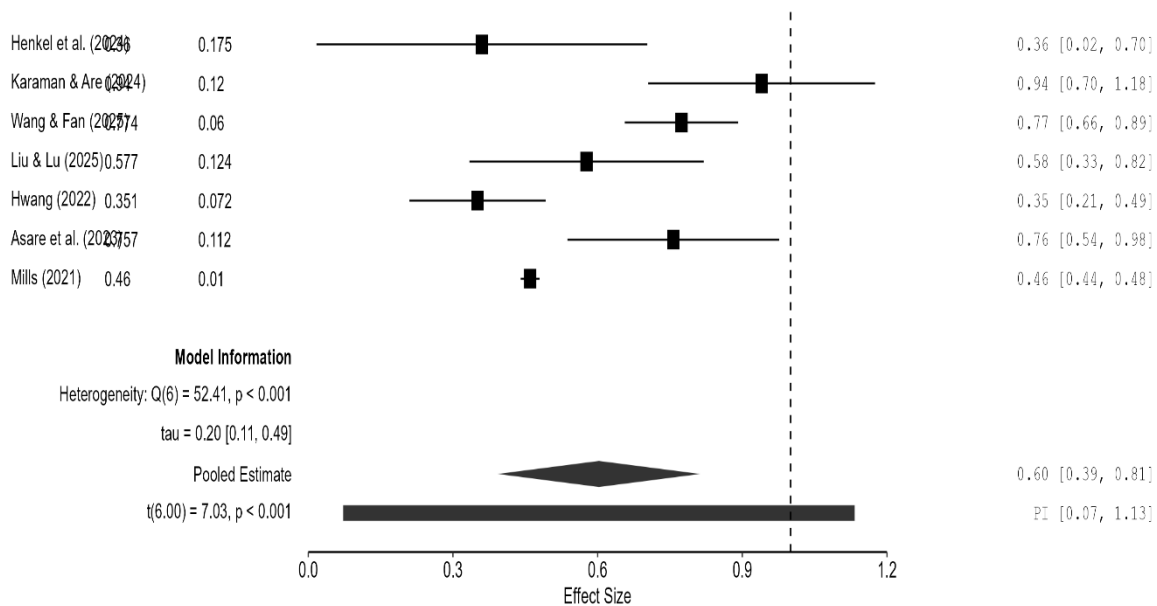


Figure 3. Forest plot showing the overall effect size of artificial intelligence on mathematics performance.

The bottom portion provides the model information. The heterogeneity test ($Q(6) = 52.41, p < 0.001$) shows significant variation between studies, supported by $\tau = 0.20$, indicating moderate between-study variance. The pooled estimate, represented by the diamond, is 0.60 with a 95% CI of [0.39, 0.81], suggesting a moderate overall effect. The prediction interval (PI) [0.07, 1.13] is wider, accounting for the expected

variability if a new study were conducted, reflecting uncertainty about true effects in different contexts. Finally, the plot indicates that while there is a consistent positive effect across studies, the strength varies substantially, warranting cautious interpretation.

Evaluation of Publication Bias

Table 3. File-Safe N for the Effect of Artificial Intelligence on Mathematics Performance.

File Drawer Analysis			
	Fail-safe N	Target Significance	Observed Significance
Rosenthal	520.124	.05	< .001

To assess the potential publication bias, a fail-safe N analysis was conducted, as shown in Table 3, focusing on the effect of the teacher-student relationship on students' mathematics performance. With 13 studies included ($k = 7$), the critical threshold for bias detection was calculated as $5k + 10 = 45$. Impressively, the analysis produced a Rosenthal's Fail-Safe N of 520.124, far surpassing the threshold and indicating that

over 2,000 additional null-result studies would be needed to nullify the observed effects. With a target significance level of $\alpha = 0.05$ and a p -value of less than 0.001, these results suggest that the findings are statistically robust and not influenced by publication bias. This meta-analysis, grounded in methodological precision and objectivity, revealed a pooled effect size of $d = 0.603$ using a random-effect model.

Table 4. Effect of Artificial Intelligence on Mathematics Performance.

Authors (source)	Year	Country	Grade level	AI intervention (type)	Outcome measure	Reported effect (β)	SE (reported)
Mills	2021	USA	High school Algebra 1	ALEKS (ITS)	LEAP math/course score	0.460	0.01
Asare et al.	2023	Ghana	High School	ChatGPT	Mathematics Performance	0.757	0.112
Hwang	2022	Korea	All levels of Education	Artificial Intelligence	Mathematics performance	0.351	0.072
Liu & Lu	2025	China	All levels of Education	ChatGPT	Mathematics Achievement	0.577	0.124
Wang & Fan	2025	China	High Education	ChatGPT	Mathematics Performance	0.774	0.060
Karaman & Are	2024	China	Middle School	ChatGPT	Mathematics Performance	0.940	0.120
Henkel et al.	2024	Ghana	Primary and lower secondary	ChatGPT	Mathematics Performance	0.360	0.175

CONCLUSION

This meta-analysis employed a rigorous and systematic methodology to synthesize recent empirical evidence on the impact of artificial intelligence (AI) on mathematics performance. Through strict eligibility criteria, comprehensive database searches, PRISMA-guided screening, and robust quality and bias checks, the study ensured reliability, transparency, and validity in its findings. By consolidating results across diverse contexts, the analysis not only quantified a moderate-to-large positive effect of AI on mathematics achievement but also highlighted the consistency of this impact despite study heterogeneity. This methodological rigor provides a strong foundation for advancing theory, guiding classroom practice, and informing policy on the integration of AI in mathematics education.

Limitations and Suggestions for Further Studies

One of the major limitations of this meta-analysis is the substantial heterogeneity observed among the included studies ($Q = 52.410$, $p < 0.001$; $I^2 = 90.08\%$). This indicates that the effect of AI on mathematics performance is not uniform but varies significantly across educational contexts, intervention types, and student populations. Such variability reduces the precision of the pooled estimate and makes it difficult to generalize findings to all teaching and learning settings. The diversity of AI tools (e.g., intelligent tutoring systems, adaptive platforms, automated assessment tools) and the differing lengths and intensities of interventions may account for these inconsistencies. Future research should conduct moderator analyses that explicitly examine how factors such as type of AI application, students' educational level, duration of use, and cultural or curricular differences influence mathematics performance outcomes. Longitudinal and experimental studies that isolate these moderators could help clarify “what works, for whom, and under what

conditions,” thereby enhancing the precision of future meta-analytical findings.

Theoretical Implications

The findings of this meta-analysis offer significant contributions to educational theory, particularly within the domains of constructivist and socio-cultural learning frameworks. The moderate-to-large pooled effect size ($g = 0.603$) provides empirical support for the argument that artificial intelligence (AI) functions as a cognitive scaffold, enabling learners to construct knowledge actively while receiving tailored guidance. The results suggest that AI is not merely a technological supplement but rather a transformative mediator that aligns with Vygotsky's concept of the Zone of Proximal Development, where adaptive feedback and individualized instruction accelerate conceptual growth. Moreover, the heterogeneity observed across studies highlights the importance of contextual and moderating factors, suggesting fertile ground for refining theoretical models of technology-enhanced learning by integrating the dynamic interplay between AI tools, learner characteristics, and instructional design.

Educational Implications

From an educational standpoint, the evidence underscores the powerful role of AI in reshaping mathematics instruction and learning outcomes. The consistent positive impact found across diverse studies suggests that AI can address some of the most persistent challenges in mathematics education, such as low student engagement, difficulties with abstract concepts, and limited access to personalized support. AI-based intelligent tutoring systems, adaptive platforms, and automated feedback mechanisms create opportunities for differentiated instruction at scale, ensuring that students across varying levels of ability can benefit from individualized learning trajectories. These results also call

attention to the need for teacher preparation programs and continuous professional development initiatives to equip educators with the knowledge and skills to integrate AI meaningfully into classroom practice, thereby bridging the gap between traditional pedagogy and modern digital innovation.

Practical Implications

The practical significance of these findings extends to policy, curriculum design, and institutional strategies aimed at improving mathematics achievement. Policymakers and educational leaders can leverage the demonstrated effectiveness of AI by embedding its applications into national curricula, supporting schools with the necessary infrastructure, and incentivizing research-driven innovations. At the classroom level, educators can harness AI tools to complement, rather than replace, their instructional expertise, allowing for more time to focus on higher-order problem-solving, collaboration, and creativity. Furthermore, the robust fail-safe N analysis (520.124) indicates that these results are not easily undermined by potential publication bias, giving stakeholders confidence in investing resources into AI-driven educational reforms. Collectively, these practical implications suggest that AI, when implemented strategically, has the potential to narrow achievement gaps, foster equity in mathematics education, and prepare learners for a future where digital literacy and mathematical competence are both indispensable.

REFERENCES

1. Asare, B., Arthur, Y. D., & Boateng, F. O. (2023). Exploring the Impact of ChatGPT on Mathematics Performance: The Influential Role of Student Interest. *Education Science and Management*, 1(3), 158–168. <https://doi.org/10.56578/esm010304>
2. Eshaq, H. A. (2024). The effect of using STEM education on students' mathematics achievement. *Journal of Pedagogical Research*, 8(1), 75–82. <https://doi.org/https://doi.org/10.33902/JPR.202423476>
3. Henkel, O., Horne-Robinson, H., Kozhakhmetova, N., & Lee, A. (2024). Effective and Scalable Math Support: Experimental Evidence on the Impact of an AI-Math Tutor in Ghana. *Communications in Computer and Information Science*, 2150 CCIS, 373–381. https://doi.org/10.1007/978-3-031-64315-6_34
4. Hwang, S. (2022). Examining the Effects of Artificial Intelligence on Elementary Students' Mathematics Achievement: A Meta-Analysis. *Sustainability (Switzerland)*, 14(20). <https://doi.org/10.3390/su142013185>
5. Karaman, M. R., & Are, I. (2024). *Are Lesson Plans Created by ChatGPT More Effective? An Experimental Study To cite this article : Are Lesson*

Plans Created by ChatGPT More Effective? An Experimental Study.

6. Liu, Z., & Lu, Y. (2025). The Impact of ChatGPT on Students' Academic Achievement: A Meta-Analysis. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.70096>
7. Mills, N. J. D. (2021). Heliyon 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>
8. Ran, H., Kim, N. J., & Secada, W. G. (2022). A meta-analysis on the effects of technology's functions and roles on students' mathematics achievement in K-12 classrooms. *Journal of Computer Assisted Learning*, 38(1), 258–284. <https://doi.org/10.1111/jcal.12611>
9. Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. *Humanities and Social Sciences Communications*, 1–21. <https://doi.org/10.1057/s41599-025-04787-y>