www.coast.oaustech.edu.ng
COAST
JOURNAL OF THE SCHOOL OF SCIENCE
Special Issue 1, September, 2025 ISSN: 2714-3716
Coast, J. Sch. Sci. 7 (2): 1351 - 1357

HARNESSING AI-POWERED TOOLS IN MATHEMATICS EDUCATION TO IMPROVE PROBLEM-SOLVING SKILLS FOR ECONOMIC AND TECHNOLOGICAL ADVANCEMENT

[1*]Sangoniyi, S. O., [2]Lukumon, G. and [3]Maharaj, A.

[1]Department of Mathematics, Emmanuel Alayande University of Education, Oyo;

[2]Centre for Artificial Intelligence and Multidisciplinary Innovations Studies, College of Accounting Sciences, University of South Africa.

[3]School of Mathematics, Computer Science and Statistics, University of KwaZulu-Natal, South Africa

Corresponding Author's Email: lukumongafaria@gmail.com

Data and analysis scripts for all the studies, as well as pre-registration documents for all experiments, can be accessed on the Open Science Framework at https://osf.io/8ewgz/

https://dx.doi.org/10.4314/coast.v7i2.11s

Abstract

Can AI tools like ChatGPT genuinely enhance how students solve mathematical problems? This study investigates the impacts of Artificial Intelligence (AI) assistance on undergraduate students' accuracy and speed during a mathematical reasoning task. A total of 92 students from Emmanuel Alayande University of Education, Oyo, were randomly assigned to one of the two groups: one used traditional resources (e.g., textbooks and class notes), while the other used ChatGPT during a test on the *sum to infinity* of a geometric series. Accuracy was scored using a 3-point rubric, and completion time was recorded in seconds. Students in the AI-assisted group demonstrated significantly higher accuracy compared to the traditional group with no significant difference with respect to the time completion. An exploratory analysis revealed a modest negative correlation between accuracy and speed, suggesting that more accurate students tended to complete the task slightly faster. These findings indicate that AI might support mathematical reasoning without promoting shortcuts or diminishing effort. This has important implications for the integration of AI in education, particularly in fostering both correctness and thoughtful engagement in problem-solving.

Keywords. ChatGPT; Mathematical problem solving; AI-assisted learning; Accuracy and speed; Undergraduate mathematics

Introduction

Mathematics holds a pivotal role in daily life, serving as a fundamental tool for tasks ranging from commerce to construction (Polya, 2014). Even individuals without formal education engage in mathematical reasoning during activities such as trading and measurement (Lukumon et al., 2024). Despite its significance, mathematics is often met with apprehension and dislike among students. Observations suggest that while one out of ten students may express a liking for mathematics, the remaining nine often harbor a strong aversion to the subject.

This dichotomy raises a pertinent question: Can artificial intelligence enhance

mathematics learning without undermining the skills it is meant to develop? Artificial Intelligence (AI) tools, such as ChatGPT, Gemini, and Grok, are increasingly being integrated into educational settings. These tools offer instant feedback, personalized learning experiences, and easy access to information, making them valuable assets in modern education (Holmes et al., 2019). Their adoption spans various educational levels, providing support for self-paced and individualized learning. For instance, studies have shown that incorporating AI-powered platforms like Kahoot (Pellas, 2024; Altawalbeh, 2023; Wang, 2020; Plump, 2017; Wang 2015) in assessments can enhance student motivation and reduce mathematics-related anxiety (Lukumon et al., 2024; Wang & Tahir, 2020).

However, the integration of AI in education presents a double-edged sword. While AI can augment learning, there is a growing concern about students becoming overly reliant on these technologies, potentially at the expense of developing critical thinking and problem-solving skills. This phenomenon, known as cognitive offloading, suggests that excessive dependence on AI may lead to diminished cognitive engagement and creativity (Risko & Gilbert, 2016). For example, a 2023 article in The Guardian highlighted concerns about AI tools like ChatGPT reducing students' ability to think independently when overused (APO Group, 2023).

In the context of Nigerian education, mathematics is a compulsory subject, essential for admission into tertiary institutions alongside English (JAMB, 2023). Mathematics education fosters precision, logical reasoning, and the ability to navigate abstract concepts, serving as a cornerstone for STEM fields and economic development (National Policy on Education, 2013).

Central to mathematical proficiency are the dual aspects of accuracy and speed (Heitz, 2014). Accuracy ensures the correct application of formulas and methods, while speed reflects efficiency in problem-solving. In real-world and educational contexts, speed is also a highly valued dimension. Whether in high-stakes standardized exams (e.g., WAEC, NECO, UTME) or competitive academic contests like Nigeria's Cowbellpedia Mathematics Competition, students are often expected to demonstrate not only correctness but also efficiency i.e. solving problems swiftly under time constraints. The ability to recall concepts quickly, recognize patterns, and work through procedures efficiently is considered a sign of fluency and mastery. The interplay between these aspects is encapsulated in the speed-accuracy tradeoff, a concept suggesting that increasing speed may compromise accuracy and vice versa (Wickelgren, 1977). The advent of AI tools like ChatGPT has the potential to alter this dynamic, possibly enabling students to achieve both speed and accuracy simultaneously (Brynjolfsson & McAfee, 2017).

Despite the growing adoption of AI-powered tools in global education, empirical research evaluating their impact on mathematics learning especially in African contexts like Nigeria remains limited. In particular, there is little evidence on how tools like ChatGPT influence core cognitive outcomes such as accuracy and efficiency in mathematical problem-solving. This study addresses this gap by experimentally testing the effect of AI assistance on undergraduate students' mathematical performance.

Accordingly, the present study asks: *Can access to an AI-powered tool like ChatGPT improve students' mathematical accuracy and efficiency compared to traditional learning resources?* To answer this, we use a randomized controlled design to compare

students' performance across two key outcomes accuracy and completion time during a structured mathematics problem-solving task. We hypothesize that students in the AI-assisted group will perform more accurately than those using traditional materials.

The current studies

Our study aimed to provide empirical evidence on how AI-powered tools *specifically* ChatGPT influence students' mathematical problem-solving in terms of speed and accuracy. This research investigates whether access to an AI-powered assistant could enhance students' performance compared to those relying solely on traditional resources such as textbooks and class notes. This research builds on ongoing efforts to understand how technological innovations are transforming mathematics learning, particularly in contexts like Nigeria where mathematics is a compulsory subject yet frequently disliked by students.

Participants were randomly assigned to one of two experimental conditions: Traditional and AI-assisted. In Traditional, students were allowed to consult class notes, textbooks, and personal materials during a mathematics test and for the AI-Assisted, students were allowed to use ChatGPT. To assess the impact of AI on mathematical reasoning, we focused on two core outcomes: accuracy (measured by the number of correct answers) and time of completion (measured in seconds). All students were given the same question involving the sum to infinity of a geometric series.

Below are the outlines of the methods and results of this investigation.

Methods

Open Science. Sample size calculations, hypotheses, and the analysis plan were pre-registered. We report all experimental conditions, measured variables, and statistical analyses conducted. All data, materials, and annotated R scripts are available at the Open Science Framework: https://osf.io/8ewgz/

Participants. Participants were undergraduate students from Emmanuel Alayande University of Education, Oyo. A priori power analysis (with 85% power, α = 0.05, and an assumed small effect size of Cohen's d = 0.30) indicated a required sample size of approximately 240 participants. Due to recruitment constraints, we obtained complete data from 92 participants. We proceeded with analysis based on this available sample.

Materials. Students were asked the following question: *"You receive ₦50,000 now, then ₦25,000 next year, ₦12,500 the following year, and so on—each payment is half of the previous one, continuing indefinitely. What is the total value of this infinite series of payments?"*

This task assessed both conceptual understanding and the ability to apply the formula for the sum of an infinite geometric series. By conceptual understanding, we refer to a learner's grasp of the underlying principles behind mathematical operations (e.g., understanding why a series converges)

Procedure. Before the experiment, students were randomly assigned to one of the two experimental groups. All participants received the same open-book mathematical task. The AI group was permitted to use ChatGPT to assist with the problem, while the control group could only use non-AI materials such as textbooks or handwritten notes. To maintain anonymity and reduce performance bias, students were instructed not to include their names. Time of completion was recorded in seconds for each participant. Their written responses were later graded using a standardized 3-point rubric assessing procedural accuracy and correct final answer.

Below headings contain more information about the measure.

Measures. We assessed two primary outcomes in this study:

Accuracy. Each student's solution was graded on a 3-point scale: one point for correctly stating the formula for the sum to infinity of a geometric series; one point for accurately identifying the relevant parameters (first term and common ratio); one point for computing the correct final answer. This scoring rubric allowed us to capture both procedural understanding and computational correctness. Procedural understanding reflects the student's ability to apply appropriate steps to solve problems, while computational correctness indicates whether those steps yield valid solutions.

Speed. Time to completion was measured in seconds, starting from the moment the task was administered to the moment the student submitted their response. This measure served as an index of students' efficiency in solving the problem under open-book conditions.

Results.

An independent-samples t-test was conducted to compare performance between students using AI assistance and those relying on traditional resources.

Accuracy. There was strong evidence that students in the AI-assisted group achieved significantly higher accuracy scores than those in the traditional group, $t(43.77) = 4.54$, $p < .001$, 95% CI [0.59, 1.53]. The AI group (M = 2.91, SD = 0.96) outperformed the traditional group (M = 1.85, SD = 1.13), suggesting that access to AI tools like ChatGPT enhanced students' ability to solve the sum-to-infinity problem correctly. This result is indicated in Figure 1 below.

Speed. There was no statistically significant difference in completion time between the groups, $t(89.82) = -0.91$, $p = .36$, 95% CI [-59.49, 21.98]. On average, the AI group (M = 372.81 seconds, SD = 106.52) completed the task slightly faster than the traditional group (M = 391.56 seconds, SD = 122.13), but this difference was not statistically meaningful. Figure 1 below shows this result.

Exploratory Analysis. Although not pre-registered, we conducted an exploratory correlation analysis to examine the relationship between speed and accuracy across all participants. A modest negative correlation was found, $r = -0.25$, $t = -0.91$, $p = .36$, indicating that students who were more accurate tended to complete the task slightly faster, though this association is weak and not statistically significant.

General Discussion

This study provides preliminary evidence that access to AI-powered tools such as ChatGPT can improve mathematical problem-solving accuracy among undergraduate students. Participants in the AI-assisted group outperformed their peers who relied solely on textbooks and class notes when solving a geometric series problem. These findings align with broader literature on AI-enhanced learning particularly studies showing that platforms like Kahoot can boost student motivation and reduce math-related anxiety by providing engaging, interactive, and responsive learning environments (Pellas, 2024; Altawalbeh, 2023; Wang, 2020; Plump, 2017; Wang & Tahir, 2020; Lukumon et al., 2024).

Interestingly, the anticipated advantage in speed was not observed. Although AI-assisted students completed the task slightly faster on average, the difference was not statistically significant. This suggests that students may be using AI not to rush through problems, but to enhance correctness potentially by reflecting more carefully on their responses. Such behavior is encouraging, as it indicates thoughtful engagement with the task rather than passive or superficial use of the technology.

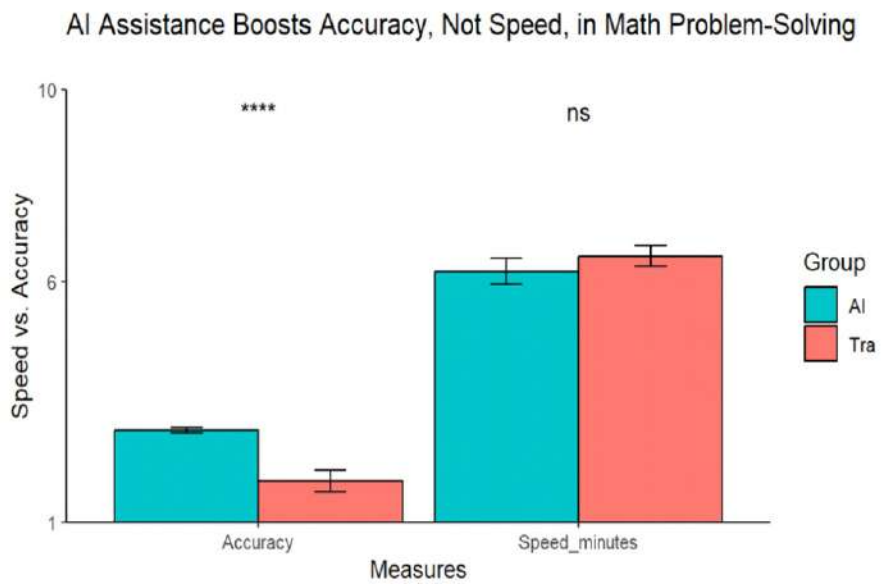

Figure 1. Accuracy and speed by experimental condition. Left panel: Mean accuracy scores (maximum = 3 points) for students using ChatGPT (green) versus traditional resources (red). Right panel: Mean task completion time (in minutes) for both groups. Students with access to ChatGPT demonstrated significantly higher accuracy, but no significant difference in completion time was observed.

While the current study did not directly assess students' reasoning processes, the higher accuracy in the AI group may reflect better execution of solution steps and clearer reasoning. These patterns align with what some frameworks describe as procedural fluency or correctness. It is also possible that access to AI responses prompted students to revisit or clarify conceptual ideas—though this remains speculative in the absence of qualitative or process data. Future research should directly investigate these dimensions by observing how students interact with AI, how they interpret its output, and whether their understanding improves as a result.

Finally, our exploratory analysis revealed a modest negative correlation between speed and accuracy: students who were more accurate tended to complete the task more quickly. While not a strong effect, it hints at variability in student strategies some may focus on precision, others on efficiency. Understanding this trade-off, and whether AI shifts how students allocate cognitive effort, is an important direction for future research.

Implications for curriculum designers and educators

These findings hold meaningful implications for curriculum designers and classroom teachers. The observed accuracy gains suggest that AI tools like ChatGPT can be thoughtfully integrated into mathematics instruction to enhance conceptual understanding and procedural accuracy. Rather than replacing traditional instruction, AI can complement existing methods particularly during practice and revision phases by offering individualized guidance and instant feedback. Curriculum designers should consider developing structured activities that incorporate AI tools in ways that promote active engagement, critical thinking, and responsible use. Similarly, teachers need targeted training on how to guide students in using AI ethically and

effectively, ensuring that such tools support deep learning rather than passive reliance.

Limitations of the study

This study has several limitations that should be acknowledged. First, the sample size was modest (n = 92) and drawn from a single institution, which limits the generalizability of the findings. While the sample allowed us to detect significant differences in accuracy, it may not have provided sufficient statistical power to detect smaller effects, such as the modest difference observed in task completion time. In other words, the absence of a significant speed effect might partly reflect limited power rather than the absence of a true effect. Second, the study focused on a single mathematical concept, sum to infinity which restricts the scope of inference. Future studies should include a broader range of mathematical tasks to evaluate whether the observed benefits generalize across topics.

Third, we did not assess students' subjective experiences (e.g., enjoyment, frustration, perceived usefulness), which could offer valuable insight into the motivational and emotional aspects of AI-assisted learning. Finally, the AI-assisted interaction was treated as a "black box". We did not monitor how students engaged with ChatGPT, what types of prompts they used, or how they incorporated AI responses into their reasoning. This limits our understanding of the cognitive processes involved and calls for future research to include qualitative or process-tracing components.

Suggestion for future studies

To build on the current findings and address the study's limitations, future research should aim to recruit larger and more diverse samples to improve statistical power and enhance the generalizability of results across different educational settings. Expanding the range of mathematical topics beyond geometric series will also help determine whether AI tools offer consistent benefits across varying domains and levels of complexity.

In addition to cognitive performance outcomes, incorporating affective measures such as students' enjoyment, motivation, perceived usefulness, and frustration will provide a more holistic understanding of how AI tools influence learning experiences. This is especially important given concerns about student engagement and overreliance on technology.

Moreover, future studies should move beyond treating AI interaction as a "black box" by integrating qualitative or process-oriented analyses. Capturing how students craft prompts, interpret AI responses, and incorporate these into their reasoning can yield valuable insights into the depth and nature of their learning. Such analyses would help distinguish between surface-level assistance and genuine cognitive support.

Finally, it would be beneficial to systematically vary task difficulty and formally examine speed-accuracy tradeoffs across conditions. This could uncover strategic differences in how students allocate time and effort when supported by AI tools, offering a deeper understanding of both the affordances and boundaries of AI-enhanced learning.

Conclusion

This study suggests that AI-powered tools like ChatGPT can enhance undergraduate students' accuracy in mathematical problem-solving without significantly altering the time required to complete tasks. Rather than serving as a shortcut, AI appears to function as a scaffold that supports students' reasoning and application of mathematical procedures. Although no significant speed advantage was observed, the improvement in accuracy is particularly encouraging in

contexts where student performance or confidence in mathematics is often low. The modest negative correlation between speed and accuracy also points to a potential tradeoff in students' strategy use, an area that merits further investigation in future studies.

References

Altawalbeh, K. (2023). Game-based learning: The impact of Kahoot on a higher education online classroom. *J. Educ. Technol. Instr.*, 2(1), 30–49.

APO Group. (2023, May 17). ChatGPT in schools: A new tool or a threat to critical thinking? *The Guardian*. Retrieved from https://guardian.ng/apo-press-releases/will-chat-gpt-impact-critical-thinking-skills-for-the-youth/

Brynjolfsson, E., & McAfee, A. (2017). The business of artificial intelligence. *Harv. Bus. Rev.*, 7(1), 1–2.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Front. Neurosci.*, 8, 150.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.

JAMB (Joint Admissions and Matriculation Board). (2023). *Admission Guidelines for Tertiary Institutions in Nigeria*. Retrieved from https://www.jamb.gov.ng

Lukumon, G., & Maharaj, A. (2024). From phobia to fun: Enhancing mathematics engagement with AI-powered tools. *OSF Preprint*. https://doi.org/10.31219/osf.io/jkt2p

National Policy on Education. (2013). *Federal Republic of Nigeria: Federal Ministry of Education*.

Pellas, N. (2024). Effects of Kahoot! on K-12 students' mathematics achievement and multi-screen addiction. *Multimodal Technol. Interact.*, 8(9), 81.

Plump, C. M., & LaRosa, J. (2017). Using Kahoot! in the classroom to create engagement and active learning: A game-based technology solution for eLearning novices. *Manag. Teach. Rev.*, 2(2), 151–158.

Polya, G. (2014). *How to Solve It: A New Aspect of Mathematical Method*. Princeton Univ. Press.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends Cogn. Sci.*, 20(9), 676–688.

Wang, A. I. (2015). The wear out effect of a game-based student response system. *Comput. Educ.*, 82, 217–227.

Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot! for learning – A literature review. *Comput. Educ.*, 149, 103818.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.*, 41