Croatian Journal of Education
Vol.28; No.3/2026, pages: 797-832
Original scientific paper
Paper submitted: 14th October 2025
Paper accepted: 4th April 2026
https://doi.org/10.15516/cje.v28i3.42192

Integrating Artificial Intelligence into College Mathematics Curricula in China

Tingting Feng
Qiongtai Normal University, College of Science

Abstract

This study analyzes the use of Google's Socratic application as a support tool for teaching mathematics at higher educational institutions in China. The study involved 400 students from four universities in Eastern China, divided into two groups: an experimental group and a control group, each comprising 200 students with comparable demographic characteristics. Over a 16-week period, students in the experimental group used Socratic for self-study and homework completion, while the control group continued using traditional learning methods. Student performance was evaluated using standardized examinations. Data on student engagement with the application were also collected, including usage frequency and feature utilization patterns. Additionally, metacognitive skills were assessed using a specialized questionnaire. The results showed that students in the experimental group significantly improved their mathematics performance compared to the control group, where no changes were observed. Statistical analysis confirmed that using the application to support the learning process positively influenced students' academic outcomes. Analysis of application usage revealed that students actively utilized various features, including step-by-step hints, explanatory support, and solution verification tools. Overall, the study's findings highlight the potential of Socratic as a tool for university-level mathematics instruction in contexts where opportunities for personalized feedback are limited.

Keywords: *educational technology; mathematical knowledge; metacognitive regulation; personalized learning; problem-solving strategies*

Introduction

Artificial intelligence (AI) refers to computer systems capable of performing tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and

© Author(s) 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). https://creativecommons.org/licenses/by-nc/4.0/

adaptation (Abar et al., 2024; Drijvers & Sinclair, 2024; Gadanidis, 2017). This paper explores AI as an educational tool that leverages machine learning and natural language processing models to offer tailored and interactive assistance to students (Engelbrecht & Borba, 2024). Specifically, AI is integrated into the learning process to improve students' academic performance, particularly in mathematics, through personalized solutions, step-by-step explanations, and instant feedback (Kissi et al., 2019).

In rapidly evolving higher education, AI integration into academic curricula represents a transformative shift in the teaching approaches for fundamental subjects such as mathematics (Engelbrecht & Borba, 2024). Colleges and universities worldwide strive to enhance learning outcomes and leverage technological advancements. In this context, AI has emerged as a potent tool for creating a more engaging and personalized learning environment (Drijvers & Sinclair, 2024; Knox, 2020).

The impetus for integrating AI into mathematics arises from several converging circumstances. The complexity of the discipline and the diversity of learning styles necessitate more personalized and adaptive strategies (Shin, 2022). AI-based tools and platforms can provide unique learning experiences by catering to individual students' needs and offering real-time feedback, which facilitates comprehension and retention (Hwang & Tu, 2021; Tay et al., 2024). Furthermore, in larger classes, it becomes challenging for instructors to accommodate the educational needs of each student (Xie & Wang, 2024).

Neuropsychological factors often present significant obstacles in the perception of mathematics. Cognitive load theory explains the burden mathematical abstraction places on working memory (Matzakos et al., 2023; Valjarević et al., 2025). In Chinese universities, the instructor-to-student ratio often exceeds 1:60, which limits opportunities for individualized support (Jianzheng & Xuwei, 2023; Knox, 2020). Mathematics instruction requires personalized cognitive support yet operates within a mass-education model, thereby necessitating technological mediation (Crompton & Song, 2021). Adaptive systems can address these challenges by tailoring instruction to individual performance levels and visualizing abstract concepts to make them more accessible (Fowler, 2023; Hwang, 2022).

In the Chinese higher education context, these challenges are particularly visible due to large cohort sizes, strong curricular expectations in mathematics, and limited opportunities for individualized feedback during classroom instruction. Recent studies indicate that Chinese universities increasingly explore digital and AI-supported learning environments as a way to complement traditional instruction and support students' independent engagement with complex mathematical material (Huang, 2021; Jianzheng & Xuwei, 2023; Knox, 2020). Consequently, AI-based tools are viewed not only as technological innovations but also as practical mechanisms for addressing structural limitations of mass higher education.

Traditional representations of mathematical concepts are static; however, computational mediation can make them explorable. Digital tools and adaptive algorithms transform

mathematical objects into constructed realities emerging through interactive processes (Betteridge et al., 2022; Birgin et al., 2015). AI-based approaches can reduce extraneous cognitive load through adaptive sequencing and optimized representations, thereby supporting conceptual integration (Romney et al., 2019; Tanveer et al., 2020). The integration of mathematical cognition theory through technological implementation provides a theoretical foundation for innovations in educational psychology (Maphosa & Maphosa, 2021).

Despite the rapid adoption of AI technologies in education, it remains unclear how these tools impact student performance in mathematics. Although prior research indicates that AI tools may enhance problem-solving skills and engagement (Mohamed et al., 2022; Richard et al., 2022; Tashtoush et al., 2024), relatively few studies have rigorously compared control and experimental groups in controlled higher education settings (Drijvers & Sinclair, 2024). Existing studies mainly focus on general AI-supported learning environments or theoretical discussions of intelligent tutoring systems, while empirical investigations of specific AI applications in university mathematics remain limited (Gao, 2020; Huang, 2021; Wu, 2021). The present study addresses this gap by examining the pedagogical effectiveness of the Socratic AI application in authentic university classrooms in China and by providing empirical evidence regarding its influence on students' mathematical achievement and engagement. There is also limited empirical evidence regarding the effectiveness of AI-based applications in university-level mathematics education (Gao, 2020; Huang, 2021).

Literature review

This study is grounded in Self-Determination Theory (SDT), which explains learning motivation through the satisfaction of three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2000; 2020). In educational contexts, these needs are supported when learners regulate their learning processes, receive feedback that strengthens their sense of competence, and interact within supportive environments. AI-based educational technologies can contribute to these conditions by providing personalized explanations, adaptive feedback, and opportunities for self-paced problem solving (Hwang & Tu, 2021; Tashtoush et al., 2024). In mathematics education, where students require step-by-step guidance and repeated practice, such features may support engagement and conceptual understanding. Therefore, SDT provides a relevant framework for examining how the Socratic application may be associated with changes in students' motivation and academic performance.

Recent studies increasingly examine AI-based learning tools in relation to cognitive and metacognitive processes. For example, Rızvı (2023) links AI-supported instruction with the development of self-regulation, while Luo (2024) shows that adaptive feedback systems can support learners' autonomy and task mastery. Panit (2025) further demonstrates that interaction with AI systems may be associated with the development of metacognitive awareness.

The use of AI tools in pedagogy has been discussed in relation to changes in how students interact with learning resources. Research on instrumental genesis suggests that digital tools can influence how learners internalize knowledge (Broutın, 2024; Lagrange & Richard, 2022). At the same time, studies on semiotic mediation show that meaning-making processes remain central, although they increasingly involve interaction with digital systems (Batsaikhan & Correia, 2024; Gupta et al., 2025). Empirical evidence indicates that AI-supported explanations and examples can facilitate the acquisition of mathematical discourse (Möller et al., 2024; Pardos & Bhandari, 2023).

Teachers' roles have also evolved in AI-supported environments, shifting from knowledge transmission toward facilitation (Hall Jr., 2020). DeFalco and Sinatra (2019) argue that interaction with AI systems may redistribute certain instructional functions, particularly those related to feedback and guidance.

Research identifies several models of AI integration in education. Supplementary models treat AI as an additional resource, while integrated models combine AI and teacher-led instruction (Hampton & Wang, 2019). Hybrid systems reorganize learning processes based on performance data. Evidence suggests that the effectiveness of these approaches depends on how well technological tools are aligned with instructional design (Luo, 2024).

Within AI in education, mathematics is frequently identified as a domain where AI-supported tools may be particularly effective due to the structured nature of problem solving and the need for step-by-step feedback.

A review of studies from 2018 to 2023 indicates that AI-supported learning can improve mathematics outcomes in specific contexts, although long-term effects remain unclear (Batsaikhan & Correia, 2024). For instance, Möller et al. (2024) report improved learning progress in AI-supported environments, while Pardos and Bhandari (2023) show that AI-generated hints enhance algebra problem solving. At the same time, these studies emphasize that AI is most effective when used as a supplementary tool rather than a replacement for teachers (Gupta et al., 2025; Seo et al., 2025).

Among specific applications, the Socratic app represents a mobile AI-based learning assistant that provides step-by-step explanations and links to additional resources. Yanuarto et al. (2021) show that Socratic supports conceptual understanding through visual recognition and guided explanations. However, existing studies primarily focus on its technical features or homework support, while evidence from controlled university-level implementations remains limited.

AI has also been applied to instructional design and learning analytics. Li and Zaki (2024) demonstrate that AI can identify effective teaching strategies based on large datasets, while Aljamali and Jabbar (2020) show that AI can support error correction by identifying knowledge gaps. AI-based simulations have been used to model complex mathematical problems, providing interactive representations that support understanding (Jaboob et al., 2024; Rawat & Mishra, 2022). Such simulations may help students engage with abstract concepts more effectively (Orhani, 2021).

The literature also highlights challenges associated with AI implementation. Privacy concerns arise due to data collection (Rane, 2023; Tomar & Soshal, 2023), and AI systems may generate incorrect explanations without supervision (Melchor et al., 2023). Consequently, researchers emphasize the need for validated and reliable AI tools (Tashtoush et al., 2024; Wang, 2023). Despite these limitations, AI-supported systems may contribute to independent learning and problem-solving by providing structured explanations and immediate feedback (Hwang & Tu, 2021).

Despite these developments, relatively few empirical studies examine the pedagogical impact of specific AI applications in authentic university mathematics settings. Therefore, the present study contributes to the literature by investigating the use of the Socratic application in Chinese universities and providing empirical evidence on its association with students' academic performance and engagement.

Problem statement

This study pursues two principal objectives: first, to evaluate the pedagogical effectiveness of integrating algorithms – particularly Socratic artificial intelligence – into mathematics curricula; and second, to establish an empirical basis for the technological transformation of quantitative education.

Research tasks: 1) to quantify differences in learning outcomes between AI-assisted instruction and traditional pedagogical approaches; 2) to identify the taxonomy of interactions emerging during algorithmic engagement by analyzing usage patterns across procedural, representational, and verificational dimensions; 3) to document cognitive changes induced by technological mediation, particularly in problem-solving topologies and metacognitive regulation; 4) to examine the domain-specific characteristics of mathematics regarding susceptibility to algorithmic enhancement.

Current educational systems encounter persistent challenges, such as fragmented conceptual understanding, sustained mathematical anxiety, and limitations in personalization within mass education. These issues necessitate an investigation of technological augmentation beyond conventional learning paradigms.

Methodology

Methodological framework

The Socratic application (Figure 1) was chosen due to its broad capabilities in providing real-time educational support through natural language processing and machine learning. The tool supports textual and visual input, offering detailed explanations, problem-solving steps, and additional educational resources (Google LLC, 2020).

With Socratic's accessible and user-friendly design, students can receive immediate assistance by scanning their educational assignments with their mobile devices. This feature is particularly beneficial in large classrooms where teacher attention must be distributed among many students. Additionally, the application's ability to deconstruct complex mathematical problems into steps makes it a valuable tool for students with

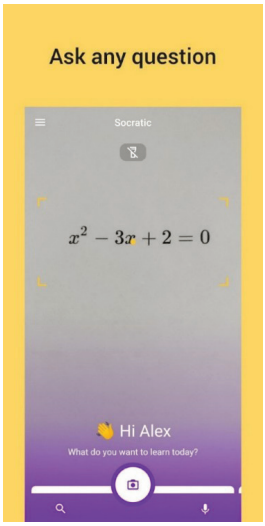


Figure 1. Student interaction with the Socratic application

varying proficiency levels. Socratic is primarily designed for practicing procedural mathematical tasks. Its Explainers feature helps deepen conceptual knowledge by providing explanatory material that supports students in solving such problems. For applied problems, however, the application only offers links to external resources related to real-world contexts that may interest the student, assisting with explanations and guiding further exploration (Martinho & Martins, 2025; Yanuarto et al., 2021).

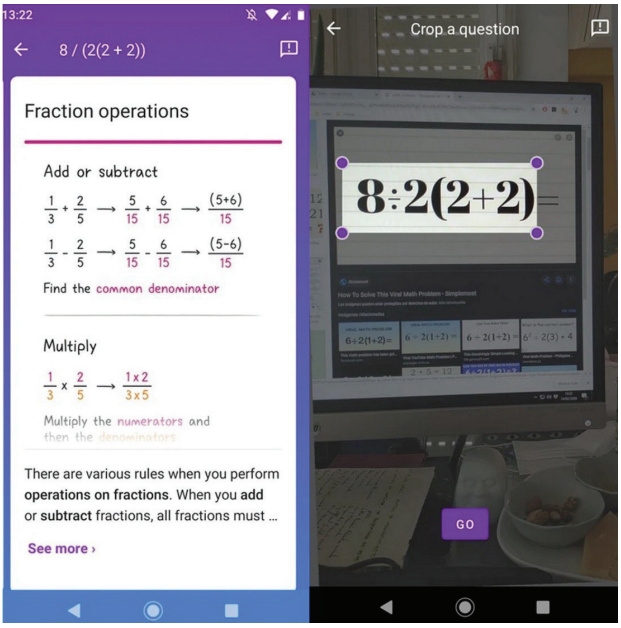


Figure 2. Interface of the Socratic application by Google

The application accepts input in text, audio (user voice), and image formats. Students can photograph a task, including handwritten text; the AI then recognizes the input, provides explanations and solutions, and links to related educational resources. It also checks whether the student understands the task content, associated theory, and the correctness of the implementation. However, the application does solve the task for the student; it supports rather than replaces the learner's own efforts (Figure 2). Feedback is delivered through infographics, visualizations of mathematical formulas and concepts, text explanations, and meaningful dialogue focused on the student's questions and responses to tasks.

Study design

The study employed a controlled-group pretest-posttest design with quantitative analysis to investigate the effectiveness of Google's Socratic AI-based educational system. The study was conducted in several sequential phases: 1) the development and validation of instruments through pilot testing (n = 42); 2) participant allocation via stratified randomization procedures; 3) baseline assessment (mathematical tests covering both procedural and conceptual skills); 4) differentiated intervention (AI-assisted instruction in the experimental group and traditional instruction in the control group); 5) systematic monitoring of interaction patterns using digital analytics (experimental group only); 6) post-intervention assessment; 7) statistical comparison of outcomes; 8) behavioral data extraction. The intervention spanned 16 academic weeks (one semester), with continuous AI system accessibility maintained throughout (mean participant interactions: 4.3±1.2 per week).

Stratified randomization was performed with respect to gender, year of study, and baseline mathematical proficiency. Participants were first allocated into strata based on these characteristics, after which random assignment to the experimental and control groups was carried out within each stratum using computer-generated random numbers. This approach ensured group comparability and reduced the risk of selection bias.

Participants

Participants were selected through stratified random sampling from mathematics departments across four universities in Eastern China (distribution by province: Jiangsu = 2, Zhejiang = 1, Shanghai = 1). The total sample comprised 400 participants (experimental = 200; control = 200) with matched demographic parameters: gender distribution (experimental: 106 female and 94 male students; control: 102 female and 98 male students). Gender composition was reported to demonstrate the demographic comparability of the experimental and control groups and to minimize potential gender-related bias in the interpretation of learning outcomes.

Year of study (second year = 58%; third year = 42%). In absolute numbers, this corresponded to 232 second-year students and 168 third-year students. Only second- and

third-year students were included because by this stage, students have already completed introductory university mathematics courses and possess sufficient disciplinary background to engage with more advanced mathematical problem-solving tasks.

Mathematical knowledge level (mean entrance examination scores: experimental = 81.3 ± 6.2; control = 80.9 ± 6.5). The selected institutions represented diverse geographic and administrative contexts, including both cities and towns. Sample size was determined via a priori power analysis ($\alpha = 0.05$; $\beta = 0.80$; anticipated effect size $d = 0.40$), indicating a minimum requirement of 156 participants per condition.

Intervention

Students primarily used Socratic outside the classroom for independent learning, reviewing material, seeking additional explanations, and receiving individualized support tailored to their understanding. They also relied on the application to complete homework assignments.

The use of the Socratic application was standardized: students were instructed to engage with the application 3–5 times per week while completing independent assignments. The actual frequency of interaction averaged 4.3 times per week (SD = 1.2). The primary component of the intervention was implemented outside of class (independent study).

The application was also used in the classroom, particularly when the teacher recommended it or lacked sufficient time to interact individually with each student. While the teacher presented material on the board and provided explanations, students could point their device's camera at the board, allowing the application to offer immediate access to detailed explanations, examples, and exercises. The teacher could either assign their own exercises or suggest those generated by the application. Students were encouraged to direct questions to the teacher. However, if time constraints prevented the teacher from answering everyone, or if a student was unsatisfied with or unclear about an explanation, they could turn to the application for further clarification. During subsequent practical tasks, the teacher verified whether students had correctly understood and internalized the material with the application's support. This approach enabled full coverage of the study group, enhanced adaptability, and allowed for an individualized learning pace while maintaining alignment with the university curriculum.

During the intervention, the roles of the instructor and the AI tool were delineated: the instructor was responsible for explaining the material and monitoring comprehension, whereas the application was used for individualized support (step-by-step explanations, solution verification, and revision). To ensure consistency in the implementation of the intervention, instructors received prior briefing with standardized guidelines stipulating that the application be used exclusively as a supplementary tool, without replacing the primary delivery of instructional content.

To safeguard the integrity of the experiment, students were asked not to use other similar educational applications on their devices during the study. Participants

consented to share their device activity logs, thereby confirming that they used only the specified application throughout the semester. Students retained the right to withdraw from this practice at any time if it interfered with their educational needs, but no participant chose to do so.

Adherence to the intervention protocol was monitored through analysis of application activity logs and instructor observations. On average, students in the experimental group engaged with the application 4.3 times per week (SD = 1.2). Instructors additionally supervised application use during practical sessions. No substantial deviations from the planned intervention format were identified.

Data collection tools

Data for the study were collected using several instruments. First, students' academic performance was measured using standardized university examination protocols based on a 100-point grading scale. These assessments were administered at two time points during the semester (pre-test and post-test), allowing the researchers to compare learning outcomes between the experimental and control groups. The examinations were conducted according to the institution's official disciplinary program and followed standard university evaluation procedures. This ensured that the assessment remained independent of the AI tool used in the experiment and reflected the full range of students' competencies expected by the curriculum, including conceptual understanding, procedural skills, and metacognitive aspects of mathematical problem solving.

The mathematics achievement tests were developed in accordance with the official university curriculum and comprised two types of tasks: (1) procedural tasks aimed at assessing computational skills (e.g., solving equations, computing derivatives and integrals), and (2) conceptual tasks evaluating understanding of mathematical concepts (e.g., analysis of function behavior, limits, and convergence).

Second, activity logs generated by the Socratic application were used to monitor the frequency and patterns of students' interaction with the AI-based learning tool during independent study. Third, classroom observation protocols were used to document students' engagement with the application during practical sessions and to verify how AI-supported explanations were integrated into the learning process.

Ethical considerations

The study was conducted in accordance with established ethical principles, ensuring the protection of participants' well-being and rights throughout the research process. All participants provided informed consent after receiving full information about the research objectives, the nature of their involvement, and the use of results. They were also assured of their right to withdraw at any time without consequence. Confidentiality and anonymity were strictly maintained. Personal information was only accessible

to the research team, and all identifying information was removed from the dataset before analysis to prevent the identification of individuals. Special attention was paid to potential risks associated with AI usage, particularly concerning data security. Additionally, the AI tool underwent thorough testing during the pilot phase to ensure its reliability and accuracy in providing explanations and solutions.

Monitoring of application use was conducted based on voluntarily provided activity logs, which reflected the frequency of interaction with the application. Data collection was limited to educational activities and did not involve access to personal device content. Participants were informed about the procedure and provided consent, with the right to withdraw at any stage without consequences.

Data analysis

The following data collection tools were used in the study. First, mathematics achievement was measured using assessments with a reliability coefficient of $\alpha = 0.89$. Validity was established through expert review, while construct validity was confirmed via confirmatory factor analysis. Second, an interaction tracking system documented behavioral parameters, including time metrics, feature usage frequency, and problem-solving pathways. Third, the Metacognitive Awareness Inventory (MAI) was administered, and the results were compared with observational data. This self-reported measure was selected for its depth, flexibility, and extensive validation in research on university mathematics education. The MAI comprises 52 items assessing knowledge of cognition (declarative, procedural, and conditional knowledge) and regulation of cognition (planning, monitoring, evaluation, and information management). Procedural validity was further established through pilot implementation with subsequent refinement (Nurjanah & Jailani, 2020; Rashidi & Fahmideh, 2020).

Statistical analysis

Statistical analysis was performed using SPSS version 27.0, with a significance level of $\alpha = 0.05$. Descriptive statistics, including measures of central tendency and dispersion, were computed for all outcome variables. To assess the intervention's effectiveness, a 2×2 mixed-design ANOVA was conducted, with Group (experimental vs. control) as a between-subject factor and Time (pre-test vs. post-test) as a within-subject factor. The primary indicator of effectiveness was the Group × Time interaction, which assessed whether performance improvements over time varied significantly between the experimental and control groups. Effect sizes were calculated using partial eta squared ($\eta^2 p$). Before conducting inferential tests, parametric assumptions such as normality (evaluated through the Shapiro–Wilk test) and homogeneity of variance (assessed via Levene's test) underwent thorough verification. Assumption testing indicated that the data met the requirements for parametric analysis: the Shapiro–Wilk test revealed no significant deviations from normality ($p > 0.05$), and Levene's test confirmed the homogeneity of variances ($p > 0.05$).

Results

Table 1 presents descriptive statistics for mathematics performance in both groups at the pre-test and post-test stages. Data analysis revealed that prior to the intervention, the groups demonstrated similar baseline levels, which is an important finding. Equivalent academic standing at the start of the semester minimized the likelihood that differences in final outcomes could be attributed to initial disparities rather than instructional conditions. By the end of the semester, the experimental group exhibited a marked shift in the distribution of achievement scores toward improvement. The change was reflected both in increased mean values and a higher median. In contrast, the control group showed stable dynamics, with negligible changes in measures of central tendency over the study period. This contrast suggests that the performance gains observed in the experimental group were not attributable to temporal changes associated with course completion, but rather to the specific pedagogical approach employed.

The data in Table 1 further confirm the stability of distributions in the control group, where variability remained virtually unchanged between testing points. In the experimental group, however, the increase in performance was accompanied by a reduction in score dispersion. This may indicate that AI-supported learning contributed not only to an overall rise in student achievement but also to a reduction in disparities among learners.

Table 1
Distribution of mean mathematics achievement scores by group type (0–100 scale)

Group	Time Point	N	Mean (M)	SD	Median
Experimental	Pre-test	200	77.60	6.77	77.00
Experimental	Post-test	200	86.94	6.45	87.00
Control	Pre-test	200	78.26	6.57	78.00
Control	Post-test	200	78.05	7.18	78.00

Although descriptive statistics indicate considerable performance gains in the experimental group, these data alone are insufficient to confirm the intervention's effectiveness. Accordingly, an inferential analysis was conducted to determine whether the rate of improvement in the experimental group was statistically higher than that in the control group. To formally examine the impact of the AI intervention, mathematics performance scores were analyzed using a 2×2 mixed-design ANOVA. Group (experimental vs. control) served as the between-subjects factor, while Time (pre-test vs. post-test) served as the within-subjects factor. This analysis enabled an assessment of the effect of instructional conditions on performance dynamics and provided a robust test of the intervention's impact through the Group × Time interaction.

The results of the analysis are presented in Table 2. Significant main effects were found for both Group and Time. The main effect of Time reflects changes between pre-test and post-test scores, indicating overall performance progression over the

semester. However, this effect alone does not constitute evidence of intervention effectiveness, as it does not capture differential improvement between groups. The most critical finding was a statistically significant Group × Time interaction, indicating that the magnitude of change over time differed between the experimental and control groups. Students who received AI-supported instruction demonstrated significantly greater performance gains compared to those who underwent traditional instruction. The mean improvement in the experimental group amounted to $\Delta M = +9.34$ points, whereas changes in the control group were minimal ($\Delta M = -0.21$), further supporting the effectiveness of the intervention.

Table 2
Mixed-design ANOVA results for mathematics achievement

Effect	df	F	p	Partial η^2
Group	1, 398	49.73	< .001	0.111
Time	1, 398	55.30	< .001	0.122
Group × Time	1, 398	100.05	< .001	0.201

The interaction between variables was accompanied by a large effect size, indicating that the improvements in the experimental group were not only statistically significant but also educationally meaningful. From a practical standpoint, these findings demonstrate that instruction incorporating AI tools facilitated more pronounced gains in academic performance over the semester compared to traditional mathematics teaching approaches. To assess effect size, partial η^2 was used, revealing that the interaction accounted for a considerable proportion of the variance in mathematics learning outcomes. This underscores the major contribution of the implemented intervention to enhancing student achievement. In the context of educational research, such results are typically regarded as compelling evidence that the experimental method yielded educational benefits beyond the typical progression observed over a semester.

Statistical analysis supports the conclusion that AI usage positively influenced the development of students' procedural accuracy and fostered more effective task-solving strategies. Immediate feedback and structured explanations provided by Socratic allowed students to identify errors more rapidly, review problems at their own pace, and reduce time spent on less productive trial-and-error approaches. This may explain the increase in mean scores within the experimental group and the observed reduction in performance variability among its participants.

An equally noteworthy finding is the stability of results in the control group, which reinforces the significance of the intervention. Had both groups demonstrated comparable progress, the attribution of improvements to AI tools would be questionable. The control group's minimal changes suggest that the substantial gains in the experimental group are highly likely attributable to the intervention rather than to natural progression over a standard academic semester.

Discussion

The results of the study indicate that the integration of AI-supported learning tools may influence students' cognitive engagement and learning processes in mathematics. The observed changes in cognitive structure are manifested through three-component mechanical mechanisms. Specifically, the algorithmic externalization of computational processes may reduce external cognitive load; mathematical working memory may be redirected from procedural execution to conceptual understanding. Individual developmental trajectories may emerge through adaptive pacing regulation mechanisms (Remoto, 2024).

Transmissive educational paradigms (knowledge transfer from expert to novice) have yielded to constructivist architectures (knowledge is actively constructed through mediated interaction). Passive information perception has transitioned to active knowledge construction, while content mastery has expanded to include cognitive process development (Jiayu, 2023). Within mathematical learning specifically, three key adaptations have emerged: theorem memorization has evolved into practical application, formula memorization now rests on conceptual visualization, and standardized examples have transformed into generative problem-solving. These learning adaptations may enhance mathematical understanding while maintaining content coverage (Yang, 2020).

Contemporary learning architectures have shifted from sequential progression to networked conceptual exploration. Traditional static pathways have become less effective than knowledge navigation (Knox, 2020). Conceptual networks are replacing linear curricula; metacognitive reflection is superseding content coverage; and multimodal representations are displacing unilateral presentations (Yang & Bai, 2020). The role of educators has also undergone a parallel transformation. In modern learning environments, instructors now primarily serve as facilitators of learning rather than as transmitters of knowledge. This shift reflects a move away from the traditional model of teachers as content authorities toward their reconceptualization as guides for metacognitive development (Woodroffe et al., 2021).

The methodological innovation lies in systemic integration that may contribute to reconfiguring educational processes within specific instructional contexts. This approach differs from supplementary implementation models (Saha et al., 2023), although the extent of transformation depends on contextual and institutional factors. Educational systems exhibit varying degrees of AI adoption – from traditional peripheral augmentation to more integrated approaches that may influence learning relationships (Jiayu, 2023).

The comparison of learning outcomes between the control and experimental groups demonstrated that students who used the AI-supported learning environment achieved higher post-intervention performance in mathematics. These findings suggest that AI-based tools may support deeper engagement with mathematical tasks and may be associated with improved academic outcomes when integrated into regular

learning activities. These results are consistent with previous studies demonstrating that AI-assisted learning environments can enhance problem-solving abilities and support individualized learning trajectories (Richard et al., 2022; Tashtoush et al., 2024; Wang, 2023).

The present findings also shed light on the scalability of AI interventions across different educational contexts. By examining the Chinese educational setting, the study extends the evidence base for AI tools, such as Socratic, beyond homogeneous Western populations (Wu, 2021; Xie & Wang, 2024). The results may indicate potential applicability of AI-supported learning in contexts similar to those examined in this study; however, generalization to other populations and educational systems should be made with caution. Demonstrating positive outcomes in this context contributes to the growing body of evidence on AI in education; however, further cross-cultural research is required before broader implementation claims can be substantiated. The marginal changes observed in the control group's scores further underscore the need for innovative educational tools to support academic improvement. Traditional methods, which lack interactive and technologically advanced systems, tend to recreate static learning environments that may limit progress, as previously demonstrated in earlier studies (Aljamali & Jabbar, 2020; Ee & Huh, 2018; Rane, 2023; Zhang, 2022).

The novelty of this article lies in its targeted assessment of the impact of the Socratic application on mathematics performance among university students in China. In contrast to broader studies on AI in education, this research evaluates a single AI tool, providing insights into its capabilities and effectiveness in enhancing learning outcomes. Additionally, the study examines the short-term effects of AI integration within a controlled educational environment, empirically demonstrating improved student performance. This study contributes to addressing a gap in the literature by providing empirical data on the use of a specific AI tool in university mathematics education.

The findings of this study carry important practical implications, particularly for colleges in rural and under-resourced areas seeking to enhance student academic performance. AI-driven educational tools such as Socratic may help mitigate some resource constraints in specific educational contexts, although their effectiveness depends on implementation conditions and institutional support. Artificial intelligence provides personalized learning materials, real-time feedback, and step-by-step guidance for problem-solving. This support is especially valuable in environments where individual teacher attention is limited due to large class sizes or geographic constraints (Mimoudi, 2024). Integrating AI technologies into the educational framework may contribute to improving learning outcomes, although the magnitude of this effect may vary depending on contextual factors. Furthermore, AI-based tools have the potential to expand access to learning support for students who lack access to high-quality tutoring or modern educational resources. Such tools may expand access to learning support; however, whether they can ensure equivalent opportunities across different regions remains an open question requiring further investigation.

Limitations

When interpreting the findings of this study, several significant methodological limitations should be considered. First, the restricted sample comprised students from only four universities in Eastern China. This limitation reduces the generalizability of the findings to other regions, educational contexts, or settings characterized by different technological and pedagogical conditions. Furthermore, although stratified sampling was employed to account for key demographic characteristics and baseline academic performance, the study was conducted in authentic classroom environments. As a result, it was impossible to completely rule out the influence of uncontrolled external variables, such as differences in student motivation, self-regulated learning skills, or attitudes toward digital technology. These factors may have contributed to variations in engagement with the intervention and, consequently, in learning outcomes.

The specific nature of the intervention itself warrants consideration as it relied on students' independent use of the Socratic application outside the classroom. This resulted in considerable variability in the intensity of interaction with the tool among participants. Although activity log data were available, these records do not fully capture qualitative aspects of engagement, such as the depth of information processing or the level of reflection during task completion. Another important factor is the potential novelty effect: students' initial interest and heightened effort may have been stimulated by the introduction of an innovative digital tool, potentially leading to temporary improvements in performance. Differences in participants' baseline digital literacy may have further influenced the effectiveness of their interaction with the AI system.

Although general guidelines were established regarding the frequency of use and the role of the application as a supplementary tool, the actual level of student engagement and the manner in which the tool was integrated into the learning process may have varied across instructors and institutions. Furthermore, despite the conceptual delineation of roles between the instructor and the AI tool, this distinction may not have been fully standardized in real educational practice, thereby complicating the precise estimation of the "pure" effect of the intervention.

An additional limitation concerns the procedure for monitoring application use. Although adherence to the intervention protocol was assessed through activity logs and participants' voluntary consent, this approach is limited in terms of the completeness of monitoring and may be sensitive to students' self-regulation. Moreover, such tracking methods may raise concerns related to privacy and ethical aspects of digital data collection, despite adherence to principles of informed consent and confidentiality.

It should also be noted that learning outcomes were assessed immediately upon conclusion of the 16-week intervention period. This makes it impossible to evaluate the long-term retention of knowledge or determine whether the observed gains represent sustained educational achievements or transient effects. Finally, although standardized examinations were employed to enhance the reliability of the findings, the resulting scores may partially reflect variations in academic standards and grading practices across the participating universities.

Conclusion

This article examined the use of Google's Socratic application in mathematics instruction at higher education institutions in China. The findings suggest that learning supported by artificial intelligence was associated with greater academic gains compared to traditional teaching methods within the context of this study. Over the course of the semester, students in the experimental group showed greater improvement in mathematics performance than their counterparts in the control group. This finding supports the potential of Socratic as a supplementary tool for university-level mathematics instruction, although further research is required to confirm its effectiveness across different contexts. From a practical standpoint, the study indicates that accessible AI-driven tutoring systems may assist universities in improving educational outcomes in large-enrollment courses, where personalized feedback from instructors is often limited.

Future research directions include longitudinal analyses of knowledge retention patterns beyond single-semester interventions. Another promising direction is analyzing domain-specific transferability in mathematical subfields. Cross-cultural comparative studies would offer deeper insights into implementation efficacy across educational systems. It would also be beneficial to develop assessment systems capturing cognitive dimensions beyond traditional performance metrics and investigate institutional readiness factors predicting successful implementation.

References

Abar, C. A. A. P., Dos Santos, J. M. D. S., & de Almeida, M. V. (2024). Contributes of the integration between computational thinking and artificial intelligence for mathematics education. In Á. Rocha, C. Ferrás, J. Hochstetter Diez, & M. Diéguez Rebolledo (Eds.), *International Conference on Information Technology & Systems* (pp. 290–299). Springer. https://doi.org/10.1007/978-3-031-54256-5_27

Aljamali, N. M., & Jabbar, A. A. H. (2020). Review on modern mathematics in the analysis of artificial intelligence data and innovative devices. *Journal of Instrumentation and Innovation Sciences, 6,* 10–19.

Batsaikhan, Z., & Correia, A.-P. (2024). The effects of generative artificial intelligence on intelligent tutoring systems in higher education: A systematic review. *Studies in Technology Enhanced Learning, 4*(1), 1–19. https://doi.org/10.21428/8c225f6e.33570bb1

Betteridge, J., Chan, E. Y., Corless, R. M., Davenport, J. H., & Grant, J. (2022). Teaching programming for mathematical scientists. In P. R. Richard, M. P. Vélez, & S. Van Vaerenbergh (Eds.), *Mathematics Education in the Age of Artificial Intelligence: How Artificial Intelligence can Serve Mathematical Human Learning* (pp. 251–276). Springer. https://doi.org/10.1007/978-3-030-86909-0_12

Birgin, O., Bozkurt, E., Gürel, R., & Duru, A. (2015). The effect of computer-assisted instruction on 7th grade students' achievement and attitudes toward mathematics: The case of the topic "Vertical Circular Cylinder". *Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje*, *17*(3), 783–813. https://doi.org/10.15516/cje.v17i3.1075

Broutın, M. S. T. (2024). Exploring mathematics teacher candidates' instrumentation process of generative artificial intelligence for developing lesson plans. *Yükseköğretim Dergisi, 14*(1), 165–176. https://doi.org/10.53478/yuksekogretim.1347061

Crompton, H., & Song, D. (2021). The potential of artificial intelligence in higher education. *Revista Virtual Universidad Catolica del Norte, 62,* 1–4. https://doi.org/10.35575/RVUCN.N62A1

DeFalco, J. A., & Sinatra, A. M. (2019). Adaptive instructional systems: The evolution of hybrid cognitive tools and tutoring systems. In R. Sottilare, & J. Schwarz (Eds.), *Adaptive Instructional Systems* (pp. 52–61). Springer. https://doi.org/10.1007/978-3-030-22341-0_5

Drijvers, P., & Sinclair, N. (2024). The role of digital technologies in mathematics education: Purposes and perspectives. *ZDM–Mathematics Education, 56,* 239–248. https://doi.org/10.1007/s11858-023-01535-x

Ee, J. H., & Huh, N. (2018). A study on the relationship between artificial intelligence and change in mathematics education. *Communications of Mathematical Education, 32*(1), 23–36. https://doi.org/10.7468/jksmee.2018.32.1.23

Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM–Mathematics Education, 56,* 281–292. https://doi.org/10.1007/s11858-023-01530-2

Fowler, D. S. (2023). AI in higher education: Academic integrity, harmony of insights, and recommendations. *Journal of Ethics in Higher Education, 3,* 127–143. https://doi.org/10.26034/fr.jehe.2023.4657

Gadanidis, G. (2017). Artificial intelligence, computational thinking, and mathematics education. *The International Journal of Information and Learning Technology, 34*(2), 133–139. https://doi.org/10.1108/IJILT-09-2016-0048

Gao, S. (2020). Innovative teaching of integration of artificial intelligence and university mathematics in big data environment. *IOP Conference Series: Materials Science and Engineering, 750*(1), 012137. https://doi.org/10.1088/1757-899X/750/1/012137

Google LLC. (2020). *Socratic by Google.* https://play.google.com/store/apps/details?id = com.google.socratic&hl = uk&gl = US

Gupta, A., Reddig, J., Calo, T., Weitekamp, D., & MacLellan, C. J. (2025). Beyond final answers: Evaluating large language models for math tutoring. In A. I. Cristea (Ed.), *International Conference on Artificial Intelligence in Education* (pp. 323–337). Springer. https://doi.org/10.1007/978-3-031-98414-3_23

Hall Jr, O. P. (2020). Achieving Bloom's two-sigma goal using intelligent tutoring systems: Application to management education. In K. Kumar, & J. P. Davim (Eds.), *Methodologies and Outcomes of Engineering and Technological Pedagogy* (pp. 133–162). IGI Global. https://doi.org/10.4018/978-1-7998-2245-5.ch009

Hampton, A. J., & Wang, L. (2019). Conversational AIS as the cornerstone of hybrid tutors. In R. Sottilare, & J. Schwarz (Eds.), *Adaptive Instructional Systems* (pp. 634–644). Springer. https://doi.org/10.1007/978-3-030-22341-0_49

Huang, X. (2021). Aims for cultivating students' key competencies based on artificial intelligence education in China. *Education and Information Technologies, 26*(5), 5127–5147. https://doi.org/10.1007/s10639-021-10530-2

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

Hwang, S. (2022). Examining the effects of artificial intelligence on elementary students' mathematics achievement: A meta-analysis. *Sustainability, 14*(20), 13185. https://doi.org/10.3390/su142013185

Jaboob, A., Durrah, O., & Chakir, A. (2024). Artificial intelligence: An overview. In A. Chakir (Eds.), *Engineering Applications of Artificial Intelligence* (pp. 3–22). Springer. https://doi.org/10.1007/978-3-031-50300-9_1

Jianzheng, S., & Xuwei, Z. (2023). Integration of AI with higher education innovation: Reforming future educational directions. *International Journal of Science and Research, 12*(10), 1727–1731. https://doi.org/10.21275/sr231023183401

Jiayu, Y. (2023). Challenges and opportunities of generative artificial intelligence in higher education student educational management. *Advances in Educational Technology and Psychology, 7*(9), 92–96. https://doi.org/10.23977/aetp.2023.070914

Kissi, P. S., Nat, M., & Idowu, A. (2019). Systematic review of web-based learning environments in high school mathematics education: Attitude, achievement, challenges, and possible solutions. *Croatian Journal of Education/Hrvatski Časopis za Odgoj i Obrazovanje*, *21*(4), 1061. https://doi.org/10.15516/cje.v21i4.3104

Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology, 45*(3), 298–311. https://doi.org/10.1080/17439884.2020.1754236

Lagrange, J. B., & Richard, P. R. (2022). Instrumental genesis in the theory of MWS. In A. Kuzniak (Ed.), *Mathematical work in educational context* (pp. 211–228). Springer. https://doi.org/10.1007/978-3-030-90850-8_9

Li, X., & Zaki, R. (2024). Harnessing the power of digital resources in mathematics education. In S. Papadakis (Ed.), *IoT, AI, and ICT for Educational Applications* (pp. 191–223). Springer. https://doi.org/10.1007/978-3-031-50139-5_10

Luo, Y. (2024). Revolutionizing education with AI: The adaptive cognitive enhancement model (ACEM). *Applied and Computational Engineering, 82,* 71–76. https://doi.org/10.54254/2755-2721/82/20240929

Maphosa, V., & Maphosa, M. (2021). The trajectory of artificial intelligence research in higher education: A bibliometric analysis and visualisation. In V. Greunen (Ed.), *2021 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)* (pp. 1–7). IEEE. https://doi.org/10.1109/icABCD51485.2021.9519368

Martinho, C., & Martins, S. (2025). Artificial intelligence in higher education mathematics teaching: Applications, challenges, and opportunities. In L. G. Chova, C. G. Martínez, & J. Lees (Eds.), *17th International Conference on Education and New Learning Technologies* (pp. 6163–6167). IATED Academy. https://doi.org/10.21125/edulearn.2025.1516

Matzakos, N., Doukakis, S., & Moundridou, M. (2023). Learning mathematics with large language models. *International Journal of Emerging Technologies in Learning, 18*(20), 51–71. https://doi.org/10.3991/ijet.v18i20.42979

Melchor, P. J. M., Lomibao, L. S., & Parcutilo, J. O. (2023). Exploring the potential of AI integration in mathematics education. *Journal of Innovations in Teaching and Learning, 3*(1), 39–44. https://doi.org/10.12691/jitl-3-1-8

Mimoudi, A. (2024). AI, personalized education, and challenges. In C. Gonçalves, & J. C. D. Rouco (Eds.), *4th International Conference on AI Research* (Vol. 4, pp. 271–280). ICAIR. https://doi.org/10.34190/icair.4.1.3133

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

Möller, M., Nirmal, G., Fabietti, D., Stierstorfer, Q., Zakhvatkin, M., Sommerfeld, H., & Schütt, S. (2024). *Revolutionising Distance Learning: A Comparative Study of Learning Progress with AI-Driven Tutoring.* arXiv preprint arXiv:2403.14642. https://doi.org/10.48550/arXiv.2403.14642

Nurjanah, R., & Jailani, J. (2020). Metacognitive awareness and mathematical problem solving ability of undergraduate students. *Journal of Physics: Conference Series, 1470*(1), 012042. https://doi.org/10.1088/1742-6596/1470/1/012042

Orhani, S. (2021). Artificial intelligence in teaching and learning mathematics. *Kosovo Educational Research Journal, 2*(3), 29–38.

Panit, N. (2025). Can critical thinking and AI work together? *Environment and Social Psychology, 9*(11), 1–16. https://doi.org/10.59429/esp.v9i11.3141

Pardos, Z. A., & Bhandari, S. (2023). *Learning Gain Differences between ChatGPT and Human Tutor Generated Algebra Hints.* arXiv preprint arXiv:2302.06871. https://doi.org/10.48550/arXiv.2302.06871

Rane, N. (2023). *Enhancing Mathematical Capabilities through ChatGPT and Similar Generative Artificial Intelligence: Roles and Challenges in Solving Mathematical Problems.* SSRN 4603237. https://doi.org/10.2139/ssrn.4603237

Rashidi, N., & Fahmideh, Z. (2020). The relationship between EFL university students' metacognitive awareness and their performance in reading comprehension. *Cogent Education, 7*(1), 1712045. https://doi.org/10.1080/2331186X.2020.1712045

Rawat, K., & Mishra, M. K. (2022). Role of mathematics in novel artificial intelligence realm. In R. Bhardwaj (Eds.), *Mathematics in Computational Science and Engineering* (pp. 211–231). Wiley. https://doi.org/10.1002/9781119777557.ch10

Remoto, J. P. (2024). ChatGPT and other AIs: Personal relief and limitations among mathematics-oriented learners. *Environment and Social Psychology, 9*(1), 1–13. https://doi.org/10.54517/esp.v9i1.1911

Richard, P. R., Vélez, M. P., & Van Vaerenbergh, S. (2022). *Mathematics Education in the Age of Artificial Intelligence.* Springer. https://doi.org/10.1007/978-3-030-86909-0

Rızvı, M. (2023). Investigating AI-powered tutoring systems. *The Eurasia Proceedings of Educational and Social Sciences, 31,* 67–73. https://doi.org/10.55549/epess.1381518

Romney, G. W., Guymon, J., Romney, M. D., & Carlson, D. A. (2019). Curriculum for hands-on artificial intelligence cybersecurity. In *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1–8). IEEE. https://doi.org/10.1109/ithet46829.2019.8937373

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54-67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Saha, S., Kamath, A., & Chaudhuri, M. S. (2023). A bibliometric analysis of the approaches of integrating artificial intelligence in higher education. In W. H. Ip (Ed.), *2023 International Conference on Intelligent Computing, Communication & Convergence (ICI3C)* (pp. 185–193). IEEE. https://doi.org/10.1109/ICI3C60830.2023.00045

Seo, S., Han, B., Harari, R. E., Dias, R. D., Zenati, M. A., Salas, E., & Unhelkar, V. (2025). *Socratic: Enhancing Human Teamwork via AI-enabled Coaching.* arXiv preprint arXiv:2502.17643. https://doi.org/10.48550/arXiv.2502.17643

Shin, D. (2022). Teaching mathematics integrating intelligent tutoring systems. *International Journal of Science and Mathematics Education, 20*(8), 1659–1676. https://doi.org/10.1007/s10763-021-10221-x

Tanveer, M., Hassan, S., & Bhaumik, A. (2020). Academic policy regarding sustainability and artificial intelligence (AI). *Sustainability, 12*(22), 9435. https://doi.org/10.3390/su12229435

Tashtoush, M. A., Wardat, Y., Ali, R. A., & Saleh, S. (2024). Artificial intelligence in education: mathematics teachers' perspectives, practices and challenges. *Iraqi Journal for Computer Science and Mathematics, 5*(1), 60–77. https://doi.org/10.52866/ijcsm.2024.05.01.003

Tay, L. Y., Chan, M., Chong, S. K., Tan, J. Y., & Aiyoob, T. B. (2024). Learning of mathematics: A metacognitive experiences perspective. *International Journal of Science and Mathematics Education, 22*(3), 561–583. https://doi.org/10.1007/s10763-023-10385-8

Tomar, J., & Soshal, M. (2023). Synergies of artificial intelligence and mathematics. *Journal Global Values, 14,* 156–162. https://doi.org/10.31995/jgv.2023.v14iS3.019

Valjarević, D., Stanojević-Ristić, Z., Petrović, M., & Tomović, K. (2025). Factors that influence pupils' engagement in learning mathematics and their effects on general academic achievement. *Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje*, *27*(3), 891–918. https://doi.org/10.15516/cje.v27i3.6263

Wang, K. (2023). Design and implementation of smart teaching. *Applied Mathematics and Nonlinear Sciences, 9*(1), 1–14. https://doi.org/10.2478/amns-2024-0503

Woodroffe, T., Wallace, R., Guthadjaka, K., Funk, J., Maypilama, E., Ireland, S., Adair, R., Ober, R., Armstrong, S., Lowell, A., & Pollard, K. (2021). Indigenous women in science: A proposed framework for leadership, knowledge, innovation, and complexity. In W. B. James, C. Cobanoglu, & M. Cavusoglu (Eds.), *Global Conference on Education and Research* (pp. 1–16). University of South Florida M3 Center. https://doi.org/10.5038/9781955833042

Wu, R. (2021). Visualization of basic mathematics teaching based on artificial intelligence. *Journal of Physics: Conference Series, 1992*(4), 042042. https://doi.org/10.1088/1742-6596/1992/4/042042

Xie, X., & Wang, T. (2024). Artificial Intelligence: A help or threat to contemporary education. *Education and Information Technologies, 29,* 3097–3111. https://doi.org/10.1007/s10639-023-11947-7

Yang, H. (2020). Calculus teaching method based on artificial intelligence application. In C. T. Yang, Y. Pei, & J. W. Chang (Eds.), *Innovative Computing: IC 2020* (pp. 139–145). Springer. https://doi.org/10.1007/978-981-15-5959-4_17

Yang, S., & Bai, H. (2020). The integration design of artificial intelligence and normal students' education. *Journal of Physics: Conference Series, 1453*(1), 012090. https://doi.org/10.1088/1742-6596/1453/1/012090

Yanuarto, W. N., Jaelani, A., Purwanto, J., & Zakaria, M. I. (2021). Socratic as mathematics learning application for differential equations concept. *Hipotenusa: Journal of Mathematical Society, 3*(1), 97–110. https://doi.org/10.18326/hipotenusa.v3i1.97-110

Zhang, F. (2022). Design and application of artificial intelligence technology-driven education system. *Computational and Mathematical Methods in Medicine, 2022,* 8503239. https://doi.org/10.1155/2022/8503239

Tingting Feng
Qiongtai Normal University, College of Science
8 Zhongshan S Rd, 571127, Haikou, China
fengtingting@qtnu.edu.cn

Integracija umjetne inteligencije u kurikule matematike na kineskim fakultetima

Sažetak

Ova studija analizira upotrebu Googleove aplikacije Socratic kao alata za podršku u nastavi matematike na visokim obrazovnim ustanovama u Kini. U studiji je sudjelovalo 400 studenata s četiriju sveučilišta u istočnoj Kini, podijeljenih u dvije skupine: eksperimentalnu i kontrolnu, od kojih je svaka obuhvaćala 200 studenata s usporedivim demografskim karakteristikama. Tijekom 16 tjedana studenti u eksperimentalnoj skupini koristili su aplikaciju Socratic za samostalno učenje i rješavanje domaćih zadaća, dok je kontrolna skupina nastavila koristiti tradicionalne metode učenja. Postignuća studenata ocijenjena su standardiziranim ispitima. Također su prikupljeni podatci o angažmanu studenata s aplikacijom, uključujući učestalost korištenja i obrasce korištenja značajki. Dodatno, metakognitivne vještine ocijenjene su pomoću specijaliziranoga upitnika. Rezultati su pokazali da su učenici u eksperimentalnoj skupini značajno poboljšali svoje rezultate iz matematike u usporedbi s kontrolnom skupinom, u kojoj promjene nisu zabilježene. Statistička analiza potvrdila je da je korištenje aplikacije za podršku procesu učenja pozitivno utjecalo na akademske rezultate učenika. Analiza korištenja aplikacije otkrila je da su učenici aktivno koristili razne značajke, uključujući upute korak po korak, pomoć pri objašnjenju i alate za provjeru rješenja. Sveukupno, nalazi istraživanja ističu potencijal platforme Socratic kao alata za poučavanje matematike na sveučilišnoj razini u kontekstima u kojima su prilike za personaliziranu povratnu informaciju ograničene.

Ključne riječi: *matematičko znanje; metakognitivna regulacija; obrazovna tehnologija; personalizirano učenje; strategije rješavanja problema*

Uvod

Umjetna inteligencija (AI) odnosi se na računalne sustave sposobne za obavljanje zadataka koji obično zahtijevaju ljudsku inteligenciju, kao što su učenje, rasuđivanje, rješavanje problema i prilagodba (Abar i dr., 2024.; Drijvers i Sinclair, 2024.; Gadanidis, 2017). U ovome radu istražuje se umjetna inteligencija kao obrazovni alat koji koristi modele strojnoga učenja i obrade prirodnoga jezika kako bi ponudio prilagođenu i interaktivnu pomoć studentima (Engelbrecht i Borba, 2024). Konkretno, umjetna

inteligencija integrira se u proces učenja kako bi se poboljšali akademski uspjesi studenata, posebice iz matematike, putem personaliziranih rješenja, detaljnih objašnjenja korak po korak i trenutačne povratne informacije (Kissi i sur., 2019).

U brzo razvijajućem visokom obrazovanju, integracija umjetne inteligencije u akademske kurikule znači transformativni pomak u pristupima poučavanju temeljnih predmeta kao što je Matematika (Engelbrecht i Borba, 2024). Visokoškolske ustanove diljem svijeta nastoje poboljšati ishode učenja i iskoristiti tehnološki napredak. U tome kontekstu, UI se nametnula kao moćan alat za stvaranje zanimljivijega i personaliziranijega okružja za učenje (Drijvers i Sinclair, 2024; Knox, 2020).

Poticaj za integraciju umjetne inteligencije u matematiku proizlazi iz nekoliko okolnosti koje se poklapaju. Složenost discipline i raznolikost stilova učenja zahtijevaju personaliziranije i adaptivnije strategije (Shin, 2022). Alati i platforme temeljeni na UI-u mogu pružiti jedinstvena iskustva učenja prilagođavanjem potrebama pojedinih učenika i nudeći povratne informacije u stvarnome vremenu, što olakšava razumijevanje i zadržavanje gradiva (Hwang i Tu, 2021; Tay i dr., 2024). Nadalje, u višim razredima nastavnicima postaje izazovno zadovoljiti obrazovne potrebe svakog učenika (Xie i Wang, 2024).

Neuropsihološki čimbenici često predstavljaju značajne prepreke u percepciji matematike. Teorija kognitivnoga opterećenja objašnjava teret koji matematička apstrakcija nameće radnoj memoriji (Matzakos i dr., 2023; Valjarević i dr., 2025). Na kineskim sveučilištima omjer nastavnika i studenata često premašuje 1 : 60, što ograničava mogućnosti za individualiziranu podršku (Jianzheng i Xuwei, 2023; Knox, 2020). Nastava matematike zahtijeva personaliziranu kognitivnu podršku, no odvija se unutar modela masovnoga obrazovanja, što nameće tehnološku medijaciju (Crompton i Song, 2021). Adaptivni sustavi mogu odgovoriti na te izazove prilagođavanjem nastave individualnim razinama postignuća i vizualiziranjem apstraktnih koncepata kako bi ih učinili pristupačnijima (Fowler, 2023; Hwang, 2022).

U kontekstu visokoga obrazovanja u Kini, ovi su izazovi osobito izraženi zbog brojnosti studentske populacije, snažnih kurikulnih očekivanja iz matematike i ograničenih mogućnosti za individualiziranu povratnu informaciju tijekom odgojno-obrazovnoga procesa. Nedavne studije pokazuju da kineska sveučilišta sve više istražuju digitalna okružja za učenje podržana umjetnom inteligencijom kao način dopunjavanja tradicionalne nastave i podržavanja samostalnoga bavljenja studenata složenim matematičkim gradivom (Huang, 2021; Jianzheng i Xuwei, 2023; Knox, 2020). Posljedično, alati temeljeni na umjetnoj inteligenciji smatraju se ne samo tehnološkim inovacijama, već i praktičnim mehanizmima za rješavanje strukturnih ograničenja masovnoga visokog obrazovanja.

Tradicionalni prikazi matematičkih koncepatauglavnom su statični, međutim, računalna medijacija može ih učiniti istraživima. Digitalni alati i adaptivni algoritmi pretvaraju matematičke objekte u konstruirane stvarnosti koje nastaju tijekom interaktivnih procesa (Betteridge i dr., 2022; Birgin i dr., 2015). Pristupi temeljeni

na umjetnoj inteligenciji mogu smanjiti suvišno kognitivno opterećenje pomoću prilagodljivoga sekvenciranja i optimiziranih prikaza, čime se podržava konceptualna integracija (Romney i dr., 2019; Tanveer i dr., 2020). Integracija teorije matematičkoga kogniranja putem tehnološke implementacije pruža teorijsku osnovu za inovacije u obrazovnoj psihologiji (Maphosa i Maphosa, 2021).

Unatoč brzom prihvaćanju tehnologija umjetne inteligencije u obrazovanju, još uvijek nije jasno kako ti alati utječu na uspjeh učenika iz matematike. Iako prethodna istraživanja upućuju na to da alati umjetne inteligencije mogu poboljšati vještine rješavanja problema i povećanju angažmana studenata(Mohamed i dr., 2022; Richard i dr., 2022; Tashtoush i dr., 2024), još uvijek postoji ograničen broj empirijskih studija koje sustavno uspoređuju učinke takvih alata u kontroliranim uvjetima visokoga obrazovanja (Drijvers i Sinclair, 2024). Dosadašnja istraživanja uglavnom su usmjerena na opća okružja za učenje podržana umjetnom inteligencijom ili na teorijske rasprave o inteligentnim sustavima za poučavanje, dok su empirijska istraživanja specifičnih primjena umjetne inteligencije u sveučilišnoj matematici i dalje ograničena (Gao, 2020; Huang, 2021; Wu, 2021). Ova studija rješava taj nedostatak ispitivanjem pedagoške učinkovitosti aplikacije Socratic AI u autentičnim sveučilišnim učionicama u Kini te pružanjem empirijskih dokaza o njezinom utjecaju na matematičko postignuće i angažman studenata. Također, postoji ograničen broj empirijskih dokaza o učinkovitosti aplikacija temeljenih na umjetnoj inteligenciji u obrazovanju iz matematike na sveučilišnoj razini (Gao, 2020; Huang, 2021).

Pregled literature

Ova je studija utemeljena na teoriji samodeterminacije (SDT), koja objašnjava motivaciju za učenje zadovoljavanjem triju osnovnih psiholoških potreba: autonomije, kompetencije i povezanosti (Ryan i Deci, 2000.; 2020). U obrazovnim kontekstima te se potrebe podržavaju kada učenici reguliraju svoje procese učenja, primaju povratne informacije koje jačaju njihov osjećaj kompetentnosti i djeluju u okružjima koja pružaju podršku. Obrazovne tehnologije temeljene na umjetnoj inteligenciji mogu doprinijeti tim uvjetima pružanjem personaliziranih objašnjenja, prilagođenih povratnih informacija i mogućnosti rješavanja problema vlastitim tempom (Hwang i Tu, 2021; Tashtoush i dr., 2024). U području matematike u kojem učenici često trebaju postupno vođenje korak po korak i ponavljanu praksu, takve značajke mogu poticati angažman i konceptualno razumijevanje. Stoga SDT pruža relevantan okvir za ispitivanje načina na koji bi aplikacija Socratic mogla biti povezana s promjenama u motivaciji i akademskim uspjehom učenika.

Nedavne studije sve više istražuju alate za učenje temeljene na umjetnoj inteligenciji u odnosu na kognitivne i metakognitivne procese. Na primjer, Rızvı (2023) povezuje nastavu podržanu umjetnom inteligencijom s razvojem samoregulacije, dok Luo (2024) pokazuje da sustavi prilagodljive povratne sprege mogu podržati autonomiju i ovladavanje zadatcima učenika. Panit (2025) dalje pokazuje da interakcija sa sustavima UI-a može biti povezana s razvojem metakognitivne svijesti.

Upotreba alata AI-a u pedagogiji raspravlja se u vezi s promjenama u načinu na koji studenti komuniciraju s resursima za učenje. Istraživanja o instrumentalnoj genezi sugeriraju da digitalni alati mogu utjecati na to kako učenici internaliziraju znanje (Broutın, 2024; Lagrange i Richard, 2022). Istovremeno, studije o semiotičkoj medijaciji pokazuju da procesi stvaranja značenja ostaju u središtu, iako sve više uključuju interakciju s digitalnim sustavima (Batsaikhan i Correia, 2024; Gupta i dr., 2025). Empirijski dokazi ukazuju na to da objašnjenja i primjeri potpomognuti umjetnom inteligencijom mogu olakšati usvajanje matematičkoga diskursa (Möller i dr., 2024.; Pardos i Bhandari, 2023).

Uloge nastavnika također su se razvile u okružjima podržanim umjetnom inteligencijom, premještajući se s prenošenja znanja na olakšavanje (Hall Jr., 2020). DeFalco i Sinatra (2019) tvrde da interakcija sa sustavima UI-a može preraspodijeliti određene nastavne funkcije, posebice one povezane s povratnim informacijama i usmjeravanjem.

Istraživanja identificiraju nekoliko modela integracije UI-a u obrazovanje. Dodatni modeli tretiraju kao dodatni resurs, dok integrirani modeli kombiniraju UI i nastavu koju vodi učitelj (Hampton i Wang, 2019). Hibridni sustavi reorganiziraju procese učenja na temelju podataka o učinku. Dokazi upućuju na to da učinkovitost ovih pristupa ovisi o tome koliko su tehnološki alati usklađeni s didaktičkim dizajnom (Luo, 2024).

U području umjetne inteligencije u obrazovanju, matematika se često navodi kao područje u kojem bi alati podržani umjetnom inteligencijom mogli biti posebno učinkoviti zbog strukturirane prirode rješavanja problema i potrebe za povratnim informacijama korak po korak.

Pregled studija iz razdoblja od 2018. do 2023. ukazuje na to da učenje podržano umjetnom inteligencijom može poboljšati rezultate iz matematike u određenim kontekstima, iako dugoročni učinci ostaju nejasni (Batsaikhan i Correia, 2024). Na primjer, Möller i dr. (2024) izvještavaju o poboljšanom napretku u učenju u okružjima podržanim umjetnom inteligencijom, dok Pardos i Bhandari (2023) pokazuju da napomene generirane umjetnom inteligencijom poboljšavaju rješavanje algebarskih problema. Istovremeno, ove studije naglašavaju da je umjetna inteligencija najučinkovitija kada se koristi kao dodatni alat, a ne kao zamjena za učitelje (Gupta i dr., 2025; Seo i dr., 2025).

Među specifičnim primjenama izdvaja se aplikacija Socratic kao alat za učenje utemeljen na umjetnoj inteligenciji koji korisnicima pruža objašnjenja korak po korak te upućuje na dodatne relevantne izvore. Yanuarto i dr. (2021) pokazuju da Socratic podržava konceptualno razumijevanje putem vizualnoga prepoznavanja i vođenih objašnjenja. Međutim, postojeća istraživanja uglavnom se usredotočuju na njegove tehničke značajke ili podršku pri rješavanju domaćih zadaća, dok su dokazi iz kontroliranih implementacija na sveučilišnoj razini i dalje ograničeni.

Umjetna inteligencija također se primijenila na didaktički dizajn i analitiku učenja. Li i Zaki (2024) pokazuju da UI može identificirati učinkovite nastavne strategije na temelju velikih skupova podataka, dok Aljamali i Jabbar (2020) pokazuju da UI

može podržati ispravljanje pogrešaka identificiranjem praznina u znanju. Simulacije temeljene na VUM-u korištene su za modeliranje složenih matematičkih problema, pružajući interaktivne prikaze koji potiču razumijevanje (Jaboob i dr., 2024.; Rawat i Mishra, 2022). Takve simulacije mogu pomoći studentima da se učinkovitije bave apstraktnim konceptima (Orhani, 2021).

U literaturi se također ističu izazovi povezani s implementacijom VUM-a. Pojavljuju se zabrinutosti u vezi s privatnošću zbog prikupljanja podataka (Rane, 2023; Tomar i Soshal, 2023), a UI sustavi mogu generirati netočna objašnjenja bez nadzora (Melchor i sur., 2023). Stoga istraživači naglašavaju potrebu za provjerenim i pouzdanim UI alatima (Tashtoush i sur., 2024; Wang, 2023). Unatoč tim ograničenjima, sustavi podržani umjetnom inteligencijom mogu doprinijeti samostalnom učenju i rješavanju problema pružanjem strukturiranih objašnjenja i trenutačne povratne informacije (Hwang i Tu, 2021).

Unatoč tim napretcima, relativno je malo empirijskih studija koje ispituju pedagoški utjecaj specifičnih primjena umjetne inteligencije u autentičnim sveučilišnim matematičkim okružjima. Stoga ova studija doprinosi literaturi istraživanjem upotrebe aplikacije Socratic na kineskim sveučilištima i pružanjem empirijskih dokaza o njezinoj povezanosti s akademskim uspjehom i angažmanom studenata.

Definiranje problema

Ova studija ima dva glavna cilja: prvi, procijeniti pedagošku učinkovitost integracije algoritama – posebice Sokratove umjetne inteligencije – u nastavne planove i programe iz matematike i drugi, uspostaviti empirijsku osnovu za tehnološku transformaciju kvantitativnoga obrazovanja.

Istraživački zadatci: 1) kvantificirati razlike u ishodima učenja između nastave potpomognute umjetnom inteligencijom i tradicionalnih pedagoških pristupa, 2) identificirati taksonomiju interakcija koje nastaju tijekom algoritamskoga angažmana analizom obrazaca korištenja u proceduralnim, reprezentacijskim i verifikacijskim dimenzijama, 3) dokumentirati kognitivne promjene inducirane tehnološkom medijacijom, posebice u topologijama rješavanja problema i metakognitivnoj regulaciji i 4) ispitati specifične značajke matematike u pogledu podložnosti algoritamskom poboljšanju.

Trenutačni obrazovni sustavi suočavaju se s postojanim izazovima, kao što su fragmentirano konceptualno razumijevanje, trajna matematička anksioznost i ograničenja u personalizaciji unutar masovnoga obrazovanja. Ovi problemi zahtijevaju istraživanje tehnološkoga pojačanja izvan konvencionalnih obrazovnih paradigma.

Metodologija

Metodološki okvir

Socratic aplikacija (Slika 1) odabrana je zbog svojih širokih mogućnosti pružanja obrazovne podrške u stvarnome vremenu putem obrade prirodnoga jezika i strojnoga učenja. Alat podržava tekstualni i vizualni unos, nudeći detaljna objašnjenja, korake za rješavanje problema i dodatne obrazovne resurse (Google LLC, 2020).

Slika 1

Zahvaljujući pristupačnom i jednostavnom dizajnu aplikacije Socratic, učenici mogu dobiti trenutačnu pomoć skeniranjem svojih obrazovnih zadataka pomoću mobilnih uređaja. Ova je značajka osobito korisna u velikim učionicama pri čemu se pozornost učitelja mora podijeliti među mnogim učenicima. Osim toga, sposobnost aplikacije da složene matematičke probleme razloži na korake čini je vrijednim alatom za učenike s različitim razinama znanja. Socratic je prvenstveno osmišljen za vježbanje proceduralnih matematičkih zadataka. Njegova značajka Explainers pomaže produbljivanju konceptualnoga znanje pružajući dodatna objašnjenja koja učenicima olakšavaju rješavanje takvih problema. Međutim, za primijenjene probleme aplikacija nudi samo poveznice na vanjske resurse povezane s kontekstima iz stvarnoga svijeta koji bi mogli zanimati učenika, pomažući s objašnjenjima i vodeći ga u daljnje istraživanje (Martinho i Martins, 2025; Yanuarto i sur., 2021).

Slika 2

Aplikacija prihvaća unose u tekstualnom, zvučnom (glas korisnika) i slikovnom formatu. Učenici mogu fotografirati zadatak, uključujući rukopisni tekst; UI zatim prepoznaje unos, pruža objašnjenja i rješenja te povezuje na povezane obrazovne resurse. Također provjerava razumije li učenik sadržaj zadatka, pripadajuću teoriju i ispravnost izvedbe. Međutim, aplikacija ne rješava zadatak umjesto učenika; ona podržava, ali ne zamjenjuje učenikove vlastite napore (Slika 2). Povratne informacije daju se putem infografika, vizualizacija matematičkih formula i koncepata, tekstualnih objašnjenja i smislenoga dijaloga usmjerenoga na pitanja učenika i odgovore na zadatke.

Dizajn studije

U studiji je korišten dizajn s kontrolnom skupinom i predtest-posttest metodom uz kvantitativnu analizu kako bi se istražila učinkovitost Googleova obrazovnoga sustava temeljenoga na Socratic UI-u. Studija je provedena u nekoliko uzastopnih faza: 1) razvoj i validacija instrumenata putem pokusnoga testiranja (n = 42), 2) raspodjela sudionika putem postupaka stratificirane randomizacije, 3) početno mjerenje (matematički testovi koji obuhvaćaju proceduralne i konceptualne vještine), 4) diferencirana intervencija (nastava pomoću VUM-a u eksperimentalnoj skupini i tradicionalna nastava u kontrolnoj skupini), 5) sustavno praćenje obrazaca interakcije pomoću digitalne analitike (samo eksperimentalna skupina), 6) mjerenje nakon intervencije, 7) statističko usporedivanje ishoda i 8) izdvajanje podataka o ponašanju. Intervencija je trajala 16 akademskih tjedana (jedan semestar), uz neprekidan pristup sustavu umjetne inteligencije tijekom cijeloga razdoblja (prosječne interakcije sudionika: 4,3 ± 1,2 po tjednu).

Provedena je stratificirana randomizacija uzimajući u obzir spol, godinu studija i početnu razinu matematičke pismenosti. Sudionici su prvo raspoređeni u podskupine prema tim karakteristikama, nakon čega je unutar svake podskupine nasumično dodjeljivanje u eksperimentalnu i kontrolnu skupinu provedeno pomoću računalno

generiranih slučajnih brojeva. Ovaj je pristup osigurao usporedivost skupina i smanjio rizik od pristranosti odabira.

Sudionici

Sudionici su odabrani stratificiranim nasumičnim uzorkovanjem s odsjeka za matematiku na četirima sveučilištima u istočnoj Kini (raspodjela po provincijama: Jiangsu = 2, Zhejiang = 1, Shanghai = 1). Ukupni uzorak obuhvaćao je 400 sudionika (eksperimentalna skupina = 200; kontrolna skupina = 200) s usklađenim demografskim parametrima: raspodjela po spolu (eksperimentalna skupina: 106 studentica i 94 studenta; kontrolna skupina: 102 studentice i 98 studenata). Distribucija sudionika prema spolu navedena je kako bi se pokazala demografska usporedivost eksperimentalne i kontrolne skupine te smanjila moguća pristranost u tumačenju ishoda učenja koja je povezana sa spolom.

Godina studija (druga godina = 58 %; treća godina = 42 %). U apsolutnim brojkama to je odgovaralo 232 studenta druge godine i 168 studenata treće godine. Uključeni su samo studenti druge i treće godine jer su do te faze studenti već završili uvodne sveučilišne kolegije iz matematike i imaju dovoljnu stručnu osnovu za bavljenje naprednijim zadatcima rješavanja matematičkih problema.

Razina matematičkoga znanja (prosječne ocjene na prijemnom ispitu: eksperimentalna = 81,3 ± 6,2; kontrolna = 80,9 ± 6,5). Odabrane institucije predstavljale su različite geografske i administrativne kontekste, uključujući i gradove i općine. Veličina uzorka određena je a priori analizom moći ($\alpha = 0{,}05$; $\beta = 0{,}80$; očekivana veličina učinka $d = 0{,}40$), što je ukazivalo na minimalni zahtjev od 156 sudionika po uvjetu.

Intervencija

Učenici su prvenstveno koristili aplikaciju Socratic izvan učionice za samostalno učenje, ponavljanje gradiva, traženje dodatnih objašnjenja i dobivanje individualizirane podrške prilagođene njihovu razumijevanju. Također su se oslanjali na aplikaciju za rješavanje domaćih zadaća.

Korištenje aplikacije Socratic bilo je standardizirano: studentima je naloženo da koriste aplikaciju 3 – 5 puta tjedno tijekom rješavanja samostalnih zadataka. Stvarna učestalost interakcije iznosila je u prosjeku 4,3 puta tjedno (SD = 1,2). Glavna komponenta intervencije provodila se izvan nastave (samostalno učenje).

Aplikacija se također koristila u učionici, osobito kada bi je nastavnik preporučio ili kada mu je nedostajalo dovoljno vremena za individualni rad sa svakim studentom. Dok je nastavnik prezentirao gradivo na ploči i davao objašnjenja, studenti su mogli usmjeriti kameru svojega uređaja prema ploči, što je aplikaciji omogućilo da ponudi trenutačan pristup detaljnim objašnjenjima, primjerima i vježbama. Nastavnik je mogao dodijeliti vlastite vježbe ili predložiti one koje je generirala aplikacija. Studenti su bili poticani na postavljaje pitanja nastavniku. Međutim, ako su vremenska ograničenja spriječila nastavnika da odgovori svima ili ako je student bio nezadovoljan, odnosno

nije razumio objašnjenje, mogao se obratiti aplikaciji za dodatno pojašnjenje. Tijekom kasnijih praktičnih zadataka, nastavnik je provjeravao jesu li studenti pomoću aplikacije ispravno razumjeli i usvojili gradivo. Ovaj je pristup omogućio potpuno pokrivanje grupe za učenje, poboljšanu prilagodljivost i individualizirani tempo učenja, uz istovremeno usklađivanje sa sveučilišnim kurikulom.

Tijekom intervencije razrađene su uloge instruktora i UI alata: instruktor je bio odgovoran za objašnjavanje gradiva i praćenje razumijevanja, dok se aplikacija koristila za individualiziranu podršku (objašnjenja korak po korak, provjera rješenja i ponavljanje). Kako bi se osigurala dosljednost u provedbi intervencije, instruktori su prethodno dobili standardizirane upute kojima je bilo propisano da se aplikacija koristi isključivo kao dodatni alat, bez zamjene primarnoga načina iznošenja nastavnoga sadržaja.

Kako bi se osigurala valjanost eksperimenta, od studenata se tražilo da tijekom studije ne koriste druge slične obrazovne aplikacije na svojim uređajima. Sudionici su pristali podijeliti zapise aktivnosti svojih uređaja, čime su potvrdili da su tijekom cijeloga semestra koristili samo navedenu aplikaciju. Studenti su zadržali pravo odustati od ove prakse u bilo kojem trenutku ako je to ometalo njihove obrazovne potrebe, ali nijedan sudionik nije iskoristio to pravo.

Poštivanje protokola intervencije pratilo se analizom dnevnika aktivnosti aplikacije i promatranjima instruktora. U prosjeku su se studenti u eksperimentalnoj skupini služili aplikacijom 4,3 puta tjedno (SD = 1,2). Instruktori su dodatno nadzirali korištenje aplikacije tijekom praktičnih sati. Nisu utvrđena značajna odstupanja od planiranoga tijeka intervencije.

Alati za prikupljanje podataka

Podatci za studiju prikupljani su pomoću nekoliko instrumenata. Prvo, akademski uspjeh studenata mjerio se pomoću standardiziranih sveučilišnih protokola za ispitivanje temeljenih na stotinskoj ljestvici ocjenjivanja. Ove su procjene provodile u dvama različitim razdobljima tijekom semestra (pretest i posttest), što je istraživačima omogućilo usporedbu ishoda učenja između eksperimentalne i kontrolne skupine. Ispiti su provedeni u skladu sa službenim disciplinarnim programom ustanove i slijedili su standardne sveučilišne procedure ocjenjivanja. Time je osigurano da procjena ostane neovisna o alatu umjetne inteligencije korištenom u eksperimentu i da odražava puni raspon kompetencija studenata koje se očekuju prema nastavnom planu i programu, uključujući konceptualno razumijevanje, proceduralne vještine i metakognitivne aspekte rješavanja matematičkih problema.

Testovi postignuća iz matematike razvijeni su u skladu sa službenim sveučilišnim nastavnim planom i programom te su se sastojali od dvije vrste zadataka: (1) proceduralnih zadataka usmjerenih na procjenu računalnih vještina (npr. rješavanje jednadžbi, izračunavanje derivacija i integrala) i (2) konceptualnih zadataka kojima se procjenjuje razumijevanje matematičkih koncepata (npr. analiza ponašanja funkcija, limita i konvergencije).

Drugo, dnevnici aktivnosti generirani aplikacijom Socratic korišteni su za praćenje učestalosti i obrazaca interakcije studenata s alatom za učenje temeljenim na umjetnoj inteligenciji tijekom samostalnoga učenja. Treće, protokoli opažanja u učionici korišteni su za dokumentiranje angažmana učenika s aplikacijom tijekom praktičnih sesija i za provjeru načina na koji su objašnjenja podržana umjetnom inteligencijom integrirana u proces učenja.

Etika

Studija je provedena u skladu s utvrđenim etičkim načelima, osiguravajući zaštitu dobrobiti i prava sudionika tijekom cijeloga istraživačkog procesa. Svi su sudionici dali informirani pristanak nakon što su primili potpune informacije o ciljevima istraživanja, prirodi svojega sudjelovanja i korištenju rezultata. Također im je zajamčeno pravo na povlačenje u bilo kojem trenutku bez posljedica. Povjerljivost i anonimnost strogo su poštovani. Osobnim podatcima mogao je pristupiti samo istraživački tim, a sve identifikacijske informacije uklonjene su iz skupa podataka prije analize kako bi se spriječila identifikacija pojedinaca. Posebna je pozornost posvećena potencijalnim rizicima povezanim s upotrebom umjetne inteligencije, osobito u pogledu sigurnosti podataka. Nadalje, alat s umjetnom inteligencijom podvrgnut je temeljitom testiranju tijekom pokusne faze kako bi se osigurala njegova pouzdanost i točnost u pružanju objašnjenja i rješenja.

Praćenje korištenja aplikacije provedeno je na temelju dobrovoljno dostavljenih dnevnika aktivnosti koji odražavaju učestalost interakcije. Prikupljanje podataka bilo je ograničeno na obrazovne aktivnosti i nije uključivalo pristup sadržaju osobnih uređaja. Sudionici su bili informirani o postupku i dali su pristanak, s pravom povlačenja u bilo kojoj fazi bez posljedica.

Analiza podataka

U studiji su korišteni sljedeći alati za prikupljanje podataka. Prvo, postignuće u matematici mjerilo se pomoću procjena s koeficijentom pouzdanosti $\alpha = 0{,}89$. Valjanost je utvrđena stručnim pregledom, dok je konstrukcijska valjanost potvrđena potvrdnom faktorskom analizom. Drugo, sustav za praćenje interakcija dokumentirao je ponašajne parametre, uključujući vremenske mjere, učestalost korištenja značajki i putanje rješavanja problema. Treće, primijenjen je Inventar metakognitivne svijesti (MAI), a rezultati su uspoređeni s opažačkim podatcima. Ova mjera na temelju samoprijave odabrana je zbog svoje dubine, fleksibilnosti i opsežne validacije u istraživanjima sveučilišnoga obrazovanja iz matematike. MAI se sastoji od 52 stavke koje procjenjuju znanje o kogniciji (deklarativno, proceduralno i kondicionalno znanje) i regulaciju kognicije (planiranje, praćenje, evaluacija i upravljanje informacijama). Proceduralna valjanost dodatno je potvrđena pokusnom provedbom, nakon koje su provedene potrebne prilagodbe. (Nurjanah i Jailani, 2020; Rashidi i Fahmideh, 2020).

Statistička analiza

Statistička analiza provedena je pomoću programa SPSS verzija 27.0, s razinom značajnosti $\alpha = 0,05$. Deskriptivna statistika, uključujući mjere središnje tendencije i raspršenosti, izračunata je za sve ishodne varijable. Kako bi se procijenila učinkovitost intervencije, provedena je 2×2 mješovita ANOVA, s čimbenikom grupa (eksperimentalna naspram kontrolne) kao međusubjektivnim čimbenikom i čimbenikom vrijeme (predtest naspram posttesta) kao unutar-subjektivnim čimbenikom. Primarni pokazatelj učinkovitosti bila je interakcija grupa × vrijeme, kojom se procjenjivalo jesu li poboljšanja u učinku tijekom vremena značajno različita između eksperimentalne i kontrolne skupine. Veličine učinka izračunate su pomoću djelomičnoga eta kvadrata (η^2p). Prije provođenja inferencijalnih testova, parametrijske pretpostavke kao što su normalnost (procijenjena Shapiro–Wilkovim testom) i homogenost varijance (procijenjena Leveneovim testom) podvrgnute su temeljitoj provjeri. Provjera pretpostavki pokazala je da su podatci ispunili uvjete za parametrijsku analizu: Shapiro–Wilkov test nije otkrio značajna odstupanja od normalne raspodjele ($p > 0,05$), a Leveneov test potvrdio je homogenost varijanci ($p > 0,05$).

Rezultati

Tablica 1 prikazuje deskriptivnu statistiku za postignuća iz matematike u objema skupinama na početnom i završnom testiranju. Analiza podataka otkrila je da su skupine prije intervencije pokazale slične početne razine, što je važan nalaz. Ekvivalentan akademski status na početku semestra umanjio je vjerojatnost da se razlike u konačnim ishodima mogu pripisati početnim razlikama, a ne uvjetima poučavanja. Do kraja semestra eksperimentalna je skupina pokazala izražen pomak u raspodjeli rezultata prema poboljšanju. Promjena se odrazila i u povećanim srednjim vrijednostima i višem medijanu. Nasuprot tome, kontrolna je skupina pokazala stabilnu dinamiku, sa zanemarivim promjenama mjera središnje tendencije tijekom razdoblja istraživanja. Ovaj kontrast upućuje na to da se napredak u postignućima eksperimentalne skupine ne može pripisati samo vermenskom učinku povezanom sa završetkom kolegija, već primjeni specifičnog pedagoškog pristupa.

Podatci u Tablici 1 dodatno potvrđuju stabilnost raspodjela u kontrolnoj skupini, pri čemu je varijabilnost ostala gotovo nepromijenjena između točaka testiranja. Međutim, u eksperimentalnoj skupini porast uspješnosti bio je popraćen smanjenjem disperzije rezultata. To može ukazivati na to da je učenje potpomognuto umjetnom inteligencijom doprinijelo ne samo sveukupnom porastu uspješnosti učenika, već i smanjenju nejednakosti među učenicima.

Tablica 1

Iako deskriptivna statistika ukazuje na znatna poboljšanja u postignućima eksperimentalne skupine, ti podatci sami po sebi nisu dovoljni za potvrdu učinkovitosti intervencije. Stoga je provedena inferencijalna analiza kako bi se utvrdilo je li stopa

poboljšanja u eksperimentalnoj skupini statistički viša nego u kontrolnoj skupini. Za formalno ispitivanje utjecaja intervencije temeljene na umjetnoj inteligenciji, rezultati iz matematike analizirani su pomoću 2 × 2 mješovite ANOVA analize. Grupa (eksperimentalna naspram kontrolne) služila je kao međusubjektni čimbenik, dok je vrijeme (predtest naspram posttesta) služilo kao unutarsubjektni čimbenik. Ova je analiza omogućila procjenu učinka uvjeta poučavanja na dinamiku postignuća i pružila robustan test utjecaja intervencije putem interakcije grupa × vrijeme.

Rezultati analize prikazani su u Tablici 2. Pronađeni su značajni glavni učinci za grupu i vrijeme. Glavni učinak vremena odražava promjene između rezultata predtestiranja i posttestiranja, što ukazuje na opći napredak u učinku tijekom semestra. Međutim, ovaj učinak sam po sebi ne predstavlja dokaz učinkovitosti intervencije jer ne obuhvaća diferencijalno poboljšanje između grupa. Najvažniji nalaz bila je statistički značajna interakcija grupa × vrijeme, što ukazuje na to da se opseg promjene tijekom vremena razlikovao između eksperimentalne i kontrolne skupine. Studenti koji su pohađali nastavu podržanu umjetnom inteligencijom pokazali su značajno veće dobitke u učinku u usporedbi s onima koji su pohađali tradicionalnu nastavu. Prosječno poboljšanje u eksperimentalnoj skupini iznosilo je $\Delta M = +9{,}34$ boda, dok su promjene u kontrolnoj skupini bile minimalne ($\Delta M = -0{,}21$), što dodatno potvrđuje učinkovitost intervencije.

Tablica 2

Interakciju između varijabli pratila je velika veličina učinka, što ukazuje na to da poboljšanja u eksperimentalnoj skupini nisu bila samo statistički značajna, već i pedagoški relevantna. S praktičnoga stajališta, ovi nalazi pokazuju da je nastava koja uključuje alate umjetne inteligencije omogućila izraženije dobitke u akademskome uspjehu tijekom semestra u usporedbi s tradicionalnim pristupima poučavanju matematike. Za procjenu veličine učinka korišten je djelomični η^2 koji je pokazao da interakcija objašnjava znatan dio varijance u ishodima učenja matematike. To naglašava značajan doprinos provedenoga programa poboljšanju postignuća učenika. U kontekstu obrazovnih istraživanja, takvi se rezultati obično smatraju uvjerljivim dokazom da je eksperimentalna metoda donijela obrazovne koristi koje nadilaze tipičan napredak zabilježen tijekom semestra.

Statistička analiza podupire zaključak da je uporaba UI-a pozitivno utjecala na razvoj proceduralne točnosti studenata i potaknula učinkovitije strategije rješavanja zadataka. Neposredna povratna informacija i strukturirana objašnjenja koje je pružao Socratic omogućili su učenicima da brže uoče pogreške, pregledaju zadatke vlastitim tempom i smanje vrijeme provedeno na manje produktivne metode pokušaja i pogrešaka. To može objasniti porast prosječnih rezultata u eksperimentalnoj skupini i zabilježenu smanjenu varijabilnost postignuća među njezinim sudionicima.

Jednako je vrijedan nalaz stabilnost rezultata u kontrolnoj skupini, što dodatno potvrđuje važnost intervencije. Da su obje skupine pokazale usporediv napredak, pripisivanje poboljšanja alatima umjetne inteligencije bilo bi upitno. Minimalne

promjene u kontrolnoj skupini upućuju na to da su značajni dobitci u eksperimentalnoj skupini vrlo vjerojatno posljedica intervencije, a ne prirodnoga tijeka standardnoga akademskog semestra.

Diskusija

Rezultati studije ukazuju na to da integracija alata za učenje podržanih umjetnom inteligencijom može utjecati na kognitivno angažiranje studenata i procese učenja matematike. Uočene promjene u kognitivnoj strukturi očituju se trokomponentnim mehaničkim mehanizmima. Konkretno, algoritamska eksternalizacija računalnih procesa može smanjiti vanjski kognitivni opterećenje, matematička radna memorija može se preusmjeriti s proceduralne izvedbe na konceptualno razumijevanje. Pojedinačne razvojne putanje proizlaze iz mehanizma prilagodljive regulacije tempa (Remoto, 2024).

Transmisivne obrazovne paradigme (prijenos znanja od stručnjaka prema početniku) ustupile su mjesto konstruktivističkim pristupima u kojima se znanje se aktivno izgrađuje putem posredovanom interakcijom. Pasivno primanje informacija zamijenjeno je aktivnom konstrukcijom znanja, dok se ovladavanje sadržajem proširilo na razvoj kognitivnih procesa (Jiayu, 2023). U području matematičkoga učenja točno su se pojavile tri ključne prilagodbe: pamćenje teorema razvilo se u praktičnu primjenu, pamćenje formula sada se temelji na konceptualnoj vizualizaciji, a standardizirani primjeri pretvorili su se u generativno rješavanje problema. Ove prilagodbe u učenju mogu poboljšati matematičko razumijevanje uz istovremeno očuvanje opsega gradiva (Yang, 2020).

Suvremena arhitektura učenja pomaknula se od sekvencijalne progresije prema umreženom konceptualnom istraživanju. Tradicionalni statični putovi postali su manje učinkoviti od fleksibilne navigacije znanjem (Knox, 2020). Konceptualne mreže zamjenjuju linearne kurikule, metakognitivna refleksija nadilazi puku pokrivenost sadržaja, a multimodalna predstavljanja istiskuju jednostrane načine prezentacije (Yang & Bai, 2020). Uloga edukatora također je doživjela paralelnu transformaciju. U modernim okružjima za učenje, nastavnici danas ponajprije djeluju kao posrednici u procesu učenja, a ne kao prenositelji znanja. Ova promjena predstavlja odmak od tradicionalnoga shvaćanja učitelja kao autoriteta za određeno područje prema novom poimanju njihove uloge kao vodiča u metakognitivnom razvoju učenika (Woodroffe i dr., 2021).

Metodološka inovacija očituje se u sustavnoj integraciji koja može pridonijeti preoblikovanju obrazovnih procesa u specifičnim nastavnim kontekstima. Ovaj se pristup razlikuje od modela dopunske implementacije (Saha i dr., 2023), iako opseg transformacije ovisi o kontekstualnim i institucionalnim čimbenicima. Obrazovni sustavi pokazuju različite stupnjeve usvajanja umjetne inteligencije – od tradicionalnoga perifernog proširenja do integriranijih pristupa koji mogu utjecati na odnose u učenju (Jiayu, 2023).

Usporedba ishoda učenja između kontrolne i eksperimentalne skupine pokazala je da su studenti koji su koristili okružje za učenje podržano umjetnom inteligencijom

postigli bolje rezultate iz matematike nakon intervencije. Ovi nalazi upućuju na to da alati temeljeni na umjetnoj inteligenciji mogu potaknuti dublje uključivanje u matematičke zadatke i da bi njihova integracija u redovite aktivnosti učenja mogla biti povezana s poboljšanim akademskim ishodima. Ovi su rezultati u skladu s prethodnim studijama koje pokazuju da okružja za učenje uz podršku umjetne inteligencije mogu poboljšati sposobnosti rješavanja problema i podržati individualizirane obrazovne putove (Richard i dr., 2022; Tashtoush i dr., 2024; Wang, 2023).

Ovi nalazi također rasvjetljavaju skalabilnost intervencija temeljenih na umjetnoj inteligenciji u različitim obrazovnim kontekstima. Istraživanjem kineskoga obrazovnog okružja, studija proširuje bazu dokaza za alate temeljene na umjetnoj inteligenciji, kao što je Socratic, izvan homogenih zapadnih populacija (Wu, 2021; Xie & Wang, 2024). Rezultati mogu ukazivati na potencijalnu primjenjivost učenja podržanoga umjetnom inteligencijom u kontekstima sličnim onima ispitivanim u ovoj studiji. Međutim, generalizaciju na druge populacije i obrazovne sustave treba provoditi s oprezom. Pokazivanje pozitivnih ishoda u ovome kontekstu pridonosi rastućem skupu dokaza o primjeni umjetne inteligencije u obrazovanju, no prije nego što se mogu potkrijepiti tvrdnje o širem implementiranju, potrebna su dodatna međukulturalna istraživanja. Uočene marginalne promjene u rezultatima kontrolne skupine dodatno naglašavaju potrebu za inovativnim obrazovnim alatima koji podupiru akademski napredak. Tradicionalne metode, kojima nedostaju interaktivni i tehnološki napredni sustavi, sklone su stvaranju statičnih okružja za učenje koja mogu ograničiti napredak, kao što je prethodno pokazano u ranijim studijama (Aljamali i Jabbar, 2020; Ee i Huh, 2018; Rane, 2023; Zhang, 2022).

Novost ovoga članka jest u ciljanoj procjeni utjecaja aplikacije Socratic na postignuća u matematici među sveučilišnim studentima u Kini. Za razliku od širega istraživanja o umjetnoj inteligenciji u obrazovanju, ovim istraživanjem procjenjuje jedan UI alat, pružajući uvid u njegove mogućnosti i učinkovitost u poboljšanju ishoda učenja. Dodatno, studija ispituje kratkoročne učinke integracije umjetne inteligencije u kontroliranom obrazovnom okružju, empirijski pokazujući poboljšane rezultate studenata. Ova studija doprinosi rješavanju praznine u literaturi pružajući empirijske podatke o korištenju specifičnoga alata umjetne inteligencije u sveučilišnoj nastavi matematike.

Nalazi ove studije imaju važne praktične implikacije, osobito za fakultete u ruralnim i područjima s manjkom resursa koji nastoje poboljšati akademski uspjeh studenata. Alati za učenje vođeni umjetnom inteligencijom, poput Socratica, mogu ublažiti neka ograničenja resursa u specifičnim obrazovnim kontekstima, iako njihova učinkovitost ovisi o uvjetima implementacije i institucionalnoj podršci. Umjetna inteligencija pruža personalizirane materijale za učenje, povratne informacije u stvarnom vremenu i postupno vođenje u rješavanju problema. Ova podrška posebno je vrijedna u okružjima gdje je individualna pozornost učitelja ograničena zbog velikih razreda ili geografskih ograničenja (Mimoudi, 2024). Integracija UI tehnologija u obrazovni okvir može doprinijeti poboljšanju ishoda učenja, iako se opseg toga učinka može razlikovati

ovisno o kontekstualnim čimbenicima. Nadalje, alati temeljeni na umjetnoj inteligenciji imaju potencijal proširiti pristup podršci u učenju za učenike koji nemaju pristup visokokvalitetnom poučavanju ili modernim obrazovnim resursima. Takvi alati mogu proširiti pristup podršci u učenju, no može li se njima osigurati jednake mogućnosti u različitim regijama, ostaje otvoreno pitanje koje zahtijeva daljnja istraživanja.

Ograničenja

Pri tumačenju nalaza ove studije treba uzeti u obzir nekoliko značajnih metodoloških ograničenja. Prvo, ograničeni uzorak obuhvaćao je studente sa samo četiriju sveučilišta u istočnoj Kini. Ovo ograničenje smanjuje općenitost nalaza na druge regije, obrazovne kontekste ili okružja koja su obilježena različitim tehnološkim i pedagoškim uvjetima. Nadalje, iako je primijenjeno stratificirano uzorkovanje kako bi se uzele u obzir ključne demografske značajke i početni akademski uspjeh, studija je provedena u autentičnim učilišnim okružjima. Kao rezultat toga, bilo je nemoguće u potpunosti isključiti utjecaj nekontroliranih vanjskih varijabli, kao što su razlike u motivaciji studenata, vještine samoreguliranoga učenja ili stavovi prema digitalnoj tehnologiji. Ti su čimbenici mogli pridonijeti varijacijama u angažmanu u intervenciji i, posljedično, u ishodima učenja.

Specifična priroda same intervencije zahtijeva razmatranje jer se temeljila na samostalnoj upotrebi aplikacije Socratic koju su studenti koristili izvan učionice. To je rezultiralo znatnom varijabilnošću u intenzitetu interakcije s alatom među sudionicima. Iako su podatci iz dnevnika aktivnosti bili dostupni, ti zapisi ne bilježe u potpunosti kvalitativne aspekte angažmana, kao što su dubina obrade informacija ili razina promišljanja tijekom izvršavanja zadatka. Još jedan važan čimbenik jest mogući učinak novosti: početni interes i povećani trud studenata mogli su biti potaknuti uvođenjem inovativnog digitalnog alata, što je moglo privremeno pridonijeti poboljšanju postignuća. Razlike u početnoj digitalnoj pismenosti sudionika mogle su dodatno utjecati na učinkovitost njihove interakcije s UI sustavom.

Iako su uspostavljene opće smjernice u vezi s učestalošću korištenja i ulogom aplikacije kao dodatnoga alata, stvarna razina angažmana studenata i način na koji je alat integriran u proces učenja mogli su varirati među nastavnicima i institucijama. Nadalje, unatoč konceptualnom razgraničenju uloga između nastavnika i alata temeljenoga na umjetnoj inteligenciji, to razgraničenje možda nije bilo u potpunosti standardizirano u stvarnoj obrazovnoj praksi, čime se komplicira precizna procjena „čistoga" učinka intervencije.

Dodatno ograničenje odnosi se na postupak praćenja korištenja aplikacije. Iako je pridržavanje protokola intervencije procjenjivano putem dnevnika aktivnosti i dobrovoljnoga pristanka sudionika, ovaj je pristup ograničen s obzirom na potpunost praćenja i može biti osjetljiv na samoregulaciju studenata. Štoviše, takve metode praćenja mogu izazvati zabrinutost vezanu uz privatnost i etičke aspekte prikupljanja digitalnih podataka, unatoč pridržavanju načela informiranoga pristanka i povjerljivosti.

Također treba napomenuti da su ishodi učenja ocijenjeni odmah po završetku 16-tjednog razdoblja intervencije. To onemogućuje procjenu dugoročnoga zadržavanja

znanja ili utvrđivanje predstavljaju li zabilježeni dobici trajna obrazovna postignuća ili prolazne učinke. Na kraju, iako su standardizirani ispiti korišteni radi povećanja pouzdanosti nalaza, dobivene ocjene mogu djelomično odražavati razlike u akademskim standardima i praksama ocjenjivanja među sveučilištima koja su sudjelovala u ovome istraživanju.

Zaključak

U ovom je članku ispitana upotreba Googleove aplikacije Socratic u nastavi matematike na visokoškolskim ustanovama u Kini. Nalazi ukazuju na to da je učenje potpomognuto umjetnom inteligencijom u kontekstu ove studije bilo povezano s većim akademskim napretkom u usporedbi s tradicionalnim metodama poučavanja. Tijekom semestra studenti u eksperimentalnoj skupini pokazali su veće poboljšanje u postignućima iz matematike od svojih kolega u kontrolnoj skupini. Ovaj nalaz podupire potencijal aplikacije Socratic kao dodatnoga alata u nastavi matematike na sveučilišnoj razini, iako su potrebna daljnja istraživanja kako bi se potvrdila njezina učinkovitost u različitim kontekstima. S praktičnoga stajališta, studija ukazuje na to da pristupačni sustavi za poučavanje vođeni umjetnom inteligencijom mogu pomoći sveučilištima u poboljšanju obrazovnih ishoda na kolegijima s velikim brojem upisanih studenata, pri čemu je personalizirana povratna informacija instruktora često ograničena.

Budući smjerovi istraživanja uključuju longitudinalne analize obrazaca zadržavanja znanja izvan intervencija ograničenih na jedan semestar. Još jedan obećavajući smjer jest analiza prenosivosti specifične za određena matematička područja. Međukulturalne komparativne studije pružile bi dublji uvid u učinkovitost primjene u različitim obrazovnim sustavima. Također bi bilo korisno razviti sustave procjene koji obuhvaćaju kognitivne dimenzije izvan tradicionalnih pokazatelja uspješnosti te istražiti čimbenike institucionalne spremnosti koji pridonose uspješnoj primjeni.