# A literature Review of the Mediators of AI-Assisted STEM Teaching on the Academic Performance of K12 Students

Anqi Liu *

Graduate School, The Education University of Hong Kong, Hongkong, China

* Corresponding Author Email: s1163597@s.eduhk.hk

**Abstract.** Under the traditional teaching mode, mathematics education has long faced multiple challenges. Due to the various practical conditions, schools are often forced to replace systematic curriculum reform with temporary and patchwork teaching activities, resulting in students' excessive cognitive load and limited learning effect. Based on the cognitive load theory, this paper systematically discusses the application potential of artificial intelligence (AI) in mathematics education and its practical impact on students' cognitive load and academic performance. Empirical research shows that AI can effectively reduce students' external cognitive load and promote more efficient allocation of cognitive resources by providing personalized learning paths, adaptive content adjustments, and instant feedback mechanisms, thereby improving the effectiveness of mathematics learning. In particular, AI has shown significant potential in narrowing the achievement gap between students of different genders, races, and socioeconomic backgrounds. However, if AI systems are improperly designed or overused, they may also add additional cognitive burden, leading to distraction and reduced teaching effectiveness. Therefore, the future development of AI in the field of education urgently needs to further integrate educational psychology theories, optimize human-machine collaboration strategies in the learning and teaching process, adhere to the principle of learner-centered design, and systematically improve its educational effectiveness.

**Keywords:** Artificial intelligence, STEM Teaching, Academic Performance, K12 Students.

## 1. Introduction

### 1.1. Research Background

In recent years, with the rise of artificial intelligence technology, its adaptive advantages have become increasingly apparent, and it has great potential in reducing the cognitive load of learners. In school education, AI acts as a cognitive companion to help reduce learners' cognitive burden by handling formatting tasks and freeing up cognitive resources for unstructured tasks. In terms of personalized learning support, intelligent tutoring systems can dynamically adjust the difficulty of learning materials based on learners' real-time performance to reduce intrinsic cognitive load. At the same time, AI can provide targeted feedback, including step-by-step guidance or Socratic dialogue, to reduce external cognitive load and enhance learning engagement. AI also can provide targeted feedback, including step-by-step guidance or Socratic dialogue, to reduce external cognitive load and enhance learning engagement. However, using AI tools also carries the risk of increasing cognitive load. AI-generated content is too fast or lacks integration, which can easily lead to learners' working memory overload. In the context of learning using AI tools, learners need to continuously judge when and how to use AI, and judge the authenticity of the content they output, which consumes their cognitive resources.

Mathematics provides the necessary analytical tools and thinking patterns for scientific, engineering, and technical problem solving, including modeling, data reasoning, quantitative analysis, etc. However, in actual teaching interventions, mathematical content is often in a subordinate position. Mathematical literacy should not be limited to arithmetic ability, but should also encompass skills such as data processing, information evaluation, inference, and decision-making in real situations. Students' ability to deal with contradictions and unreliable information relies heavily on mathematical reasoning training, which also prepares them for solving complex real-world problems. Therefore, in STEM integration, mathematics must shift from an "implicit tool" to an "explicit goal", highlighting its role through deliberately designed teaching activities such as modeling, data reasoning, cost-benefit analysis (English, L. D., 2016).

AI technology has potential in supporting the improvement of academic performance in mathematics, but its application in the K12 stage faces multiple challenges such as lack of research, stage mismatch, single strategy, data limitations, and goal deviation. Future research needs to focus on developing AI technologies and application models suitable for K12 mathematics teaching scenarios, which can be deeply integrated with inquiry-based teaching methods, and can effectively promote the development of higher-order thinking, and conduct rigorous empirical verification of the effect (Xu, W., & Ouyang, F., 2022).

### 1.2. Concept Definition

Cognitive load theory states that students' learning effects are limited by working memory capacity. When learners process new, unorganized information, working memory capacity (about 4±1 unit of information) and duration (about 20 seconds) are extremely limited. This limitation is particularly pronounced when dealing with multiple elements of information that need to be associated simultaneously (i.e., high "element interactivity"). Working memory is primarily responsible for the temporary storage and manipulation of information, and its limited resources are a key bottleneck in learning (Sweller, 2011). It is primarily responsible for the temporary storage and manipulation of information, and its limited resources are a key bottleneck in learning. The effective transfer of processed information into long-term memory is crucial for meaningful learning.

Cognitive load theory has guiding significance in mathematical knowledge learning. The teaching design should be adapted to the learner's prior knowledge level. Learners with low prior knowledge can reduce their external cognitive load and promote the construction of basic schemas through complete problem-solving examples (showing all steps). Learners with advanced prior knowledge stimulate the genitative cognitive load by completing examples (leaving some steps blank) to deepen their understanding and transfer ability. At the same time, the difficulty of the task needs to optimize the allocation of cognitive resources. The intrinsic cognitive load is determined by the intrinsic difficulty of the task, and it is necessary to release working memory resources by reducing the extrinsic load, ensure the space for generative load, and improve effective cognitive investment in combination with learning motivation.

### 1.3. Research Questions

The impact of artificial intelligence (AI) on assisted teaching and learning is mostly focused on higher education or general education contexts, and mainly focuses on its direct effects on cognitive load, while empirical discussions in primary schools, especially mathematics STEM (science, technology, engineering, and mathematics) curriculum, are still insufficient. To this end, this paper aims to analyze previous articles and test whether cognitive load plays a mediating role between AI-assisted teaching and students' math performance.

## 2. The Impact of AI-assisted STEM Instruction on Math Achievement

In traditionally teacher-dominated teaching environments, students' math performance is significantly influenced by interactions such as their race, gender, and school socioeconomic status (SES). Smart tutoring systems demonstrate unique potential in empowering disadvantaged student populations by providing personalized feedback and adaptive learning pathways, thereby achieving more equitable educational outcomes.

Artificial intelligence has a modest yet significant positive effect on primary school students' math performance, but its effect is moderated by math learning topics and grade level, while other factors such as intervention duration, AI type, and learning organization form are not significant.

## 2.1. Research Design and Methodology

Hwang's study was based on 30 independent samples from 21 studies, which used a random-effects model and calculated the effect size, which provided a strong generalization of the macro trend for the development of AI-assisted technology in the field of mathematics (Hwang et al., 2021). However, the conclusions of the study were highly dependent on the quality and completeness of the reports of existing studies, and lacked original analysis. In contrast, Huang's study used a quasi-experimental design with a sample of sixth-grade students in the United States and paid special attention to race, gender, and socioeconomic status (SES) differences in the selection process to improve representativeness and inference strength. In terms of data analysis, the study tested the math performance of the subjects before and after the experiment, and used analysis of covariance (ANCOVA) to test the interaction between factors, thus providing empirical support for "AI-assisted teaching has a positive impact on mathematics performance" from a micro-level mechanism. In summary, two studies show that the positive impact of AI on math performance is not only universal across contexts, but also shows significant effects in specific practical situations (such as reducing performance differences between different groups).

## 2.2. Differences between AI Types and Teaching Scenarios

Hwang’s research covers a variety of AI educational applications, including intelligent learning systems (ITS), adaptive learning systems (ALS), and educational robots, and conducts empirical analysis in various scenarios such as classroom teaching, after-school projects, and personalized learning (Hwang et al., 2021). Through meta-analysis, the study found that the effect of artificial intelligence intervention in mathematics learning peaked in short-term intervention of 1 to 5 hours, and the longer duration was accompanied by a decrease in effect. At the same time, AI did not bring significant differences in effect between personalized learning and group learning, indicating that AI type did not become a significant moderating variable in this study.

In contrast, Huang et al.'s study focused on a single type of AI system, the ITS (ALEKS system), and deployed it in an after-school learning context to completely replace teacher teaching. In this quasi-experimental design, the system provides a fully individualized learning path, with the instructor playing only a supporting role. The empirical results show that even when the intervention period is extended to 48 days, ITS maintains a stable and significant positive effect, especially in narsrowing the gap in mathematics achievement among students of different gender, races, and socioeconomic status.

Both studies show that the role of artificial intelligence in education is highly dependent on application scenarios and implementation methods. Hwang's research reveals the overall trends and boundary conditions of AI intervention from multiple types and a wide range of scenarios, while Huang's research highlights the micro-mechanisms of ITS in reducing performance gaps through high-precision contextual control (Hwang et al., 2021). This complementary evidence system shows that different types of AI and their applicable models can be organically combined in educational practice to achieve more comprehensive and sustainable learning effect improvement.

## 2.3. Key Factors Affecting the Effect and Educational Equity

In terms of moderating variables, Hwang's research showed that the improvement effect of AI-assisted teaching on mathematics performance was significantly moderated by mathematics topics and grade level, indicating that the effect varied according to subject content and learner development stage (Hwang et al., 2021). Huang et al. found that AI-assisted teaching was significantly moderated by variables such as race, gender, and school socioeconomic status (SES) in reducing the gap in math achievement.

From the perspective of educational equity, Hwang's research does not directly assess the gap reduction, but points out that AI may have potential benefits for disadvantaged learners; Instead, Huang's study provides clear evidence that ITS interventions can significantly reduce achievement disparities due to race, gender, and SES (Hwang et al., 2021)..

In terms of mechanism explanation, Hwang emphasized the role of personalized feedback and constructivist learning mechanism in improving mathematics performance from the cognitive and emotional dimensions. Huang's research based on the perspective of social equity, proposed an AI mechanism based on accurate knowledge status diagnosis and unbiased feedback, which effectively promoted the realization of educational equity (Hwang et al., 2021).

Together, both studies show that AI not only improves overall academic performance but also has significant potential to promote educational equity with long-term implications for narrowing systemic disparities. Hwang's research explains the universal benefits of AI from the perspective of cognitive and emotional mechanisms, while Huang's research provides empirical support for AI to reduce systemic achievement gaps from the social level, reflecting an explanatory framework for multi-level mechanisms complementary to each other (Hwang et al., 2021).

# 3. Mediating Role of Cognitive Load

## 3.1. The Degree and Method of AI Technology Intervention

Hwang developed an adaptive learning system (ACALS) based on a fuzzy expert system, which dynamically adjusts learning materials by analyzing students' cognitive performance (e.g., correct answering rate) and emotional state (e.g., concentration, patience, willingness to learn). The system uses three versions of materials with different difficulty levels (basic, standard, and advanced) and uses fuzzy reasoning rules for real-time decision-making (Hwang et al., 2020). AI technology is the core driving mechanism in this system and is deeply integrated into the learning process.

Van de Weijer-Bergsma & Van der Ven adopted a content personalization strategy, dividing students into personalized and non-personalized groups through an early interest questionnaire, and students in the personalized group saw content related to their interests (such as favorite objects, names) in math problems (Van de Weijer-Bergsma & Van der Ven, 2021).The study does not use complex AI algorithms, but achieves personalization through simple conditional assignment, and the degree of AI technology is less involved, closer to situational manipulation in traditional experimental design.

Sweller systematically reviewed and expanded the cognitive load theory (CLT), highlighting principles such as modality, segmentation, and worked-example effects that inform AI instructional design, emphasizing the mediating pathway of instructional design affecting learning effects by regulating cognitive load (Sweller et al., 2019). Although this study does not directly involve artificial intelligence technology, the various cognitive load effects (such as work sample effect, modal effect, segmentation effect, etc.) provide a theoretical basis for the design of AI-assisted teaching. For example, AI can effectively reduce external cognitive load and optimize resource allocation through adaptive content presentation, multimedia integration, and personalized feedback, thereby improving learning efficiency.

## 3.2. Validation Results of Cognitive Load

Hwang found that although the system of emotion and cognitive adaptation significantly improved students' math performance and reduced math anxiety, there was no significant difference in cognitive load between the three groups (Hwang et al., 2020). The researchers believe that the moderate difficulty design and material arrangement of the system keep the cognitive load within a reasonable range without overload, so the cognitive load does not show a significant mediating effect in this system.

Van de Weijer-Bergsma & Van der Ven examined the mediating effect of cognitive load through two studies (N=238 and N=149). The results showed that although cognitive load could significantly negatively predict math performance (path B was significant), personalized intervention did not significantly reduce cognitive load (path A was not significant), so cognitive load did not play a mediating role (Van de Weijer-Bergsma & Van der Ven,2021). The researchers pointed out that it may be due to the low difficulty of the questions, the lack of personalization or the limitation of the measurement method.

Endres directly explored the mechanism of prior knowledge on cognitive load through two empirical studies (Endres et al., 2023). The study found that in some contexts (such as mathematical problems with apparent simplicity but actually complex mathematical problems), high prior knowledge learners reported higher intrinsic cognitive load due to their ability to recognize the true complexity of the task. This result supports the mediating pathway of "prior knowledge-cognitive load-task performance" and is validated by between-group comparison and repeated measures ANOVA. Although the study did not directly use AI intervention, its pre-training design (e.g., video-guided instruction) can be seen as a preliminary "intelligent" teaching support, suggesting that AI systems can indirectly modulate cognitive load levels by activating prior knowledge.

### 3.3. Research Methods and Sample Characteristics

Hwang employed a quasi-experimental design with a sample of Taiwanese sixth-grade students (N=162) and divided them into three groups (dual adaptation group, cognitive adaptation only group, traditional system group), relying on system log data for behavioral sequence analysis, combined with ANCOVA to test the differences between groups (Hwang et al., 2020).

Van de Weijer-Bergsma & Van der Ven used a randomized experimental design with a sample of sixth grade students in the Netherlands (Study 1: N=238; Study 2: N=149), using PROCESSES macros for moderated mediation analysis, controlling mathematics, reading, and working memory abilities, and having strong causal inference ability (Van de Weijer-Bergsma & Van der Ven, 2021).

Sweller relied on years of empirical research for theoretical integration, covering multidisciplinary and multi-age samples, with strong generalization and generalization (Sweller et al., 2019). Endres used a specific experimental design, with samples from German adults and Australian primary school students, focusing on the two fields of forestry and mathematics, providing more direct evidence of mediating effects (Endres et al., 2019). Both studies used subjective cognitive load scales, which guaranteed consistency of measurement tools.

### 3.4. Summary

Cognitive load plays a key mediating role in mathematics learning, and its specific role is significantly moderated by intervention methods, task difficulty and individual differences. Sweller provide a broad theoretical framework and cognitive load design principles (Sweller et al., 2019). Endres further provide empirical support in mathematical learning contexts, which together lay a theoretical and methodological foundation for understanding the mediating mechanism of cognitive load in AI-assisted teaching (Endres et al., 2023). These studies suggest that AI can effectively regulate students' cognitive load levels through personalized content delivery, complexity recognition, and adaptive teaching, thereby affecting their mathematics academic performance. Future research needs to combine high-intervention AI technology with more refined mediating variable measurement methods to further clarify the specific role paths and boundary conditions of cognitive load in personalized learning, offering practical insights for instructional design.

## 4. Discussion and Revelation

### 4.1. Theoretical Reinterpretation of Mediation Mechanism

In recent years, the application of generative artificial intelligence in the field of education has spread rapidly, but most of its practices are still fragmented and unstructured, and even increase the cognitive burden of students to a certain extent. Due to the ability of artificial intelligence to generate a large amount of complex information in a short period of time, and the lack of effective knowledge organization and structured support, novice learners are prone to cognitive overload when processing such information, which affects their absorption and integration of knowledge.

From the perspective of cognitive load theory, cognitive load in the learning process can be divided into three types: intrinsic cognitive load, extrinsic cognitive load and related cognitive load. The intrinsic cognitive load is determined by the complexity and difficulty of the learning material itself. The extrinsic cognitive load stems from improper instructional design and information presentation, such as redundant or chaotic content organization. Related cognitive load refers to the active cognitive input used to construct schemas and deepen understanding, which helps learners consolidate and transfer knowledge.

In response to the cognitive overload problem that generative AI may cause, load reduction pedagogy provides a structured pedagogical intervention path. This pedagogy systematically reduces the extrinsic cognitive load and optimizes the relevant cognitive load through phased strategies such as difficulty adjustment, providing structured support, establishing multiple connections, providing corrective feedback, and guiding forward, so as to support learners to gradually transition from relying on external assistance to independent inquiry. Empirical studies have shown that structured and phased teaching support can significantly alleviate the cognitive overload caused by generative AI and promote more efficient knowledge absorption and ability development.

### 4.2. Practical Enlightenment

In order to effectively integrate generative AI into the teaching environment, developers need to systematically optimize the educational support functions of AI according to the principle of Load Reduction Instruction (LRI) (Martin et al., 2025). Based on students' real-time academic performance, AI can dynamically adjust the difficulty of questions, build personalized learning paths, and provide structured scaffolding through Socratic dialogue, effectively suppressing the additional external cognitive load caused by improper teaching organization and promoting knowledge absorption. At the same time, as a "helper" of students' cognitive activities, generative AI can also assist learners in establishing connections between knowledge and promoting their schema construction and integration, which is embodied in generating adaptive exercises, providing instant feedback, and simulating teacher correction. In particular, AI should gradually guide students to formulate effective prompts and cultivate their independent inquiry ability in a subtle way.

However, the effective implementation of technical functions is highly dependent on the acceptance and integration ability of teachers. Pörn & Braskén's research shows that teachers' attitudes towards AI are a key factor in the success of their educational applications; Even if the technical tools are perfect, if teachers lack the willingness and integration ability to use AI, it is still difficult to truly embed AI in teaching practice (Pörn, R., Braskén, M., Wingren, M., & Andersson, S., 2024). Therefore, teachers should play the role of "intermediary" to carefully evaluate the quality and timing of AI-generated content to ensure that its output meets teaching objectives. In subjects such as mathematics, teachers must especially clarify the positioning of AI assistance - that is, to help students understand concepts and improve their reasoning skills, and be wary of over-reliance on technology leading to deviations from the core literacy goals of the subject.

In order to achieve a wide and deep integration of generative AI and education, relevant institutions and enterprises should strengthen the practical training of teachers and improve their AI tool design capabilities, output monitoring and interpretation capabilities. Teachers also need to coordinate the use of AI and core literacy training in curriculum design, strengthen data ethics awareness, and effectively protect students' privacy and educational data security.

## 5. Conclusion

As an auxiliary teaching tool, the effectiveness of generative AI is highly dependent on reasonable application strategies. Empirical studies have shown that it shows more significant effects on specific subject topics (such as decimals, geometry, etc.) and in senior teaching environments, mainly because senior students usually have some cognitive self-regulation skills and can benefit from the personalized and inquiry-based support provided by AI. However, in the lower grades, due to students' short attention spans, executive functions and self-management skills, teachers should use AI prudently to avoid distracting students due to premature or excessive technical intervention, and replace necessary teacher-student interaction and embodied participation. Instead, students should be more guided to integrate into structured classroom activities and maintain their focus through moderate "freshness" design, and limited and precise AI interventions in the short term are often more effective than long-term continuous use.

Generative AI does have the potential to become a powerful educational tool, but its design and application must be systematically integrated into educational psychology theories (such as cognitive load theory, scaffolding teaching, etc.) to effectively reduce external cognitive load and support learning paths that conform to individual cognitive differences. Future R&D should pay more attention to education-oriented prompt engineering, design highly structured interactive processes, and enhance AI's ability to discriminate and adapt to students with different cognitive levels, so as to promote its transformation from technical function to educational effectiveness on an empirical basis.

## References


[1] Endres, T. Lovell, O., Morkunas, D., Rieß, W., & Renkl, A. (2023). Can prior knowledge increase task complexity? – Cases in which higher prior knowledge leads to higher intrinsic cognitive load. British Journal of Educational Psychology, 93 Suppl. 2), 305–317.

[2] English, L. D. (2016). STEM education K-12: Perspectives on integration. International Journal of STEM Education, 3(3), Article 3. 1–8.

[3] Hwang, G.-J., Sung, H.-Y., Chang, S.-C., & Huang, X.-C. (2020). A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. Computers and Education: Artificial Intelligence, 1, Article 100003.

[4] 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.

[5] Huang, X., Craig, S. D., Xie, J., Graesser, A., & Hu, X. (2016). Intelligent Tutoring Systems Work as a Math Gap Reducer in 6th Grade After-School Program. Computers in Human Behavior, 59, 339–346.

[6] Martin, A. J., Collie, R. J., Kennett, R., Liu, D., Ginns, P., Sudimantara, L. B., Dewi, E. W., & Rüschenpöhler, L. G. (2025). Integrating generative AI and load reduction instruction to individualize and optimize students' learning. Learning and Individual Differences, 121, 102723.

[7] Pörn, R., Braskén, M., Wingren, M., & Andersson, S. (2024). Attitudes towards and expectations on the role of artificial intelligence in the classroom among digitally skilled Finnish K-12 mathematics teachers. LUMAT: International Journal on Math, Science and Technology Education, 12(3), 53–77.

[8] Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. Educational Psychology Review, 31(2), 261–292.

[9] Van de Weijer-Bergsma, E., & Van der Ven, S. H. G. (2021). Why and for whom does personalizing math problems enhance performance? Testing the mediation of enjoyment and cognitive load at different ability levels. Learning and Individual Differences, 87, 101982.

[10] Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. International Journal of STEM Education, 9(1), Article 59