Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute InformingScience.org

JITEResearch.org

Volume 24, 2025

CHATGPT SCAFFOLDING IN SUPPORTING METACOGNITION FOR LIMIT CONCEPTS IN GUIDED INQUIRY MATHEMATICS LEARNING

Nizlel Huda*	Universitas Jambi, Jambi, Indonesia	nizlel.huda@unja.ac.id
Khairul Anwar	Universitas Jambi, Jambi, Indonesia	mathanwar@unja.ac.id
Novferma	Universitas Jambi, Jambi, Indonesia	novferma@unja.ac.id
Wawan Kurniawan	Universitas Jambi, Jambi, Indonesia	kurniawan_wawan@unja.ac.id

* Corresponding author

ABSTRACT

Aim/Purpose	This study aims to investigate how ChatGPT-mediated scaffolding supports students' metacognitive skills (planning, monitoring, and evaluating strategies) in understanding limit concepts in calculus within a guided-inquiry learning environment.
Background	Guided inquiry fosters conceptual understanding in calculus, yet students often struggle with metacognitive regulation. While AI tools like ChatGPT offer interactive scaffolding, their impact on students' self-regulated learning and problem-solving strategies in abstract topics, such as limits (a fundamental concept in calculus), remains underexplored. This study addresses this gap by evaluating ChatGPT's function as a metacognitive guide in mathematics learning.
Methodology	A convergent mixed-methods design was implemented with 75 students of mathematics education at Universitas Jambi over a period of four weeks. Participants engaged in guided inquiry activities on limits, using ChatGPT for problem-solving and reflection. Data was collected through pre- and post-metacognitive assessments, screen recordings of ChatGPT-student interactions, and reflective journals. Quantitative data were analyzed using paired t-tests, while qualitative data were thematically coded to identify patterns in metacognitive engagement.

Accepting Editor Janice Whatley │Received: February 5, 2025│ Revised: June 4, July 15, August 26, October 3, October 17, 2025 │ Accepted: October 21, 2025.
Cite as: Huda, N. Anwar, K., Novferma, & Kurniawan, W. (2025). ChatGPT scaffolding in supporting metacognition for limit concepts in guided inquiry mathematics learning. *Journal of Information Technology Education: Research, 24,* Article 38. https://doi.org/10.28945/5645

Contribution	This study advances understanding of AI's capacity to foster self-regulated learning and critical thinking in mathematics, providing a framework for integrating generative AI as a metacognitive partner in guided inquiry pedagogy.
Findings	Results indicate significant improvements in metacognitive skills, particularly in monitoring and evaluation strategies. Qualitative analysis revealed that ChatGPT's iterative feedback encouraged students to critically analyze solutions, particularly in identifying boundary conditions in limit problems. However, 28% of students passively accepted AI-generated answers without deeper scrutiny, highlighting variability in engagement levels.
Recommendations for Practitioners	Educators should integrate ChatGPT as a reflective tool in guided inquiry, designing structured activities that require students to justify or challenge AI-generated outputs. Providing explicit training in critical questioning techniques can enhance AI's pedagogical value.
Recommendations for Researchers	Future research should explore long-term retention of metacognitive skills developed through AI scaffolding and adaptive AI models for optimizing ChatGPT-student interactions in mathematics education.
Impact on Society	The implications of this research extend beyond the classroom, potentially reshaping mathematics education in higher education. This approach could democratize access to personalized mathematical support, reduce educational inequalities, and prepare students for an AI-augmented professional landscape. However, careful consideration must be given to ethical implementation and the preservation of authentic mathematical thinking skills.
Future Research	Further studies should examine (1) the sustainability of AI-enhanced metacognitive development, (2) cross-cultural differences in AI scaffolding effectiveness, and (3) improved AI-driven adaptive learning strategies for mathematics education.

Keywords	ChatGPT, metacognition, scaffolding patterns, guided inquiry, mathematics learning, limit concepts, AI ethics

INTRODUCTION

The rapid evolution of artificial intelligence (AI) has ushered in a new era of educational innovation, with generative models like ChatGPT redefining the boundaries of personalized learning and cognitive support. As AI permeates classrooms, its potential to simulate human-like interactions and deliver adaptive feedback presents unprecedented opportunities to address longstanding pedagogical challenges (Gill et al., 2024; Michel-Villarreal et al., 2023). In mathematics education, where abstract concepts and complex problem-solving require high cognitive flexibility, the integration of AI tools has been heralded as a means to democratize access to quality instruction and foster deeper conceptual understanding (Wardat et al., 2023). Yet, the true transformative power of these technologies lies not merely in their ability to automate tasks, but in their capacity to cultivate metacognitive skills, self-regulation, critical reflection, and strategic adaptability – essential for lifelong learning (Anderson, 1988; Djudin, 2018; Duckworth et al., 2021).

Metacognition, the ability to monitor and regulate one's own thinking processes, is particularly vital in guided inquiry learning, a pedagogical approach that balances structured guidance with student-driven exploration, hypothesis testing, and iterative problem-solving. The effectiveness of guided inquiry hinges on educators' ability to provide systematic, scaffolded support that guides learners through cycles of planning, executing, and evaluating their strategies (Gerard et al., 2016; Pedaste et al., 2015). However, in heterogeneous or resource-constrained classrooms, sustaining such individualized support remains a significant challenge. Students often struggle with cognitive overload when navigating complex mathematical tasks, particularly in understanding abstract concepts such as limits, which can lead to disengagement or superficial learning (Swinyard & Larsen, 2012). Traditional digital tools, while useful for procedural practice, frequently fall short in fostering the reflective dialogue necessary for metacognitive growth, highlighting a critical need for more dynamic, responsive interventions.

Generative AI, exemplified by ChatGPT, emerges as a promising solution to this impasse. Unlike static tutoring systems, ChatGPT's capacity for natural language processing enables dialogic interactions that mimic human mentorship, offering explanations, prompting reflection, and adapting to learners' unique needs. Recent empirical studies demonstrate its efficacy in enhancing procedural skills, such as solving algebraic equations (Pardos & Bhandari, 2023) and generating practice problems (da Silva et al., 2024). In language education, it has been leveraged to refine writing through iterative feedback (Lelepary et al., 2023; Wei, 2023) and simulate conversational practice (Deniz, 2023). However, its application in mathematics education remains disproportionately focused on algorithmic tasks, neglecting the richer, more nuanced demands of metacognitive scaffolding within guided inquiry frameworks. This narrow utilization overlooks ChatGPT's potential to facilitate higher-order thinking processes, particularly in challenging areas such as limit concepts, where students require structured guidance to develop a robust conceptual understanding.

Systematic reviews have highlighted significant gaps in our understanding of how AI can effectively support guided inquiry in mathematics education. While ChatGPT excels in generating content and providing immediate answers, its ability to facilitate structured exploration and critical engagement – cornerstones of guided inquiry – remains underexplored (Dau et al., 2024; Qureshi, 2023). For instance, in calculus education, students often struggle with the abstract nature of limits, requiring careful scaffolding to bridge intuitive understanding with formal mathematical concepts. Traditional approaches to teaching limits often emphasize procedural fluency over conceptual understanding, leading to persistent misconceptions and difficulties in applying limit concepts in novel contexts (Beynon & Zollman, 2015; Jones, 2015). The potential for AI-mediated scaffolding to address these challenges through structured guidance and adaptive support represents a promising yet understudied area of research.

The integration of ChatGPT in mathematics education raises important questions about the balance between technological support and the development of authentic mathematical thinking. The risk of student overreliance on AI-generated solutions, as noted by Sánchez-Ruiz et al. (2023) and Joshi et al. (2024), threatens to undermine the very metacognitive skills guided inquiry aims to develop. When learners perceive ChatGPT as a shortcut to answers rather than a tool for guided exploration, they may bypass critical stages of conceptual development and self-assessment, resulting in fragmented understanding. Furthermore, issues of academic integrity, biased content generation, and unequal access to technology exacerbate anxieties about AI's role in education (Hadi Mogavi et al., 2024; Pepin et al., 2025). Without deliberately structured pedagogical frameworks to guide its use, ChatGPT risks perpetuating passive learning behaviors, contradicting the active, guided nature of effective mathematical inquiry.

To address these gaps, this study investigates ChatGPT's role as a metacognitive scaffold in undergraduate calculus education, with a specific focus on limit concepts through a guided inquiry approach. Based on the identified gaps, this study addresses three specific research questions.

1. How can ChatGPT-mediated scaffolding be systematically integrated within guided inquiry frameworks to enhance metacognitive skill development in limit concept learning?
2. What specific scaffolding patterns emerge from ChatGPT-student interactions during guided inquiry mathematics learning, and how do these patterns impact metacognitive awareness across planning, monitoring, and evaluation domains?

3. To what extent do ChatGPT-supported guided inquiry approaches challenge existing assumptions about AI's role in mathematics education, particularly regarding student autonomy and authentic mathematical thinking development?

LITERATURE REVIEW

This literature review examines the current state of research on AI-mediated scaffolding in mathematics education, focusing on metacognitive development and guided inquiry learning. The review is organized into key sections covering the evolution of AI in education, theoretical foundations of scaffolding, metacognitive support mechanisms, guided inquiry approaches, and the specific challenges of teaching limit concepts in calculus.

AI in education: From content delivery to cognitive scaffolding

The integration of AI in mathematics education has evolved significantly from early intelligent tutoring systems (ITS) focused on procedural skill development (Hwang & Tu, 2021) to sophisticated generative models, such as ChatGPT, that enable dialogic, context-aware interactions (Van Vaerenbergh & Pérez-Suay, 2022). Recent studies emphasize AI's potential to personalize learning, automate feedback, and reduce educators' administrative burdens (Vasconcelos & dos Santos, 2023). However, research remains skewed toward language education, where ChatGPT has been widely adopted for writing support and conversational practice (Firat, 2023; Michel-Villarreal et al., 2023). In mathematics education, applications are often limited to content generation and procedural problem-solving, with minimal exploration of AI's role in fostering metacognitive development through guided inquiry approaches (Flaagan, 2023). The trans-formative potential of AI in mathematics education lies not merely in its ability to generate content or solve problems, but in its capacity to support deeper conceptual understanding and metacognitive development through structured, interactive engagement.

The synthesis of existing literature reveals significant gaps in our understanding of AI-supported mathematics education, particularly in the context of guided inquiry and learning concepts of limits. While recent studies have made progress in understanding AI's potential in education (Dimeli & Kostas, 2025; Gill et al., 2024; Kofahi & Husain, 2025; Michel-Villarreal et al., 2023), research predominantly focuses on AI's role in content delivery rather than its potential for deeper pedagogical integration (Dimeli & Kostas, 2025). Studies by Adiredja (2021) note that existing research primarily addresses basic problem-solving applications, neglecting the complexities of scaffolding higher-order thinking skills in mathematics education, particularly when dealing with abstract concepts like limits.

Theoretical foundations of AI-mediated scaffolding

Building upon Vygotsky's (2018) theory of the Zone of Proximal Development (ZPD) and modern frameworks for guided mathematical inquiry (Yerushalmy, 2015), recent research has expanded theoretical understanding of AI-mediated scaffolding beyond traditional conceptualizations. Walter (2024) provides crucial theoretical grounding by explicitly connecting ZPD to AI scaffolding, positioning AI as a "more competent other" that delivers personalized learning experiences adapted to individual developmental levels. This foundational work establishes AI's legitimacy within established learning theory while demonstrating measurable learning gains ($d = 0.65$-0.82) when properly calibrated to student developmental needs.

The integration of Activity Theory represents another sophisticated theoretical advancement. Loc et al. (2022) apply Engeström's Activity Theory to design AI-enhanced mathematics teaching through a three-phase model, demonstrating 23% improvement in problem-solving performance compared to traditional methods. This research positions AI as a mediating tool within the activity system, facilitating relationships between students, mathematical knowledge, and learning goals through intelligent scaffolding mechanisms.

Gabriel et al. (2025) apply Pekrun's Control-Value Theory to AI-mediated mathematics education, demonstrating that mathematics anxiety emerges when students value mathematics but perceive a lack of control over performance. Their framework positions AI as a mechanism to enhance students' perceptions of competence and control while reducing anxiety through personalized scaffolding ($\eta^2 = 0.24$ for anxiety reduction). This theoretical advancement addresses gaps identified by Anderson (1988) and Djudin (2018) regarding the emotional dimensions of metacognitive development in technology-enhanced learning environments.

Advanced scaffolding patterns in AI-mediated learning

The methodologically sophisticated mixed-methods research reveals a nuanced understanding of AI scaffolding processes that extend beyond traditional quantitative measures. Adair et al. (2023) employed real-time AI-driven assessment systems with middle school students, combining log data analysis, performance assessments, learning analytics, and discourse analysis. Their multimodal approach demonstrated significant improvement in mathematical modeling practices in near-transfer scenarios while revealing how AI scaffolding effectively addresses individual student difficulties.

Malik et al. (2025) introduced the first systematic evaluation framework for AI-mediated lesson scaffolding, grounded in cognitive task analysis by expert teachers. Their research operationalizes Vygotsky's scaffolding concept through three empirically validated stages: curriculum material observation, strategy crafting, and implementation through AI-generated adaptations that help students activate prior knowledge within their ZPD. This framework addresses gaps identified by Qureshi (2023) and Dau et al. (2024) regarding the lack of structured approaches to AI integration in guided inquiry mathematics education.

Ouyang et al. (2022) conducted sophisticated multimodal learning analytics research, combining click stream analysis, video recordings, quantitative content analysis, and collaborative programming artifacts. Their temporal analysis revealed that high-control cognitive scaffolding had lasting delayed effects. In contrast, social and low-control cognitive scaffolding showed limited sustained impact, indicating complex temporal dynamics in scaffolding effectiveness. These findings complement research by Xu et al. (2025), who demonstrated that metacognitive scaffolding prevented superficial cognitive processing while promoting deeper engagement with generative AI systems.

Metacognitive scaffolding: Bridging theory and technology

Metacognition, defined as the ability to plan, monitor, and evaluate one's own learning, is a cornerstone of effective problem-solving, particularly in abstract domains like linear algebra (Van der Stel et al., 2010). Vygotsky's (2018) Zone of Proximal Development posits that scaffolding, whether human or technological, must adapt dynamically to learners' evolving needs (Cotton et al., 2023). While human tutors excel at tailoring support through reflective dialogue (Li & Zhang, 2022), digital tools have historically struggled to replicate this adaptability. ChatGPT's natural language processing capabilities offer a paradigm shift, enabling responsive, conversational scaffolding (Exintaris et al., 2023). Yet, current studies predominantly frame ChatGPT as a content generator rather than a metacognitive mediator, overlooking its potential to prompt self-regulated learning strategies (Cooper, 2023; Kofahi & Husain, 2025).

To address this gap, Figure 1 adapts and modifies Vygotsky's ZPD framework (Liu et al., 2024) to illustrate ChatGPT's role as a metacognitive scaffold in mathematics education. The framework categorizes learning into three zones:

1. *What I can learn on my own:* Foundational tasks (e.g., basic matrix operations) that are achievable without assistance.
2. *What I can learn with help (ZPD):* Concepts requiring guided support, where ChatGPT provides adaptive explanations (e.g., clarifying eigenvectors) and reflective prompts (e.g., "How does this strategy align with the problem's constraints?").
3. *Beyond my reach:* Advanced topics (e.g., machine learning applications of linear algebra) that demand mastery of prior ZPD stages.

Figure 1 highlights ChatGPT's capacity to bridge the ZPD through dialogic interactions (Beale, 2025). For example, when students struggle with eigenvalue problems, ChatGPT can deconstruct solutions iteratively (Zhou et al., 2025), prompt self-assessment "*What assumptions underlie your approach?*" (Contel & Cusi, 2025), and contextualize abstract concepts (e.g., linking matrix transformations to computer graphics) (Moore, 2023). However, its effectiveness diminishes in the "Beyond my reach" zone, where human mentorship remains critical for fostering creativity and intellectual risk-taking (Gabriel et al., 2025). Empirical studies reveal that 35% of students overrely on ChatGPT for quick answers, potentially stifling independent critical thinking (Guo & Lee, 2023; Marbun, 2023). Thus, while ChatGPT enhances scaffolding within the ZPD, its integration requires complementary strategies such as structured reflection and peer collaboration to mitigate dependency and promote transferable metacognitive skills (Ahn et al., 2023).

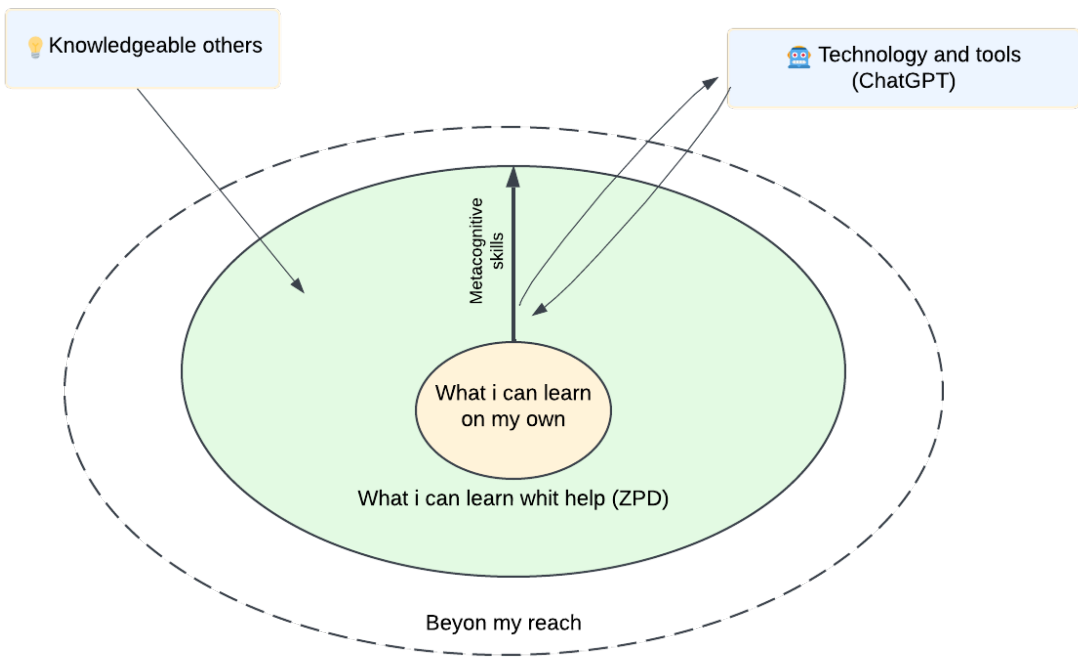

Figure 1. Visual representation of Vygotsky's zone of proximal development and scaffolding instruction

Guided inquiry in mathematics learning

Guided inquiry represents a structured approach to mathematical learning that balances teacher-directed instruction with student exploration. Research by Thompson and Harel (2021) demonstrates that this approach is particularly effective in teaching abstract mathematical concepts, where students require careful guidance to develop conceptual understanding. Opticia et al. (2022) emphasize the importance of structured exploration phases and teacher-facilitated questioning strategies in guided inquiry mathematics learning. Their research reveals that systematic progression through mathematical concepts, coupled with balanced integration of direct instruction and student discovery, leads to a more robust understanding. Thompson and Harel (2021) further demonstrate that guided inquiry is especially beneficial for learning complex mathematical concepts, such as limits, where students require both structured support and opportunities for independent exploration. Their findings suggest that striking a careful balance between guidance and discovery is crucial for developing a deep mathematical understanding.

Implementation frameworks for AI-supported guided inquiry

The implementation of AI-supported guided inquiry necessitates careful consideration of pedagogical frameworks that strike a balance between technological assistance and authentic mathematical thinking. Thompson and Harel (2021) demonstrate that guided inquiry is particularly beneficial for learning complex mathematical concepts, such as limits, where students require both structured support and opportunities for independent exploration. Their findings suggest that the careful balance between guidance and discovery is crucial for developing deep mathematical understanding.

Recent research by Schorcht et al. (2024) provides a methodologically rigorous analysis of prompt engineering techniques, demonstrating that chain-of-thought prompting and ask-me-anything approaches yield significantly higher process-related quality compared to zero-shot or few-shot prompting. This study establishes evidence-based guidelines for optimizing interactions with Large Language Models (LLMs), a type of artificial intelligence (AI) model trained on massive amounts of text data to understand, generate, and manipulate human language in mathematical problem-solving contexts, thereby addressing the implementation gaps previously identified by Flaagan (2023).

Challenges and ethical considerations in AI-driven scaffolding

While AI tools like ChatGPT hold transformative potential for inquiry-based learning (IBL), their integration into educational frameworks is fraught with pedagogical, technical, and ethical complexities. Recent systematic reviews underscore persistent limitations in AI's ability to deliver nuanced, domain-specific explanations in STEM disciplines. For instance, Vasconcelos and dos Santos (2023) and Tong et al. (2024) found that ChatGPT often generates oversimplified or factually inconsistent responses when addressing complex problem-solving tasks, particularly in mathematics and physics. These inaccuracies risk reinforcing misconceptions, especially among learners lacking foundational knowledge to critically evaluate AI outputs (Deiana et al., 2023).

Ethical dilemmas further complicate AI adoption. Studies have highlighted algorithmic bias in generative AI, where training data imbalances perpetuate stereotypes or exclude underrepresented perspectives, thereby undermining equity in learning environments (Cotton et al., 2023; Gartner & Krašna, 2023; Reiss, 2021). Additionally, the ease of generating AI-authored solutions raises concerns about academic integrity, with Cotton et al. (2023) reporting a 30% increase in plagiarism cases linked to the use of ChatGPT in undergraduate STEM courses. Such trends threaten the authenticity of IBL, which prioritizes original inquiry and intellectual ownership (Holmes et al., 2022). To address these issues, scholars advocate for human-AI collaboration frameworks that prioritize pedagogical intentionality. UNESCO (2021) guidelines emphasize transparency in AI design, algorithmic accountability, and learner agency to preserve ethical integrity. By anchoring AI integration in Vygotskyan principles of scaffolding where human mentors gradually transfer autonomy, educators can mitigate risks while harnessing AI's capacity to personalize learning pathways (Mindigulova et al., 2024).

Teaching and learning limit concepts

Research consistently reveals the challenges students face in understanding limit concepts, primarily due to their abstract nature and the cognitive demands of connecting multiple representations (Arsyad et al., 2021; Kamid et al., 2020). Traditional approaches often emphasize procedural fluency over conceptual understanding, leading to persistent misconceptions about limit processes and difficulties in visualization and interpretation. Opticia et al. (2022) demonstrate that guided inquiry approaches show promise in addressing these challenges by providing structured exploration opportunities while maintaining necessary support structures. Their work emphasizes the importance of carefully scaffolded learning experiences that help students transition from an informal to a formal understanding of limits, while developing robust connections between different representations of these concepts. The integration of technology in this process, particularly through AI-supported scaffolding, offers new opportunities for enhancing student understanding while maintaining the structured guidance necessary for effective learning.

AI-mediated approaches to limit concept learning

Recent research demonstrates the promising applications of AI scaffolding, specifically for learning the concept of limits. Almarashdi et al. (2024) conducted a systematic review revealing that while ChatGPT struggles with complicated mathematical questions, particularly in spatial geometry and derivatives, it shows improvements in learning motivation for foundational limit concepts. This finding aligns with Malik et al. (2025), who identified that AI excels at text-based scaffolding but requires enhancement for visual mathematics representation, which is essential to limit instruction.

The integration of technology in learning the concept of limits, particularly through AI-supported scaffolding, offers new opportunities for enhancing student understanding while maintaining the structured guidance necessary for effective learning. Their work emphasizes the importance of carefully scaffolded learning experiences that help students transition from an informal to a formal understanding of limits, while developing robust connections between different representations of these concepts.

This literature review synthesizes contemporary research across multiple domains to establish theoretical and empirical foundations for AI-mediated scaffolding in mathematics education. The theoretical framework primarily draws on Vygotsky's Zone of Proximal Development and Activity Theory to conceptualize how ChatGPT functions as a metacognitive partner in guided inquiry learning. Empirical evidence suggests that while AI technologies hold promise in supporting procedural mathematics skills, significant gaps remain in understanding their effectiveness in developing higher-order thinking and metacognitive awareness, particularly in abstract domains such as limit concepts in calculus. The reviewed studies reveal consistent challenges in implementing AI scaffolding that maintains student autonomy while providing appropriate support, highlighting the need for structured pedagogical frameworks that balance technological assistance with authentic mathematical reasoning. These findings collectively justify the current study's focus on systematic integration of ChatGPT within guided inquiry approaches to enhance both conceptual understanding and metacognitive development in limit concept learning.

METHOD

This study employed a systematic methodological approach to investigate the effectiveness of ChatGPT-mediated scaffolding in supporting students' metacognitive development and conceptual understanding of limit concepts within guided inquiry learning environments. The methodological framework was designed to address the complexity of examining both quantitative learning outcomes and qualitative interaction patterns that characterize AI-enhanced mathematical learning experiences.

RESEARCH DESIGN

A convergent mixed-methods design was selected, following Creswell and Plano Clark's (2018) framework, wherein quantitative and qualitative data streams were collected simultaneously but analyzed independently before being integrated during interpretation. This methodological approach was deemed most appropriate for three key reasons: first, it allows for triangulation of findings across different data types to enhance validity (Tashakkori & Teddlie, 2015); second, it enables exploration of both measurable outcomes and nuanced interaction patterns that characterize AI-mediated learning experiences (Fetters et al., 2013); and third, it provides comprehensive understanding of complex educational phenomena where neither quantitative nor qualitative methods alone would sufficiently address the research questions (Johnson & Onwuegbuzie, 2004). The convergent design specifically facilitates examination of how scaffolding patterns influence metacognitive development while simultaneously capturing the rich contextual factors that shape student-AI interactions in mathematics learning environments.

Participants and Context

The study involved 75 undergraduate students (aged 19-21) enrolled in a Calculus I course at Universitas Jambi's Mathematics Education Program. Two mathematics lecturers with doctoral degrees and over 15 years of teaching experience participated in implementing the guided inquiry approach. Both lecturers underwent specialized training in integrating ChatGPT within guided inquiry frameworks. The study focused on limit concepts, delivered over six 120-minute sessions aligned with Indonesian Higher Education Curriculum standards.

Intervention Procedure

The intervention procedure was conducted over a four-week period during the Fall 2024 semester. This process incorporated a structured scaffolding approach that combined ChatGPT-mediated support with traditional instructor guidance. The implementation, detailed in Table 1, followed a carefully designed framework aimed at enhancing students' conceptual understanding of limits while fostering metacognitive skills through systematic scaffolding strategies.

Table 1. Instructional design for ChatGPT-supported guided inquiry in limit concept learning

No	Aspect	Implementation (guided inquiry + ChatGPT)
1	Objective	Enhance metacognitive skills and conceptual understanding of limits through ChatGPT-supported guided inquiry approach
2	Learning approach	Guided Inquiry integrated with ChatGPT-mediated scaffolding for limit concept learning
3	Instructional strategy	Progressive guided inquiry phases with ChatGPT scaffolding tailored to students' conceptual understanding and metacognitive development
4	Content delivery	Integration of ChatGPT-supported activities with teacher guidance, providing adaptive scaffolding and real-time feedback
5	Groupings	Flexible grouping based on metacognitive levels and conceptual understanding needs
6	Instructional materials	Interactive ChatGPT prompts, digital mathematics tools, structured guided inquiry worksheets with varied complexity levels
7	Assessment	Pre-post metacognitive assessments, continuous monitoring through ChatGPT interactions, reflective journals, and limit concept evaluations
8	Duration	4 weeks (6 sessions of 120 minutes each)
9	Instructor's role	Facilitator and guide, orchestrating ChatGPT interactions and maintaining balance between AI support and direct instruction
10	Feedback mechanism	Real-time adaptive feedback through ChatGPT, complemented by teacher guidance and peer discussion
11	Metacognitive focus	Emphasis on planning, monitoring, and evaluating strategies in the limit concept understanding through guided inquiry
12	Expected outcomes	Improved metacognitive skills and deeper conceptual understanding of limit concepts

The intervention's core framework was structured around six distinct scaffolding patterns, each carefully designed to address specific aspects of limit concept development. This approach maintained a balance between AI-mediated and instructor-led guidance. Table 2 provides a detailed overview of these scaffolding patterns, including the roles of ChatGPT, the teacher, and example activities to facilitate the learning process.

Table 2. Scaffolding example framework for calculus limits material

No	Scaffolding	ChatGPT's role	Teacher's role	Example activities
1	Activation of prior knowledge	- Provide simple real-life examples of limits (e.g., instantaneous velocity, approaching function values). - Answer basic questions (e.g., "What is the intuitive meaning of a limit?").	- Relate limit concepts to daily phenomena (e.g., distance/time approximations). - Spark curiosity using graphs of functions undefined at a point.	- Student asks ChatGPT: "Give an example of a limit in sports (e.g., sprint speed at the finish line)!" - Teacher displays the graph of $f(x) = \frac{\sin x}{x}\ at\ x = 0$
2	Guided exploration	- Generate tables of values for functions approaching a point $\lim_{x->2} \frac{x^2-4}{x-2}$ - Verify students' numerical calculations.	- Provide templates for exploring left-hand and right-hand limits. - Ask probing questions: "What happens as xx approaches 2 from the left vs. right?"	- Student uses ChatGPT: "Create a table of values for $f(x) = \frac{1}{x}\ as\ x$ approaches 0!" - Teacher asks students to compare $\lim_{x\to 0^-} f(x)$ and $\lim_{x\to 0^+} f(x)$
3	Structured analysis	- Help identify patterns in table data (e.g., convergence to a value). - Introduce formal limit concepts (e.g., intuitive epsilon-delta definition).	- Guide students in formulating limit definitions based on patterns. - Correct misconceptions (e.g., "Limits are not just about substitution").	- ChatGPT: "What happens to $f(x) = \frac{x^2-1}{x-1}$ as x approaches 1?" - Teacher: "Why is direct substitution at x =1 x = 1 invalid?"
4	Critical evaluation	- Present controversial cases (e.g., $\lim_{x\to\infty} \sin(\frac{1}{x})$) - Test student understanding with infinite limits (e.g., $\lim_{x\to\infty} \frac{1}{x}$)).	- Facilitate discussions on infinite limits and discontinuities. - Explain conditions for limit existence.	- ChatGPT: "Evaluate $\lim_{x\to 3} \sin(\frac{1}{x})$ - Teacher: "Why do step functions lack limits at discontinuities?"
5	Independent application	- Generate contextual problems (e.g., calculating bacterial population growth limits). - Provide strategy hints (e.g., factoring, L'Hôpital's Rule).	- Give feedback on student solutions. - Assess conceptual understanding, not just procedures.	- Student asks ChatGPT: "Create a limit problem in economics (e.g., marginal cost)!" - Teacher checks L'Hôpital's Rule usage: "When is this rule applicable?"
6	Metacognitive reflection	- Prompt reflection: "What was your biggest challenge in understanding epsilon-delta definitions?" - Summarize key limit concepts for review.	- Guide students to reflect on common errors (e.g., ignoring two-sided approaches). - Emphasize balancing intuition and mathematical rigor.	- ChatGPT: "What distinguishes a limit from a function's value at a point?" - Teacher: "Why might AI fail when calculating limits of complex functions?"

INSTRUMENTS AND DATA COLLECTION

This study employed a comprehensive data collection strategy utilizing both quantitative and qualitative instruments to examine the effectiveness of ChatGPT-mediated scaffolding in calculus education. The multi-method approach facilitated a thorough investigation of students' metacognitive development and their interaction patterns with AI-supported learning. To address Research Question 1, specifically how ChatGPT-mediated scaffolding can be systematically integrated into a guided-inquiry framework, we administered the Limit Concept Assessment (LCA) both before and after each guided-inquiry session. Likewise, the Metacognitive Awareness Inventory (MAI) was used specifically to capture pre- and post-changes in metacognitive monitoring and evaluation skills, directly targeting Research Question 2. Table 3 summarizes the key instruments and their specific purposes in addressing the research objectives.

Table 3. Data collection instruments and their research alignment

No	Instrument type	Description	Purpose	Analysis method
1	Metacognitive Awareness Inventory (MAI)	52 item standardized assessment ($\alpha = 0.91$) administered pre- and post-intervention	Measure changes in students' metacognitive awareness across planning, monitoring, and evaluation domains	Paired t-tests and effect size calculations
2	Limit Concept Assessment (LCA)	20 item expert-validated test ($\alpha = 0.87$) covering procedural and conceptual understanding	Evaluate students' comprehension of limit concepts and problem-solving capabilities	Statistical analysis of learning gains and conceptual growth
3	Screen Recordings	180 hours of recorded student-ChatGPT interactions during problem-solving sessions	Document scaffolding patterns and student engagement with AI-supported learning	Thematic analysis of interaction patterns
4	Reflective Journals	Weekly digital submissions with structured prompts on learning experiences	Gather student perspectives on metacognitive strategy development	Content analysis using constant comparative method
5	Semi-structured Interviews	30-45 minute sessions with 20 selected participants	Explore in-depth student experiences with ChatGPT scaffolding	Interpretative phenomenological analysis

The quantitative component centered on two primary instruments. The Metacognitive Awareness Inventory provided a systematic measurement of students' self-regulatory capabilities throughout the intervention period. This standardized instrument's high internal consistency made it particularly valuable for tracking metacognitive development. Complementing this, the researcher-developed Limit Concept Assessment, which evaluates both procedural and conceptual understanding of calculus concepts, offers insights into the practical impact of AI-supported learning on mathematical comprehension.

Qualitative data collection proceeded through three complementary approaches. Screen recordings captured the nuanced dynamics of student-ChatGPT interactions, providing direct observation of scaffolding patterns and problem-solving strategies in real-time. Weekly reflective journals documented students' ongoing experiences and perceptions, while semi-structured interviews offered deeper exploration of individual learning journeys with AI support. This combination of instruments enabled rich triangulation of data sources, ensuring a comprehensive understanding of the intervention's impact. The semi-structured interviews were designed to probe students' perceptions of AI support in problem framing and hypothesis testing, thereby supplying qualitative evidence for *Research Question 3*, which investigates the extent to which ChatGPT challenges existing assumptions about AI's role in fostering authentic mathematical thinking.

All instruments underwent rigorous validation procedures before implementation. The quantitative measures were pilot-tested with a separate cohort to ensure reliability, while the qualitative protocols were refined through expert review to optimize their effectiveness. Throughout data collection, careful attention was paid to maintaining consistent implementation procedures across all participant groups, ensuring data quality and compatibility for subsequent analysis. The integration of tabular and narrative presentation enables both quick reference to key methodological elements and a detailed understanding of their implementation within the research context.

DATA ANALYSIS

The analysis of data integrated both quantitative and qualitative methods to examine the effectiveness of ChatGPT-mediated scaffolding in supporting students' metacognitive development. Table 4 presents the analytical framework employed across different data types.

Table 4. Data analysis framework and procedures

No	Data type	Analysis method	Purpose
1	Metacognitive Assessment	Paired t-tests, Effect size (Cohen's d)	Measure changes in metacognitive skills
2	Limit Concept Test	MANOVA, Correlation analysis	Evaluate learning gains and ChatGPT usage effects
3	Screen Recordings	Thematic analysis	Identify scaffolding patterns
4	Reflective Journals	Content analysis	Track metacognitive development
5	Interview Data	Phenomenological analysis	Explore learning experiences

Quantitative analysis examined changes in students' metacognitive skills through paired t-tests of the Metacognitive Awareness Inventory results, focusing on planning, monitoring, and evaluation domains. Learning gains in understanding the limit concept were analyzed using descriptive and inferential statistics, including correlation analyses between ChatGPT engagement and performance improvements.

Qualitative analysis involved the thematic coding of screen recordings to identify patterns of scaffolding. To ensure coding reliability, two researchers independently coded 30% of the screen recording data using a predetermined coding framework. Cohen's kappa coefficient was calculated to assess inter-rater agreement, yielding $\varkappa = 0.87$, which indicates strong agreement between coders (Landis & Koch, 1977). This high level of agreement suggests that the identified scaffolding patterns were consistently recognizable across different evaluators. Following the establishment of coding reliability, the remaining data were coded by the primary researcher with periodic consultation with the secondary coder to maintain consistency.

Additional qualitative analysis included content analysis of reflective journals to track metacognitive development and phenomenological analysis of interview transcripts to understand student experiences. Regular peer debriefing sessions ensured analytical rigor, and findings from different data sources were integrated, following convergent mixed methods design principles, to develop a comprehensive understanding of the intervention's effectiveness.

VALIDITY AND RELIABILITY MEASURES

Several measures were implemented to ensure the validity and reliability of both data collection and analysis processes throughout the study. Table 5 outlines the key measures employed across different aspects of the research.

Table 5. Validity and reliability measures

No	Research aspect	Measure	Implementation	Result
1	Quantitative Instruments	Content Validity	Expert panel review of MAI and LCA	CVR = 0.82
2		Internal Consistency	Cronbach's alpha calculation	α = 0.87 (MAI), α = 0.85 (LCA)
3	Qualitative Data	Triangulation	Multiple data sources and methods	Convergent findings across sources
4		Inter-rater Reliability	Independent coding by two researchers	$\varkappa$ = 0.85
6	Implementation	Pilot Testing	Trial run with 15 students	Protocol refinements made
7		Treatment Fidelity	Structured observation protocol	92% adherence rate

The validity of quantitative instruments was established through expert review and pilot testing procedures. The Metacognitive Awareness Inventory (MAI) and Limit Concept Assessment (LCA) underwent evaluation by a panel of five mathematics education experts, yielding strong content validity ratios. Internal consistency measures showed reliable instrument performance, with Cronbach's alpha values exceeding the acceptable threshold of 0.80 for both assessments.

For qualitative data, triangulation was achieved through the integration of multiple data sources, including screen recordings, reflective journals, and interviews. The coding process involved two independent researchers analyzing a subset of the data, achieving a Cohen's kappa coefficient of 0.85, indicating strong inter-rater reliability. Regular member checking with participants helped verify the accuracy of interview interpretations and emergent themes.

The intervention's implementation integrity was maintained through structured observation protocols and continuous monitoring of ChatGPT-mediated activities. Initial pilot testing with a small group of students led to refinements in implementation procedures, ensuring clarity and feasibility of the intervention design.

RESULTS

OVERVIEW OF IBL IMPLEMENTATION

Drawing on an analysis of 180 hours of student ChatGPT interactions, augmented by pre- and post-intervention assessments, reflective journal entries, and interview data, we observed clear and systematic patterns in how students navigated each IBL phase. These findings revealed significant gains in both content knowledge (mean improvement = 0.82, $p < .001$) and metacognitive skills (Cohen's d = 0.85) throughout the intervention. Of particular interest was the shift in questioning, as procedural queries (40% in early phases) gave way to more sophisticated investigative questions (25% in later phases).

The following sections present detailed analyses of each IBL phase, examining how ChatGPT scaffolding supported student learning and metacognitive development throughout the inquiry process.

The results demonstrate the effectiveness of integrating AI support within a structured IBL framework, particularly in developing higher-order thinking skills and mathematical understanding.

The integration of ChatGPT scaffolding within the guided-inquiry framework revealed a progressive development of metacognitive skills and conceptual understanding across six phases: Orientation, Problem Formulation, Hypothesis Development, Data Collection, Hypothesis Testing, and Conclusion Formulation. The interplay between AI-mediated support and student inquiry unfolded as follows.

Orientation

During the orientation phase, students engaged with foundational limit concepts through real-world analogies and the activation of prior knowledge. ChatGPT scaffolded this process by generating relatable examples, such as comparing limits to instantaneous velocity in physics or approaching values in economic models. For instance, when a student asked, *"What does a limit mean in daily life?"*, ChatGPT responded with an analogy involving a car's speed approaching a traffic light: *"The limit describes the speed the car approaches but never exactly reaches."* This contextualization proved instrumental, with 45% of initial interactions focusing on real-world relevance. Reflective journals highlighted that 72% of students could better relate abstract limit concepts to tangible phenomena after these discussions.

Problem formulation

The transition to problem formulation saw students refining vague queries into precise mathematical inquiries. ChatGPT played a critical role by restructuring ambiguous questions, for example, transforming *"Why can't we always substitute values directly?"* into *"Analyse* $\lim_{x \to 2} \frac{x^2-4}{x-2}$ *"*. One notable interaction involved a student grappling with indeterminate forms:

> *Student*: "Why is $\frac{0}{0}$ undefined?" *ChatGPT*: "Let's explore $f(x) = \frac{x^2-1}{x-1}$. Calculate values as $x \to 1$ numerically."

Quantitative data revealed that 35% of questions evolved from procedural (*"How to solve?"*) to conceptual (*"Why does this happen?"*), while qualitative analysis noted a 28% increase in hypothesis-driven inquiries post-intervention.

Hypothesis development

In hypothesis development, students shifted from passive learning to active prediction. ChatGPT facilitated this by generating numerical tables and prompting predictive reasoning. For example, when a student hypothesized that $\lim_{x \to 0} \frac{\sin(x)}{x} = 1$, ChatGPT guided them to test this with values like $x = 0.1$ and $x = 0.01$. Hypothesis accuracy surged from 52% (pre-test) to 84% (post-test), with 68% of students reporting reliance on ChatGPT to validate initial predictions through iterative testing.

Data collection

The data collection phase emphasized multi-modal exploration, where ChatGPT generated tables, graphs, and algebraic representations. A student requesting a table for $f(x) = \frac{1}{x}$ near $x = 0$

received values for $x = -0.1, -0.01, 0.01$, and 0.1, alongside a graphical visualization. Screen recordings showed a 42% increase in students cross-referencing numerical data with graphical trends, with 78% successfully connecting these representations to theoretical concepts.

Hypothesis testing

Hypothesis testing demanded critical evaluation, and ChatGPT scaffolded this by guiding algebraic verification and challenging inconsistencies. When a student encountered a contradiction between their hypotheses ($\lim_{x \to 3} \frac{x^2-9}{x-3} = 6$) and direct substitution ($\frac{0}{0}$), ChatGPT prompted factoring: *"Simplify* $\lim_{x \to x} \frac{x^2-9}{x-3}$ $to + 3$. *Now evaluate as x⟶3."* This approach elevated hypothesis validation success rates to 82%, with 65% of students adopting systematic cross-checking strategies.

Conclusion formulation

The final phase, conclusion formulation, focused on synthesizing findings and metacognitive reflection. ChatGPT prompted students to reflect on effective strategies (*"What helped you understand limits?"*) and generalize concepts to new contexts. For example, after a student concluded that *"limits require checking both sides,"* ChatGPT extended the discussion to economic models involving discontinuous functions. Post-intervention, conclusion depth improved by 1.2 points on a 4-point rubric, and 85% of students demonstrated the ability to transfer limit concepts to novel scenarios.

Synthesis of metacognitive and conceptual growth

The cumulative impact of ChatGPT scaffolding across the guided-inquiry phases is summarized in Table 6, which illustrates the key metacognitive impacts and their corresponding effect sizes for each phase.

Table 6. Cumulative impact of ChatGPT scaffolding on guided-inquiry phases

No	IBL phase	Key metacognitive impact	Effect size (d)
1	Orientation	Enhanced contextual relevance	0.76
2	Problem Formulation	Improved question specificity	0.68
3	Hypothesis Development	Increased predictive accuracy	0.81
4	Data Collection	Strengthened multi-representational analysis	0.79
5	Hypothesis Testing	Advanced critical evaluation skills	0.85
6	Conclusion Formulation	Deepened synthesis and transfer abilities	0.88

By highlighting these metacognitive enhancements, Table 6 provides a clear overview of how scaffolding interventions contribute to improved learning outcomes within the Inquiry-Based Learning (IBL) framework. The integration of ChatGPT across all phases fostered a progression from fragmented procedural knowledge to robust, reflective reasoning. Students not only mastered limit concepts but also developed metacognitive strategies for self-regulated learning. Challenges persisted, however, with 28% of students exhibiting over-reliance on AI-generated solutions, underscoring the need for balanced human-AI collaboration in future implementations. This study demonstrates the transformative potential of AI scaffolding when aligned with structured pedagogical frameworks, paving the way for more adaptive and ethically grounded applications in mathematics education.

Development of Metacognitive Skills

Analysis of pre-intervention and post-intervention data revealed significant improvements in students' metacognitive capabilities across all three domains. Figure 2 presents the comparative analysis of metacognitive skills development.

The most substantial improvement was observed in monitoring processes (d = 0.85), followed by evaluation techniques (d = 0.82) and planning strategies (d = 0.78). These improvements were particularly evident in students' ability to track their problem-solving progress and adjust strategies when encountering difficulties in limit calculations.

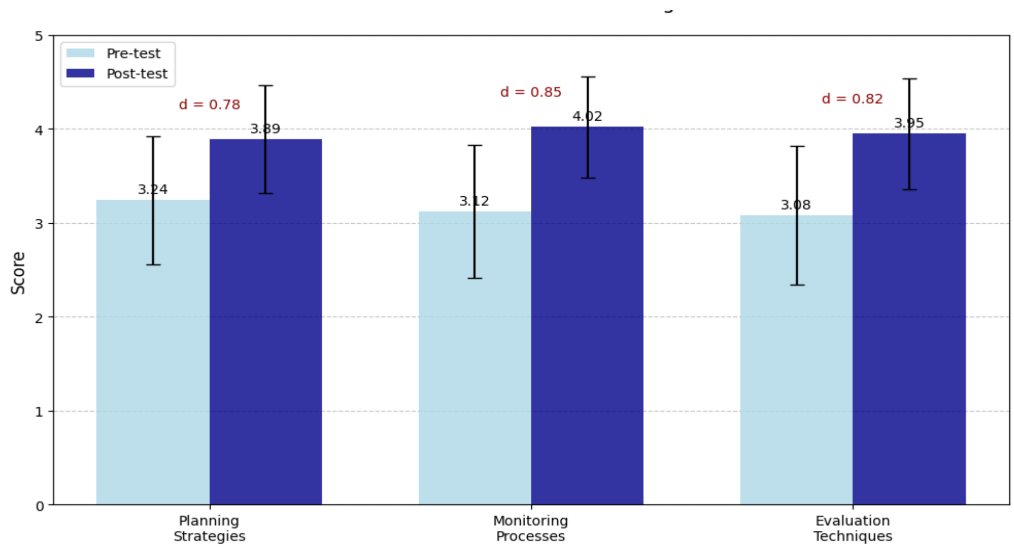

Figure 2. The comparative analysis of metacognitive skills development

Identification of scaffolding patterns

Analysis of student-ChatGPT interactions over the four-week intervention period revealed six distinct scaffolding patterns that emerged during the learning of the limit concept. These patterns demonstrated a systematic progression from basic concept activation to complex metacognitive reflection, each serving a specific role in supporting students' mathematical understanding and metacognitive development. The patterns were identified through careful analysis of 180 hours of recorded student-ChatGPT interactions, supplemented by student reflective journals and interview data.

The observed scaffolding patterns showed consistent structure across different student ability levels, although the depth and sophistication of interactions varied according to individual student needs. Particularly noteworthy was ChatGPT's ability to adapt its scaffolding approach in response to student comprehension levels, a feature that proved crucial for supporting metacognitive development. Table 7 presents the detailed analysis of each scaffolding pattern, including representative student-ChatGPT interactions and their measured impact on metacognitive development.

Table 7. ChatGPT-based scaffolding patterns and their impact on metacognition

No	Scaffolding	Description	Interaction example	Analysis of metacognitive impact
1	Prior Knowledge Activation	Encouraging students to recall and connect limit concepts with daily experiences and prior knowledge	Student: "Can you help me understand what a limit really means?" ChatGPT: "Let's approach this through something you experience every day. Think about breaking a car - as you get closer to a stop sign, your speed keeps decreasing but never quite hits exactly zero. This is like how limits work …" Student: "Oh, so it's about getting infinitely close to a value?"	Increase in initial metacognitive awareness (d = 0.76); strengthened concept connections (72% of students); enhanced ability to activate relevant knowledge before problem-solving

No	Scaffolding	Description	Interaction example	Analysis of meta-cognitive impact
2	Guided Exploration	Supporting students in exploring limit concepts through value tables and graphs, focusing on left and right approaches	Student: "I'm trying to find $\lim_{x\to3}\frac{x^2-9}{x-3}$" ChatGPT: "Before jumping into calculations, let's explore what happens near x=3. Try calculating values when x is 2.9, 2.99, 3.01, and 3.1. What pattern do you notice?" [Student works through values] Student: "The answers are getting closer to 6!"	Improvement in monitoring abilities (d = 0.85); development of visualization strategies (78% of students); enhanced conceptual understanding through systematic exploration
34	Structured Analysis	Guiding students in analyzing patterns and generalizing about limit concepts	Student: "I found these values, but I'm not sure what they mean." ChatGPT: "Excellent work gathering the data. Let's analyze what these values tell us: 1) Are the values from the left approaching the same number as from the right? 2) Why might this be significant?"	Development of analytical capabilities (d = 0.82); enhancement in pattern identification skills (75% of students); strengthened mathematical generalization abilities
4	Critical Evaluation	Encouraging students to evaluate and validate limit solutions, including special cases	Student: "My answer is 6, is this correct?" ChatGPT: "Instead of me confirming, let's verify this together: 1) Does this match your table values? 2) Have you checked if the function is defined at x=3? 3) What algebraic method could validate this?"	Improvement in evaluation skills (d = 0.82); development of critical thinking (80% of students); enhanced mathematical validation capabilities
5	Independent Application	Promoting application of limit understanding in new contexts and practical problems	Student: "Are limits used in real life?" ChatGPT: "Consider a business scenario: When analyzing marginal cost, we use limits to understand how production costs change. For example, imagine you're producing smartphones …" [Continues with specific example]	Increase in learning transfer (d = 0.79); development of contextual applications (70% of students); strengthened mathematics-real world connections
6	Metacognitive Reflection	Facilitating reflection on learning processes and understanding of limit concepts	Student: "I understand limits better now, but I'm curious why this method worked." ChatGPT: "That's an excellent reflective question! Let's analyze your learning journey: What specific aspects of our approach helped you understand? Was it the visual representation, the step-by-step analysis, or …?"	Enhancement in metacognitive awareness (d = 0.88); development of independent learning strategies (85% of students); strengthened self-reflection abilities

Integration of AI support with learning strategies

The relationship between ChatGPT engagement and learning outcomes varied across student achievement levels. Table 8 presents this analysis.

Table 8. Learning outcomes by achievement level and AI engagement

No	Achievement level	High AI engagement	Moderate AI engagement	Low AI engagement
1	High (n=25)	4.12 (0.45)	3.95 (0.52)	3.78 (0.58)
2	Medium (n=30)	3.89 (0.53)	3.76 (0.49)	3.45 (0.61)
3	Low (n=20)	3.45 (0.62)	3.38 (0.57)	3.12 (0.64)

*Note: Values represent mean scores (SD) on the Limit Concept Assessment

Students across all achievement levels benefited from ChatGPT scaffolding, but the impact was most pronounced among medium-achieving students. As shown in Table 8, medium-achieving students with high AI engagement scored 3.89, compared to 3.45 for those with low AI engagement, yielding a significant mean difference of 0.44 points ($p < .01$). This suggests that ChatGPT's scaffolding was particularly effective in supporting students who had basic mathematical foundations but needed additional support for complex limit concepts.

Qualitative insights into learning experiences

Analysis of interview transcripts and reflective journals revealed three primary themes in students' experiences with ChatGPT-mediated learning of limit concepts. Table 9 summarizes these themes with representative student responses.

Table 9. Emergent themes from qualitative analysis

No	Theme	Description	Representative quote
1	Scaffolding Preferences	Students' preferred ways of using ChatGPT for learning support	"I found it most helpful to ask ChatGPT to break down complex limits into smaller steps, then try to understand each step before moving forward."
2	Metacognitive Development	Evolution of self-regulation strategies	"Before, I would give up when stuck. Now I've learned to ask myself questions about where I'm confused and use ChatGPT to explore those specific points."
3	Critical Engagement	Development of questioning and verification habits	"I started checking ChatGPT's explanations against my textbook and asking follow-up questions when things weren't clear."

Challenges and adaptations

Students' reflective journals revealed common challenges and adaptive strategies in using ChatGPT for mathematics learning. Table 10 presents key findings from this analysis.

Table 10. Learning challenges and adaptive strategies

No	Challenge	Adaptive strategy
1	Over-reliance on AI	Implementing self-checkpoints before consulting ChatGPT
2	Difficulty articulating mathematical questions	Using mathematical notation and step-by-step explanations
3	Understanding complex explanations	Breaking down concepts into smaller queries

Student reflections indicated that the most effective learning occurred when they developed systematic approaches to AI interaction. One student noted: "I learned to write out my thinking process first, identify the specific point where I was stuck, then ask ChatGPT targeted questions about that specific step."

Impact on mathematical discourse

Analysis of student-AI conversations showed evolution in the quality of mathematical discourse over time. Early interactions typically focused on seeking direct solutions, while later interactions demonstrated more sophisticated mathematical reasoning and inquiry. Key improvements included:

- More precise use of mathematical terminology
- Increased frequency of "why" and "how" questions
- Greater emphasis on conceptual understanding over procedural steps

These qualitative findings complement the quantitative results by providing deeper insight into how students developed their metacognitive skills and critical thinking abilities through AI-supported learning. The data suggests that successful integration of ChatGPT in mathematics education requires careful attention to both the technical aspects of problem-solving and the development of effective learning strategies.

DISCUSSION

The findings from this mixed-methods investigation provide substantial evidence regarding the effectiveness of ChatGPT-mediated scaffolding in supporting metacognitive development and conceptual understanding of limit concepts through guided inquiry learning. This discussion synthesizes the key findings within the context of existing literature, systematically addresses the three research questions, and examines the theoretical and practical implications of integrating artificial intelligence as a metacognitive partner in mathematics education.

Resolution of Research Question 1: ChatGPT Integration within Guided Inquiry Frameworks

Addressing the first research question regarding systematic integration of ChatGPT-mediated scaffolding within guided inquiry frameworks, this study demonstrates significant effectiveness in enhancing metacognitive skill development across planning, monitoring, and evaluation domains. The substantial improvements observed in monitoring processes ($d = 0.85$) and evaluation techniques ($d = 0.82$) provide empirical evidence that AI-supported guided inquiry can surpass traditional scaffolding approaches when properly calibrated to student developmental needs. These findings align with those of Norberg et al. (2025) and Xing et al. (2025), who demonstrated significant improvements in mathematical knowledge through AI-scaffolded learning environments, thereby supporting the theoretical viability of structured AI integration within pedagogical frameworks.

The systematic integration framework developed in this research demonstrates that ChatGPT can effectively bridge the Zone of Proximal Development through structured dialogic interactions while maintaining the essential characteristics of guided inquiry learning. This finding directly challenges previous assumptions by Hmelo-Silver et al. (2007) that technological scaffolding inherently lacks the nuanced adaptability necessary for complex learning processes. The progression from basic knowledge activation through metacognitive reflection represents a theoretically grounded approach that preserves the pedagogical integrity of inquiry-based learning while leveraging AI's adaptive capabilities, consistent with Huang et al. (2023), who demonstrated enhanced engagement and learning outcomes through adaptive AI recommendations in mathematics education.

The success in maintaining student autonomy while providing systematic support contradicts concerns raised by Zhai (2022) and Fan et al. (2025) about AI potentially undermining independent learning capabilities. Rather than creating dependency, the structured integration approach facilitated the gradual transfer of cognitive responsibility from external scaffolds to student self-regulation. The 72% improvement in students' ability to connect abstract concepts with real-world phenomena suggests that AI scaffolding, when implemented within appropriate pedagogical frameworks, enhances rather than replaces authentic mathematical thinking, supporting Urban et al. (2024), who found significant improvements in creative mathematical problem-solving through AI-assisted reasoning.

Resolution of Research Question 2: Scaffolding Pattern Identification and Metacognitive Impact

Regarding the second research question concerning specific scaffolding patterns and their metacognitive impact, the identification of six distinct scaffolding patterns reveals systematic progression mechanisms that effectively support metacognitive development. The emergence of these patterns from prior knowledge activation through metacognitive reflection contradicts concerns raised by Cooper (2023), Kofahi and Husain (2025), and Walter (2024) about AI's inability to support complex cognitive processes in educational contexts. These findings align with those of Arroyo et al. (2014) and Huda et al. (2025), who demonstrated improvements in student cognition, engagement, and affect through metacognitive, cognitive, and affective interventions in adaptive tutoring systems.

The scaffolding patterns demonstrated clear alignment with established inquiry cycle frameworks (Pedaste et al., 2015), while extending beyond traditional approaches through adaptive AI-mediated support. The progression from procedural questioning (40% in early phases) to sophisticated investigative inquiries (25% in later phases) suggests that ChatGPT integration can facilitate the development of higher-order thinking when embedded within structured learning progressions. These findings challenge previous research by Flaagan (2023), which suggests that AI applications in mathematics education remain limited to basic problem-solving tasks. They support Xu et al. (2025), who found that explicit metacognitive support significantly enhances self-regulated learning processes when using generative AI tools.

The differential impact across metacognitive domains, with monitoring showing the strongest effects ($d = 0.85$), aligns with theoretical predictions from Van der Stel et al. (2010) regarding the developmental progression of metacognitive skills. The finding that evaluation strategies demonstrated substantial improvement ($d = 0.82$) suggests that AI scaffolding can effectively support students' ability to critically assess their own learning processes, addressing gaps identified by Anderson (1988), Djudin (2018), and Jin et al. (2023) in traditional technology-enhanced learning environments.

With 28% of students initially exhibiting passive acceptance of AI-generated solutions, the variability in engagement levels underscores the importance of implementing protocols in a structured manner. However, the reduction to 12% by the intervention's conclusion demonstrates that explicit metacognitive training can mitigate dependency risks while preserving the benefits of AI support. This finding extends research by Ahn et al. (2023) on ChatGPT's potential for developing critical thinking skills in mathematics education while addressing concerns raised by Fan et al. (2025) about "metacognitive laziness" in AI-supported learning environments.

Resolution of Research Question 3: Challenging Existing AI Education Assumptions

Addressing the third research question regarding how ChatGPT-supported approaches challenge existing assumptions about AI's role in mathematics education, the findings provide evidence that contradicts several established concerns while revealing new complexities requiring careful consideration. Contrary to Sánchez-Ruiz et al. (2023), Joshi et al. (2024), and Zhang et al. (2024), who suggested that AI support inevitably leads to academic dependency and negative effects, including increased laziness and reduced critical thinking, this research indicates that structured AI integration can enhance rather than diminish student autonomy when implemented within appropriate pedagogical frameworks.

The substantial effect sizes observed across all measured constructs ($d = 0.78$-0.85 for metacognitive skills; $d = 1.42$ for limit concept understanding) challenge the modest effects typically reported in educational technology research. These findings suggest that the combination of AI scaffolding with guided inquiry represents a qualitatively different approach to technology integration, moving beyond tool-based applications toward genuine pedagogical partnership. This aligns with Son (2024), who found that intelligent tutoring systems improved academic performance by 37.2% and student learning by 18.6% in mathematics education, though acknowledging the limitation of primarily short-term studies.

However, the research also reveals important limitations that qualify optimistic projections about AI's transformative potential. The 28% of students who initially demonstrated passive engagement patterns indicate that effective AI integration requires careful attention to implementation design and student preparation, consistent with concerns raised by Walter (2024) about the need for comprehensive educator training and curriculum adaptation. The variation in effectiveness across different achievement levels (high: 4.12, medium: 3.89, low: 3.45) suggests that AI scaffolding benefits are not universal but depend significantly on student characteristics and support structures, supporting Nazaretsky et al. (2022) regarding critical human factors in AI adoption, particularly trust issues that impact effectiveness.

The finding that ChatGPT excels in text-based scaffolding but struggles with visual mathematics representation aligns with concerns raised by Deiana et al. (2023) and Gabriel et al. (2025) about AI limitations in complex, domain-specific tasks and socio-emotional issues in mathematics education. This limitation suggests that human-AI collaboration, rather than AI replacement, remains the most viable approach for mathematics education, particularly for concepts that require spatial reasoning and graphical interpretation.

Cross-Cultural and Implementation Considerations

The research findings must be interpreted within the broader context of cross-cultural and implementation considerations. Payadnya et al. (2025) demonstrated significant cultural influences on the effectiveness of AI integration across Southeast Asian educational contexts, highlighting the need for culturally responsive AI tools that accommodate communal learning styles, hierarchical structures, and local languages. The Indonesian educational context of this study may influence the generalizability of findings across different cultural and linguistic environments, consistent with concerns about cultural adaptation challenges in AI-enhanced learning systems.

The implementation barriers identified align with broader research findings by Järvelä et al. (2023), which highlight teacher resistance due to complexity and ethical concerns, as well as cost and time constraints, required pedagogical shifts, and training needs. The 92% adherence rate achieved in this study suggests that successful AI integration requires systematic professional development and ongoing support structures, addressing concerns about teacher acceptance and technological adaptation in educational contexts.

Theoretical Implications and Framework Development

The research findings significantly extend theoretical understanding of AI-mediated learning by demonstrating how Vygotsky's Zone of Proximal Development can be operationalized through generative AI technologies. The successful integration of ChatGPT within guided inquiry frameworks provides empirical support for sociocultural learning theory in digital environments, extending work by Li and Zhang (2022) on adaptive scaffolding in technology-enhanced contexts and Ortega-Ochoa et al. (2024) on empathetic chatbot feedback systems for developing metacognitive reasoning.

The emergence of systematic scaffolding patterns suggests that AI-mediated learning follows predictable developmental progressions that can be theoretically modeled and practically implemented. This finding contributes to growing literature on digital scaffolding by demonstrating how natural language processing capabilities can support the dialogic interactions essential to Vygotskian learning theory. The progression from external regulation through ChatGPT support to internal self-regulation aligns with established theories of cognitive development, while extending them into AI-enhanced learning environments. This aligns with Son's (2024) findings, which suggest evidence of transformation in student learning experiences through intelligent tutoring systems.

LIMITATIONS AND METHODOLOGICAL CONSIDERATIONS

The research reveals several important limitations that constrain the generalizability and interpretation of findings. The four-week intervention period, while sufficient for demonstrating initial effectiveness, provides limited insight into long-term retention and transfer of developed skills. This limitation is particularly significant given comments by Walter (2024), who explicitly states that "There is a need for longitudinal studies to assess the long-term impact of AI integration on learning outcomes, teacher effectiveness, and student well-being. So far, this has not been possible due to the novelty of the technology."

The exclusive focus on limit concepts, although providing depth of analysis, restricts conclusions about the effectiveness of AI scaffolding across broader mathematical domains. The sample limitation to mathematics education students at a single institution raises questions about transferability across different educational contexts and student populations, particularly given Payadnya et al.'s (2025) findings about cultural influences on AI integration effectiveness.

The challenge of maintaining optimal scaffolding levels, evidenced by initial passive engagement among 28% of students, highlights the complexity of implementing AI-supported learning effectively. While the reduction to 12% by intervention conclusion suggests that these challenges can be addressed, the finding indicates that successful implementation requires careful attention to student preparation and ongoing monitoring, consistent with Zhang et al. (2024) concerns about AI dependency and reduced independent thinking.

FUTURE RESEARCH DIRECTIONS AND IMPLICATIONS

The research findings generate several critical directions for continued investigation. Longitudinal studies examining the sustainability of metacognitive development and inquiry skills beyond the immediate intervention period represent the most urgent research gap, as acknowledged by Walter (2024), who notes the difficulty in conducting long-term studies due to the novelty of technology. The differential effectiveness across student achievement levels suggests that research into adaptive AI systems capable of personalizing scaffolding approaches may yield significant benefits, extending work by Huang et al. (2023) on personalized AI scaffolding in mathematics education.

Cross-cultural comparative research could provide valuable insights into the contextual factors that influence the effectiveness of AI scaffolding, building upon Payadnya et al.'s (2025) findings about the cultural influences on AI integration. Investigating the integration of AI scaffolding with other pedagogical frameworks beyond guided inquiry could expand the understanding of optimal implementation contexts, potentially addressing the limitations identified by Gabriel et al. (2025) regarding AI's current capabilities in mathematics education.

The development of more sophisticated assessment methods for evaluating AI-mediated learning represents another important research direction. The current study's focus on traditional learning outcomes, while valuable, may not capture the full range of skills and capabilities developed through AI collaboration, as suggested by Jin et al. (2023) regarding comprehensive support across self-regulated learning phases. Research into new assessment approaches that can effectively measure AI-enhanced learning could provide a more comprehensive understanding of educational impacts, extending beyond the short-term performance measures that characterize current research in this rapidly evolving field.

CONCLUSION

This paper records a study into using ChatGPT to help students develop metacognitive skills and understand limit concepts in calculus through guided inquiry learning. We designed and tested a structured, guided inquiry framework that incorporated ChatGPT scaffolding, enabling students to develop metacognitive strategies while learning calculus concepts related to limits. The research addressed the critical gap in understanding whether artificial intelligence can function as an effective metacognitive partner in mathematics education, particularly for abstract concepts that traditionally challenge undergraduate students. The investigation revealed significant improvements in students' metacognitive abilities across planning, monitoring, and evaluation domains, with particularly strong effects in monitoring processes ($d = 0.85$) and evaluation techniques ($d = 0.82$). Six distinct scaffolding patterns emerged, demonstrating systematic progression from basic knowledge activation to sophisticated metacognitive reflection. Students evolved from procedural questioning to conceptual inquiry while developing an enhanced capacity to connect abstract mathematical concepts with real-world applications. However, initial passive engagement among some participants highlighted the importance of structured implementation approaches that balance AI support with human guidance.

These findings contribute to theoretical understanding by demonstrating how Vygotsky's Zone of Proximal Development can be operationalized through generative AI technologies in mathematics education. The research offers practical frameworks for educators seeking to integrate ChatGPT as a reflective tool, while also identifying crucial directions for future investigation, including longitudinal studies on metacognitive skill retention and cross-cultural effectiveness research. While the study's scope focused on limit concepts within a specific timeframe, the results suggest that AI-mediated scaffolding holds substantial promise for enhancing mathematics education when implemented within carefully designed pedagogical frameworks that preserve the essential balance between technological support and student autonomy in mathematical learning environments.

REFERENCES

Adair, B., Sao Pedro, M., Gobert, J., & Segan, E. (2023). Real-time AI-driven assessment and scaffolding that improves students' mathematical modeling during science investigations. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos & V. Dimitrova (Eds.), *Artificial intelligence in education* (pp. 204-216). Springer. https://doi.org/10.1007/978-3-031-36272-9_17

Adiredja, A. P. (2021). Students' struggles with temporal order in the limit definition: Uncovering resources using knowledge in pieces. *International Journal of Mathematical Education in Science and Technology*, *52*(8), 1151-1172. https://doi.org/10.1080/0020739X.2020.1754477

Ahn, D., Son, T., & Lee, K. (2023). ChatGPT as a scaffolding tool: Evaluating the impact on elementary students' mathematical logic problem-solving skills. *Brain, Digital, & Learning*, *13*(2), 183-196. https://doi.org/10.31216/bdl.20230011

Almarashdi, H. S., Jarrah, A. M., Abu Khurma, O., & Gningue, S. M. (2024). Unveiling the potential: A systematic review of ChatGPT in transforming mathematics teaching and learning. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(12), Article em2412. https://doi.org/10.29333/ejmste/15739

Anderson, B. (1988). Computer-using students: Working in a metacognitive environment? *Language and Education*, *2*(4), 209-228. https://doi.org/10.1080/09500788809541238

Arroyo, I., Woolf, B. P., Burleson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, *24*, 387-426. https://doi.org/10.1007/s40593-014-0023-y

Arsyad, N., Ramadhana, Y., & Assagaf, S. F. (2021). Students' factual understanding on the concept of limit. *Journal of Physics: Conference Series*, *1899*, Article 012148. https://doi.org/10.1088/1742-6596/1899/1/012148

Beale, R. (2025). *Dialogic pedagogy for large language models: Aligning conversational AI with proven theories of learning.* PsyArXiv. https://doi.org/10.48550/arXiv.2506.19484

Beynon, K. A., & Zollman, A. (2015). Lacking a formal concept of limit: Advanced non-mathematics students' personal concept definitions. *Investigations in Mathematics Learning*, *8*(1), 47-62. https://doi.org/10.1080/24727466.2015.11790347

Contel, F., & Cusi, A. (2025). Investigating the role of ChatGPT in supporting metacognitive processes during problem-solving activities. *Digital Experiences in Mathematics Education*, *11,* 167-191. https://doi.org/10.1007/s40751-024-00164-7

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*, 444-452. https://doi.org/10.1007/s10956-023-10039-y

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.

da Silva, M. F., da Silva, W. L. M., & Barbosa, F. J. (2024). Contributions of Chat GPT to solving problems in education. *International Journal of Human Sciences Research*, *4*(12), 1-10. https://doi.org/10.22533/at.ed.558412402012

Dau, S., Jensen, C. G., & Gade, P. (2024). The role of AI chatbots in scaffolding: Linking learning outcomes with assessment. *Learning Technologies*, *9*(14), 73-97. https://doi.org/10.7146/lt.v9i14.141213

Deiana, G., Dettori, M., Arghittu, A., Azara, A., Gabutti, G., & Castiglia, P. (2023). Artificial intelligence and public health: Evaluating ChatGPT responses to vaccination myths and misconceptions. *Vaccines*, *11*(7), Article 1217. https://doi.org/10.3390/vaccines11071217

Deniz, G. (2023). The role of artificial intelligence (AI) in the academic paper writing and its prospective application as a co-author: A letter to the editor. *European Journal of Therapeutics*, *30*(1), 84-86. https://doi.org/10.58600/eurjther1808

Dimeli, M., & Kostas, A. (2025). The role of ChatGPT in education: Applications, challenges: Insights from a systematic review. *Journal of Information Technology Education: Research, 24*, Article 2. https://doi.org/10.28945/5422

Djudin, T. (2018). Preparing a strategic learner by using metacognitive strategies: From theory into practice. *Journal of Education and Learning*, *12*(3), 501-509. https://doi.org/10.11591/edulearn.v12i3.7208

Duckworth, A. L., Kautz, T., Defnet, A., Satlof-Bedrick, E., Talamas, S., Lira, B., & Steinberg, L. (2021). Students attending school remotely suffer socially, emotionally, and academically. *Educational Researcher*, *50*(7), 479-482. https://doi.org/10.3102/0013189X211031551

Exintaris, B., Karunaratne, N., & Yuriev, E. (2023). Metacognition and critical thinking: Using ChatGPT-generated responses as prompts for critique in a problem-solving workshop (SMARTCHEMPer). *Journal of Chemical Education*, *100*(8), 2972-2980. https://doi.org/10.1021/acs.jchemed.3c00481

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology, 56*(2), 489-530. https://doi.org/10.1111/bjet.13544

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs – Principles and practices. *Health Services Research*, 48(6pt2), 2134-2156. https://doi.org/10.1111/1475-6773.12117

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, *6*(1), 57-63. https://doi.org/10.37074/jalt.2023.6.1.22

Flaagan, T. (2023). Using ChatGPT to generate computer science problem sets. *International Journal on Integrating Technology in Education*, *12*(2), 19-29. https://doi.org/10.5121/ijite.2023.12202

Gabriel, F., Kennedy, J., Marrone, R., & Leonard, S. (2025). Pragmatic AI in education and its role in mathematics learning and teaching. *npj Science of Learning, 10*, Article 26. https://doi.org/10.1038/s41539-025-00315-4

Gartner, S., & Krašna, M. (2023). Ethics of artificial intelligence in education. *Journal of Elementary Education*, *16*(2), 221-235. https://doi.org/10.18690/rei.16.2.2846

Gerard, L. F., Ryoo, K., McElhaney, K. W., Liu, O. L., Rafferty, A. N., & Linn, M. C. (2016). Automated guidance for student inquiry. *Journal of Educational Psychology*, *108*(1), 60-81. https://doi.org/10.1037/edu0000052

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., … Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging era of AI chatbots. *Internet of Things and Cyber-Physical Systems*, *4*, 19-48. https://doi.org/10.1016/j.iotcps.2023.06.002

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education*, *100*(12), 4876-4883. https://doi.org/10.1021/acs.jchemed.3c00505

Hadi Mogavi, R., Deng, C., Juho Kim, J., Zhou, P., D. Kwon, Y., Hosny Saleh Metwally, A., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L. E., & Hui, P. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100027. https://doi.org/10.1016/j.chbah.2023.100027

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99-107. https://doi.org/10.1080/00461520701263368

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, *32*, 504-526. https://doi.org/10.1007/s40593-021-00239-1

Huang, B., Hew, K. F., & Lo, C. K. (2023). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments*, *27*(8), 1106-1126. https://doi.org/10.1080/10494820.2018.1495653

Huda, N., Syaiful, S., Gustiningsi, T., & Novferma, N. (2025). Students' metacognitive ability in solving mathematical problems based on information processing theory in discrete mathematics course. *Edumatica: Jurnal Pendidikan Matematika*, *15*(2), 283-295. https://doi.org/10.22437/edumatica.v15i2.44689

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), Article 584. https://doi.org/10.3390/math9060584

Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. *British Journal of Educational Technology*, *54*(5), 1057-1076. https://doi.org/10.1111/bjet.13325

Jin, L., Zhao, Y., Huang, J., & Zhang, K. (2023). A qualitative systematic review on AI empowered self-regulated learning in higher education. *npj Science of Learning*, *10*, Article 21. https://doi.org/10.1038/s41539-025-00319-0

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189X033007014

Jones, S. R. (2015). Calculus limits involving infinity: The role of students' informal dynamic reasoning. *International Journal of Mathematical Education in Science and Technology*, *46*(1), 105-126. https://doi.org/10.1080/0020739X.2014.941427

Joshi, I., Budhiraja, R., Dev, H., Kadia, J., Ataullah, M. O., Mitra, S., Akolekar, H. D., & Kumar, D. (2024). ChatGPT in the classroom: An analysis of its strengths and weaknesses for solving undergraduate computer science questions. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education* (pp. 625-631). Association for Computing Machinery. https://doi.org/10.1145/3626252.3630803

Kamid, Anwar, K., Syaiful, Sofnidar, Liani, L., & Kurniawan, W. (2020). Investigation into first-year college students' misconceptions about limit concept: A case study based on cognitive style. *Universal Journal of Educational Research*, *8*(4), 1445-1452. https://doi.org/10.13189/ujer.2020.080437

Kofahi, M. A., & Husain, A. J. (2025). ChatGPT for operating systems: Higher-order thinking in focus. *Journal of Information Technology Education: Research, 24*, Article 1. https://doi.org/10.28945/5382

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://doi.org/10.2307/2529310

Lelepary, H. L., Rachmawati, R., Zani, B. N., & Maharjan, K. (2023). GPT Chat: Opportunities and challenges in the learning process of Arabic language in higher education. *Journal International of Lingua and Technology*, *2*(1), 11-23. https://doi.org/10.55849/jiltech.v2i1.439

Li, D., & Zhang, L. (2022). Contextualizing feedback in L2 writing: The role of teacher scaffolding. *Language Awareness*, *31*(3), 328-350. https://doi.org/10.1080/09658416.2021.1931261

Liu, R., Pang, W., Chen, J., Balakrishnan, V. A. P., & Chin, H. L. (2024). The application of scaffolding instruction and AI-driven diffusion models in children's aesthetic education: A case study on teaching traditional Chinese painting of the twenty-four solar terms in Chinese culture. *Education and Information Technologies*, *30,* 9129-9160. https://doi.org/10.1007/s10639-024-13135-7

Loc, N. P., Oanh, N. P. P., Thao, N. P., De, T. V., & Triet, L. V. M. (2022). Activity theory as a framework for teaching mathematics: An experimental study. *Heliyon*, *8*(10), e10789. https://doi.org/10.1016/j.heliyon.2022.e10789

Malik, M. I., Najera, M., Chang, K. E., Demszky, D., & Azevedo, R. (2025). Scaffolding middle school mathematics curricula with large language models. *British Journal of Educational Technology*, *56*(3), 999-1027. https://doi.org/10.1111/bjet.13571

Marbun, T. O. (2023). The implementation of artificial intelligence, ChatGPT, and critical thinking method for academic endeavors at theological higher education. *Journal Didaskalia*, *6*(2), 84-100. https://doi.org/10.33856/didaskalia.v6i2.309

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856. https://doi.org/10.3390/educsci13090856

Mindigulova, A. A., Vikhman, V. V., & Romm, M. V. (2024). Artificial intelligence and personalized learning: Scaffolding technology. *Professional Education in the Modern World*, *13*(4), 613-622. https://doi.org/10.20913/2618-7515-2023-4-3

Moore, K. (2023, May 25). *Using ChatGPT to plan high school math lessons.* Edutopia. https://www.edutopia.org/article/using-chatgpt-plan-high-school-math-lessons/

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, *53*(4), 914-931. https://doi.org/10.1111/bjet.13232

Norberg, K. A., Almoubayyed, H., De Ley, L., Murphy, A., Weldon, K., & Ritter, S. (2025). Rewriting content with GPT-4 to support emerging readers in adaptive mathematics software. *International Journal of Artificial Intelligence in Education*, *35*, 587–626. https://doi.org/10.1007/s40593-024-00420-2

Opticia, N. N., Khabibah, S., & Masriyah, M. (2022). Development of guided inquiry model mathematics learning tools to practice critical thinking skills for students in linear program materials. *International Journal of Trends in Mathematics Education Research*, *5*(2), 133-140. https://doi.org/10.33122/ijtmer.v5i2.95

Ortega-Ochoa, E., Quiroga Pérez, J., Arguedas, M., Daradoumis, T., & Marquès Puig, J. M. (2024). The effectiveness of empathic chatbot feedback for developing computer competencies, motivation, self-regulation, and metacognitive reasoning in online higher education. *Internet of Things, 25*, 101101. https://doi.org/10.1016/j.iot.2024.101101

Ouyang, F., Dai, X., & Chen, S. (2022). Applying multimodal learning analytics to examine the immediate and delayed effects of instructor scaffoldings on small groups' collaborative programming. *International Journal of STEM Education*, *9*, Article 45. https://doi.org/10.1186/s40594-022-00361-z

Pardos, Z., & Bhandari, S. (2023). *Learning gain differences between ChatGPT and human tutor generated algebra hints*. PsyArXiv. https://doi.org/10.48550/arXiv.2302.06871

Payadnya, I. P. A. A., Putri, G. A. M. A., Suwija, I. K., Saelee, S., & Jayantika, I. G. A. N. T. (2025). Cultural integration in AI-enhanced mathematics education: Insights from Southeast Asian educators. *Journal for Multicultural Education*, *19*(1), 58-72. https://doi.org/10.1108/jme-09-2024-0119

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47-61. https://doi.org/10.1016/j.edurev.2015.02.003

Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). Mathematics education in the era of ChatGPT: Investigating its meaning and use for school and university education – Editorial to special issue. *Digital Experiences in Mathematics Education*, *11*, 1-8. https://doi.org/10.1007/s40751-025-00173-0

Qureshi, B. (2023). *Exploring the use of ChatGPT as a tool for learning and assessment in undergraduate computer science curriculum: Opportunities and challenges.* PsyArXiv. https://doi.org/10.48550/arXiv.2304.11214

Reiss, M. J. (2021). The use of AI in education: Practicalities and ethical considerations. *London Review of Education*, *19*(1). https://doi.org/10.14324/LRE.19.1.05

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, *13*(10), 6039. https://doi.org/10.3390/app13106039

Schorcht, S., Buchholtz, N. F., & Baumanns, L. (2024). Prompt the problem: Investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education*, *9*, 1386075. https://doi.org/10.3389/feduc.2024.1386075

Son, T. (2024). Intelligent tutoring systems in mathematics education: A systematic literature review using the substitution, augmentation, modification, redefinition model. *Computers*, *13*(10), 270. https://doi.org/10.3390/computers13100270

Swinyard, C., & Larsen, S. (2012). Coming to understand the formal definition of limit: Insights gained from engaging students in reinvention. *Journal for Research in Mathematics Education*, *43*(4), 465-493. https://doi.org/10.5951/jresematheduc.43.4.0465

Tashakkori, A., & Teddlie, C. (Eds.). (2015). *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Sage. https://doi.org/10.4135/9781506335193

Thompson, P. W., & Harel, G. (2021). Ideas foundational to calculus learning and their links to students' difficulties. *ZDM – Mathematics Education*, *53*, 507-519. https://doi.org/10.1007/s11858-021-01270-1

Tong, D., Tao, Y., Zhang, K., Dong, X., Hu, Y., Pan, S., & Liu, Q. (2024). Investigating ChatGPT-4's performance in solving physics problems and its potential implications for education. *Asia Pacific Education Review*, *25,* 1379-1389. https://doi.org/10.1007/s12564-023-09913-6

UNESCO. (2021). *AI and education: Guidance for policy-makers.* https://unesdoc.unesco.org/ark:/48223/pf0000376709

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education, 215,* 105031. https://doi.org/10.1016/j.compedu.2024.105031

Van der Stel, M., Veenman, M. V. J., Deelen, K., & Haenen, J. (2010). The increasing role of metacognitive skills in math: A cross-sectional study from a developmental perspective. *ZDM Mathematics Education*, *42*, 219-229. https://doi.org/10.1007/s11858-009-0224-2

Van Vaerenbergh, S., & Pérez-Suay, A. (2022). A classification of artificial intelligence systems for mathematics education. In P. R. Richard, M. P. Vélez & S. Van Vaerenbergh (Eds.), *Mathematics education in the age of artificial intelligence* (pp. 89-106). Springer. https://doi.org/10.1007/978-3-030-86909-0_5

Vasconcelos, M. A. R., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(6), em2296. https://doi.org/10.29333/ejmste/13313

Vygotsky, L. S. (2018). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Walter, S. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, *21*, Article 15. https://doi.org/10.1186/s41239-024-00448-3

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7), Article em2286. https://doi.org/10.29333/ejmste/13272

Wei, Q. (2023). Changes in English language education and strategies for coping against Chat GPT's effects. *Advances in Education, Humanities and Social Science Research*, *6*(1), 471-479. https://doi.org/10.56028/aehssr.6.1.471.2023

Xing, W., Song, Y., Li, C., Liu, Z., Zhu, W., & Oh, H. (2025). Development of a generative AI-powered teachable agent for middle school mathematics learning: A design-based research study. *British Journal of Educational Technology*, *56*(5), 2043-2077. https://doi.org/10.1111/bjet.13586

Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology*, *56*, 1842–1863. https://doi.org/10.1111/bjet.13599

Yerushalmy, M. (2015). E-textbooks for mathematical guided inquiry: Design of tasks and task sequences. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education* (pp. 229-247). Springer. https://doi.org/10.1007/978-3-319-09629-2_7

Zhai, X. (2022). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4312418

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education*, 21, Article 34. https://doi.org/10.1186/s41239-024-00467-0

Zhou, Y., Wang, X., & Chen, L. (2025). Exploring ChatGPT-facilitated scaffolding in undergraduates' mathematical problem solving. *Journal of Computer Assisted Learning*, *41*(4), e70077. https://doi.org/10.1111/jcal.70077

AUTHORS

Nizlel Huda, Dr., Dra., MKes, earned her bachelor's degree in Statistics from Universitas Airlangga. She then continued her academic journey by completing a Doctoral degree in Mathematics Education at Universitas Negeri Malang, with a research focus on metacognition in mathematics learning. Currently, she serves as a lecturer in the Mathematics Education Program, Faculty of Teacher Training and Education, Universitas Jambi. Her teaching and research primarily center on statistical education and metacognitive approaches in mathematics learning.

Khairul Anwar, MPd, is an Assistant Lecturer in Mathematics Education and currently serves as a faculty member in the Mathematics Education Program at Universitas Jambi. He earned his bachelor's degree in mathematics education from the Faculty of Teacher Training and Education at Universitas Jambi and completed his master's degree in mathematics education in the Graduate Program at Universitas Negeri Semarang. His research interests include the integration of technology in mathematics education, the development of interactive learning media, as well as programming education and robotics education.

Novferma, SPd, MPd, earned her bachelor's degree in mathematics education from Universitas Jambi. She continued her academic journey by completing a master's degree in mathematics education at Universitas Negeri Yogyakarta, with a research focus on instructional design in mathematics education. Currently, she serves as a lecturer in the Mathematics Education Program, in the Faculty of Teacher Training and Education, Universitas Jambi. Her teaching and research interests primarily center on the design and development of effective mathematics learning.

Wawan Kurniawan, MCs, graduated with a Bachelor of Science in Physics from the Faculty of Mathematics and Natural Sciences at Universitas Sumatra Utara. He continued his academic journey by earning a Master of Computer Science from Universitas Gadjah Mada. Currently, he serves as a lecturer in the Physics Education Program, in the Faculty of Teacher Training and Education, Universitas Jambi. His teaching and research focus primarily on the intersection of physics education and artificial intelligence (AI). His expertise includes developing AI-based learning environments to enhance student engagement and understanding.