Exploring Potential Uses and Concerns of GenAI Use in Problem Solving in the Mathematics Classroom

Nirmala Naresh[1], **Zuhal Yilmaz**[2], **Iris DeLoach Johnson**[3], & **Deborah Cockerham**[3]
[1]University of North Texas, [2]North Carolina State University, [3]LSU Shreveport, Retired

Abstract

This two-part longitudinal study explored the integration of generative artificial intelligence (GenAI) in mathematics education by examining the perspectives, knowledge, and experiences of in-service and preservice secondary mathematics teachers. The study began with a 7-item questionnaire completed by in-service secondary mathematics teachers in the West South-Central region of the United States to identify their initial views and concerns regarding GenAI. A volunteer subgroup then participated in an exploratory task-based activity, completing one of eight structured mathematical problems using a GenAI tool. Insights from this exploratory work informed the second part of the study: a case study involving preservice teachers who engaged in open-ended mathematical modeling tasks with GenAI. Analysis of the teachers' engagement focused on their teacher-AI interactions, such as the prompt types in communicating with the AI Tool (e.g., Ask, Affirm/Seek Validation, Guide) and ways their interactions addressed concerns highlighted in previous research. Further analysis provided a deeper awareness of teachers' reactions and beliefs regarding conceptual understanding that had been supported during the GenAI problem-solving experience and related GenAI issues. This study concluded with the authors' pedagogical reflections regarding challenges, limitations, and plans for future use of AI tools.

Keywords: generative AI, mathematics education, teacher-AI interaction, learner-AI engagement framework

1 Introduction

Innovations in digital technology often impact teaching and student learning with implementation varying in scope, depth, speed, and fidelity (Organization for Economic Co-operation & Development [OECD], 2021; Timotheou, 2023). For example, the proper implementation of calculator use in mathematics education continues to spark some element of debate after more than three decades (Banilower et al., 2013; National Advisory Committee on Mathematical Education [NACOME], 1975; National Council of Teachers of Mathematics [NCTM], 1978, 1980; National Research Council [NRC],1989; Paige et al., 2006; Ronau et al., 2011; U.S. Department of Education [USDOE], 2001). However, society's widespread adoption of a relatively new technological advance known as artificial intelligence (AI) has rapidly expanded into our daily lives and mathematics classrooms (The School Superintendents Association [AASA] et al., 2023; Chen, Chen, & Lin, 2020; Center on Reinventing Public Education [CRPE], 2024b; Hwang & Tu, 2021; USDOE, 2023). AI, a branch of computer science that engages machines in a powerful mimicry of human intelligence, has evolved for more than seven decades (McCarthy et al., 2006; Mohamed et al., 2022; Turing, 1950). It is "not a single [technological] innovation, but an umbrella term" (USDOE, 2023, p. 11). This article will share information about one type of AI that may be employed during mathematical problem-solving, interactions between the learners and AI, cognitive demands on learners, and the need to actively explore AI tools without the fear of making mistakes. Broadly, there are three main types of AI. Reactive AI (e.g., voice assistants like Alexa, automated vacuum cleaners) typically responds with specialized functions or tasks requested by a person or other entity (AASA et al., 2023; Alsharidah & Alkramiti, 2024; Milicevic et al., 2024). Predictive AI (e.g., Netflix, early warning systems for flood-prone areas) typically makes decisions about future use based upon data patterns gathered about habits of a person, group, or other entity (AASA et al., 2023; Alsharidah & Alkramiti, 2024). Generative AI (GenAI) (e.g., ChatGPT, Khanmigo, self-driving cars) typically seems to synthesize and learn from data to create new products that previously could only be produced by humans (AASA et al., 2023; Adžić et al., 2024; Asamoah et al., 2024; Baidoo-Anu & Ansah, 2023; Haddley & Ardito, 2024; Yoon et al., 2024). This article focuses on GenAI as an area with the most potential and challenge for our mathematics classrooms (Asamoah et al., 2024). Analysis of questionnaire responses and observations of teacher-AI interactions while engaging in mathematical problem-solving tasks led to the following research questions:

RQ1: How do specific learner-AI interactions during mathematical problem-solving support or challenge teachers' concerns about using GenAI to do mathematical problem-solving in their classrooms?

RQ2: What types of cognitive demand are exhibited by learners (prospective teachers, hereafter referred to as PSTs) as they engage in different levels of cognitive engagement when interacting with GenAI to do mathematical problem-solving?

2 Literature Review

Many teachers may be surprised that their students have more experience—favorable and unfavorable—with today's widely accessible GenAI tools than they do (AASA et al., 2023; Baidoo-Anu & Ansah, 2023; CRPE, 2024a; Yonder Consulting, 2025). With all the power that GenAI brings, there is often a dual nature associated with it: promising vs. potentially threatening, safe vs. unsafe, and informative vs. misleading (Lim, 2023; Milicevic et al., 2024; USDOE, 2023). However, the prevalence of AI in the everyday lives of society, combined with potential future impact on every aspect of our lives, assures us that we can not go back to the pre-AI world (Cacho, 2024; Dykema, 2023; Fabijanić Gagro, 2024; Lim et al., 2023). The ominous message delivered from the National Research Council (1989) almost four decades ago seems more formidable now than then:

> Priorities for mathematics education must change to reflect the way computers [technology] are used in mathematics ... establish[ing] new ground rules for mathematics education ... (T)he stable diet ... will diminish under the assault of machines that specialize in mimicry. (p. 63)

The magnitude of concerns for future use of AI regarding individual privacy, beneficence, autonomy, fairness, bias, and transparency is concerning (CRPE, 2024a; USDOE, 2023). There is an urgency for teachers, educational leaders, policymakers, and government officials to know AI well to "harness the good" to serve educational and societal priorities (USDOE, 2023, p. 2). Given these realities, more proactive work lies ahead for international, national, and regional policymakers to involve all stakeholders in preparing guidelines and regulations for proper use. Various professional educational organizations released position statements challenging teachers to actively explore AI and to teach students to be strategic users of AI while remaining current about AI trends to support student learning (CRPE, 2024b; NCTM, 2024b; USDOE, 2023).

We should move forward "now to realize key opportunities, prevent and mitigate emergent risks, and tackle unintended consequences" (USDOE, 2023, p. 3). Furthermore, to harness the power of AI we must explore GenAI without fear of "making mistakes as we learn and experience the capabilities and limitations of AI tools" (AASA et al., 2023, p. 2). This study investigated a move forward via unfettered exploration of GenAI and participants' reflections upon its use while engaging in mathematical problem-solving.

To explore how generative AI (GenAI) can support mathematical problem-solving, it is useful to briefly revisit the historical evolution of problem-solving in mathematics education, the essential skills required by human learners, pedagogical principles that promote effective problem-solving, and the capabilities of AI to mimic human processes required to solve mathematical problems successfully.

Prior to 1989, mathematical problem-solving was often relegated to solving word problems using arithmetic, algebraic, or geometrical procedures. Instructional models during this time often focused on Polya's (2014) four-step problem-solving model: understand the problem, devise a plan, execute the plan, and reflect upon the solution. Educators taught students to rely on reading comprehension and critical thinking to interpret keywords, symbols, and problem contexts—strategies that AI can now replicate with remarkable speed and high incidences of accuracy. Since 1989, problem-solving has been identified as a foundational process in mathematics learning beyond merely solving word problems (NCTM, 1989, 2000, 2024a). Mathematical problem-solving requires learners to use a variety of cognitive skills such as interpreting language (provided in words or symbols), evaluating solution strategies, and communicating their reasoning, which are each within the domain of AI-powered tools. However, the potential of GenAI must also be critically examined for alignment of pedagogical goals to foundational instructional frameworks such as the Effective Mathematics Teaching Practices, which highlight the pedagogical importance of "setting meaningful goals, implementing reasoning-based tasks, … facilitating discourse, … developing procedural fluency from conceptual understanding, … and using evidence of student thinking" (NCTM, 2024a, p. 3). The Standards for Mathematical Practice also highlight pedagogical roles specifically for mathematical problem-solving: "making sense of problems, reasoning abstractly, … [and] using tools [including AI] strategically…" (NGA & CCSSO, 2010, pp. 6-8). AI should support—never replace—these essential competencies.

Effective problem-solving involves reasoning through mathematical situations, interpreting symbolic language, evaluating the plausibility of solutions, articulating strategies, and making connections to prior knowledge (Ahn et al., 2024; NCTM, 2000; Rane, 2023; Wardat et al., 2023; Yoon et al., 2024). If GenAI tools bypass these processes by generating multi-step solutions without learner interaction, their educational value diminishes (Ahn et al., 2024). Various GenAI tools (e.g., ChatGPT, Claude.ai, Khanmigo, SNORKL, MagicSchool, Knowledgehook) are currently being used to support mathematical problem-solving. Open-ended tools like ChatGPT may require more oversight to ensure educational goals are met, while tutorial-based tools (e.g., Khanmigo, SNORKL) typically offer structured scaffolding, feedback loops, and checks for understanding that mimic teacher-student interactions (Ahn et al., 2024; Lawasi et al., 2024).

Prompt engineering—the ability to effectively communicate with AI—is a critical skill (Biton & Segal, 2025; Fagbohun et al., 2024). Superficial engagement with AI, in which students accept responses without critique, risks mirroring the services of human tutors that lack either content knowledge or pedagogical rigor (Ahn et al., 2024; Opesemowo & Ndlovu, 2024). Evidence suggests that students using GenAI for practice may perform well initially, but without deeper understanding, long-term outcomes may be no better—or worse—than those of students who practiced without AI (Bastani et al., 2024). Learning gains are more pronounced when students question AI-generated feedback, attempt problems independently, and reflect critically on solutions (NCTM, 2024a). Despite frequent disclaimers about potential inaccuracies, users often ignore these warnings. Some GenAI systems can misinterpret problems, provide incorrect solutions, or alter problem statements without user awareness (Ahn et al., 2024; Bang et al., 2023). Effective use of GenAI in mathematical problem-solving hinges on three factors: (a) leveraging AI tools that support tutorial-style learning rather than simple answer generation (Gilbert et al., 2015; Pepin et al., 2025), (b) training users to use prompts strategically and critically (Park & Choo, 2024), and (c) fostering human oversight to ensure that AI serves pedagogical—not just computational—goals (Authors et al., 2025; Gabriel et al., 2025).

3 Methodology

For this study, we employed a longitudinal design that collected data over the course of an academic year. All procedures were approved by the first author's Institutional Review Board, and informed consent was obtained from all participants. Participants included in-service secondary mathematics teachers from the West South-Central region of the United States and preservice teachers (PSTs) enrolled at a Hispanic-Serving Institution in the same area. This study incorporated three key sources of data, each building on the previous to investigate how educators engage with generative AI (GenAI) in mathematical problem solving. Table 1 provides an overview of data collected and analyzed along with the purposes for each.

Table 1. *Data Overview*

Data Source	Participants	Purpose
Teacher questionnaire (Qualitative survey)	50 inservice Teachers	Identify teacher concerns and guide next steps
Exploratory task-based interview (Qualitative data)	Ten inservice teachers (One featured in this paper)	Explore AI use in problem solving; inform future design
Case Study (Qualitative data)	22 preservice teachers (Two featured in this paper)	Investigate deeper engagement and adaptive supports

In the following sections, we provide detailed descriptions of the data collected and the corresponding analysis techniques. Quantitative analysis methods used for the teacher questionnaire are described within that section. For the exploratory task-based interviews and the case study, qualitative analysis was conducted using a three-component prompt categorization framework, which is introduced prior to the presentation of those findings.

3.1 Teacher Questionnaire

We designed and administered a 7-item questionnaire (see Appendix A) to gather perspectives and experiences with AI in the mathematics classrooms of secondary mathematics teachers (n = 50). The questionnaire examined participants' familiarity with AI, including exposure to AI applications, use of AI in the classroom, and awareness of AI curricula or tools. We included open-ended questions to capture teachers' views on AI in mathematics education, including concerns, benefits, and suggestions for future use.

We used two stages of coding to inform our subsequent research cycle and develop targeted recommendations for addressing teacher concerns. In the first cycle, analytic memos and in vivo coding (Saldaňa, 2012) were used to document patterns and identify emerging themes, especially around teachers' concerns on AI usage, such as potential decline in critical thinking, superficial engagement, and accuracy of AI responses. In the second cycle, pattern coding (Saldaňa, 2012) was applied to refine and group these themes. Through this process, individual concerns were organized into broader categories and situated within the larger context of AI in teaching and learning. Some quotes were multi-dimensional in that certain parts of the shared insight were coded under multiple concerns. Table 2 illustrates teacher concerns from the survey with examples for each concern.

Table 2. *In-service Teachers' Concerns about AI Use in Mathematics Classrooms*

Concern	Example Quotes
Superficial engagement with mathematical concepts at the expense of deeper conceptual understanding ($n = 11$)	only see AI being used in my classroom by students as a quick way to get answers for problems. … "just give them the answers" without learning the concepts"
Potential exploitation of AI for cheating ($n = 8$)	The kids who want to cheat, they are going to cheat, no matter what, but like as a student who wants to learn, I think this [Khanmigo] would be a useful tool.
Generation of mathematically inaccurate responses ($n = 2$)	There were a couple of inaccuracies that it pointed me in the wrong direction because I gave it incorrect instructions, and I recognize that as a teacher. But I'm not sure the students would recognize that… It doesn't start pointing you in the right direction until you prompt it and push it a couple of times…I can tell you that the students are gonna get frustrated…
AI Literacy ($n = 7$)	Students "may not be prepared or educated on how to ask AI questions, at least in this AI's format" There were a couple of inaccuracies that it pointed me in the wrong direction because I gave it incorrect instructions, and I recognize that as a teacher. But I'm not sure the students would recognize that…

These findings directly guided the design of the exploratory task-based interviews, and the case study specifically constructed to address the concerns identified by the teachers. While the task-based interviews explore the engagement of in-service teachers with AI tools in a structured task-based setting, the Case Study shifts the focus to PSTs who, despite having less exposure to AI, were also tasked with mathematical problem-solving in a similar format.

3.2 Exploratory Task-based Interviews with Inservice Teachers

Of the 50 secondary mathematics teachers who completed the questionnaire, 10 volunteered to participate in task-based interviews. These in-service teachers were asked to use Khanmigo, an AI powered tool which assists in learning by providing personalized learning experiences, acting as a virtual tutor that helps students understand complex concepts, practice problems, and giving real-time feedback (Ofgang, 2023). Teachers engaged with Khanmigo for 45-60 minutes in an unstructured setting to complete at least one problem from the set of eight problem-solving tasks (see Appendix B for example tasks). Each task had a single correct solution and addressed topics such as algebra, geometry, and number sense. The tasks were selected from a publicly available collection of rich mathematical tasks developed by the Virginia Department of Education (VDOE, n.d.), which were designed to align with the 2016 Mathematics Standards of Learning and promote reasoning, problem solving, and deep mathematical thinking. Teachers engaged in think-aloud protocols (Lajoie, 2008) while interacting with the AI. To facilitate natural use of the tool, the researchers provided no guidance or intervention during real-time use of AI.

For discussion here, we present the Bake Sale Fundraiser task (see Appendix B), which is one of the eight problem-solving tasks requiring multiple steps of numerical reasoning and interpretation of contextual information. Conceptually, the task includes basic operations with common fractions, including calculations with decimals and percents characteristic of the Numbers & Operations and Algebra Standards (NCTM, 1989). The problem could be solved using a linear equation or informal arithmetic strategies, making it useful for observing teacher interaction with AI support during real-world problem-solving.

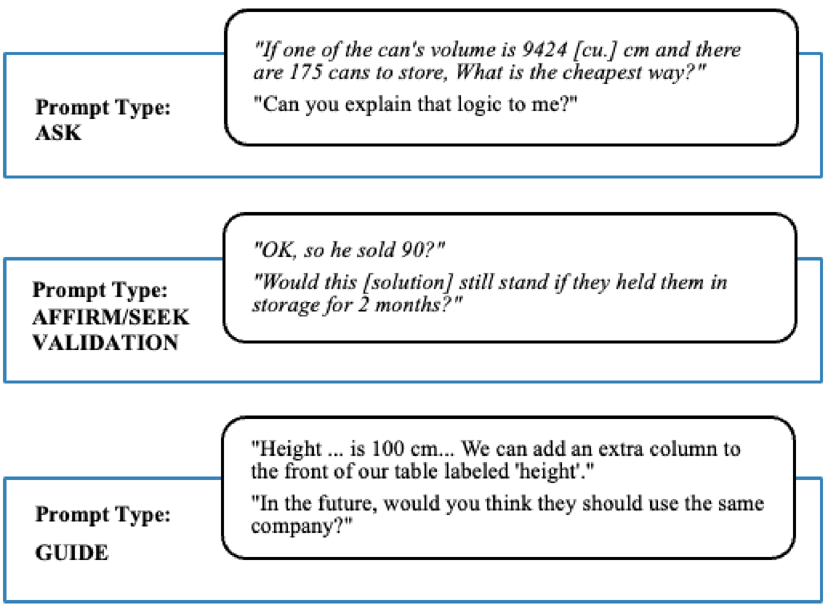
3.3 Qualitative Data Analysis

In analyzing the interactions of both in-service teachers and PSTs with the AI tools, our goal was to examine how these interactions either aligned with or diverged from teacher concerns (see Table 2), particularly regarding AI accuracy, engagement, and problem-solving. This layered approach to data collection enabled us to capture how AI-supported problem-solving unfolded in practice, highlighting specific ways AI tools can address or exacerbate concerns related to student engagement and mathematical learning. To inform our approach to analyzing teacher-AI interactions, we reviewed relevant literature in human-AI interaction (e.g., Baker et al., 2004; Holstein et al., 2019), dialogue analysis in tutoring systems (Graesser et al., 1995), and teacher discourse practices (O'Connor & Michaels, 2019). This literature helped us better understand common patterns of interaction and consider how a framework might capture different types of engagement. Drawing on these insights, we examined the prompts generated during problem-solving and, through close reading and collaborative discussion, identified recurring patterns. Focusing on the types of prompts the participants created during their AI interactions, the prompts were categorized into three broad types:

1. Ask: The teacher restates the question, requests help with strategies or definitions or asks for step-by-step assistance.
2. Affirm/Seek Validation: The teacher confirms the AI's response, performs calculations, or agrees with the AI's output.
3. Guide: The teacher corrects the AI's mistakes, proposes alternative strategies, or extends the conversation with their own reasoning.

These categories emerged inductively as the researchers sorted prompts based on their function and learner intent within the problem-solving process (See Figure 1). This coding scheme allowed us to interpret how the teachers engaged with the AI, and how those patterns related to concerns about accuracy, engagement, and critical thinking.

Figure 1
Examples for Each Prompt Type



3.3.1 Task-based Interview Analysis with Inservice Teacher Jana

We feature the work of one teacher, Jana (pseudonym), whose response reflects a pattern we observed across several participants. Jana approached the Bake Sale Fundraiser Task by breaking it into smaller, manageable parts and sequentially asking the AI for help with each. Jana began by identifying key details from the problem statement and posed a series of ask-type prompts to confirm these details. We noticed that Jana relied on the AI to extract numerical information and perform calculations, such as determining the total cost of ingredients or calculating revenue. Jana checked in with the AI frequently after each step, often seeking confirmation before moving forward. Table D1 (Appendix D) illustrates the type of prompts used and how they were used. We noticed that Jana approached the problem tentatively and leaned heavily on the AI to validate and guide the solution process. We found evidence of limited engagement with the underlying mathematics. Jana used the AI to move through the task and followed its suggested steps without pausing to reason through or justify the approach. This pattern was consistent across other participants in this study, where we observed that many teachers were engaging with the AI tool in a pragmatic, task-focused manner. Their interactions often followed a step-by-step pattern, with limited elaboration on strategy or conceptual reasoning. Based on our findings from the task-based interviews with in-service teachers, we made several changes to the Case Study to support more meaningful engagement for the PSTs. First, we introduced open-ended mathematical modeling tasks that allowed for multiple strategies and interpretations. Second, we facilitated a learning session to support PSTs in reflecting on how to engage with AI tools in more meaningful ways, focusing on productive ways to engage with AI during problem-solving. Finally, we allowed participants to choose Khanmigo or ChatGPT as the generative AI tool they preferred to better reflect the kind of tools they might encounter or use in their own classrooms.

3.3.2 Case Study with Prospective Teachers (PSTs)

In this case study, we focused on PSTs ($n = 22$) who engaged AI with mathematical modeling tasks. The PSTs, mostly sophomores and juniors, were enrolled in mathematics content courses and were majoring in mathematics to obtain certification to teach middle or high school mathematics. These PSTs had not yet completed any methods courses related to teaching and had little to no prior experience using AI in their mathematics coursework.

To support deeper engagement with mathematics and AI, we intentionally chose mathematical modeling tasks because they support sense-making, allow for multiple correct solutions, and promote generalizations. While our initial selection was not explicitly guided by the Effective Mathematics Teaching Practices (NCTM, 2014a, 2024), the modeling tasks we used align with several of those practices, including supporting productive struggle, facilitating meaningful mathematical discourse, and using and connecting mathematical representations (NGA & CCSSO, 2010). The PSTs worked with the How to Store Task (Erbaş et al., 2016; see Appendix C), which was designed to encourage creative problem-solving to find the most cost-effective way to store 175 cylindrical cans. A table identifying varying sizes and rental costs of storage cabinets was included with the task. The objective was to minimize costs while ensuring the cans were stored upright for safety. The PSTs were instructed to attempt to solve problems independently before consulting AI for alternative approaches or suggestions, then use AI as a collaborative thinking partner rather than merely a tool for generating answers. The PSTs were instructed to choose ChatGPT or Khanmigo based on their preferences. This flexibility provided insight into how the different tools impacted the problem-solving process. The goal was to explore how PSTs, with little to no prior experience with AI, navigated mathematical problem-solving with an AI tool. To illustrate how the mathematical engagement with the AI tool developed after we adapted the task design, we share two examples from a case study (featuring PSTs with pseudonyms of Pari and Vira). Both PSTs used ChatGPT to engage with the modeling task.

3.3.3 Analysis of Preservice Teacher Pari's Engagement with ChatGPT

Pari input the task, including the table, unchanged, into ChatGPT, prompting it to find a solution. The table indicated the width, length, and rental cost per month for each storage unit, but did not specify the height of the unit. ChatGPT, in response, calculated the volume of each can using the cost per month as the height, albeit incorrectly (referred to as an AI hallucination; Bang et al., 2023). Pari did not challenge this mistake.

Excerpt:

- Pari: If one of the can's volume is 9424 cm and there are 175 cans to store. What is the "way?"
- ChatGPT: Responds with cost per unit volume based on flawed cabinet volume calculations.

During the interaction, Pari did not question the generated response but proceeded to the next step, prompting, "What would be the cheapest way to store 175 cans?" In response, ChatGPT suggested a strategy focused on calculating the cost per unit volume and comparing it for each storage unit. Without questioning the AI's assumptions or identifying the flawed cabinet volume calculations, Pari carried out calculations based on the suggested strategy and concluded that Storage Unit 3 would be the best option to minimize cost.

Excerpt:

- Pari: So, Number 3, the best price?
- ChatGPT: Affirms conclusion.

Toward the end, Pari shifted focus slightly, asking more open-ended questions related to future business decisions and cost volatility, like "Do you think the price will change?" and "In the future, would you think they should use the same company or change?" Table D2 (Appendix D) provides examples of the categorization of the various prompts that Pari provided AI.

This interaction unfolded into a dialogue shaped by the learner's trust in the AI's authority. Pari followed the AI's lead, used its strategy without question, and rarely pushed back. The exchange was mostly linear: Pari asked, the AI responded; and Pari moved forward based on the AI's advice. However, in the last few prompts, we notice a shift from procedural problem-solving towards more strategic engagement characteristic of business decision-making.

3.3.4 Preservice Teacher Vira's Engagement with ChatGPT

Vira initiated the interaction with ChatGPT by summarizing the task information. "There will be three columns. The first column will be called width, the second column will be called length, and the third column will be called rental cost per month. Table D3 (Appendix D) shares the categorization of examples of the various prompts that Vira communicated to AI. Vira's interaction with the AI tool indicates a trend from basic inquiry to deeper cognitive engagement. Early in the conversation, Vira primarily used Ask-type prompts, entering task details and requesting calculations or strategies (e.g., asking how to determine the best cabinet or how much space a can occupies). These prompts reflected an information-seeking approach typical at the beginning of problem-solving.

Vira progressively shared additional information from the task, guiding the AI tool in constructing a table collaboratively. This approach, distinct from foundational engagement, demonstrated that the learner was taking an active role in shaping the engagement by making sure that all relevant information was included in the table. We noticed a brief Affirm/Seek Validation phase where the student accepted the AI's reasoning and proceeded with calculations.

A specific prompt used by Vira was, "The height of each storage cabinet is 100 cm. Therefore, we can add an extra column to the front of our table titled 'height'." This proactive approach prevented the ChatGPT from generating a response using an incorrect height. Notably, in this instance, Vira assumed ownership of directing the mathematical engagement. When Vira posted the following prompt, "We need to determine which storage cabinet would be the most cost-effective to store these cans. How could we do this?" ChatGPT responded by calculating the cost per can for each cabinet. During this process, Vira actively checked the mathematical accuracy of the calculations rather than assuming their correctness. Upon review, Vira discovered that ChatGPT had incorrectly calculated the volume of one can as 3,140 cm^3 instead of the correct value, 9,420 cm^3. Despite this discrepancy, the AI initially determined Storage Unit 3 to be the most economical choice, overlooking 175 as the number of cans to be stored.

Midway through the interaction, a notable shift occurred: Vira began to reframe the problem. Instead of comparing all cabinet volumes, Vira guided the discussion toward identifying the cheapest cabinet that could still store 175 cans. This is where Guide-type prompts and higher-order engagement prompts began to happen. Vira further clarified their objective, proposed simplified criteria, asked about long-term cost implications, and took ownership over the problem-solving process. To summarize, Vira's AI-interaction shows a progression from passive information gathering to active problem solving.

4 Discussion

This study explored two central research questions:

RQ1: How do specific learner-AI interactions during mathematical problem-solving support or challenge teachers' concerns about using GenAI to do mathematical problem-solving in their classrooms?

RQ2: What types of cognitive demand are exhibited by learners (prospective teachers, hereafter referred to as PSTs) as they engage in different levels of cognitive engagement when interacting with GenAI to do mathematical problem-solving?

To address these questions, we analyzed the interactions of three participants—Jana, Pari, and Vira—who engaged with GenAI tools during mathematical problem-solving tasks. Their varied approaches offer insight into both the instructional affordances and limitations of AI-supported problem-solving, particularly in relation to teacher concerns and learners' levels of cognitive engagement. Table 3 provides a comparison of the three participants' interactions with key patterns of engagement. Additionally, it demonstrates the varied cognitive demands the participants placed on AI based on their engagement types, and it provides insight into how those interactions either support or challenge teacher concerns regarding AI's role in mathematical problem-solving. We draw from prior research on an analysis of learner engagement with AI (Authors et al., 2024b, 2024c) to categorize each participant's level of engagement with AI. This allows for a nuanced exploration of how various learners and teachers engage with AI.

Table 3. *An Overview of Teachers' Engagement with GenAI tools*

Dimension	**Jana**	**Pari**	**Vira**
Prompt Type	Mostly *Ask*-type prompts	Primarily *Ask*-type, with some validation prompts	Mix of *Ask*, *Affirm*, and *Guide* prompts
Cognitive engagement	Surface-level, basic task-oriented engagement	More task complexity and validation, but remained procedural	Some depth, but mostly focused on procedural steps with occasional reflection
Follow-Up on AI's Responses	Minimal follow-up or questioning	Moderate follow-up, seeking confirmation	More active follow-up, with occasional shifts in direction
Exploration	One related problem posed (but still directed at AI)	Exploration in form of validation and planning	Some exploration, but still procedural in nature
Level of Engagement (Authors et al., 2024a)	Foundational (task-focused, procedural)	Foundational, but moving towards constructive (more elaborate task-based interactions, a greater need for validation due to task complexity)	Constructive (reframes problem, awareness of task constraints, changes strategies based on newer insights)

These comparisons provide direct insight into RQ2 by illustrating the range of cognitive demands exhibited during AI engagement.

4.1 Task Design and Learner Engagement

This section directly addresses both research questions by examining how task structure shaped AI interactions (RQ1) and how these interactions revealed varying levels of cognitive engagement (RQ2). Jana's reliance on AI to guide each step in solving the problem reflects the procedural nature of her engagement. This type of engagement may have been influenced by the well-structured, single-solution task. In contrast, both Pari and Vira encountered tasks that were more complex and required deeper decision-making. While Pari sought validation at various stages of the task, indicating a foundational engagement, this was partly due to the task's complexity, which elicited need for more AI-based verification. This alignment between task structure and engagement style is consistent with findings that task specificity can lead to more foundational interactions, where learners focus on applying known procedures rather than conceptualizing or exploring alternative strategies (Authors et al., 2024b; Baker et al., 2022).

Moreover, for straightforward or routine tasks, our participants saw little need to interrogate or expand on the AI's responses. This pattern echoes earlier findings where learners engaged with AI tools primarily for convenience, specifically when teachers perceived the framing of the task or classroom norms as emphasizing correct answers over reasoning (Holmes et al., 2019; Khosravi et al., 2022). Jana's engagement style may reflect contextual affordances rather than a limitation in her reasoning abilities. This highlights the importance of designing AI-integrated tasks that explicitly support and invite deeper conceptual engagement.

4.2 Task Facilitation and the Learning Environment

To further explore RQ1, we consider how changes in the learning environment and scaffolding influenced the nature of teacher-AI interactions, either reinforcing or challenging teachers' concerns about GenAI use. Our findings suggest that the structuring of the AI-engagement influenced how teachers engage with AI tools during mathematics tasks. In Jana's case, she worked on single-solution tasks with minimal scaffolding. By contrast, Pari and Vira were introduced to modeling tasks along with a brief orientation session on interacting productively with AI. Vira's engagement demonstrated a constructive approach where he adapted AI's suggestions and critically reflected on the proposed solutions, adjusting his strategy throughout the task. This type of engagement could be viewed as a productive struggle (Boaler, 2016). It highlights the potential for AI to support constructive learning by enabling teachers to generate alternative solutions and rethink strategies based on emerging information (Author et al., 2025).

While the study focused primarily on foundational and constructive engagement, the lack of creative engagement in the task warrants attention. Creative engagement with AI involves learners using the tool to explore new ways of problem-solving and to generate innovative solutions (Author et al., 2024), which were not evident in this study. In future work, it would be beneficial to rethink the structuring of the learning engagement to allow for more open-ended problems that encourage exploration, experimentation, and creative use of AI tools.

4.3 Gaps in AI Literacy

This theme intersects with RQ1 by illuminating gaps between teachers' concerns and their own interactions with AI, and with RQ2 by highlighting missed opportunities for deeper cognitive engagement. One consistent theme across all three teachers' interactions was the reliance on AI for validation and problem-solving, with limited engagement in verifying the AI's logic. This points to a critical gap in AI literacy, as the teachers, especially Jana and Pari, did not consistently challenge or verify the AI's outputs. As Ng et al. (2021) argue, AI literacy involves not only using AI tools effectively but also critically assessing the responses generated by these tools. In the cases of Pari and Vira, this lack of verification highlights the need for targeted professional development on how to interact meaningfully with AI (Walkington, 2025).

Interestingly, many teachers expressed concerns about students using AI in procedural ways. Yet their own interactions often mirrored the same approach. This points to a clear need for intentional AI literacy–based professional development (Sperling et al., 2024; Walkington, 2025). This must focus more on how to use it effectively during mathematical problem-solving. Understanding how to frame productive prompts and engage in iterative conversations with AI tools is still emerging (Fagbohun et al., 2024; Jatin, 2024). These early findings signal the need to explore how prompt engineering and AI-supported reasoning can be taught and modeled in ways that go beyond answer-seeking.

5 Implications

While our observations revealed that participants—such as Jana—often engaged with the AI tool in a procedural, task-focused manner, we hesitate to conclude that this reflects a general devaluing of conceptual understanding among teachers. Rather, we believe several contextual factors shaped this pattern of interaction. First, the AI platform was unfamiliar to many participants, and their engagement may reflect a cautious, exploratory approach rather than deliberate avoidance of deeper reasoning. Second, the research setting (working independently in a constrained time frame without student-facing instructional goals) may have influenced how teachers interacted with the tool, prioritizing task completion over in-depth exploration. Some teachers may still be developing understanding of how GenAI might support conceptual thinking and thus defaulted to procedural exchanges. These factors suggest that limited conceptual engagement in this study reflects the early and evolving nature of teacher-AI interactions, rather than lack of commitment to conceptual understanding in their broader teaching practice.

Nonetheless, our findings highlighted in the discussion strongly support the recommendation for teachers to explore AI freely without fear of mistakes so they can learn the "capabilities and limitations of AI tools" (AASA et al., 2023, p. 2). The findings inform the following implications for future practice and future research regarding AI-supported mathematical problem-solving experiences: (a) task design and learner engagement, (b) task facilitation and the learning environment, (c) gaps in AI literacy, (d) learner engagement is developmental, and (e) learner engagement—whether teacher or student—highlight similar issues and concerns.

With regards to task design, we found support for addressing the specific learner objectives as they relate to conceptual understanding and procedural fluency of the mathematical engagement (NCTM, 2023). Single-solution tasks compared to multiple-solution tasks will require different levels of engagement between the learner and AI. Task design should support goals for both conceptual understanding and procedural fluency to avoid under-utilizing the capabilities of AI and fueling the old fears about technology undermining student thinking. More open-ended and complex tasks will provide an avenue for exploration, reflection, and decision-making to offset the learner's tendency to simply permit AI to find answers for them.

Tasks should be structured to facilitate an engaging learning environment that involves learners in divergent thinking. An introduction to productive AI prompts as a component of AI literacy can support goals for critically considering responses provided by AI, for using prompts that support more meaningful AI responses, and for enticing AI to respond with more depth to support human-AI interaction. With appropriate scaffolding and guided reflections, learners become more enabled to mine AI for what it can offer. Future practice should explore more comprehensive support for creating better prompts, both when using AI as well as during the teacher-classroom learning experiences. When teachers engage AI as learners, they will gain hands-on experience with AI similar to the experiences that lie ahead for their students. The beliefs and concerns teachers had prior to preparing AI experiences for their students will be informed by their experiences. As teachers' experience evolves, so will their beliefs and their knowledge and skills for preparing productive AI experiences for their students.

5.1 Conclusion

When learners engage in well-designed mathematical problem-solving tasks, the effective prompts used will support productive AI interactions. Teachers should explore various aspects of AI interactions, such as constructing tasks and solutions themselves, engaging AI as a collaborator rather than merely as an answer provider, and critically questioning the responses provided by AI. Although our study did not explicitly focus on ethical issues, we recognized that as AI literacy develops over time, through experience, teachers will become more aware of the possible openings for breaches of integrity and ethical concerns. Whereas the potential for good from AI use in mathematical problem-solving is great, the potential for harm is more serious than any other technological innovation that has approached our classrooms. Professional development for teachers, free exploration of AI meant to bring discovery of capabilities and limitations without fear of making mistakes, should support learning about the safe and potentially unsafe realities for the mathematical classroom.

References

Ahn, J., Verma, R., Lou, R., Liu, D., Zhang, R., & Yin, W. (2024). Large language models for mathematical reasoning: Progresses and challenges. arXiv https://arxiv.org/abs/2402.00157

Alsharidah, M. A., & Alkramiti, A. M. (2024). Teachers' perceptions towards using artificial intelligence in mathematics education. *Amazonia Investiga, 13*(84), 43-60. https://doi.org/10.34069/AI/2024.84.12.3

Asamoah, P., Zokpe, D., Boateng, R., Mafo, J. S., Boateng, S. L., Asamoah, D., Muntaka, A. S., & Manso, J. F. (2024). Domain knowledge, ethical acumen, and query capabilities (DEQ): A framework for generative AI use in education and knowledge work. *Cogent Education, 11*(1), https://doi.org/10.1080/2331186X.2024.2439651

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52-62.

Baker, R.S., Corbett, A.T., Koedinger, K.R. (2004). Detecting Student Misuse of Intelligent Tutoring Systems. In: Lester, J.C., Vicari, R.M., Paraguaçu, F. (eds) Intelligent Tutoring Systems. ITS 2004. Lecture Notes in Computer Science, vol 3220. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30139-4_50

Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V., Xu, Y., & Fung, P. (2023). A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. arXiv. https://arxiv.org/abs/2302.04023

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education.* Chapel Hill, NC: Horizon Research.

Bastani, H., Bastani, O., Sungu, Al., Haosen, G., Kabakci, Ö., & Mariman, R. (2024). Generative AI can harm learning. The Wharton School Research Paper Bastani, Hamsa and Bastani, Osbert and Sungu, Alp and Ge, Haosen and Kabakcı, Özge and Mariman, Rei, Generative AI Can Harm Learning (July 15, 2024). The Wharton School Research Paper, http://dx.doi.org/10.2139/ssrn.4895486

Biton, Y., & Segal, R. (2025). Learning to craft and critically evaluate prompts: The role of generative AI (ChatGPT) in enhancing pre-service mathematics teachers' TPACK and problem-posing skills. International Journal of Education in Mathematics, Science, and Technology (IJEMST), 13(1), 202-223. https://doi.org/10.46328/ijemst.4654

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Jossey-Bass/Wiley.

Cacho, R. (2024). Integrating generative AI in university teaching and learning: A model for balanced guidelines, Online Learning, 28(3), 55-81. https://doi.org/10.24059/olj.v28i3.4508

Center on Reinventing Public Education (CRPE). (2024a). *Wicked opportunities: Leveraging AI to transform education.* A report from CRPE's Think Forward: AI Learning Forum.

—. (2024b). *AI and education policy 101: The evolving landscape and examples from early adopters.* https://crpe.org/focus_areas/ai-education/

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510

Dykema, K. (2023, November). Math and artificial intelligence. President's message. National Council of Teachers of Mathematics, https://www.nctm.org/News-and-Calendar/Messages-from-the-President.

Erbaş, A. K., Çetinkaya, B., Alacacı, C., Çakıroğlu, E., Aydoğan Yenmez, A., Şen Zeytun, A., Korkmaz, H., Kertil, M., Didiş, M. G., Baş, S., & Şahin, Z. (2016). *Everyday life modeling questions for high school math subjects.* TÜBA, Turkish Academy of Sciences.

Fabijanić Gagro, S. (2024). Artificial intelligence in education – Current challenges. Annals of the Faculty of Law in Belgrade, 72(4), 405–440. https://doi.org/10.51204/Anali_PFBU_24405A

Fagbohun, O., Harrison, R. M., & Dereventsov, A. (2024). An empirical categorization of prompting techniques for large language models: A practitioner's guide. arXiv preprint arXiv:2402.14837. https://arxiv.org/abs/2402.14837

Gilbert, S., Blessing, S., & Guo, E. (2015). Authoring effective embedded tutors: An overview of the extensible problem specific tutor (xPST) system. International Journal of Artificial Intelligence in Education, 25(3), 428–454

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. Applied Cognitive Psychology, 9(6), 495–522. https://doi.org/10.1002/acp.2350090604

Haddley, J. A., & Ardito, C. G. (2024). Generative AI in assessment: Towards understanding the student view. MSOR Connections, 23(1). http://doi.org/10.21100/msor.v23i1.1515

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Holstein, Kenneth & Mclaren, Bruce & Aleven, Vincent. (2019). Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms. https://doi.org/10.1007/978-3-030-23204-7_1

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

Jatin. (2024, July 25). Understanding prompt engineering: Creating effective interactions with AI. Medium. https://medium.com/@jatin2707/understanding-prompt-engineering-crafting-effective-interactions-with-ai-fc5b7e91fc00

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education, Computers and Education: Artificial Intelligence, 3, https://doi.org/10.1016/j.caeai.2022.100074.

Lawasi, M., Rohman, V., & Shoreamanis, M. (2024). The use of AI in improving students' critical thinking skills. Proceedings Series on Social Sciences & Humanities. 18, 366-370. https://doi.org/10.30595/pssh.v18i.1279

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarok or reformation? A paradoxical perspective from management educators. The International Journal of Management Education, 21, Article 100790. https://doi.org/10.1016/j.ijme.2023.100790

McCarthy, J., Minsky, M. I., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955. AI Magazine, 24(4), 12–14. https://doi.org/10.1609/aimag.v27i4.1904

Milicevic, V., Lazarova, L. K., & Pavlovic, M. J. (2024). The application of artificial intelligence in education–The current state and trends. International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 12(2), 259-272.

Mohamed, M., Hidayat, R., Suhaizi, N., Mahmud, M., Baharuddin, S. (2022). Artificial intelligence in mathematics education: A systematic literature review, International Electronic Journal of Mathematics Education, 17, em0694.

Naresh, N., Yilmaz, Z., & Galanti, T. (2024a). Leveraging the Potential of AI as a Partner in Teaching. *Connections, Artificial Intelligence in Mathematics Teacher Education, 33* (4). Association of Mathematics Teacher Educators.

Naresh, N., Yilmaz, Z., & Li, E. (2024b). Exploring Learner-AI Engagement in Mathematics Education: An Emerging Framework for Inquiry. In Kosko, K. W., Caniglia, J., Courtney, S. A., Zolfaghari, M., & Morris, G. A., (Eds.) *Proceedings of the forty-sixth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kent State University. Cleveland, OH.

National Advisory Committee on Mathematical Education (NACOME). (1975). *Overview and analysis of school mathematics, Grades K-12.* Conference Board of the Mathematical Sciences, Washington, DC: Author.

National Council of Teachers of Mathematics (NCTM). (1978). A position statement on calculators in the classroom. Reston, VA: Author.

—. (1980). *An agenda for action: Recommendations for school mathematics of the 1980s.* Reston, VA: Author.

—. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

—. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

—. (2014a). *Principles to actions: Ensuring mathematical success for all.* Reston, VA: Author.

—. (2014b). *Principles to actions: Ensuring mathematical success for all.* [Executive Summary]. Reston, VA: Author.

—. (2023). *A position statement on procedural fluency: Reasoning and decision-making, not rote application of procedures position.* Reston, VA: Author.

—. (2024a). *Principles to actions: Ensuring mathematical success for all (10th Anniversary Edition).* Reston, VA: Author.

—. (2024b). *Artificial intelligence and mathematics teaching*. Position statement. Reston, VA: Author.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics.* Washington, DC: Authors.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education.* Washington, DC: The National Academies Press. https://doi.org/10.17226/1199.

Ng, D., Leung, J., Chu, S., & Qiao, M. (2021). Conceptualizing AI literacy: An exploratory review, Computers and Education: Artificial Intelligence, 2, 100041, https://doi.org/10.1016/j.caeai.2021.100041.

O'Connor, C., & Michaels, S. (2017). Supporting teachers in taking up productive talk moves: The long road to professional learning at scale. International Journal of Educational Research. 97. https://doi.org/10.1016/j.ijer.2017.11.003

Ofgang, E. (2023, April 17). What is Khanmigo? The GPT-4 learning tool explained by Sal Khan. Tech & Learning. https://www.techlearning.com/news/what-is-khanmigo-the-gpt-4-learning-tool-explained-by-sal-khan

Opesemowo, O. A. G., & Ndlovu, M. (2024). Artificial intelligence in mathematics education: The good, the bad, and the ugly. Journal of Pedagogical Research, 8(3), 333-346. https://doi.org/10.33902/JPR.202426428.

Organization for Economic Co-operation & Development (OECD). (2021). *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots.* https://www.oecd-ilibrary.org/education/oecd-digital-education-outlook-2021_589b283f-en

Paige, R., Seshaiyer, P., & Toda, M. (2006). Student misconceptions caused by misuse of technology. International Journal for Technology in Mathematics Education, 14(4), 189-195.

Park, J., & Choo, S. (2024). Generative AI prompt engineering for educators: Practical strategies. Journal of Special Education Technology, 0(0). https://doi.org/10.1177/01626434241298954

Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. Digital Experiences in Mathematics Education, 1, 9–41. https://doi.org/10.1007/s40751-025-00172-1

Polya, G. (2014). *How to solve it.* Princeton University Press.

Rane, N. (2023) Enhancing mathematical capabilities through ChatGPT and similar generative artificial intelligence: Roles and challenges in solving mathematical problems. http://dx.doi.org/10.2139/ssrn.4603237

Ronau, R. N., Rakes, C. R., Bush, S. B., Driskell, S., Niess, M. L., & Pugalee, D. (2011). Using calculators for teaching and learning mathematics. (Research Brief). Karen D. King, Series Editor, NCTM, Reston, VA.

Saldaňa, J. (2012). *The coding manual for qualitative researchers* (2nd ed.). Thousand Oaks, CA: Sage.

Sperling, K., Stenberg, C., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review, Computers and Education Open, 6, 100169, https://doi.org/10.1016/j.caeo.2024.100169.

The School Superintendents Association (AASA), Association for Supervision and Curriculum Development (ASCD), International Society for Technology in Education (ISTE), National Association of Elementary School Principals (NAESP), & National Association of Secondary School Principals (NASSP). (2023). *Bringing AI to school: Tips for school leaders.* https://cdn.iste.org/www-root/2023-07/Bringing_AI_to_School-2023_07.pdf

Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino, S., Giannoutsou, N., Cachia, R., Monés, A., & Ioannou, A. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. Educational Information Technology, 28, 6695–6726. https://doi.org/10.1007/s10639-022-11431-8

Turing, A. M. (1950). I.–Computing machinery and intelligence. Mind, 59(236), 433-460, https://doi.org/10.1093/mind/LIX.236.433

U.S. Department of Education [USDOE]. (2023). *Artificial intelligence and future of teaching and learning: Insights and recommendations*, Washington, DC: Author. https://tech.ed.gov

U.S. Department of Education [USDOE]. (2001). *The nation's report card: Mathematics 2000.* NCES 2001–517. By J. S. Braswell, A. D. Lutkus, W. S. Grigg, S. L. Santapau, B. S.-H. Tay-Lim, & M. S. Johnson. Washington, DC: National Center for Education Statistics

Virginia Department of Education (VDOE). (n.d.). Rich mathematical tasks. https://www.doe.virginia.gov/teaching-learning-assessment/k-12-standards-instruction/mathematics/instructional-resources/rich-mathematical-task

Walkington, C. (2025). The implications of generative artificial intelligence for mathematics education. School Science and Mathematics, 1–10. https://doi.org/10.1111/ssm.18356

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

Yilmaz, Z., Naresh, N., & Kanbir, S. (2024). Pre-service Teachers' Perceptions on Exploring Number Theory Concepts Using Khanmigo: Benefits and Challenges. In Kosko, K. W., Caniglia, J., Courtney, S. A., Zolfaghari, M., & Morris, G. A., (Eds.) *Proceedings of the forty-sixth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kent State University. Cleveland, OH.

Yonder Consulting. (2025). Chegg 2025 global student survey. [Report/Data set]. Chegg. https://www.chegg.org/global-student-survey-2025p

Yoon, H., Hwang, J., Lee, K., Roh, K. H., & Kwon, O. N. (2024). Students' use of generative artificial intelligence for proving mathematical statements. ZDM Mathematics Education, 56, 1531-1551. https://doi.org/10.1007/s11858-024-01629-0

Nirmala Naresh is an Associate Professor in the Department of Mathematics at the University of North Texas. Her research focuses on ethnomathematics, AI in mathematics Education, and the teaching and learning of probability and statistics. She is a coPI on NSF-funded Noyce teaching fellowship grant STEM). Dr. Naresh is passionate about connecting mathematics to culture and everyday life and about preparing teachers to do the same in their own classrooms.

Zuhal Yilmaz is a Research Associate in the College of Education at North Carolina State University and an Evaluation Consultant at Louisiana State University. She has more than twelve years of experience in mathematics education as a researcher, teacher, and consultant, with interests in students' mathematical thinking and teachers' professional learning. Her current work explores connections between computer science and mathematics education, focusing on computational thinking integration.

Iris DeLoach Johnson is a retired Professor of Mathematics at Louisiana State University in Shreveport, where she served in a variety of teaching and administrative roles, including program leadership. She has over forty years of experience in education. Her work focuses on mathematics education, teacher preparation, and educational leadership. Dr. Johnson continues to consult, publish, and collaborate internationally through the Real Life Global Themes project of the Math in the 21st Century initiative.

Deborah Cockerham, Clinical Assistant Professor in the University of North Texas Learning Technologies department, also serves as Director of the Learning Technologies Masters of Science Program. Dr. Cockerham founded and directed the Research and Learning Center at the Fort Worth Museum of Science and History and continually works to build community and school relationships with higher education.

Appendix A

Participant Questionnaire

1. What uses of AI have you learned about or observed others using?

2. Are you aware of the development or implementation of an AI curriculum/tools for math in any grades K–12 in your school district?

3. Have you received any professional development/training on AI tools, uses, or methodologies in classrooms? Yes/No

4. What teaching methodologies are emphasized in using AI tools that you either have in your curriculum and/or have learned about? Check all that apply.
   - Lecture or instruction
   - Blended learning (e.g., Learning takes place partly face-to-face and partly online)
   - Remote learning
   - Group work
   - Project-based learning (e.g., Learners leverage their skills and competencies in interdisciplinary collaborations. They work together to identify and/or respond to a real-world challenge over an extended period of time.)
   - Activity-based learning (e.g., activities are facilitated by teachers; learners progress through activities at their own pace.)
   - I do not use AI tools

5. What AI tools, if any, do you currently use for teaching math in your classroom?

6. What are your thoughts on the use of AI in mathematics classrooms? Please discuss questions, concerns, and potential benefits for enhancing learning.

7. Do you have any additional information and/or questions on the use of AI tools in teaching/teaching mathematics that you would like to share?

Appendix B

Example Task for In-Service Teachers

The Bake Sale Fundraiser Task[1]

Brady and Jaquan were selling cupcakes together at a bake sale. They hope to make $100 so they can both go on the band field trip to New York City.

In the first hour, Brady sold $\frac{1}{3}$ of the cupcakes and Jaquan sold $\frac{3}{8}$ of the cupcakes.

- During the second hour, they sold two cupcakes.
- During the third hour, they sold 75% of the remaining cupcakes.
- During the fourth hour, they sold the remaining 3 cupcakes.

If they sold each cupcake for $2.75, will they make enough money to go on the field trip? If so, how much money would be left over for spending money? Explain how you know.

Appendix C

How to Store Task[2]

A company that produces canned food needs short-term storage to store the cylinder-shaped can it produces. The company wants to do this with the least possible cost. Each of the right circular cylinders can be kept 10 cm in radius and 30 cm in height. The company plans to store 175 cans for two (2) months. There are three (3) different sizes of storage cabinets that the company can store. The rental costs are shown in Table C1 according to the dimensions of base of the storage cabinets, each of which is 100 cm high.

Table 4. *Size of the Storage Cabinets and Costs*

Width (cm)	Length (cm)	Rental Cost per Month ($)
110	110	100
110	220	150
110	330	200

a. If you were the company owner, in which way(s) would you use which storage cabinet to minimize the cost?

b. The company may need to store different numbers of cans in future productions. For this, would it be appropriate for the company to always use the same type of storage cabinets? What do you suggest? Why?

Please note that keeping the cans in an upright position is important for the safety of the storage.

[1]Taken from Virginia Department of Education (VDOE). (n.d.). *Rich mathematical tasks.* https://www.doe.virginia.gov/teaching-learning-assessment/k-12-standards-instruction/mathematics/instructional-resources/rich-mathematical-task

[2]Taken from Erbaş, A. K., Çetinkaya, B., Alacacı, C., Çakıroğlu, E., Aydoğan Yenmez, A., Şen Zeytun, A., Korkmaz, H., Kertil, M., Didiş, M. G., Baş, S., & Şahin, Z. (2016). *Everyday life modeling questions for high school math subjects.* TÜBA, Turkish Academy of Sciences.

Appendix D

Analysis of Participants-AI Interactions

Table 5. *Prompt Analysis for Jana's Engagement with AI*

Type	Contextual Info with Prompt	Explanation
Ask	He made 120 and sold 3/4 of them, how many did he sell?	The learner is seeking a direct answer from the AI, showing foundational engagement where the AI is the primary source of information.
Affirm	Ok, so he sold 90?	The learner asks for confirmation of their calculation, validating the AI's response without doing the cognitive work themselves.
Ask	How much would he sell one of them for to make $60 profit?	The learner is asking a question that still relies on the AI to provide an answer without fully engaging in the reasoning process.
Ask	Can you explain that logic to me?	The learner asks for further explanation but doesn't show much ownership in trying to engage with the solution independently.
Ask	So I divide 150 by 90?	The learner asks for confirmation of a step in the solution, showing dependence on the AI for guidance.
Ask	Right. But then what do I subtract?	The learner asks the AI to continue guiding the process, rather than actively thinking through the next step themselves.

Table 6. *Prompt Analysis for Pari's Engagement with AI*

Type	Contextual Info with Prompt	Explanation
Ask	Pari restates the original questions in simplified form and asks, "What would be the cheapest way to store 175 cans?"	Asks the AI to work towards a solution
Ask	Pari provides additional information: "If one of the cans' volume is 9424 and there are 175 cans to store. What is the cheapest way?"	Seeks step-by-step help using can volume to compare storage costs.
Ask	After Pari gets a response, he asks: "So which one the bes[t] price?"	Requests clarification or confirmation from previous steps.
Affirm/Validate	So, Number 3, the best price?	Checks if they followed the AI's logic correctly.
Affirm/Validate	Do you think the price will change?	Seeks agreement about the price (trusts AI)
Guide	In the future, would you think they should use the same company?	Moves toward planning based on emerging understanding.

Table 7. *Prompt Analysis for Vira's Engagement with AI*

Prompt Type	Contextual Info & Prompt	Explanation
Ask	After entering the cabinet data, Vira asked, "We need to determine which storage cabinet would be the most cost-effective to store these cans. How could we do this?"	Seeks a strategy from the AI rather than proposing one — early-stage reasoning support.
Ask	Once the can size was established, Vira asked, "How much space does one can take up?"	Procedural question relying on the AI to compute cylinder volume.
Ask	After that, Vira followed up with, "So what would the volume be for each of the cabinets?"	Step-by-step inquiry prompting ChatGPT to complete the calculation.
Affirm/Validate	After the AI suggested computing can volume, Vira responded: "Okay, I see what you're thinking! Let's say we have 175 cans with the dimensions I had mentioned earlier..."	Shows affirmation and willingness to follow the AI's proposed strategy.
Affirm/Validate	After the AI begins computing the total volume, Vira says: "I see you already have an idea... Let's go ahead and perform that calculation."	Vira endorses the plan and invites the AI to proceed (validation + calculation).
Affirm/Validate	After concluding which cabinet is most efficient per volume, Vira asks: "Would this still stand if they held them in storage for 2 months?"	Confirms the earlier conclusion holds under different conditions.
Guide	After the volumes are known, Vira changes direction: "We might not need the extra space since we only have 175 cans. How many cans would you predict each cabinet can hold?"	Introduces a new constraint, shifting from efficiency to capacity.
Guide	Vira then asks: "Considering that we don't need to hold that many cans, just 175, which cabinet would be the cheapest to fit all the cans in?"	Reframes objective: cost minimization based on sufficiency, not per-unit efficiency.
Guide	Vira clarifies: "Let's say we don't care about the cost per can... Which one would be the cheapest in this case?"	Dismisses the earlier idea and provides a new suggestion.