Supporting Responsive Mathematics Teaching in Elementary Classrooms Using ChatGPT

Sarah Selmer
West Virginia University

Abstract

This article examines prospective teachers' reported experiences and perceptions as they used a standardized prompt with the free version of ChatGPT (ChatGPT-3.5, OpenAI, 2024) to analyze and respond to elementary students' work on high cognitive demand mathematics tasks. The purpose of this study is to contribute to understanding how AI tools, rather than replacing teacher expertise, might act as pedagogical support to enhance prospective teachers' abilities to analyze and respond to student thinking in ways that maintain cognitive demand and extend mathematical understanding (National Council of Teachers of Mathematics, 2024). Results indicated that prospective teachers perceived that ChatGPT did support their analysis and responses in ways responsive to their students' thinking. Results also demonstrated three distinct types of questions generated by ChatGPT that emerged from the prospective teachers' experiences working with elementary students: advancing, clarifying, and extending. Implications and discussion focus on the potential of using ChatGPT to support specific aspects of responsive teaching.

Keywords: Mathematics education, Artificial intelligence

1 Introduction

Responsive mathematics teaching prioritizes understanding student thinking over adhering to scripted instruction (Coffey et al., 2011). In this approach, teachers work to uncover students' reasoning about mathematical ideas and use those insights to guide their instructional decisions. To develop engaging in responsive teaching practices, teachers can utilize professional noticing of children's mathematical thinking, which involves attending, interpreting, and then using these interpretations to decide how to respond to students' thinking (Jacobs et al., 2010). Teachers' responses are considered truly responsive when they maintain the cognitive demand and focus on understanding how students make sense of and extend mathematical concepts, rather than simply correcting errors, telling, or explaining (Selmer et al., 2022). As artificial intelligence (AI) tools become increasingly sophisticated, questions arise about how these technologies might support teachers in developing these critical professional noticing skills that support responsive teaching practices.

1.1 Research Purpose and Questions

This article examines 28 prospective teachers' reported experiences and perceptions as they used a standardized prompt with the free version of ChatGPT (ChatGPT-3.5, OpenAI, 2024), a sophisticated large language model developed by OpenAI, to analyze and respond to 56 elementary students' work on high cognitive demand mathematics tasks.

Specifically, this research addresses the following questions:

RQ1: How do prospective teachers experience and perceive ChatGPT's effectiveness in supporting the "attending to" and "interpreting" components of professional noticing when analyzing elementary students' mathematical work on high cognitive demand tasks?

RQ2: Based on prospective teachers' experiences, what patterns emerge in the types of questions ChatGPT generates when responding to elementary students' mathematical work on high cognitive demand tasks?

RQ3: How do prospective teachers perceive ChatGPT's effectiveness in supporting the "deciding how to respond" component of professional noticing when responding to elementary students' mathematical work on high cognitive demand tasks?

This evidence-based study contributes to understanding how AI tools, rather than replacing teacher expertise, might act as pedagogical support to enhance prospective teachers' abilities to make sense of and respond to student thinking in ways that enhance students' problem-solving abilities (National Council of Teachers of Mathematics, 2024).

1.2 Study Overview

A review of the literature on high cognitive demand tasks, responsive teaching framed through professional noticing, and AI in education provides context for understanding how prospective teachers might experience using ChatGPT to support their professional noticing development. The study presents illustrative examples of how ChatGPT is used in practice, including prospective teachers' uploading images of student work on a high cognitive demand mathematics task, with ChatGPT analyzing the work to suggest instructional responses. The paper concludes by discussing prospective teacher experiences and perspectives on integrating AI into responsive instruction and implications for teacher education programs.

2 Literature

2.1 High Cognitive Demand Tasks

High cognitive demand tasks are challenging mathematics problems that are open-ended and require students to think beyond procedures and algorithms (Smith & Stein, 1998). These tasks allow teachers to facilitate students' engagement in analyzing what questions are being asked, often integrate conceptual understanding of mathematical topics, and develop strategies for solution finding (National Council of Teachers of Mathematics, 2000). The National Council of Teachers of Mathematics' Principles and Standards for School Mathematics (National Council of Teachers of Mathematics, 2000) recommends incorporating such problems into classroom curricula. Investigating how prospective teachers implement elementary grade-level high cognitive demand tasks involving addition, subtraction, multiplication, and division of whole numbers, with multiple entry points, is an interesting area of exploration because they naturally elicit a wide range of student thinking (Van de Walle et al., 2019). For example, the "Sharing Stickers Problem" asks students to determine how to fairly distribute 12 stickers among three friends so that each friend receives the same number (Van de Walle et al., 2019). While seemingly straightforward, this problem encourages students to develop division concepts through multiple approaches, including drawing pictures, using counters, or creating equal groups, rather than applying a memorized algorithm. The use of high cognitive demand tasks is especially well-suited for teachers to promote teaching practices that emphasize understanding how students make sense of mathematical concepts, rather than simply identifying and correcting errors (Huinker & Bill, 2017). Research demonstrates that the cognitive demand of tasks is frequently diminished—often unintentionally—during lesson implementation, as teachers' instructional practices, particularly the level of support and guidance they provide, serve as critical factors that either maintain or reduce the cognitive demand originally intended by the task designers (Henningsen & Stein, 1997; Stein et al., 2007). One approach to developing supportive instructional practices is responsive teaching, which, when framed through the lens of professional noticing, offers a promising framework for helping prospective teachers preserve the cognitive demand of tasks throughout implementation (Selmer et al., 2022).

2.2 Responsive Teaching

Responsive teaching includes both a stance and a set of practices used in classroom settings where teachers recognize the importance of using the substance of student thinking to guide instructional decisions (Coffey et al., 2008; Coffey et al., 2011). To engage in this kind of teaching, educators must view students as active learners with prior knowledge developed both inside and outside the classroom. This perspective represents the foundational stance of responsive teaching. Teachers can then leverage this knowledge through specific teaching practices to support and extend student thinking.

Responsive teaching emphasizes understanding how students make sense of mathematical concepts rather than focusing on identifying and correcting errors (Coffey et al., 2011). In responsive teaching, teachers pay attention to students' ideas and use them to guide instructional decisions (Hammer et al., 2016). Unlike scripted curricula with predetermined teacher moves, responsive teaching allows classroom discourse to emerge naturally from students' contributions, creating a dynamic and adaptable curriculum shaped by the understandings students bring to the classroom. Dyer and Sherin (2016) describe three practices that characterize responsive teaching. First, teachers probe students' ideas to better understand their thinking. Second, they invite students to comment on and react to each other's mathematical contributions. Third, teachers take up student ideas and incorporate them into ongoing instruction. When teachers use these practices, they create opportunities to build on what students are thinking rather than following a preset plan.

However, implementing responsive teaching practices can be challenging for teachers (Hammer et al., 2016; Sherin & van Es, 2005). Teachers must attend to the diverse and emerging ideas of many students in real time (Jacobs et al., 2010). Given the challenges of implementing responsive teaching practices, the professional noticing framework provides a guide for teacher development. The three components of professional noticing—attending to student thinking, interpreting mathematical significance, and deciding how to respond—offer specific skills that can be developed and supported (Jacobs et al., 2010). This framework is particularly relevant for prospective teachers who are learning to move beyond simply identifying correct or incorrect answers toward understanding the mathematical reasoning behind student work. As AI tools become available to support teacher development, examining how these technologies might support professional noticing becomes important for teacher education.

2.3 Artificial Intelligence in Education

As artificial intelligence tools become increasingly available in educational settings, researchers have begun exploring whether technologies like ChatGPT might support educators in teaching students mathematics. Recent research has demonstrated that ChatGPT and other large language models can support educators in enhancing various aspects of student learning and thinking across educational contexts. Studies have shown that when educators integrate ChatGPT into their teaching, it can help their students develop critical thinking skills, with research showing improvements in students' ability to analyze information and engage with complex concepts (Suriano et al., 2025). Additionally, research has demonstrated that ChatGPT can assist educators in supporting student learning through general uses such as providing personalized feedback to learners and helping teachers create educational materials (Mai et al., 2024). Interestingly, both teachers and students generally hold a positive perception of the use of ChatGPT in teaching and learning (Limna et al., 2023).

However, these studies have primarily focused on general educational contexts, not contexts specific to mathematics education. The limited research that has examined ChatGPT specifically in mathematics education reveals mixed findings. Studies examining ChatGPT's performance on mathematical assessments show that effectiveness varies considerably based on problem difficulty and complexity (Dao & Le, 2023). For instance, research using the Vietnamese National High School Graduation Examination found that ChatGPT achieved 83% accuracy on knowledge-level questions but only 10% accuracy on high-application problems, suggesting substantial limitations in complex mathematical reasoning. In contrast, research examining educators' perspectives found that participants believed that ChatGPT improved students' mathematical capabilities by sharing knowledge across challenging topics and conveying this information in an understandable way (Wardat et al., 2023).

Lee and Yeo (2022) developed an AI-based chatbot through iterative design-based research to provide preservice mathematics teachers with practice opportunities for responsive teaching, creating a virtual student with fraction misconceptions that could effectively respond to teachers' diagnostic questions. The study found that design features such as programmed sequential responses improved the chatbot's ability to cover preservice teachers' questioning patterns and enhanced user perceptions across two implementation cycles. Supporting these findings, Son and colleagues (2024) concluded that overall, AI-based chatbots hold promise for enhancing prospective teachers' responsive teaching skills.

While these studies demonstrate the potential of individually built educational chatbots, the effectiveness of existing general-purpose AI systems like ChatGPT in authentically analyzing students' mathematical thinking requires further research to establish. Given the current state of research, as a first step, this study examines prospective teachers' perceptions and experiences using ChatGPT to analyze elementary students' mathematical thinking on high cognitive demand tasks. The focus on prospective teachers' experiences and perceptions provided valuable data for understanding how pedagogical tools might be integrated into practice (Biesta, 2007).

3 Methodology

3.1 Setting and Data Sources

This study explores 28 prospective teachers who each conducted one-on-one problem-solving sessions with two K-5 elementary students ($N = 56$ total), facilitating high cognitive demand tasks. The prospective teachers began by preparing to work with a student by choosing a task from several high cognitive demand options spanning grade levels (primarily used or adapted from National Council of Teachers of Mathematics [NCTM], 2025), anticipating potential student work, and developing questions. Next, they facilitated mathematical problem-solving sessions, encouraging students to share their thinking and solution strategies. As students worked, the prospective teachers practiced professional noticing by maintaining running records of student thinking and photographing initial student work. They then used a computer in the classroom to access ChatGPT in real time, utilizing a prompt to analyze the collected student work samples. In the interpreting phase, teachers compared their initial observations with ChatGPT's analysis, identifying similarities and differences to inform their instructional decisions. The responding phase involved implementing ChatGPT-generated questions with students, prompting them to extend their initial work. Finally, teachers engaged in reflection, analyzing their experiences throughout the process to deepen their understanding of AI-supported responsive teaching practices. The standardized prompt is shown in Figure 1.

Figure 1
ChatGPT prompt.

Please analyze this student work using the Professional Noticing Protocol below. Complete all three parts: (1) Attending to the Student's Strategy, (2) Interpreting the Student's Mathematical Understanding, and (3) Deciding How to Respond.

Part 1: Attending to the Student's Strategy

Describe what the student did—only what is visible.

- Note exactly what is written, drawn, or labeled
- Record mathematically significant details:
  - How quantities are shown (tallies, drawings, symbols, groupings)
  - Which numbers appear, and where
  - The order or sequence of steps
  - Tools, diagrams, or organization used
- Use neutral, descriptive language ("The student drew…" not "The student understands…")
- Do not add missing steps, assume reasoning, or judge correctness

Part 2: Interpreting the Student's Mathematical Understanding

Use the evidence from Part 1 to make sense of what the student knows.

- What does the strategy show about the student's understanding of:
  - Number relationships and operations
  - Place value or grouping
  - Decomposing or combining numbers
  - Concrete vs. abstract strategies
- What understandings are not yet evident?
- Stay specific to this work; avoid broad generalizations

Part 3: Deciding How to Respond

Provide 2–3 questions that are rooted in the student's current strategy. These should build directly on what the student has done or extend the student's thinking by guiding them to clarify, deepen, or connect ideas without replacing their approach.

Note. This prompt is based on the professional noticing framework from Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169–202.

The data sources included running records of student thinking, photographs of student work samples (both initial and extended work), and ChatGPT analyses generated using the standardized prompt provided by the teacher educator. Additional data were collected from the prospective teachers as they documented their initial interpretations of student work, their comparative analyses between their observations and ChatGPT's responses, and their reflective analyses of their experiences with the high cognitive demand mathematical tasks and AI-supported responsive teaching process provided by the teacher educator. All prospective teachers were required to submit this documentation.

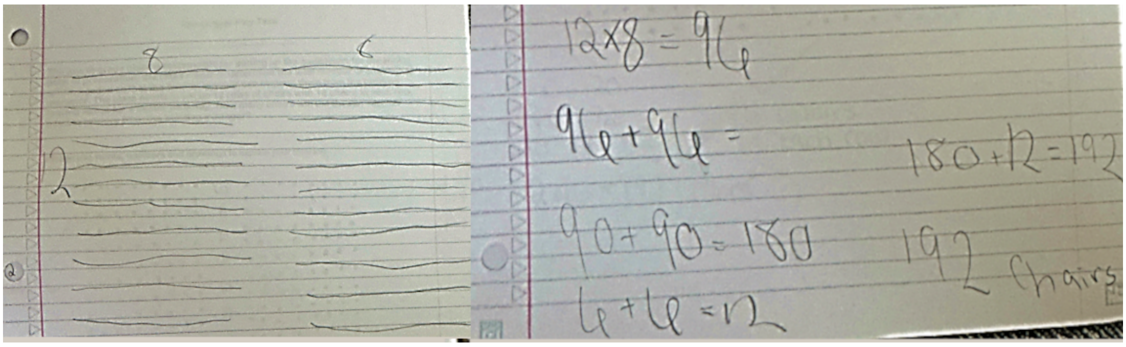
The following illustrative example of one prospective teacher, Sylvia, who worked with an elementary student, Elizabeth, is representative of a problem-solving session. Pseudonyms are used for all participants throughout this paper.

3.2 Illustrative Example: Sylvia & Elizabeth

Prospective teacher Sylvia prepared for her upcoming one-on-one problem-solving session with Elizabeth, a fourth-grade elementary student, focusing on a high cognitive demand task about a class responsible for setting up 12 rows of chairs, with 16 chairs in each row, leaving space for a center walkway (Huinker & Bill, 2017, p. 11). Sylvia arranged manipulatives, blank paper, and pencils at their workspace, and planned to attend to and interpret Elizabeth's work to gain insight into Elizabeth's current mathematical understandings.

Elizabeth modeled the problem, drawing lines to represent 12 rows split into two columns of eight (see Figure 2). She then multiplied 12×8, arriving at the answer of 96. Elizabeth added $96 + 96$ by decomposing 96 into 90 and 6, adding $90 + 90$ and $6 + 6$. She then added those sums together, finding $180 + 12 = 192$.

Figure 2
Elizabeth's visual model and calculations for the band concert task.



Note. *Elizabeth decomposed the multiplication into two groups of 8 and used partial products strategy to solve* $12 \times 16 = 192$.

Next, Sylvia uploaded a digital image of Elizabeth's work to ChatGPT and inputted a prompt (see Figure 1) used by all 28 teachers to analyze the work. The prompt was guided by professional noticing tenets, and the version presented in this article represents a refined version based on insights from the study.

In response, ChatGPT provided an accurate description of Elizabeth's work, including discussion of the visual model and related calculations. Additionally, ChatGPT interpreted her work, noting, "While their math work was clear and accurate, the student would benefit from adding written explanations to connect their strategy and drawings more clearly to the context of the problem."

As Sylvia reviewed ChatGPT's analysis of the problem, she noted that she had described and interpreted Elizabeth's work similarly to ChatGPT, recognizing her use of a direct model and a partial products strategy. Regarding how to respond to Elizabeth's thinking, Sylvia noticed that the questions generated by ChatGPT were similar to what she would have asked but included more details specific to the student work. For example, Sylvia shared that she would have asked, "Can you explain how you solved the problem to me?" In comparison, ChatGPT suggested specific questions to guide Sylvia's response to Elizabeth's thinking:

1. Can you tell me why you decided to split the 16 chairs into two groups of 8?
2. Can you explain what your drawing shows me about the rows and chairs?
3. If someone else looked at your drawing, what might confuse them? How could you make it clearer?

Drawing on the professional noticing framework and guided by ChatGPT's suggestions, Sylvia responded to Elizabeth's mathematical thinking through targeted questioning. She first asked Elizabeth why she divided the 16 chairs into two groups of 8, following ChatGPT's first suggested question, and learned that Elizabeth found multiplication by 8 more manageable than by 16. When Sylvia probed the visual representation using ChatGPT's second suggestion, Elizabeth clarified that her lines indicated rows while the 8s represented chairs positioned on either side of a central walkway.

Finally, Sylvia implemented ChatGPT's third suggested question, asking, "If someone else looked at your drawing, what might confuse them? How could you make it clearer?" This responsive question prompted Elizabeth to add labels and written descriptions to her work, enhancing the clarity and communication of her mathematical reasoning.

In summary, Sylvia reflected, "I didn't immediately notice that the student's work lacked clear written explanations and labels, but now, after seeing ChatGPT's description and suggested questions, it is clear to me." Through the professional noticing framework, ChatGPT enhanced Sylvia's ability to attend to, interpret, and respond to Elizabeth's mathematical thinking by providing detailed descriptions of her work, interpret her reasoning more deeply by identifying areas for improvement, and responding with questions explicitly focused on Elizabeth's specific strategies and representations. This support allowed Elizabeth to articulate her mathematical reasoning in ways she might not have achieved without the targeted, responsive questioning guided by ChatGPT's analysis.

3.3 Data Analysis

To address research questions one and three regarding prospective teachers' perceptions of ChatGPT's effectiveness in supporting the "attending to" and "interpreting" components (RQ1) and the "deciding how to respond" component (RQ3) of professional noticing, I analyzed participants' written reflections for statements indicating whether they perceived ChatGPT as beneficial for analyzing student mathematical work and generating instructional responses. For both research questions, responses were coded as perceiving benefit if prospective teachers explicitly noted that ChatGPT had accurately described the student's mathematical thinking, if they indicated that ChatGPT's interpretation or suggested questions aligned with their own understanding, or if they agreed with ChatGPT's analysis or instructional suggestions even when it differed from or extended beyond their initial interpretation. Conversely, responses were coded as not perceiving benefit if prospective teachers expressed concerns about ChatGPT's accuracy or indicated significant discrepancies between ChatGPT's analysis and their own interpretation that they found problematic. Of the 28 prospective teachers in the study, 22 indicated that ChatGPT was beneficial in identifying and interpreting student mathematical strategies (RQ1), while 6 did not perceive this benefit. Similarly, for RQ3, 23 prospective teachers perceived that ChatGPT posed questions that advanced their students' thinking in ways directly rooted in their current work, while 5 did not perceive this benefit. Percentages were calculated by dividing the number of teachers who perceived benefit by the total number of participants ($N = 28$).

To examine patterns in the types of questions ChatGPT generated when responding to elementary students' mathematical work (RQ2), I collected all questions that prospective teachers reported ChatGPT had generated and submitted as part of their assignments. Following Hsieh and Shannon's (2005) methodology for inductive content analysis, I analyzed these questions without predetermined categories, allowing patterns to emerge from the data itself. Through iterative reading and comparison of the questions, three distinct categories emerged based on the apparent purpose of each question and the stage of problem-solving it addressed: advancing questions, clarifying questions, and extending questions. Each question generated by ChatGPT and reported by prospective teachers was then coded into one of these three categories.

4 Results

4.1 Research Question 1

In response to the first research question examining how prospective teachers experienced and perceived ChatGPT's effectiveness in supporting the "attending to" and "interpreting" components of professional noticing when analyzing elementary students' mathematical work on high cognitive demand tasks, 78% ($n = 22$) reported that ChatGPT successfully identified key mathematical strategies and reasoning patterns that students employed. These findings suggest that prospective teachers generally perceived ChatGPT as effective and largely accurate in describing and analyzing elementary students' written mathematical work. Only six teachers noted inaccuracies in ChatGPT's descriptions, most commonly due to unclear handwriting or the absence of contextual information that a teacher would typically gather through direct interaction with the student.

For example, Katy, one of the participants, described a case in which her student had written "35" on her paper. Katy understood from their interaction that the student meant "3 + 5" and had used her fingers to find the sum. This nuance, rooted in the moment of instruction, was not captured in the written work—and therefore was not available to ChatGPT. Katy reflected that prompting her student to write her thinking more clearly could have helped ChatGPT provide a more accurate analysis. Her experience echoed that of other teachers, who noted that while ChatGPT was helpful in attending to what was visible on paper, it could not account for the richness of real-time dialogue and observation.

For the 'interpreting' component, teachers appreciated how ChatGPT connected student strategies to mathematical concepts, though five prospective teachers noted limitations in depth. Rather than connecting student work to specific strategies or mathematical concepts, the feedback remained surface-level. Alexandra, for example, described how her student solved an equal groups problem from the NCTM resource: A group of leprechauns is planting clover patches across a meadow. Each patch has the same number of clovers. There are 2 patches, and together they have 12 clovers. How many clovers are in each patch? While ChatGPT correctly described what the student did, it did not identify the underlying concept of equal groups where the size of group is unknown. Alexandra noted, "I wish AI would have gone into more detail about what knowledge the student represented having." While ChatGPT could identify that a student used direct modeling or decomposition strategies, teachers sometimes wanted deeper analysis of the mathematical content. This finding suggests that AI tools are most effective when combined with teacher expertise in mathematical content knowledge.

Together, these reflections highlight that while the prospective teachers felt that ChatGPT offers useful and generally accurate descriptions of student work, its effectiveness is bounded by the quality of student writing and its inability to capture interactive dimensions of teaching. Overwhelmingly, the prospective teachers shared that pairing ChatGPT with their own professional judgment and classroom knowledge offered the most meaningful insights.

4.2 Research Question 2

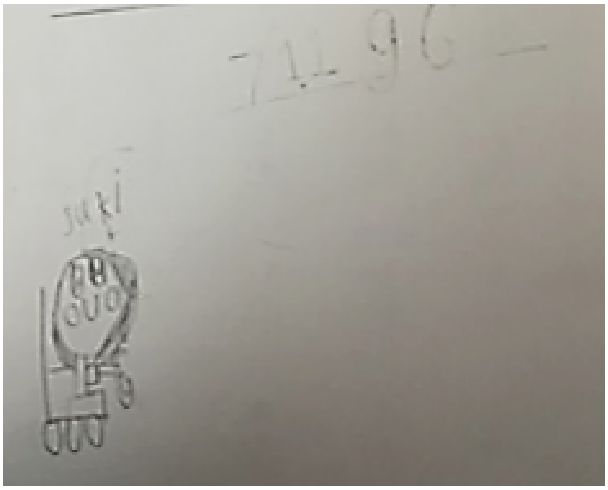
Regarding the second research question examining patterns in the types of questions ChatGPT generates when responding to elementary students' mathematical work on high cognitive demand tasks, inductive categorization of question types suggested by ChatGPT resulted in three distinct categories. Each of the following categories of question corresponded to how far the student had progressed on the problem: 1) Advancing questions encourage students to continue solving problems when they are stuck or have incomplete work by directing attention back to their current representations and helping them move forward in their problem-solving process. 2) Clarifying questions prompt students to clarify and elaborate on their thinking when they have developed a solution strategy but their explanations lack clarity or comprehensiveness, making their mathematical reasoning more explicit and complete. 3) Extending questions deepen students' thinking and promote strategic flexibility by challenging them to consider alternative approaches and make connections between different mathematical strategies when they have moved beyond a completed solution.

4.2.1 Advancing Questions

The first category, advancing questions, encourages students to persist in solving problems despite incomplete work, helping students re-engage with the problem and push forward even when their initial work is incomplete or confusing. These questions allow teachers to be responsive by attending to and building on students' partial understandings. ChatGPT often suggested prompts such as: "What do the numbers you wrote down tell you about the problem?" and "Can you go back and check what the problem is really asking?" These questions direct attention back to the student's current representations and promote perseverance in solving the problem.

For example, prospective teacher Genna worked with Molly, a first-grade student, on a dice game problem (NCTM, 2025) "Suki and John are playing a game of dice. Suki gets a 7 and 11. John gets a 9 and 6. Can you tell who has more altogether?" Molly initially wrote "7, 11, 9, 6" and drew a picture of Suki (Figure 3).

Figure 3
Molly's initial work showing only the numbers 7, 11, 9, and 6 with a drawing of Suki.



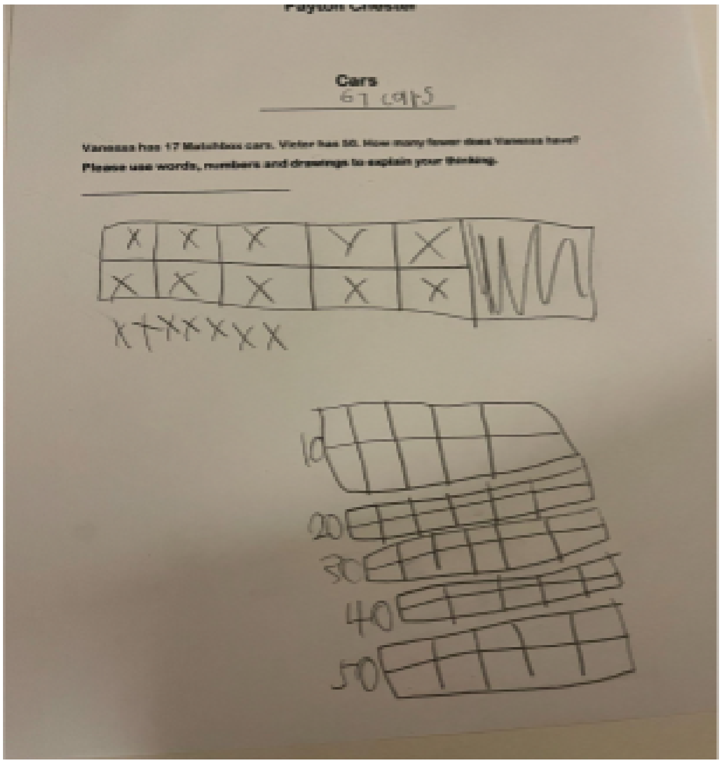
Note. *Molly's initial work showed confusion about the word "altogether" before ChatGPT's advancing question helped her clarify the problem.*

ChatGPT suggested the question: "What do the numbers 7, 11, 9, and 6 represent?" This helped Genna realize Molly was confused by the word "altogether." After clarification, Molly added tally marks to represent her calculations and arrows to compare totals.

Similarly, prospective teacher Patricia's second grade student Felix worked on this problem (NCTM, 2025): "Vanessa has 17 Matchbox cars. Victor has 50. How many fewer does Vanessa have? Please use words, numbers, and drawings to explain your thinking." Felix initially drew a ten frame with seven additional X's below it to represent Vanessa's 17 cars, and five additional ten frames labeled 10, 20, 30, 40, 50 for Victor's cars (see Figure 4).

Figure 4

Felix's representation of Vanessa's 17 cars and Victor's 50 cars using ten frames.



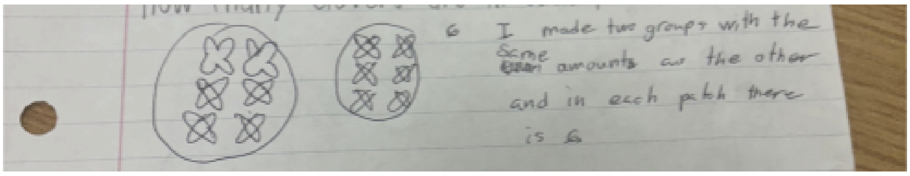
Note. *Felix's model showed clear understanding of the quantities but needed support connecting his visual representation to the difference between 50 and 17.*

ChatGPT suggested: "Can you explain how your drawings represent the difference between Vanessa's and Victor's cars?" This question prompted Felix to build on his model and calculate the difference as 33, showing how these prompts help move students forward.

4.2.2 Clarifying Questions

The second category of questions, clarifying questions, focuses on pushing students to make their ideas clearer—both for the teacher and for the students themselves—with the difference from the first type of question being that the student has completed a viable solution but could enhance their communication. They often include prompts like: "Can you explain what you did as if you were helping another student?", "How does your drawing connect to the numbers you used?", and "Can you show how your model matches your calculations?" For instance, prospective teacher Betsy worked with Leah, a fourth-grade student solving the equal groups problem described earlier involving 12 clovers and 2 patches. Leah drew two circles with six clovers each and wrote, "I made two groups with the same amounts." ChatGPT suggested asking: "Can you explain how you knew each group had to have 6 clovers?" This encouraged Leah to articulate, "Because the problem says that each patch has the same number... 6 plus 6 is 12." This interaction supported both clarity and justification of her reasoning (Figure 5).

Figure 5
Leah's representation of two clover patches with 6 clovers in each patch.



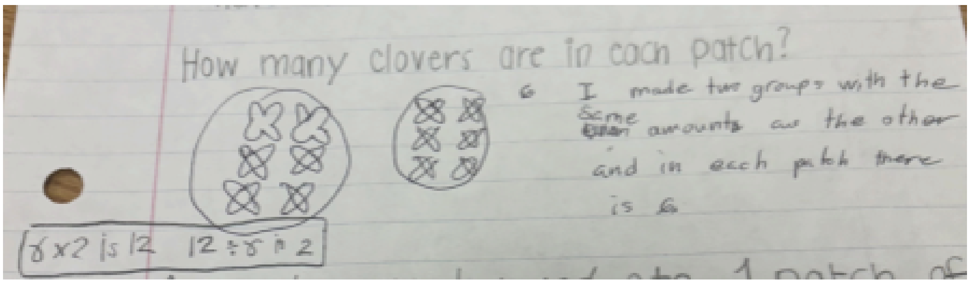
Note. *Leah's visual model effectively represented the equal groups concept, and ChatGPT's clarifying question helped her verbalize her reasoning.*

4.2.3 Extending Questions

Building on the foundation of clear explanations, the next category of questions, extending questions, challenges students to expand their mathematical thinking. These questions encourage students to consider alternate approaches, make connections between strategies, and generalize their reasoning. Prompts often include: "Can you solve this in a different way?", "What's another problem that's like this one?", and "How does subtraction help you check your answer?"

In Leah's continued work on the clover problem, ChatGPT also suggested: "You used a picture with Xs [to represent the clovers] to solve the problem—what other way could you show or explain this problem to someone without using a drawing? Could you use numbers or a number sentence instead?" The purpose, shared by ChatGPT, was to promote flexible thinking and invite Leah to explore alternative strategies focused on sharing a related equation. Leah shared in response: "I would tell them, this is a division problem and to do 12 divided by 2. Since we all know that 6 plus 6 is 12, we also know that 6 times 2 is 12 too. Hopefully the person would know that 6 times 2 is 12 which means that 12 divided by 6 is 2." She added "$6 \times 2 = 12$" and "$12 \div 6 = 2$" to her paper, which didn't align with her spoken words, leading to further opportunity for Betsy to engage responsively with Leah's thinking (see Figure 6).

Figure 6
Leah's work showing her alternate representation using number sentences.



Note. *Leah's extended work showed her ability to think flexibly about division, though the mismatch between her verbal explanation and written equations created an opportunity for further responsive teaching.*

The consistency of these question categories across different mathematical tasks and student grade levels suggests that ChatGPT's analysis follows recognizable patterns aligned with responsive teaching principles. Teachers' understanding these patterns might help them anticipate the types of questions that might be most useful at different points in student problem-solving, potentially supporting their own question-generation skills over time.

4.3 Research Question 3

Examining prospective teachers' reported experiences with ChatGPT-generated questions for supporting the 'deciding how to respond' component of professional noticing (RQ3), 82% of the 28 participating teachers reported that they felt that the AI tool posed questions that advanced their students' thinking in ways directly rooted in their current work. Five of the prospective teachers reported that these AI-generated questions supported their development of the 'responding' component of professional noticing by providing specific, targeted prompts rooted in student work. Rather than asking generic questions like 'Can you explain your thinking?' Overwhelmingly, the prospective teachers felt that the ChatGPT responses helped them learn to use responsive questions that built on student reasoning.

Betsy, a prospective teacher, shared that "ChatGPT showed me how important it is to use probing questions that center student thinking and connect math, specifically in word problems, to the situation." Meredith found that "AI worded the questions a little better." However, as Quinn reflected, AI works best as a supplementary tool rather than a replacement for teacher judgment: "AI doesn't know your students like you do" and should serve as "a base" for conversations that teachers then enhance with their own questioning.

5 Conclusions, Limitations and Implications for Practice

This evidence-based study examined prospective teachers' perceptions and experiences of using ChatGPT to support professional noticing in responsive mathematics teaching. The professional noticing framework provides a useful structure for introducing AI tools. It helps prospective teachers understand how technology can support specific aspects of responsive teaching while maintaining focus on student thinking.

Results revealed that seventy-eight percent of the prospective teachers found ChatGPT effective for attending to and interpreting student work. At the same time, some prospective teachers noted its limitations in interpreting handwriting, capturing the nuances of in-person interactions, and providing in-depth analysis of the mathematical knowledge specific to teaching. The findings suggest that prospective teachers believe that AI tools can effectively support specific components of professional noticing while highlighting the limitations of AI tools and the continued importance of teacher expertise and judgment.

Regarding the responding component, the three types of questions provide a useful framework for understanding how AI might support responsive questioning practices. When teachers understand these patterns, they might anticipate the types of questions that could be most useful at different points in student problem-solving, potentially supporting their own question-generation skills over time. Notably, eighty-two percent of the prospective teachers indicated that the AI-generated questions advanced their students' thinking.

This study utilized ChatGPT (ChatGPT-3.5, OpenAI, 2024), which was a no-cost option available to participants at the time of data collection. However, the constantly changing nature of large language models presents several limitations to generalizability. ChatGPT-3.5 has since been superseded by multiple versions with potentially different analytical capabilities, situating our study as a snapshot bounded by time. While newer free versions of ChatGPT have been released, it remains unclear whether any advancements translate to the specialized teaching skill of noticing and interpreting elementary students' mathematical thinking—a practice that requires both mathematical and pedagogical expertise.

Further, we provided prospective teachers with an initial prompt guided by tenets of professional noticing, including the responding component (Jacobs et al., 2010). While participants were asked to use this given prompt, we cannot confirm with certainty that all participants used it exactly as provided, as this study focused on examining prospective teachers' perspectives rather than tracking the precise prompts used. This potential prompt variation, while authentic to how educators might use AI tools in practice, limits our ability to isolate the specific impact of ChatGPT's analytical framework. Future research could examine prompt development and AI's ability to professionally notice students' mathematical thinking in comparison to mathematics educator experts, specifically assessing ChatGPT's accuracy in identifying strategies, recognizing misconceptions, and understanding developmental progressions within established frameworks. This paper explored prospective teachers' perspectives on using ChatGPT to professionally notice students' mathematical thinking.

Teacher educators must think deeply about how to leverage AI tools to support and enhance mathematics teacher education. When integrating AI tools like ChatGPT into teacher preparation programs, teacher educators should consider prospective teachers' existing ability and experiences with professional noticing, their access to ChatGPT for efficiently analyzing student work, and their capability to consider AI's use reflectively. Prospective teachers should learn to engage with AI as a supplement to, not a replacement for, teacher expertise and judgment. As one prospective teacher, Molly, reflected, she found that the experience demonstrated AI's potential as a useful tool for brainstorming and refining questioning techniques, while emphasizing that it required her input to ensure questions truly supported her students' learning.

AI Use Statement

Generative AI tools (e.g., ChatGPT, Claude) were used for editing, organization, and brainstorming during the writing process. All AI-generated suggestions were critically reviewed, revised, and integrated by the authors. The final manuscript reflects the authors' original ideas, interpretations, and academic contributions.

References

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1), 1–22.

Coffey, J. E., Edwards, A. R., & Finkelstein, C. (2008). Engaging with student thinking: The work of teaching in the elementary mathematics classroom. In M. Lampert & K. Kazemi (Eds.), *Proceedings of the 32nd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 853–861). PME-NA.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109–1136. https://doi.org/10.1002/tea.20440

Dao, X. Q., & Le, N. B. (2023). Investigating the effectiveness of ChatGPT in mathematical reasoning and problem solving: Evidence from the Vietnamese national high school graduation examination. arXiv. https://doi.org/10.48550/arXiv.2306.06331

Dyer, E. B., & Sherin, M. G. (2016). Instructional reasoning about interpretations of student thinking that supports responsive teaching in secondary mathematics. *ZDM, 48*(1), 69–82.

Hammer, D., Gold, D., & Ferguson, K. (2016). *Responsive teaching in the 21st century.* Cambridge University Press.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524–549.

Huinker, D., & Bill, V. (2017). *Taking action: Implementing effective mathematics teaching practices in K-grade 5* (M. S. Smith, Series Ed.). National Council of Teachers of Mathematics.

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169–202. https://doi.org/10.5951/jresematheduc.41.2.0169

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education, 191*. https://doi.org/10.1016/j.compedu.2022.104646

Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching, 6*(1), 64–74.

Mai, D. T. T., Da, C. V., & Hanh, N. V. (2024). The use of ChatGPT in teaching and learning: A systematic review through SWOT analysis approach. *Frontiers in Education, 9*, Article 1328769. https://doi.org/10.3389/feduc.2024.1328769

National Council of Teachers of Mathematics. (2025). *Problems of the week.* https://www.nctm.org/pows/

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* NCTM.

National Council of Teachers of Mathematics. (2024). *Artificial intelligence and mathematics teaching.* https://www.nctm.org/standards-and-positions/Position-Statements/Artificial-Intelligence-and-Mathematics-Teaching/

Selmer, S., Lindstrom, D., & Lampen, E. (2022). A case study of prospective teachers engaged in professional noticing of their students' mathematical thinking. *Education Sciences, 12*(10), 656. https://doi.org/10.3390/educsci12100656

Sherin, M. G., & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475–491.

Smith, M. S., & Stein, M. K. (1998). Reflections on practice: Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344–350.

Son, T., Yeo, S., & Lee, D. (2024). Exploring elementary preservice teachers' responsive teaching in mathematics through an artificial intelligence-based Chatbot. *Teaching and Teacher Education, 146.* https://doi.org/10.1016/j.tate.2024.104640

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–370). Information Age Publishing.

Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction, 95.* https://doi.org/10.1016/j.learninstruc.2024.102011

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2019). *Elementary and middle school mathematics: Teaching developmentally* (10th ed.). Pearson.

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). https://doi.org/10.29333/ejmste/13272

Sarah Selmer is an associate professor of mathematics education in the Department of Curriculum and Instruction at West Virginia University. Dr. Selmer is a current Fulbright Specialist where she pursues her interests in mathematics education through international collaborations. Her research interests include elementary teachers' noticing and sensemaking of students' mathematical thinking and the development of responsive teaching practices.