Assessing Concepts, Procedures, and Cognitive Demand of ChatGPT-generated Mathematical Tasks

Bima Sapkota
The University of Texas Rio Grande Valley, USA

Liza Bondurant
Mississippi State University, USA

To cite this article:

Assessing Concepts, Procedures, and Cognitive Demand of ChatGPT-generated Mathematical Tasks

Bima Sapkota, Liza Bondurant

Article Info	Abstract
	In November 2022, ChatGPT, an Artificial Intelligence (AI) large language model (LLM) capable of generating human-like responses, was launched. ChatGPT has a variety of promising applications in education, such as using it as thought-partner in generating curricular resources. However, scholars also recognize that the use of ChatGPT raises concerns, such as outputs that are inaccurate, nonsensical, or vague. We, two mathematics teacher educators, engaged in a collaborative self-study using qualitative descriptive approaches to investigate the procedures, concepts, and cognitive demand of ChatGPT-generated mathematical tasks focused on fraction multiplication using the area model approach. We found that the ChatGPT-generated tasks were mostly procedural and not cognitively demanding. Moreover, despite ten variations of input prompts, ChatGPT did not produce any tasks that used the area model approach for fraction multiplication. Rather, it generated tasks focused on procedural approaches. Alarmingly, some tasks were conceptually and/or procedurally inaccurate and vague. We suggest that educators cannot fully rely on ChatGPT to generate cognitively demanding fraction multiplication tasks using the area model. We offer recommendations for educators' strategic use of ChatGPT to generate cognitively demanding mathematical tasks.

Introduction

In the United States, elementary teachers are challenged to make informed decisions when selecting materials from the vast array of online and print resources (Doherty et al., 2022). Elementary teachers in an urban U.S. elementary school used up to 11 distinct mathematics curricular resources, many of which required them to pay additional out-of-pocket fees, to create engaging mathematical learning opportunities for their students (Doherty et al., 2022). In the Global context, many teachers and schools cannot afford to pay for curricular resources, which creates unfair distributions of curricular resources (Maringe et al., 2013). For example, based on the first author's personal experience as an elementary teacher, many elementary teachers in Nepal only use one mandated textbook as they do not have access to online resources. Additionally, as a teacher in rural and urban districts in New York, United States, the second author spent hundreds of dollars purchasing materials each year, because her students were not engaged by the lessons in the school-adopted textbook. Educators (e.g., Yu, 2023) have suggested that ChatGPT could play a crucial role in addressing disproportionate distributions of curricular resources, as it has

the potential to generate multiple lessons, tasks, problems, as well as different solution strategies using a variety of representations and embedded in various culturally relevant context to optimize elementary mathematics lessons. In this context, many teachers and educators have begun using it to create, proctor, and grade assignments (Trust et al., 2023). Because of its user-friendly interface (Education Week, 2023; Grassini, 2023), it is accessible to many teachers.

Even though ChatGPT has the potential to contribute to a fairer distribution of educational resources, educators (e.g., Farrokhnia et al., 2023) have identified its limitations and suggested exercising caution. Farrokhnia and colleagues studied the potential educational implications of ChatGPT and suggested the lack of in-depth understanding and higher-order thinking as two weaknesses. The same author team outlined that ChatGPT could be a threat to education as it might contribute to declining high-order cognitive skills. Similarly, Wardat and colleagues (2023) found that ChatGPT lacks an in-depth understanding of Geometry. Additionally, Shakarian et al. (2023) found that the chances of producing correct responses to mathematics word problems were lower when there was an increased number of equations, unknowns, and division and multiplication operations. This finding suggests a specific limitation of ChatGPT with mathematics teaching. Moreover, Onal and Kulavuz-Onal (2023) found that mathematical tasks generated by ChatGPT lacked real-world context.

Given the studies surfacing the limitations of ChatGPT in producing and solving mathematics problems, we became interested in investigating the extent to which ChatGPT could produce the tasks that engage students in problem-solving and non-algorithmic thinking, which are commonly referred to as *cognitively demanding tasks* (Stein et al., 2000), *rich mathematical tasks, or group worthy tasks* (Boaler, 2016; Leinwand & Wiggins, 1991; Liljedahl, 2020). For this study, we will use the term *cognitively demanding tasks* (Stein et al., 2000). In this manuscript, we delve into a collaborative self-study (Butler & Bullock, 2022) of two mathematics teacher educators (MTEs) using LLM-driven (Large Language Model) tools to create cognitively demanding elementary mathematics instructional tasks for university elementary mathematics-focused methodology and content courses. By assessing the cognitive demand, concepts, and procedures of ChatGPT-generated mathematical tasks focused on fraction multiplication using an area model approach, we guide how to effectively use ChatGPT as a resource for generating cognitively demanding tasks. We also provide the field with a cautionary tale of the limitations of using ChatGPT to produce cognitively demanding tasks. The research question guiding this study was: How do ChatGPT-generated tasks on fraction multiplication using an area model vary in terms of concepts, procedures, and cognitive demand?

Review of Relevant Literature

Use of ChatGPT as Mathematics Curricular Resources: Strengths, Limitations, and Potential Threats

Among many platforms and tools of AI, many mathematics teachers and educators have begun using ChatGPT for a variety of purposes (Wardat et al., 2023). Trust and colleagues (2023) outlined the following nine potential areas in which educators can use ChatGPT: (a) provide support with teaching; (b) provide support with student assessment; (c) support student learning; (d) offer suggestions for improving teaching; (e) support teacher-parent, and teacher-student communication; (f) personalized learning support; (g) creative thinking support; (h)

assessment support; and (i) reading and writing comprehension support. We suggest that these applications could be synthesized into three main categories: writing, providing feedback, and generating instructional tasks. Many educators appreciate ChatGPT's strength as it can provide personalized and adaptive outputs (Qadir, 2022). A plethora of studies examined the use of ChatGPT to generate curricular resources from a critical lens (e.g., Onal & Kulavuz-Onal, 2023). As such, we found some recent studies that critically evaluated the limitations and threats to educators as a result of the widespread use of ChatGPT.

Farrokhnia and colleagues conducted a study by analyzing strengths, weaknesses, opportunities, and threats (SWOT) associated with the use of ChatGPT in education. They suggested that ChatGPT is sophisticated enough to generate personalized, real-time, and plausible responses. Aligning with this finding, Li and colleagues (2019) and Kasneci and colleagues (2023) had previously suggested that relevant ChatGPT responses are possible due to their capacity to capture a range of linguistic patterns and relationships. Farrokhnia and colleagues suggested that educators could benefit from these strengths to reduce their workload. Aligning with these suggestions, we assessed whether these suggested strengths could be evidenced when educators use ChatGPT to fulfill curricular needs, specifically in generating cognitively demanding mathematical tasks.

Educators have suggested several limitations of ChatGPT and other AI platforms. Farrokhnia et al. (2023) listed the following four potential weaknesses of using ChatGPT in education: (a) lack of deep understanding, (b) difficulty in evaluating the quality of responses, (c) the risk of biases and discrimination, and (d) lack of higher-order thinking skills. The author team also stated the potential threats of ChatGPT in education including lack of contextual outputs, increased plagiarism, discrimination, and reduced higher-order cognitive skills. These suggested limitations and threats are broadly in the educational contexts. Jeon and Lee (2023) suggested that teachers and ChatGPT could complement each other when generating and using curricular resources. Even though teachers can use ChatGPT to create curricular resources efficiently, they need to examine those resources critically both from the perspective of pedagogy and content. The authors also suggested developing specific professional development for teachers to effectively navigate the use of ChatGPT for instruction. Cooper (2023) also highlighted that educators should critically evaluate any ChatGPT-generated resources for their accuracy and make them contextual. In our study, we will investigate the potential affordances and limitations of ChatGPT in generating mathematical tasks.

Because our focus was on assessing the cognitive demand of mathematical tasks (which fall under the realm of curricular resources), we reviewed studies with a similar focus. Onal and Kulavuz-Onal (2023) examined the mathematical word problems generated by ChatGPT to teach a college-level course (i.e., Manufacturing Process). They found that the generated problems were appropriate for teaching the topic. However, the problems lacked contextual factors, making them less engaging for students. Additionally, those tasks were not open-ended.

Yu (2023) found that ChatGPT lacks an in-depth understanding of geometry, which could be a limitation in generating mathematical tasks and/or asking for solutions to geometry problems. Yu also found that ChatGPT cannot effectively correct misconceptions related to several geometry concepts/topics. Gattupalli and colleagues (2023) compared teacher and ChatGPT-generated strategies and hints for fourth-grade mathematics word

problems. They found that the problem-solving strategies generated by ChatGPT were text-heavy and mostly procedural whereas teachers wrote more contextual strategies.

Since studies on the use of ChatGPT for mathematics instruction are still evolving, we found only a few studies on the potential and limitations of ChatGPT in mathematics teaching. The limited studies we cited above suggest that ChatGPT may lack in producing contextual mathematics problems and related solutions. Building on this literature, we used the concept of cognitively demanding tasks (Stein et al., 2000) to investigate the extent to which ChatGPT could generate tasks that have a potential to engage students in non-algorithmic thinking and problem-solving.

Concept of Fraction Multiplication Using Area Model

According to the Common Core State Standards for Mathematics (CCSSM), the concept of a fraction is introduced in the third grade in the United States (National Governors Association Center for Best Practices [NGA]). In grade three, students are introduced to the concept of unit fractions. In grade four, students are expected to learn to add and subtract fractions with the same denominator, and they are introduced to the multiplication of fractions by whole numbers. In fifth and sixth grades, students develop a conceptual understanding of how the algorithm of fraction multiplication makes sense and are expected to interpret the multiplication of fraction by fraction with physical and graphical representations (NGA, 2010). Overall, students are expected to develop a conceptual understanding of how the fraction multiplication algorithm works by the end of sixth grade.

Van De Walle and colleagues (2023) illustrated that "ideas of iterating (counting) fractional parts and partitioning are foundational to understanding multiplication of fractions" (p. 394). Focusing on the unit is also essential in understanding the fraction multiplication concept (Beckmann, 2022a). Beckmann (2022a) suggested developing students' conceptual understanding of fraction multiplication using rectangular area model representation. For example, when students are asked to represent the yellow shaded shell by $\frac{a}{b}$ of $\frac{c}{d}$ (i.e., $\frac{a}{b} \times \frac{c}{d}$), they begin to understand that the shaded part represents a new fraction resulting from a multiplication of two fractions. The area model representation can aid student understanding of the abstract concept of fraction multiplication (Cope, 2015).
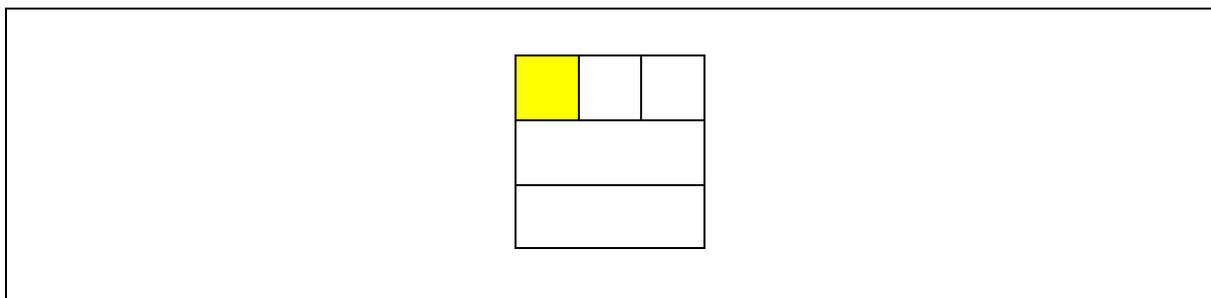

Figure 1. A Rectangular Model in which the Shaded Area Represents $\frac{1}{3}$ of $\frac{1}{3}$ of a Unit (area of a rectangle)

More preciously, using the area model, students are involved in partitioning and sub-partitioning of the rectangular unit model. In particular, students use the process below to find the product of $\frac{2}{3}$ of $\frac{5}{8}$.
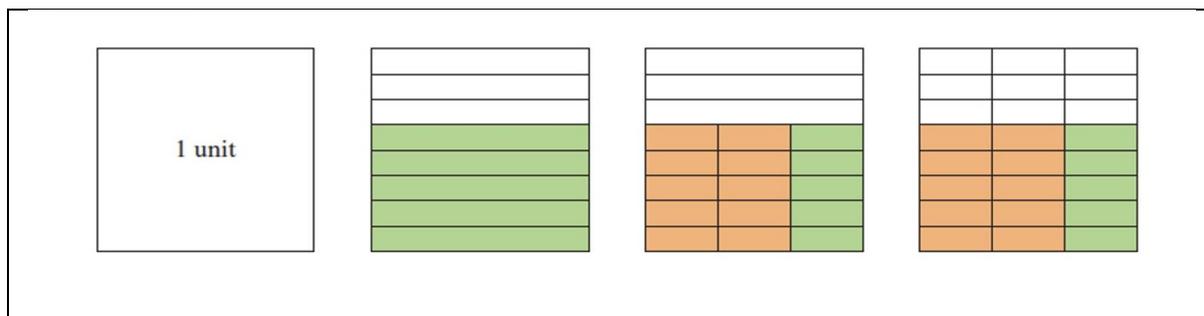
Figure 2. A Rectangular Model in which the Orange Shaded Area Represents $\frac{2}{3}$ of $\frac{5}{8}$ of a Unit (Beckmann, 2022b, p. 112)

Van De Walle and colleagues (2022) also emphasized that using story problems in which students use the area of rectangle and partitioning enhances students' understanding of fraction multiplication. In this context, we investigated whether ChatGPT was sophisticated enough to generate these types of problems. We were interested in generating cognitively demanding problems that have the potential to engage students in multiplying fractions using an area model. An example of the problem that we were looking for is as follows: *Harry ate ⅓ of a cake and kept the leftovers in the fridge. His brother, James, ate ⅖ of the leftovers. How much of the cake did James eat? Please show your process through a mathematical representation and explain how the mathematical process you selected connects with your mathematical representation.* This problem is a high-cognitive demand task for 4th or 5th-grade students because it does not indicate which procedure they need to use. In addition, it asks students to justify their reasonings with multiple representations.

Teachers must understand the nature of problems when teaching fraction multiplication through a rectangular area model because some fractions require students to subdivide fractions and some do not (Van De Walle et al., 2023). For example, $\frac{1}{2}$ of $\frac{2}{5}$ does not require students to subdivide because they have two partitions and take $\frac{1}{2}$ of those two partitions. Problems like $\frac{1}{2}$ of $\frac{1}{4}$ require subdivision as students need to split each portion of the unit rectangle into two equal parts. Problems that require subdivision are more difficult than the problems that do not require subdivision. We were interested in investigating the types of problems (requiring subdividing and not requiring subdividing) ChatGPT generates when providing relevant inputs (prompts).

Conceptual Framework: Cognitive Demand of Mathematical Tasks

We used Stein and colleagues' (2000) *Task Analysis Guide* (TAG) to analyze the cognitive demand of mathematical tasks generated by ChatGPT. TAG can be used to evaluate whether mathematical tasks are at a lower or higher level of cognitive demand. The cognitive demand of a task refers to the mental processing required to solve a task (Stein et al., 2000; Wilhelm, 2014). Tasks that are considered to have a lower level of cognitive demand typically focus on memorization, performance of algorithms, and procedural computations absent of reasoning or sense-making, whereas those with a higher level of cognitive demand are commonly rooted in conceptual understanding and emphasize the why and how behind the procedures.

Lower-level tasks are subcategorized into *Memorization* and *Procedure Without Connection.* Higher-level tasks

are subcategorized into *Procedures with Connections* and *Doing Mathematics* tasks (see Table 1). It should be noted that rating tasks may be highly subjective as scholars have found that educators may use different terms to describe the cognitive demand of tasks (Sapkota, 2022). Moreover, some educators may perceive a task as having a higher level of cognitive demand while other educators may perceive the same task as having a lower level of cognitive demand (Sapkota, 2022).

Table 1. Cognitive Demand of Mathematical Tasks

Lower Level	Higher Level
Memorization Tasks ● Demand recalling previously learned facts, formulas, and procedures. ● Are not ambiguous and lack connections to mathematical concepts and procedures due to various factors (e.g., time limitation, length of solutions). Procedure without Connection ● Follow established algorithms, where the use of a procedure is either explicitly instructed or inherently clear. ● Are minimally ambiguous and lack connections with mathematical concepts and reasonings.	Procedure with Connection ● Require using procedures in connection with mathematical concepts and big ideas. While general procedures can be followed, students must actively engage with the conceptual ideas that underpin the procedures to complete tasks and cultivate a deeper understanding. ● Typically, involves multiple representations and establishing connections among those representations and concepts. Doing Mathematics ● Involve complex, non-algorithmic thinking, with no predefined or well-established procedures and/or approaches provided by tasks, instructors, or worked examples. ● Requires devising own problem-solving strategies, which involves connections among multiple mathematical ideas.

Note. The table is adapted from Stein et al. (2000). We paraphrased the description of each category of tasks from Stein and colleagues (2000).

Understanding the cognitive demands of mathematical tasks is critical because tasks demanding different cognitive levels tend to create levels of student engagement (Sapkota, 2022; Hiebert & Wearne, 1993; Wilhelm, 2014). For example, *Doing Mathematics* tasks provides students with opportunities to devise their strategies, which might lead to meaningful discussions whereas *Procedure Without Connections* tasks only require known procedures thus these types of tasks might not provide meaningful discussion opportunities. Therefore, it is critical to assess what types of tasks are produced before implementing them.

Methods and Methodology

Researcher Positionality

As researchers in the field of education, we believe that we must practice reflexivity regarding our identities to appropriately position ourselves within the research context. Our identities and experiences shape our commitments to and challenges in undertaking this research project. Reflexivity enhances our awareness of how our positions may influence our perspectives and biases, impacting aspects such as our research design, data analysis, and interpretation of findings (Johnson & Fonbuena, 2023; Milner, 2007; Yogeeswaran et al., 2016). Although we researched, read, and incorporated the perspectives of diverse scholars in this manuscript, we recognize the potential for oversight in our analyses due to the limitations of our own identities and lived experiences.

With the growing popularity of AI, we were both curious about how ChatGPT could help us generate mathematical tasks for the courses we teach. In our prior experiences, many preservice elementary mathematics teachers struggle with fraction multiplication. However, we have found that teaching fraction multiplication using an area model helps preservice elementary mathematics teachers (PSTs) gain conceptual understanding and procedural fluency. Moreover, we have experienced success using gradually more cognitively demanding tasks to assess PSTs' knowledge and skills. Therefore, we decided to research how ChatGPT-generated tasks on fraction multiplication using an area model vary in terms of concepts, procedures, and cognitive demand and the characteristics of the input that generates the optimal ChatGPT output.

In the context of this manuscript, we consider our educational experiences as students and educators most relevant to our analysis. Both authors have doctorates in mathematics education and have experience analyzing the cognitive demands of mathematical tasks. Moreover, recently both authors have embarked on research into the optimal integration of ChatGPT for educators. The first author completed her master's and taught in elementary and secondary classrooms in Nepal for six years. In addition, she obtained her Ph.D. and has been teaching elementary and secondary mathematics methods courses in the United States for six years. The second author completed her K-12, bachelor's, master's, and Ph.D. in upstate New York. She has over 18 years of teaching experience.

Study Design

We used a descriptive qualitative research approach to investigate the characteristics of mathematical tasks generated by ChatGPT (Elliott & Timulak, 2005). This approach was appropriate because our research question was exploratory. As such, we were interested in investigating "kinds or varieties" of tasks and how the inputs influenced variations in the outputs (Elliott & Timulak, 2005, p. 149). Furthermore, this approach allowed us to use two rounds of meaning-making (Sapkota et al., in press; Smith & Osborn, 2008). First, ChatGPT made sense of our inputs. Second, we interpreted the mathematical concepts, procedures, and cognitive demands involved in the ChatGPT-generated outputs.

For this study, we embraced a collaborative self-study approach as detailed by Butler and Bullock (2022). Kosnik and colleagues (2009) outlined the key attributes of self-study methodology, encompassing elements such as openness, the incorporation of multiple viewpoints, collaboration, embracing paradoxical aspects, reframing, and adopting a postmodern perspective. Throughout our self-study, we actively encouraged one another to remain receptive to perspectives that deviated from our initial positions. Our collaborative efforts played a pivotal role in facilitating our examination of concepts from diverse viewpoints. The essence of this self-study carried a paradoxical quality since, as "critical friends," we prompted each other to question and reevaluate our initial perspectives.

Procedures for Selecting and Using Prompts

We used ChatGPT to generate fraction multiplication mathematical tasks that require an area model. As we mentioned earlier, we aimed to generate high-cognitive demand tasks, most preferably *Doing Mathematics* tasks. Our initial inputs were more generic. Then, based on our analyses of the outputs, we refined the sensitivity of our inputs (e.g., by including key terms like cognitively demanding tasks, problem-solving tasks). We used multiple phrases that had the potential to provide cognitively demanding tasks as outputs. For example, we used the phrases "problem-solving tasks" (Stein et al., 2000), "cognitively demanding tasks" (Stein et al., 2000), "rich mathematical tasks" (Boaler, 2016; Leinwand & Wiggins, 1991), "high-level cognitive demand task" (Stein et al., 2000), and "rigorous tasks" (Leinwand & Wiggins, 1991; Sapkota, 2022). We chose these phrases purposefully as they were used in the literature to indicate cognitively demanding tasks or problems (Boston & Smith, 2011).

Table 2. List of Representative Inputs

1. Generate one instructional task **to teach** fraction multiplication with an area model approach.
2. Generate one **problem-solving task** related to fraction multiplication using an area model.
3. Generate one **rigorous task on fraction multiplication** with an area model.
4. Generate one **cognitively demanding mathematical question** on fraction multiplication with an area model.
5. Generate one **high-level cognitively demanding task** for fraction multiplication using an area model.
6. Generate one **rich mathematical task** for fraction multiplication using an area model.
7. Generate a **cognitively demanding task** for fraction multiplication. Cognitively demanding tasks engage students in complex non algorithmic thinking.
8. Generate a **cognitively demanding task** for fraction multiplication. Cognitively demanding tasks engage students in complex non algorithmic thinking. Cognitively demanding tasks also require students to explore and understand the nature of mathematical concepts, processes, or relationships.
9. Generate a **cognitively demanding fraction multiplication task** that requires partitioning the area of a rectangle.
10. Generate a **cognitively demanding fraction multiplication task** that requires **partitioning the area of a rectangle.** Consider the rectangle as a unit.

After trying several inputs (i.e., prompts), we realized that ChatGPT was not generating tasks that were in the category of *Doing Mathematics*. Thus, we tried to provide input by including the descriptions of cognitive demand. As you can note in Table 2, in prompt 7 we provided a brief explanation of the cognitive demand of tasks but still did not receive our intended output. Thus, we provided a more detailed description of cognitively demanding tasks to examine whether or not that description might change the output (prompt 8).After trying many inputs, we also found that ChatGPT was *not* producing tasks that required students to engage in calculating areas of the rectangular area model. Recall that we aimed to produce tasks that require the partitioning of areas of a rectangle (considering the area of the rectangle as a unit). Thus, we tried with other inputs by including words such as partitioning and unit. Inputs 9 and 10 are examples of those prompts. We both tried with several prompts independently and discussed the outputs. We then decided to include only 10 representative inputs and their outputs here and in the findings section.

Analytic Process

We coded the cognitive demand, mathematical procedures, and mathematical concepts of ChatGPT-generated tasks. First, we applied descriptive and open coding (Saldaña, 2016) to assess mathematics concepts and procedures used in tasks. For example, task 1 requires students to use a formula (i.e., length × width) to find the area. Thus, it involves concepts and processes related to fraction multiplication and the area of a rectangle.We then used Stein and colleagues' (2000) TAG to sort the tasks into four categories of cognitive demand (i.e., Memorization, Procedure with Connection, Procedure without Connections, and Doing Mathematics). Specifically, we used the descriptor of each type of task from Table 1 to characterize the tasks. We wrote a justification for why we sorted each task into a certain category in parenthesis in the fourth column (see Table 3). We independently coded the data first and then met to discuss our coding and resolve any discrepancies.

Findings

In the section, we first present the list of ChatGPT-generated tasks, a brief description of those tasks in terms of concepts used in the tasks, required procedures to solve those tasks, and cognitive demands of tasks. We then discuss the key patterns we noticed when analyzing the cognitive demand and the nature of those tasks. Note that the outputs in this table are associated with the inputs in Table 2. For example, Task 1 in Table 3 is associated with the first input in Table 2.

Table 3. Concepts, Procedures, and Cognitive Demand of ChatGPT generated Tasks

Output (Tasks)	Concepts	Procedures	Cognitive Demand
Task 1: You are an architect working on a house design project. The kitchen floor has a length of 5/6 meters and a width of 3/4 meters. Using an area model, calculate the total area of	Area of a rectangle Fraction multiplication	Applying a formula to calculate the area of a rectangle. Doubling the	Procedure without Connections (involves only using procedures without an explicit connection to the

Output (Tasks)	Concepts	Procedures	Cognitive Demand
the kitchen floor in square meters. Next, the living room has the same dimensions as the kitchen. How can you represent the combined floor area of the kitchen and living room using the area model?		area of a rectangle	conceptual understanding)
Task 2: You are planning to create a rectangular garden with a length of 2/3 meters and a width of 3/4 meters. Each square meter of the garden requires 2 bags of soil. Using an area model, determine how many bags of soil you'll need to cover the entire garden. Additionally, if you want to extend the garden by doubling both the length and width, how many bags of soil will be required for the larger garden? Represent your solutions using the area model.	Area of a rectangle Fraction multiplication	Applying a formula to calculate the area of a rectangle. Multiplying the area by 2 to find the total bags of soil. Doubling the length and width of the original rectangle and finding the area of the new rectangle	Procedure with Connections (requires that students use procedures to find area of rectangle and connecting with the concept that the area is calculated in unit square to find the total number of bags)
Task 3: You are working on a city planning project to design a park. The park has a complex layout with multiple sections, each with varying dimensions. One section is 3/8 meters long and 5/6 meters wide, another is 4/7 meters long and 2/3 meters wide, and a third is 1/5 meters long and 3/4 meters wide. Using an area model, calculate the total area of each section. Then, determine the combined area of all three sections.	Area of a rectangle Fraction multiplication Addition of fractions	Applying a formula to calculate the area of three rectangles. Adding areas of three rectangles	Procedure without Connections (requires engaging in the same procedure of calculating the area model twice and adding areas)
Task 4: You are an engineer	Area of a	Applying a	Procedures without

Output (Tasks)	Concepts	Procedures	Cognitive Demand
tasked with designing a rectangular greenhouse with intricate sections. One section is 7/9 meters long and 2/3 meters wide, another is 3/5 meters long and 4/7 meters wide, and a third is 4/9 meters long and 5/8 meters wide. Each section needs different quantities of glass for its construction, which is sold in square meters. Using an area model, determine the total area of glass required for each section, and then calculate the total area of glass needed for the entire greenhouse, considering these diverse section dimensions.	rectangle Fraction multiplication Addition of fractions	formula to calculate the area of a rectangle and a square. Adding fractions to calculate the total area.	Connections (requires using procedures related to calculating areas of rectangles and adding those areas)
Task 5: You are designing a stained-glass window for a church. The window is shaped like a triangle and measures 1 1/3 meters in height and 1 2/5 meters in base width. The central design, in the shape of a rectangle, will occupy 2/3 of the window's total area, excluding the triangular frame. The rectangular design's height is 3/4 of the window's height. a) Using an area model, illustrate the dimensions of the rectangular central design within both the central design and the triangular frame and (b) calculate the area of the rectangular model.	Fraction multiplication Area of a rectangle	Calculating one dimension of the rectangle using fraction multiplication. Calculating the area of a rectangle by applying the formula	Procedures without Connections (requires using multiple steps and needs some level of connections with the concept of area model)
Task 6: Imagine you're planning a pizza party, and you want to make sure everyone gets a fair share of delicious pizza. You	Area of a rectangle Addition of whole	Representing given context using rectangles	We did not rate this task because it presents conceptually and

Output (Tasks)	Concepts	Procedures	Cognitive Demand
have two rectangular pizzas, Pizza A and Pizza B, each divided into a grid of squares. Pizza A is divided into 6 rows and 4 columns, and Pizza B is divided into 3 rows and 5 columns. Draw a diagram to represent Pizza A and Pizza B as rectangles divided into squares. Label the rows and columns to indicate the number of rows and columns in each pizza. Each square in Pizza A represents 1/24 of the pizza, and each square in Pizza B represents 1/15 of the pizza. Calculate the total number of squares in each pizza. Now, calculate the area of each pizza by finding the product of the number of rows and the number of columns. Calculate the total area of Pizza A and Pizza B in terms of fractions. How many "pizza units" are there in each pizza? Your task is to find out how much pizza you have in total for the party. To do this, you need to multiply the area of Pizza A by the area of Pizza B. Use the area model for fraction multiplication to find the product. Finally, express the total area of pizza for the party as a simplified fraction. What is the total area in terms of fractions, and how many pizza units do you have in all? This task encourages students to use the area model for fraction multiplication to find the total amount of pizza for the party,	numbers This task presents conceptually incorrect information by asking to multiply the areas of two rectangles to find the total area.	Calculating the area of given rectangles Adding those areas This task presents procedurally incorrect information by asking to calculate the area of each rectangular pizza in terms of fractions, which is not possible because its length and width are whole numbers.	procedurally incorrect information

Output (Tasks)	Concepts	Procedures	Cognitive Demand
making it an engaging and practical application of mathematical concepts.			
Task 7: Using an area model, determine the product of 5/6 and 7/8. Visualize and represent the multiplication of these fractions as a complex area model, without providing a numerical answer.	Area of Rectangle Fraction Multiplication	Using a formula to calculate the area of the given rectangle. Representing area using a rectangle	Procedures without Connections (requires multiple presentations and how dimensions of rectangles are represented in rectangles)
Task 8: You are designing a unique pattern for a quilt. The pattern consists of several rectangular patches, each with its own fractional dimensions. The first patch is 2/5 meters long and 1/3 meters wide, the second is 3/7 meters long and 2/5 meters wide, and the third is 1/4 meters long and 3/8 meters wide. Your goal is to determine the combined area of all these patches, which will help you calculate the amount of fabric you need to create the quilt. You must use your understanding of fraction multiplication and visualization to calculate the total area accurately, as well as explain how you arrived at your solution.	Area of a rectangle Fraction multiplication Adding fractions	Calculate areas of three rectangles separately Adding fractions	Procedures with Connections (requires to multiple representations and justify reasoning)
Task 9: You have a rectangle with the following properties: The length is divided into 5 equal sections. The width is divided into 4 equal sections. Inside this rectangle, there is a square with sides representing ⅖ of the length of the rectangle and ¼ of the width. Calculate the area of this	Area of a rectangle and square Fraction multiplication Subtraction of fractions	Find the length and width of the square. Calculate the area of a square Calculate the proportion of the	We did not classify this task as it presents conceptually vague and/or inaccurate information.

Output (Tasks)	Concepts	Procedures	Cognitive Demand
square. Question: What is the area of the square inside the rectangle, given that the side length of the square represents ⅖ of the length of the rectangle and ¼ of the width? To solve this task, the student will need to: Calculate the area of the entire rectangle, which can be considered as 1 whole unit. Determine the side length of the square within the rectangle, which requires multiplying fractions to find ⅖ of the length and ¼ of the width. Calculate the area of the square using the side length. Express the area of the square as a fraction of the area of the entire rectangle.	This task presents conceptually vague and/or incorrect information because there is no clear indication of the unit of length and width of the original rectangle, moreover, if the dimension of each section is the same, the inside rectangle cannot be a square	given square in relation to the area of the given rectangle	
Task 10: You have a large rectangle that represents one whole unit. Inside this large rectangle, there are two smaller rectangles. The first smaller rectangle occupies ⅜ of the length and ⅖ of the width of the large rectangle. The second smaller rectangle occupies ⅚ of the length and ¼ of the width of the large rectangle. Calculate the area of the space that is not covered by the smaller rectangles within the large rectangle, expressed as a fraction of the large rectangle's area.	Area of rectangle Fraction multiplication Addition of fractions Subtraction of fractions	Calculate areas of two small rectangles Add those areas of two rectangles. Subtract areas of the rectangle	Procedures with connections (requires calculating area and connecting it with the concept of rectangular area)

Our analysis (see Table 3) indicated that even though the tasks contained many words, *all* the tasks focused on procedural algorithms, and *only a few* required students to demonstrate their conceptual understanding. No tasks required students to demonstrate their problem-solving skills. We also found some tasks were procedurally or conceptually inaccurate or vague. Additionally, although some tasks asked students *what* to do by telling them to

multiply the length by the width of a rectangle, none of the tasks connected *what* students were doing when they multiplied the length by the width to *how* to model it on a rectangle by partitioning the length into equal segments according to the denominator and shade the segments according to the numerator or *why* they were doing it by explaining that the overlapping shaded region represents the product of the fractions. We elaborate on these findings in the subsections below. We have numbered the tasks as Task 1, Task 2, etc. in Table 3 for our and readers' convenience.

Nature of ChatGPT-Generated Tasks: Similar Tasks on Multiplying Fractions

As we described in our methods section, we were interested in generating cognitively demanding fraction multiplication tasks that use the area model approach. It was alarming that none of the tasks generated by ChatGPT required students to demonstrate their conceptual understanding of the area model approach. Despite our ten different variations of inputs, most of the tasks were similar. Almost all involved students perform the procedure of multiplying the length by the width to determine the area. Some tasks were multi-step problems requiring students to add or subtract fractions after multiplying the fractional length and width to determine the area. For example, in Task 3 students are asked to calculate areas of three rectangles and add those areas to determine the combined area. Similarly, Task 8 requires students to calculate areas of three rectangles separately and add those fractions. It appears that ChatGPT made the reasonable assumption that students should have prior knowledge related to fraction addition when they engage in fraction multiplication. Finally, some of the tasks involved whole numbers multiplication instead of fractions. As an example, in Task 6, students are required to multiply 6 and 4 to find the area of a given rectangle. This output suggests that ChatGPT may sometimes produce irrelevant tasks.

While a few tasks instructed students to calculate the area of a rectangle by multiplying its length and width, none of the tasks required students to explain *how* or *why* they were partitioning the length into equal segments based on the denominator, shading these segments according to the numerator, or the overlapping shaded region illustrated the product of the fractions. Although our inputs for Tasks 9 and 10 specifically prompted ChatGPT to generate a partitioning task, the output tasks focused on multiplying the length and the width and were absent of any partitioning. Based on these findings we wondered if ChatGPT did not understand the meaning of partitioning fractions. We tested this hypothesis by asking ChatGPT the definition of partitioning in the context of fractions. ChatGPT's output of "Partitioning in fractions involves dividing a whole or a region into smaller, equal parts, and it's a fundamental concept in understanding and working with fractions" (OpenAI, 2023) suggested that it does understand the meaning of partitioning fractions. We speculate that ChatGPT might not have incorporated this definition into the tasks it generated. Further investigation would be needed to determine if the pattern we observed is a common occurrence.

Cognitive Demand of ChatGPT-Generated Tasks: Focus on Procedures and Absence of Doing Mathematics Tasks

Most of the ChatGPT-generated tasks focused on procedures without connections to the area model. However, some connected the procedures to the concept of the area of rectangles. Almost all the tasks provided the

dimensions of the length and the width and asked students to determine the area. For example, in Task 4, students are prompted to calculate the area of three rectangular sections of glass. Even though Tasks 4 and 9 are multi-step problems, the tasks focus on the what and do not require students to explain the why or how behind the algorithm. Even when we explicitly asked for "rich mathematical tasks," "problem-solving tasks," or "cognitively demanding tasks," (e.g., Tasks 6 and 7), ChatGPT continued to generate tasks focused on procedures. Our findings suggest that ChatGPT may not be capable of generating conceptual tasks that move beyond the procedures to focus on the why and how of the procedures. The results suggest that ChatGPT interpreted "cognitively demanding" to mean tasks that involved multiple procedural steps. However, we argue that multiple procedural steps are not considered cognitively demanding. Further investigation would be needed to test this hypothesis.

Other Characteristics of Tasks: Real-world Context, Mathematical Concepts, and Procedures

ChatGPT appears to be good at connecting tasks to real-world contexts. As seen in Table 3, the connections included the areas of a variety of rectangular surfaces including kitchen floors, gardens, parks, and pizzas. Connecting tasks to real-world contexts may motivate students to engage in the tasks. However, these contexts might not be relatable to every student's culture. Investigating the extent to which these tasks represent diverse cultural contexts would be a separate study.

Most of the ChatGPT-generated tasks presented correct mathematical concepts and procedures and could be used for teaching. Almost all the tasks asked for procedures that we could follow and understand. However, some tasks misrepresented mathematical concepts, and/or they were presented very vaguely that we could not follow. As an example, we found several issues in Task 6. First, it asks to "calculate the area of each pizza by finding the product of the number of rows and the number of columns." This part does not involve the concept of fraction multiplication. The second part, it asks to "calculate the total area of Pizza A and Pizza B in terms of fractions. How many 'pizza units' are there in each pizza." This part does not make sense because the area of Pizza A and Pizza B cannot be in the fractions as they both have whole number length and width. In the third part, it asks "to find out how much pizza you have in total for the party. To do this, you need to multiply the area of Pizza A by the area of Pizza B. Use the area model for fraction multiplication to find the product." This part presents conceptually inaccurate information as it asks to multiply the areas of two rectangular pizzas to find the total area instead of adding two areas.

We also found similar inaccuracy and vagueness in Task 9. The output was "You have a rectangle with the following properties: the length is divided into 5 equal sections. The width is divided into 4 equal sections. Inside this rectangle, there is a square with sides representing ⅖ of the length of the rectangle and ¼ of the width." Based on this information, the square has a length ⅖ of 5 sections of length and a width ¼ of 4 sections of width. It is not clear if the size of each section in the partitioning along length and width were equal. If so, the inside quadrilateral cannot be a square. If the length and width sections are not equal, the inside quadrilateral could be a square, but we do not know the unit of length and width. Given that the unit of length and width of the inside quadrilateral are in equal length and width sections, respectively, we cannot calculate area unless and until those length and width sections of the large rectangle are equal. Either way, this task presents conceptually inaccurate,

misleading, and vague information.

Discussion and Limitations

In this section, we first discuss our findings in light of the potential limitations, weaknesses, and threats raised by researchers in the field (e.g., Farrokhnia et al., 2023). We found some of our findings aligned with the limitations and threats suggested in prior studies, while other findings did not align with the suggested limitations and threats. Our findings support, with specific examples from mathematics education, the claim that ChatGPT-generated outputs often lack higher-order thinking and deep understanding. Even though there was some improvement in the output when we provided more specific prompts (which was suggested as a strength in the prior literature), we did not obtain the desired cognitively demanding tasks with the specific prompts that included phrases such as "cognitively demanding," "rich tasks," and other similar phrases. Thus, we argue that ChatGPT might not have the ability to self-improve with differentiation in prompts, specifically in the context of generating mathematical tasks. Based on these findings, we urge AI specialists to evaluate ChatGPT's ability to self-improve, specifically in the context of cognitively demanding mathematics education tasks that use the area model to represent fraction multiplication.

As we reported in the Findings section, the mathematical tasks generated by ChatGPT had some potential in that they were connected to the real world, involved multiple steps, and assessed students' procedural fluency. However, the outputs did not require students to demonstrate a conceptual understanding of the how or why behind the procedures. Moreover, several tasks were vague or inaccurate. Based on these findings we argue that although ChatGPT may be a helpful tool, it does not replace the need for a human to read the output and make edits. Moreover, the findings suggest that ChatGPT is better at generating procedural tasks and may be unable to produce conceptual and problem-solving tasks. This finding is consistent with Gattupalli and colleagues (2023), who found that problem-solving strategies generated by AI were mostly procedural.

As reported above, ChatGPT produced some inaccurate or vague tasks. Therefore, the user must carefully and critically examine the output. We found the incorrect outputs occurred when ChatGPT was attempting to generate multi-step tasks. This finding supports Shakarian and colleagues' (2023) findings that the chances of producing correct outputs decrease when ChatGPT is asked to generate multi-step outputs. Overall, our findings reinforce the conclusion that ChatGPT can never replace the human educator. Rather, it can be viewed as a tool that can be used as a thought-partner (Cooper, 2023; Jeon & Lee, 2023).

Our study has several limitations. First, we only focused on one topic area (i.e., fraction multiplication using the area model). Thus, we cannot generalize our findings as applicable to all the ChatGPT tasks. Second, the prompts and words used in the input were based on our own experiences. Thus, educators might be able to generate more relevant tasks on fraction multiplication (with the area model) using alternative words and phrases. Third, we are not certain whether the issues we found with the outputs should be attributed to our inputs, the AI algorithm, the data the algorithm draws from, or some combination of these three factors. We acknowledge that the variations we wrote in our inputs might have limitations and alternative inputs may have resulted in outputs aligned with our

objectives. As MTEs, we do not know what the AI algorithm is or what data the algorithm draws from. Future research could explore if the algorithm asks for and is only able to output the most common tasks from all tasks published on the internet. If this is the case, it makes sense that ChatGPT was unable to produce a cognitively demanding task, because procedural approaches dominate most instructional materials (EdReports, 2023; The New Teacher Project, 2018). Fourth, we generated around 50 tasks and analyzed 10 distinct tasks in this study. Thus, it is possible to receive a different result from a quantitative study of a larger number of tasks generated through different prompts.

Conclusion

We, two MTEs engaged in a collaborative self-study aimed at assessing the variation in cognitive demand and evaluating the accuracy and relevance of ChatGPT-generated tasks about fraction multiplication using an area model. The findings of this study highlight the potential of ChatGPT in contributing to the fair distribution of curricular resources, albeit with certain limitations. ChatGPT's ability to generate reliable and challenging tasks proved inconsistent, raising concerns about its dependability. While it demonstrated utility in deriving procedural tasks, it appeared less effective in generating problem-solving tasks that require deeper cognitive engagement. Despite its potential to provide valuable curricular resources, ChatGPT fell short of delivering consistent and reliable results in our study. As educators, the challenge lies in the temptation to utilize ChatGPT's output without rigorous editing. However, our research underscores the importance of editing to enhance the accuracy, relevancy, and cognitive demand of the generated tasks. This underscores the need to remain critical and emphasizes the indispensability of human involvement in the task-generation process. Without human involvement, generating tasks on ChatGPT may not be distinguishable from an internet search or using unvetted platforms, such as Teachers Pay Teachers. We believe AI could produce more promising outputs if mathematics educators and AI experts partner in addressing how and why the current outputs are missing the mark. While ChatGPT may generate valid and reliable outputs to some inputs, it did not produce valid or reliable cognitively demanding fraction multiplication tasks using the area model. These findings prompt a reevaluation of the role of ChatGPT in mathematics education and call for a more judicious and critical approach to its utilization.

References

Beckmann, S. (2022a). *Mathematics for elementary and middle school teachers with activities (6th edition)*. Pearson.

Beckmann, S. (2022b). *Mathematics for elementary and middle school teachers, activities manual (6th edition)*. Pearson.

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

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119-156. https://www.jstor.org/stable/40539329

Butler, B. M., & Bullock, S. M. (2022). *Learning through Collaboration in Self-Study*. Springer Singapore.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*(3), 444-452. https://doi.org/10.1007/s10956-023-10039-y

Cope, L. (2015). Math manipulatives: Making the abstract tangible. *Delta Journal of Education*, *5*(1), 10-19. https://www.academia.edu/19886655/Math_Manipulatives_Making_the_Abstract_Tangible

Doherty, K., Baniahmadi, M., Sapkota, B.., Adjei, E. K., Brown, J. E., Newton, J. A., Wood, M. B., Olson, A. M., and Drake, C. (2022). Exploring dissonance and harmony in elementary mathematics teachers' curricular use, autonomy, decision-making, and coherence. In A. Lischka, J. Strayer, J. Lovett, R. Jones, & E. Dyer (Eds.), *Proceedings of the forty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 237-238). Nashville, TN.

Education Week (2023). *More Teachers are embracing ChatGPT, Students not so much.* https://www.edweek.org/technology/more-teachers-are-embracing-chatgpt-students-not-so-much/2023/07

EdReports (2023). *The state of the market.* https://www.edreports.org/resources/article/state-of-the-instructional-materials-market-use-of-aligned-materials-in-2022

Elliott, R., & Timulak, L (2005). Descriptive and interpretive approaches to qualitative research. In J. Miles, & P. Gilbert (Eds), *A handbook of research methods for clinical and health psychology* (pp. 147–159). Oxford University Press.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1-15. https://doi.org/10.1080/14703297.2023.2195846

Gattupalli, S., Maloy, R. W., & Edwards, S. (2023). *Comparing teacher-written and AI-generated math problem solving strategies for elementary school students: Implications for classroom learning.* University of Massachusetts Amherst. https://doi.org/10.7275/8sgx-xj08

Grassini, S. (2023). Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), 1-13. https://doi.org/10.3390/educsci13070692

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, *30*(2), 393-425. Advanced online publication. https://doi.org/10.3102/00028312030002393

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies,* 1-20. https://doi.org/10.1007/s10639-023-11834-1

Johnson, K. R., & Fonbuena, L. C. (2023). Positionalities in Our Practices and Papers. *Mathematics Teacher Educator*, *11*(3), 145-154. https://doi.org/10.5951/MTE.2023.0007

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Kasneci, E., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A. . . . Kasneci, G. (2023). ChatGPT for good? on opportunities and challenges of large language models for education. *Learning and Individual Differences, 103,* 1-9. https://doi.org/10.1016/j.lindif.2023.102274

Kosnik, C., Lassonde, C., & Galman, S. (2009). What does self-study research offer teacher educators? In C.

Lassonde & S. Galman (Eds), *Self-study research methodologies for teacher educators* (pp. 225-239). Brill.

Leinwand, S., & Wiggins, G. (1991). Criteria for performance tasks. In J. K. Stenmark (Ed.), *Mathematics assessment: Myths, models, good questions, and practical suggestions* (pp.16–16). National Council of Teachers of Mathematics.

Li, Z., Niu, C., Meng, F., Feng, Y., Li, Q., & Zhou, J. (2019). Incremental transformer with deliberation decoder for document grounded conversations. *arXiv preprint arXiv:1907.08854*. https://arxiv.org/pdf/1907.08854.pdf

Liljedahl, P. (2020). *Building thinking classrooms in mathematics, grades K-12*. SAGE.

Maringe, F., Foskett, N., & Woodfield, S. (2013). Emerging internationalization models in an uneven global terrain: Findings from a global survey. *Compare: A Journal of Comparative and International Education, 43*(1), 9–36. https://doi.org/10.1080/03057925.2013.746548

Milner, H. R. (2007). Race, Culture, and Researcher Positionality: Working Through Dangers Seen, Unseen, and Unforeseen. *Educational Researcher*, *36*(7), 388-400. https://doi.org/10.3102/0013189X07309471

National Governors Association Center for Best Practices and the Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Authors.

Onal, S., & Kulavuz-Onal, D. (2023). A cross-disciplinary examination of the instructional uses of ChatGPT in higher education. *Journal of Educational Technology Systems*, 1-24. Advanced online publication. https://doi.org/10.1177/00472395231196532

OpenAI. (2023). *ChatGPT* [Large language model]. https://chat.openai.com/

Qadir, J. (2022). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. *TechRxiv.* https://doi.org/10.36227/techrxiv.21789434.v1

Saldaña, J. (2016). *The coding manual for qualitative researchers*. SAGE.

Sapkota, B. (2022). Preservice teachers' conceptualizations of mathematical tasks. *The Mathematics Educator, 30*(2), 3-32. https://files.eric.ed.gov/fulltext/EJ1339807.pdf

Sapkota, B., Xuwei, L., Sapkota, M., Akarsu, M., Deogratias, E., Fauber, D., Mbewe, R., Mumba., F., Panthi, R., Newton, J., & Phillion, J. (to be Published in April 2024). Exploring social justice issues through cross cultural lenses, In A. Slapac & C. Huertas-Abril (Eds.), *Encouraging transnational learning through telecollaboration in global teacher education, --, --.* IGI.

Shakarian, P., Koyyalamudi, A., Ngu, N., & Mareedu, L. (2023). An Independent evaluation of ChatGPT on mathematical word problems (MWP), in A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle, & F. van Harmelen (Eds.), *Proceedings of the AAAI 2023 spring symposium on challenges requiring the combination of machine learning and knowledge engineering* (AAAI-MAKE 2023), San Francisco Airport.

Smith, J. A., & Osborn, M. (2008). Interpretative phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (2nd ed., pp. 53–80). SAGE.

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2000). *Implementing standards-based math instruction: A casebook for professional development.* Teachers College Press.

The New Teacher Project (2018). *The opportunity myth: What students can show us about how school is letting them down—and how to fix it.* https://opportunitymyth.tntp.org/

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education, 23*(1), 1-23. https://www.learntechlib.org/primary/p/222408/

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2023). *Elementary and middle school mathematics: Teaching developmentally* (11th edition). 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), 1-18. https://doi.org/10.30935/cedtech/13417

Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education*, *45*(5), 636-674. https://doi.org/10.5951/jresematheduc.45.5.0636

Yogeeswaran, K., Devos, T., & Nash, K. (2016). Understanding the nature, measurement, and utility of implicit intergroup biases. In C. Sibley & F. Barlow (Eds.), *The Cambridge Handbook of the Psychology of Prejudice* (pp. 241–266). Cambridge University Press. https://doi.org/10.1017/CBO9781316161579

Yu, H. (2023). Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, *14*, https://doi.org/10.3389/fpsyg.2023.1181712

Author Information

Bima Sapkota

https://orcid.org/0000-0003-4992-2666

The University of Texas Rio Grande Valley

Edinburg, Texas

USA

Contact e-mail: *bima.sapkota@utrgv.edu*

Liza Bondurant

https://orcid.org/0000-0001-8069-2382

Mississippi State University

Mississippi State, MS

USA