International Electronic Journal of Mathematics Education
2024, 19(1), em0766
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article

OPEN ACCESS

Pre-service teachers and ChatGPT in multistrategy problem-solving: Implications for mathematics teaching in primary schools

Seyum Getenet [1]*

[1] School of Education, University of Southern Queensland, Toowoomba, QLD, AUSTRALIA
***Corresponding Author:** Seyum.Getenet@usq.edu.au

Citation: Getenet, S. (2024). Pre-service teachers and ChatGPT in multistrategy problem-solving: Implications for mathematics teaching in primary schools. *International Electronic Journal of Mathematics Education, 19*(1), em0766. https://doi.org/10.29333/iejme/14141

ARTICLE INFO

Received: 24 Aug. 2023

Accepted: 16 Jan. 2024

ABSTRACT

This study compared the problem-solving abilities of ChatGPT and 58 pre-service teachers (PSTs) in solving a mathematical word problem using various strategies. PSTs were asked to solve a problem individually. Data was collected from PSTs' submitted assignments, and their problem-solving strategies were analyzed. ChatGPT was also given the same problem to solve with various prompts, and the correctness of its solutions and problem-solving strategies were assessed alongside those of PSTs. The results indicated that PSTs used diverse strategies and achieved accurate solutions, but not always relevant strategies to children's level of understanding. ChatGPT employed similar strategies to PSTs but mostly produced incorrect solutions, and its performance needed to be contextualized in the primary school context. The study highlights the potential of ChatGPT in mathematics teaching and informs teacher education programs about the possibility of using it in teaching problem-solving strategies.

Keywords: artificial intelligence-based systems, ChatGPT, mathematics education, pre-service teachers, primary school, problem-solving strategies

INTRODUCTION

Problem-solving skill is an important competency that children need to develop across a range of contexts to facilitate their understanding, connecting, and using mathematical concepts. As a result, teachers are encouraged to develop children's problem-solving skills during their early years of children's education (Szabo et al., 2020). With the advancement of technology, artificial intelligence (AI) tools have emerged as a potential for assisting children to learn and teachers to teach problem-solving skills. AI-based systems have been used to support learning and teaching processes, especially in mathematics education. Studies have shown that using AI-based systems in mathematics education increased student learning outcomes, including developing problem skills (Hwang & Tu, 2021; Van Vaerenbergh & Pérez-Suay, 2022). For example, a study by Van Vaerenbergh and Pérez-Suay (2022) found that using an AI-based system in teaching algebraic equations improved students' performance and problem-solving skills. Another study by Rosé et al. (2019) reported that an AI-based system improved students' learning outcomes in geometry. Moreover, using AI-based systems in mathematics education has also enhanced teachers' pedagogical content knowledge. A study by Csíkos and Szitányi (2020) explored the impact of an AI-based system on teachers' mathematical knowledge and pedagogical practices. The findings indicated that the AI-based system enhanced teachers' content knowledge and improved their ability to identify and address children's misconceptions. Additionally, pre-service teachers (PSTs) have been found to have positive attitudes towards integrating AI-based systems in teaching. In this regard, a study by Lee and Yeo (2022) found that PSTs perceived using AI-based systems in mathematics education as effective in supporting students' learning and promoting their engagement in learning.

Even though multiple advantages have been claimed to using AI in mathematics education, some challenges associated with using AI-based systems have been identified (Hwang & Tu, 2021; Lee & Yeo, 2022). For example, teachers' lack of appropriate training and support in using AI-based systems can hinder their effective integration into teaching (Chen et al., 2020). Furthermore, the potential overreliance on AI-based systems, including ChatGPT in problem-solving, can also negatively affect children's critical thinking and problem-solving skills (Lee, 2023). Hence, the challenges associated with integrating AI into teaching practices need to be addressed to ensure their effective use in learning and teaching.

Recently, ChatGPT has gained significant attention and generated widespread dialogue, reaching 100 million users just two months after its launch (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2023). The rapid evolution of technology, such as ChatGPT, and changes in the educational environment (Wardat et al., 2023a) are compelling teachers globally at all levels to reassess their teaching and learning approaches, especially regarding the use of AI (Strzelecki, 2023). ChatGPT utilizes deep learning algorithms trained on a vast dataset of internet text to generate natural language responses to text-based prompts. While there have been a few accounts of ChatGPT being used in education, particularly in mathematics teaching, large-scale studies are yet to be conducted. A recent interview study results by Wardat et al. (2023b) showed that ChatGPT provides immediate assistance and personalised support in mathematics, but concerns exist about its accuracy, with some instances of incorrect or incomplete solutions reported. The integration of ChatGPT into specific concepts of mathematics teaching and its effectiveness in problem-solving remains unclear. This study aims to contribute to this ongoing conversation by examining the effectiveness of ChatGPT in supporting problem-solving skills among primary school pre-service teachers (PSTs) in mathematics education. Specifically, the research aims to compare the problem-solving strategies employed by ChatGPT and 58 PSTs in solving a mathematical word problem.

Research Questions

1. What problem-solving strategies do PSTs employ in solving a mathematical problem?
2. How does ChatGPT's problem-solving approach compare to that of primary school PSTs for a single mathematical problem?

The results of this study might provide valuable insights into the effectiveness of ChatGPT in supporting primary school mathematics education. Moreover, the study's findings might contribute to developing teaching strategies that incorporate AI-based language models in primary school mathematics education. The findings can inform the development of effective professional development programs for PSTs, ultimately leading to improved mathematics learning outcomes for primary school children.

LITERATURE REVIEW

This literature review examines problem-solving in primary schools, strategies for solving problems and the existing research on the use of AI in mathematics problem-solving in primary schools.

Problem-Solving in Mathematics Education

Problem-solving is a crucial aspect of mathematics education as it involves applying mathematical knowledge and concepts to real-world contexts. It requires a combination of skills, including logical thinking, analytical reasoning, and creative problem-solving, to solve complex problems. As children progress through their mathematics education, they encounter increasingly complex and abstract problems that require higher-level thinking and problem-solving skills. Bourke and Stacey (1988) define a mathematical problem as a question that is not reliant on specific syllabus content and is new enough for the student to preclude a solution by a previously known method. This definition highlights the importance of problem-solving in mathematics education, as it encourages children to think critically and creatively to develop innovative solutions to unfamiliar problems. Word problems, in particular, are often challenging for children as they require the application of multiple skills, including reading comprehension, mathematical reasoning, and problem-solving strategies (Verschaffel et al., 2020). Therefore, teaching mathematics through problem-solving is crucial as it provides children with the opportunity to develop and practice a wide range of mathematical skills.

To effectively teach mathematics through problem-solving, teachers need to research, plan, and coordinate to ensure that children encounter a variety of problems and strategies (Reys et al., 2021). This approach enables children to better understand mathematical concepts and processes while developing their problem-solving skills. Research on problem characteristics, successful and unsuccessful problem-solvers, teaching strategies, and classroom conditions has identified broader skills required for effective problem-solving (Schoenfeld, 2016). These skills include identifying relevant information, planning and executing a problem-solving strategy, monitoring progress, and evaluating solutions.

Given its importance, the Australian National Curriculum recognizes problem-solving as integral to the process of doing mathematics, interweaving with the other proficiency strands of the Australian curriculum: mathematics, namely understanding, fluency, and reasoning (The Australian Curriculum, Assessment and Reporting Authority [ACARA], 2023). As children gain experience in problem-solving, they learn mathematical content as well as develop an awareness of and expertise in the processes of doing mathematics. For example, from foundation to year two, children solve problems using mathematics to represent unfamiliar or meaningful situations. Similarly, in years three to six, children solve problems using mathematics to represent unfamiliar or meaningful situations and plan their approaches. By providing opportunities for students to engage in problem-solving, teachers can help children develop the skills and knowledge required for success in mathematics and beyond.

Problem-Solving Strategies

Problem-solving in mathematics is a complex process that requires children to go through several stages, including problem representation, planning, execution, and evaluation (Schoenfeld, 2016). To succeed, children need a range of strategies they feel comfortable using. Having a wide range of strategies to draw from allows children to approach a variety of problems and helps them develop confidence and resilience as they find a path to a solution (Reys et al., 2021). However, research indicates that teachers and PSTs often lack the necessary problem-solving skills and strategies to teach mathematics effectively (Krawec & Montague, 2014). To address this issue, researchers have suggested various strategies for solving mathematical problems, such as enhancing teachers' skills to use various strategies to solve a single problem (Duru et al., 2011; Reys et al., 2021). Reys et al. (2021) identified seven strategies that can help children solve a wide variety of problems. These are: act it out, make a drawing or diagram, look for a pattern, construct a table, guess, check, and improve, work backward, and solve a similar but simpler problem. Duru et al. (2011) also identified six strategies PSTs used to solve problem-solving tasks, as shown in **Table 1**.

Table 1. Categories of strategies for problem solving by pre-service teachers (Duru et al., 2011, p. 3465)

Number	Strategy category	Definition
1	Arithmetic	PSTs write down a mathematical statement involving one or more operations on numbers given in problem.
2	Algebraic	PSTs select one or more unknowns as variables and translate the verbal problem statement into an algebraic statement or equation. The resulting equation is then solved to obtain the answer.
3	Model	A common approach is to transform the word problem into a visual representation such as a picture, figure, diagram, or graph, which serves as a model for solving the problem.
4	Guess & check	PSTs adopt a guessing approach and subsequently verify whether it leads to a solution.
5	Find a pattern	In order to solve the problem, PSTs search for patterns within the provided data and identify repeated numbers, relationships, or regularly recurring sequences.
6	Model & algebraic	Initially transformed the word problem into a model and subsequently formulated equations based on the model, and ultimately solved these equations to obtain the solution.

Table 1 categorizes different problem-solving strategies employed by PSTs. The strategies include arithmetic, algebraic, modelling, guess-and-check, finding patterns, and a combination of modelling and algebraic approaches. Additionally, Böckmann and Schukajlow (2018) and Hegarty and Kozhevnikov (1999) suggested that graphical representations help in the initial phases of the solution process of a word problem, providing an additional relevant source of information to solve the given problem.

Having a range of problem-solving strategies is essential for children to become successful in solving mathematical problems. Teachers need to provide opportunities for children to experience a variety of problems and strategies to help them develop their problem-solving skills. Similarly, PSTs must have the necessary problem-solving skills and strategies to teach mathematics effectively. By implementing effective teaching strategies and providing opportunities for children to solve problems, teachers can help children develop their mathematical content knowledge and their expertise in the processes of doing mathematics.

Teachers play a crucial role in supporting children's problem-solving skills by providing a variety of strategies and explicit instruction (Ramnarain, 2014; Reys et al., 2021). Csíkos and Szitányi's (2020) study investigated the pedagogical content knowledge of Hungarian pre-service and in-service teachers regarding teaching word problem-solving strategies. The results suggested that explicit teaching of a step-by-step algorithm for problem-solving is feasible and desirable, even as early as the first grade. Teachers must have a repertoire of problem-solving strategies and knowledge to support children in using multiple strategies (Reys et al., 2021). Encouraging children to generate their own ideas about how to approach new problems is important so that they are not limited to using only one set of strategies. When children successfully use an alternative strategy, it is essential to encourage them to share their strategy with others and motivate them to develop their own problem-solving strategies and share them with others (Reys et al., 2021). This way, children will learn many ways to solve a problem, which will help them gain more insight into problem-solving.

Role of Artificial Intelligence in Mathematics Problem-Solving

The connection between AI and mathematics teaching has been an area of interest for researchers for several decades. The discussion about this relationship dates back to at least the 1980s, as evidenced by scholarly literature (Schoenfeld, 1985). This highlights that the integration of AI into mathematics teaching has been a topic of interest for many years, and researchers have explored various ways in which AI can be used to enhance the teaching and learning of mathematics. However, despite the capabilities of AI and its discussion in mathematics education, its application in using school mathematics has been relatively limited to robotics and intelligent tutoring (Lee & Yeo, 2022). In their findings, bin Mohamed et al. (2022) indicated that robotics was the most popular approach to AI in mathematics education, among other approaches, and to enhance collaboration in primary schools. Other AI-based educational technologies, such as intelligent tutoring systems and educational games, have also been used in mathematics teaching to provide students with individualized feedback and support that can enhance their problem-solving skills (Mills et al., 2013). A similar study by Qiu et al. (2022) showed the role of AI in mathematics development and the education of intelligent tutoring systems for Internet educational experiences and concluded that with the implementation of Internet education, students' learning capacities improved tremendously. In general, several studies have shown that AI has the potential to enhance problem-solving skills in mathematics by providing students with personalized learning experiences, adaptive feedback, and opportunities for collaborative learning (Lee et al., 2022). For instance, Lee et al. (2022) developed an AI-based collaborative learning system that enabled students to solve problems together and receive immediate feedback on their problem-solving strategies. The results showed that the AI-based collaborative learning system improved students' problem-solving abilities and promoted social learning by providing them with real-time feedback.

A recent study by Pelton and Pelton (2023) and Shakarian et al. (2023) investigated ChatGPT's performance in mathematics and supporting teacher education in mathematics. Their findings suggested that ChatGPT's performance is highly influenced by the requirement to show its work. ChatGPT failed 20% of the time when it had to show its work compared to 84% when it did not. Additionally, the studies found that the number of unknowns and the number of operations involved in mathematics word problems also significantly affect ChatGPT's performance. The study observed that the probability of failure increases linearly with the number of addition and subtraction operations, across all experiments. Another study by Frieder et al. (2023) evaluated ChatGPT's mathematical abilities by testing it on publicly available datasets as well as hand-crafted ones. They compared ChatGPT's performance with other models trained on a mathematical corpus, such as Minerva. The results indicated that ChatGPT's mathematical capabilities are significantly below those of an average mathematics graduate student. The study found that while ChatGPT often uses the query correctly as input, it frequently fails to provide correct solutions. It is important to note that ChatGPT's training data only goes up to 2021, and it has no knowledge about any developments or creations that have occurred since then. As a result, the validity of the information provided by ChatGPT should be verified and checked.

In light of the potential risks associated with using ChatGPT for teaching, UNESCO (2023) created a flow chart to guide its safe use. The flow chart outlines the circumstances under which ChatGPT may be safely used, as well as the limitations of ChatGPT's reliability and verification of information provided. The flow chart is illustrated in **Figure 1** (UNESCO, 2023).

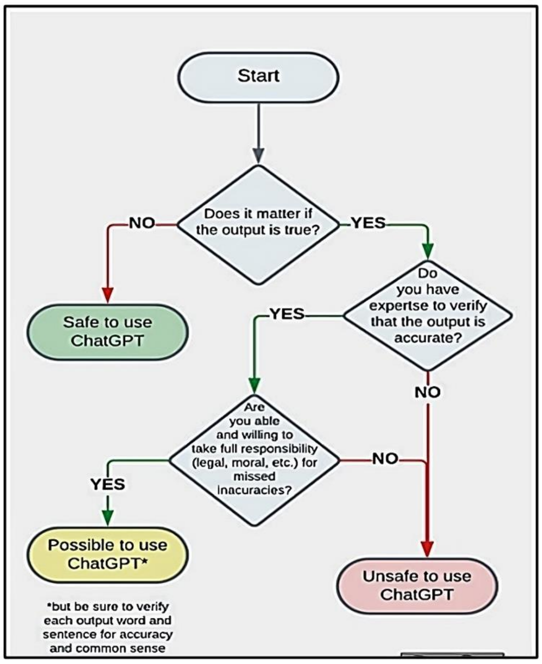


Figure 1. A flow chart to use ChatGPT safely (UNESCO, 2023, p. 6)

Teachers can also benefit from utilizing this flow chart to facilitate and encourage children's problem-solving skills. It is possible that the strategy employed to arrive at a solution may not be appropriate for the children's level of understanding, indicating that the information provided by ChatGPT may be unreliable. Therefore, it is essential for teachers to verify the accuracy of ChatGPT-generated information before incorporating it into their teaching (UNESCO, 2023).

METHOD OF THE STUDY

The study was conducted in two phases. First, PSTs were asked to solve the mathematical word problem individually using various strategies such as visualization, table, and guess, check, and refine techniques suitable for primary school children aged between five and 11. In the second phase, the same mathematical problem was presented to ChatGPT, with prompts specifically tailored to reflect the context of primary school children.

The Study Context

This study was conducted within the school of education at an Australian regional university that prepares early childhood, primary, and secondary school teachers. The study focused on PSTs who were in the final year of their degree program and were being prepared to teach children from foundation (with an average age of five years) to year 6 (with an average age of 11 years). PSTs enrolled in the primary specialization were required to complete two core mathematics curriculum and pedagogy courses as part of their degree program. One of these courses emphasized the use of problem-solving strategies in teaching mathematics. In this course assignment, PSTs were tasked with solving various mathematical problems using multiple strategies within the primary school context. For this study, a single mathematical problem was selected from the primary school mathematics curriculum. Participants were instructed to use strategies relevant to the content and children's context, particularly for upper primary year levels (age nine-age 11). The problem selected for this study was, as follows:

> In the zoo are a total of 29 lions and water birds, all in perfect condition. If these have a total of 92 legs, how many of each are there?

Participants & Data

In this study, the participants consisted of 58 PSTs who were enrolled in a mathematics education course. The course originally had 59 PSTs enrolled. Of these, 58 agreed to participate in the study, while one PST opted out of having their work included in the study. The data were collected from the 58 PSTs' assignment submissions of the course. All the submissions were from one cohort of PSTs. The relevant University and School authorities provided ethical protocols to collect PSTs' archived assignment data with approval number ETH2023-0208. The participants included PSTs with low socio-economic status (n=6), Aboriginal and Torres Strait Islanders (n=2), PSTs with English as an additional language (n=4) and PSTs with disability (n=9).

Data Analysis

The data analysis process involved two distinct stages. Initially, the assignment submissions from PSTs were evaluated to ensure that they met the course requirements, and notes were taken for the second analysis stage. During the first stage, the assignments were appraised based on the criteria specified in the marking rubric. This rubric, which comprises marking criteria and grading standards, requires using relevant strategies to solve a problem and describing problem-solving strategies, among other things. Following this assessment, the assignment submissions were further scrutinized to identify the strategies employed by each PST to solve the problem, and these strategies were subsequently grouped according to common themes. A list of strategies reviewed in the literature (such as **Table 1**) guided this process. In the second phase of the study, ChatGPT was presented with the same problem and tasked with solving it using various prompts followed by the problem. The author carefully designed prompts for ChatGPT, considering various contextual factors. These prompts were tailored to reflect primary school mathematics problems, incorporating the use of diverse resources like manipulative tables and a range of problem-solving strategies. This design of the prompts was informed by the problem-solving methods employed by PSTs in the study. The aim of considering these contexts to design the prompts was to facilitate a comparison between the strategies used by ChatGPT and those employed by PSTs. The prompts given and the number of strategies employed by ChatGPT to solve the problem are presented in **Table 2**.

Table 2. Prompt & number of strategies generated

Prompt	Number of strategies
Solve the following problem in multiple ways for primary school children	4
Solve the following problem by using manipulatives	1
Solve the following problem using pictures for primary school children	1
Solve the following problem using a table for primary school children	1

The experiment for this study was conducted following the announcement of ChatGPT 3.5 and its improved performance on mathematical tasks on January 30, 2023. The final experiment was done on 2 May 2023. This version (3.5) was selected for two reasons. Firstly, it was the most recent version available for free. Secondly, the updated mathematics performance was claimed to be more efficient and reliable, making it highly probable for teachers to incorporate this free version into their teaching. The experiment involved five trials, with ChatGPT being required to generate solutions using four given prompts. The data collected from both ChatGPT and PSTs working on the problem were then compared. The performance of both ChatGPT and PSTs was evaluated based on the accuracy of their solutions and problem-solving strategies. This was done to determine how various problem-solving strategies were used to arrive at the correct solution.

RESULTS

The performance of PSTs and ChatGPT in solving mathematical problems are presented below in relation to strategies used to solve the problem and the correctness of the solution.

Problem-Solving Strategies by Pre-Service Teachers

The solutions provided by PSTs were all correct except for a few PSTs not using a strategy to the level of children in the primary school context, including using algebraic equations (n=5). According to the national Australian curriculum v9.0, children in primary school, for example, year 6, should be able to find unknown values in numerical equations involving brackets and combinations of arithmetic operations, using the properties of numbers and operations but not solving equations (ACARA, 2023).

In relation to the strategies used, PSTs used a variety of problem-solving strategies, including creating tables, guess, check and refine, drawing diagrams, and algebraic equations. The summary of the strategies used by PSTs with their frequency (n) and description is shown in **Table 3**.

Table 3. Strategies used by PSTs with their description & frequency

Strategies	Description	n
Drawing or graphic representation	Using pictures or drawings to represent the problem visually	7
Algebra–equation	Using algebraic equations to solve the problem	5
Creating a table	Creating a table to organize and visualize information	7
Guess check & refine	Making guesses and checking them until a solution is found	33
Guess check & refine followed by an equation	Making guesses, checking them and using an equation to verify the solution	1
Logical reasoning	Using reasoning skills to solve the problem	3
Looking at patterns	Observing patterns to solve the problem	2

Table 3 shows that a considerable number of PSTs (n=33) employed the guess, check, and refine strategy to solve the problem. This strategy was utilized in different ways, as illustrated by PSTs' sample work presented in **Figure 2** and **Figure 3**, where they used various forms of the guess, check and refine strategy.

In **Figure 2**, it can be observed that PST took a more meaningful approach in using this strategy. PST has made multiple guesses and checks before arriving at the correct solution. On the other hand, PST in **Figure 3**, it appears that PSTs have a more structured approach, breaking down the problem into smaller, manageable steps. PST provided explicit descriptions of the steps, making it easier for others to follow and understand their problem-solving process.

29 = Lions + Water Birds
92= Total legs
Lions = 4 legs
Water Birds = 2 legs

Guess check and refine

Guess 1:
80 / 4 = 20
12/ 2 = 6

20+6= 26

26 does not equal 29 total legs.

Guess 2:
72/ 4 = 18
20/ 2 = 10

18+10 = 28
28 does not equal 29 total legs.

Guess 3
68/ 4 = 17
24/ 12= 12

17+12= 29

Therefore 17 lions and 12 waterbirds equals 92 legs.

Figure 2. A PST uses guess, check, & refine strategy with lengthy steps (Source: A PST)

When completing this problem, I followed the four-step process and used the guess, check and refine strategy that we learnt in our lecture. Firstly, I spilt the number of animals as close to half as I could to determine a starting off point for the problem:

15 lions = 60 legs
14 birds = 28
= 88 legs

From here I calculated the difference in legs for me to adjust my working.

92 – 88 = 4 legs
At this point, I knew that increasing by 1 lion (4 legs) would not give me the total legs I needed because I would need to minus 1 bird (2 legs), so I increased my lions by 2 to increase my total legs by 4.

17 lions = 68 legs
12 birds = 24 legs
= 92 legs

Figure 3. A PST uses guess, check, & refine strategy with shorter steps (Source: A PST)

Table 3 displays that PSTs utilized drawing or graphic representation and drawing a table as the second most employed problem-solving strategy. In each strategy, a total of 7 PSTs used in their work, revealing its popularity among the other strategies.

Figure 4 illustrates the work done by two PSTs, who utilized drawing or graphic representation as their problem-solving strategy. This strategy allowed them to visually represent the solution, breaking it down into smaller, more manageable parts. This problem-solving method can be beneficial for children who are visual learners, as it provides a clearer understanding of the problem and its components.

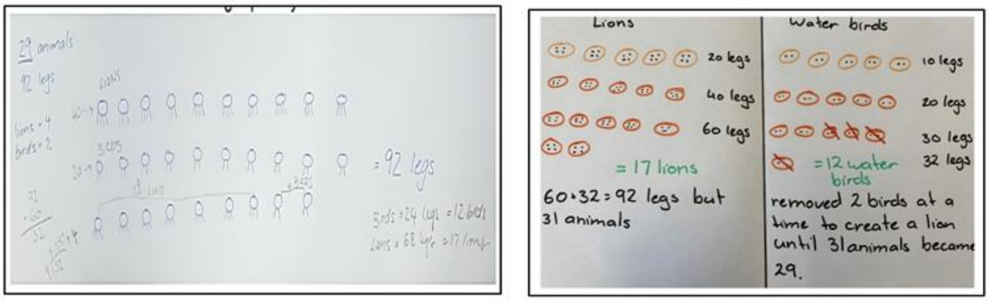


Figure 4. PSTs used a figure to solve problem (Source: A PST)

Seven PSTs utilized creating a table to solve the given problem. This approach was demonstrated in various formats, including a table coupled with pictorial representations (see **Figure 5**) and creating a table independently (see **Figure 6**). The creation of a table proved to be a valuable and effective tool for these PSTs, as it enabled them to organize and visualize complex information in a structured manner.

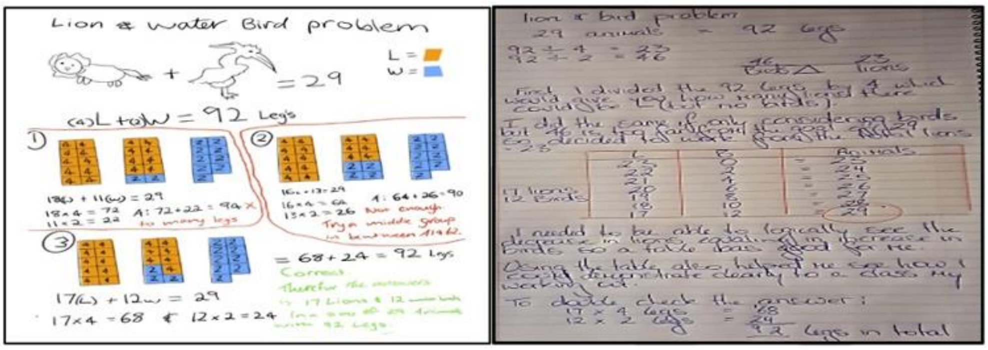


Figure 5. PSTs used graphs & a table to solve a problem with a description (Source: A PST)

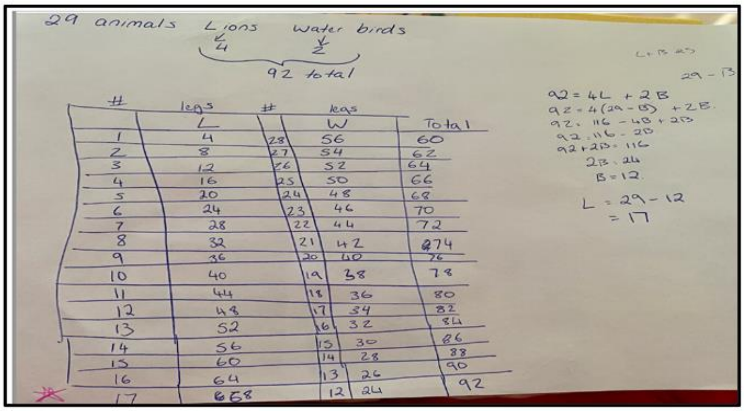


Figure 6. A PSTs used a table to solve problem (Source: A PST)

The results showed that three PSTs employed logical reasoning, a problem-solving approach closely resembling the guess, check, and refine strategy. The problem-solving technique of one of PSTs using logical reasoning is depicted in **Figure 7**.

> I decided to start by splitting the number in half. 25 can't split evenly so I just assumed 15 lions (60 legs) and 14 birds (28 legs), making 88 legs in total. I was 4 legs short so 2 lions short. I did the same maths to check this: 17 lions (68 legs) and 12 birds (24 legs) and got 92 legs

Figure 7. A PST using logical reasoning to solve problem (Source: A PST)

Problem-Solving Strategies by ChatGPT

While ChatGPT mainly relied on algebraic equations and formulas to derive correct solutions, its utilization of other problem-solving strategies did not consistently produce the correct answers. After conducting five trials for each prompt, the effectiveness of ChatGPT's strategies was assessed. **Table 4** summarizes the prompt, the strategy employed by ChatGPT, the corresponding description, and the number of correct responses obtained (n).

Table 4. Strategies used by ChatGPT with their description & frequency of correct answers (n)

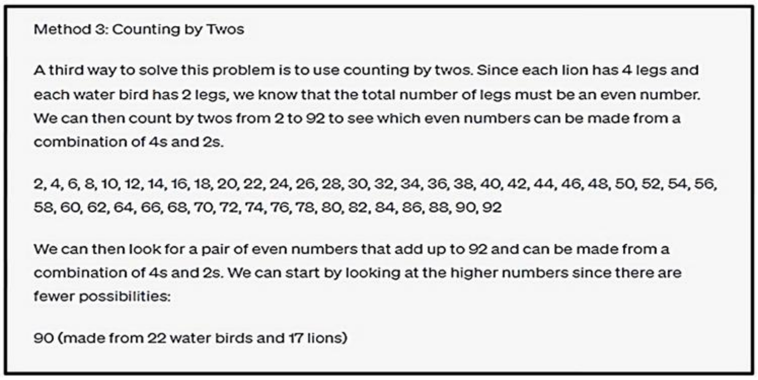
Prompt	Strategies	Description by ChatGPT	n
Solve following problem in multiple ways for primary school children	Guess & check	Guess number of lions & calculate number of water birds required to reach a total of 29 animals. Check if total number of legs adds up to 92. Repeat until a working combination is found.	0
	Algebra	Start by using variables to represent number of lions and water birds.	5
	Visualization	Create an illustration of the zoo animals using pictures, with four legs for each lion and two legs for each water bird. Count the total number of legs and determine the number of lions and water birds.	0
	Counting by twos	Count even numbers between two to 92 using combinations of 4s and 2s (from lions and water birds with four and two legs, respectively).	0
Solve following problem by using manipulatives for primary school children	Using manipulatives	Use blocks to represent animals and their legs; one color block for lions and another for water birds. A small block can represent each leg.	0
Solve following problem using pictures for primary school children	Pictures	To solve this problem using pictures, we can draw images of lions and water birds and count their legs to determine the number of each animal.	1
Solve following problem using a table for primary school children	Tables	A table can show the number of lions, water birds, and the total number of legs for each combination of animals.	1

As shown in **Table 4**, the first strategy used by ChatGPT was "guess and check", where it starts by guessing the number of lions and water birds and then checking whether the total number of legs adds up to the correct amount. But this strategy did not result in any correct answers for all five trials. An example generated by ChatGPT using guess and check strategy is shown in **Figure 8**.

> Method 1: Guess and Check
>
> One way to solve this problem is through guess and check. We can start by guessing how many lions there are and then figuring out how many water birds there must be to make the total number of animals add up to 29. Then we can check if the number of legs adds up to 92. We can repeat this process until we find a combination that works.
>
> Let's start by guessing that there are 20 lions. Then there must be 9 water birds to make a total of 29 animals. The number of legs for the lions would be 80 (20 lions x 4 legs each), and the number of legs for the water birds would be 36 (9 water birds x 2 legs each). In total, this would be 116 legs, which is too many. Let's try another guess.
>
> Let's guess that there are 15 lions. Then there must be 14 water birds to make a total of 29 animals. The number of legs for the lions would be 60 (15 lions x 4 legs each), and the number of legs for the water birds would be 28 (14 water birds x 2 legs each). In total, this would be 88 legs, which is the correct answer!
>
> Therefore, there are 15 lions and 14 water birds in the zoo.

Figure 8. Guess & check strategy generated by ChatGPT (OpenAI, 2023)

The second strategy generated by ChatGPT was “algebra”, which used variables to represent the number of lions and water birds and then solved for the variables using equations. Interestingly, this strategy resulted in five correct answers. The third strategy generated by counting in twos generated no correct response. The sample work is shown in **Figure 9**.


Method 3: Counting by Twos

A third way to solve this problem is to use counting by twos. Since each lion has 4 legs and each water bird has 2 legs, we know that the total number of legs must be an even number. We can then count by twos from 2 to 92 to see which even numbers can be made from a combination of 4s and 2s.

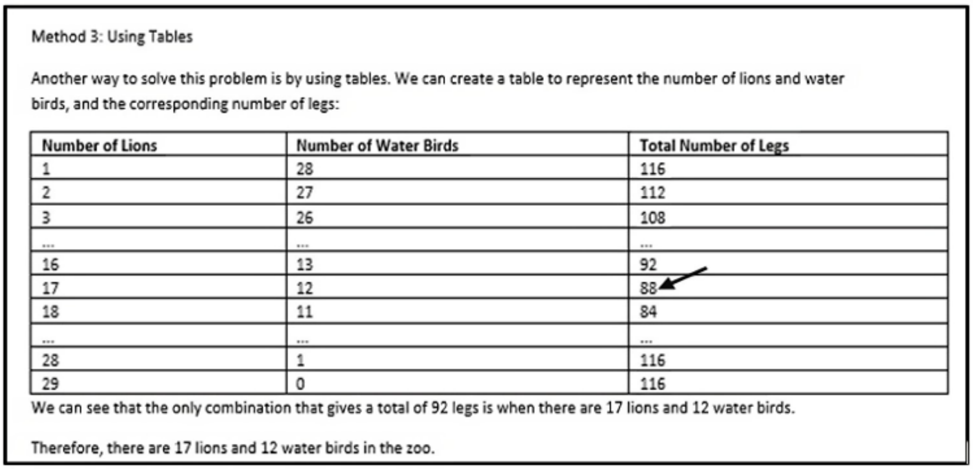
2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92

We can then look for a pair of even numbers that add up to 92 and can be made from a combination of 4s and 2s. We can start by looking at the higher numbers since there are fewer possibilities:

90 (made from 22 water birds and 17 lions)

Figure 9. Counting by twos strategy generated by ChatGPT (OpenAI, 2023)

ChatGPT produced two additional strategies, namely a table and a picture, each containing one correct response. Unfortunately, as shown in **Figure 10**, ChatGPT misinterpreted the table and incorrectly reported the total number of legs for 17 lions and 12 water birds as 88.


Method 3: Using Tables

Another way to solve this problem is by using tables. We can create a table to represent the number of lions and water birds, and the corresponding number of legs:

Number of Lions	Number of Water Birds	Total Number of Legs
1	28	116
2	27	112
3	26	108
...	...	...
16	13	92
17	12	88
18	11	84
...	...	...
28	1	116
29	0	116

We can see that the only combination that gives a total of 92 legs is when there are 17 lions and 12 water birds.

Therefore, there are 17 lions and 12 water birds in the zoo.

Figure 10. A table used by ChatGPT to solve problem (OpenAI, 2023)

While ChatGPT provided a correct response using pictures, the image generated was in JavaScript, as depicted in **Figure 11**, which may be challenging to comprehend and require an additional application to produce an accurate image.

Figure 11. A picture used by ChatGPT to generate a correct solution (OpenAI, 2023)

DISCUSSION

The results of the study demonstrated that PSTs used a variety of problem-solving strategies. These strategies are consistent with the literature on effective problem-solving approaches in mathematics education (Böckmann & Schukajlow, 2018; Duru et al., 2011; Reys et al., 2021; Verschaffel et al., 2020). As Csíkos and Szitányi’s (2020) suggested, PSTs might use their pedagogical content knowledge to generate multiple approaches to solving the problem.

Regarding PSTs’ problem-solving strategies, the most frequently used strategy was guess, check, and refine. This strategy was utilized differently, as demonstrated by PSTs’ sample work in **Figure 2** and **Figure 3**. This finding is consistent with previous research indicating that guess, check, and refine is a common problem-solving strategy children and teachers use (Duru et al., 2011; Verschaffel et al., 2020). This approach involves making an educated guess, checking the validity of the guess, and refining the solution until an acceptable solution is reached. PSTs employed this strategy in different ways, with some taking a more extensive approach with multiple guesses and checks before finally arriving at the correct solution, while others used a more structured approach, breaking down the problem into smaller, manageable steps. This strategy can be valuable in developing children's problem-solving skills, as it encourages them to make informed guesses and critically evaluate their solutions (Schoenfeld, 2016). Guess, check, and refine can effectively solve mathematical problems, particularly when children are unsure of how to proceed or do not have a specific method or formula to follow (Verschaffel et al., 2020). The preference for these strategies among PSTs could be attributed to their training in mathematics education and their exposure to pedagogical practices that emphasize visual representations and problem decomposition.

The use of drawing diagrams and creating tables was also prevalent among PSTs, indicating the value of visualization in problem-solving. PSTs' use of these strategies allowed them to visually represent and organize the problem, facilitating a step-by-step approach towards a plausible solution. Using tables allowed PSTs to break down the problem into manageable components, facilitating a step-by-step approach towards a plausible solution. Using diagrams and tables can be particularly helpful for visual learners, as it allows them to visually represent the problem and break it down into smaller, more manageable parts. This finding is also consistent with previous research that suggests that visual representations and decomposition are effective problem-solving strategies in mathematics and representation as cognitive tools in problem-solving (Böckmann & Schukajlow; 2018; Hegarty & Kozhevnikov, 1999). In addition, the creation of tables proved to be an effective tool for these PSTs, enabling them to organize and visualize complex information in a structured manner. Visual aids, such as diagrams and tables, can be beneficial for visual learners, as it provides a clearer understanding of the problem and its components (Reys et al., 2021). Logical reasoning was also evident in some PSTs' problem-solving approaches, where they applied logical thinking to carefully examine each element, seeking connections and relationships that would aid in their understanding of the problem.

On the other hand, ChatGPT relied heavily on algebraic equations and formulas to derive accurate solutions in the given context (as restricted by the prompts). According to the National Australian Curriculum v9.0, Year 6 children should be able to find unknown values in numerical equations involving brackets and combinations of arithmetic operations, using the properties of numbers and operations but not solving equations. Its utilization of other problem-solving strategies did not consistently produce the correct answers, as shown in **Table 4**. ChatGPT could not solve the problems correctly, which could be attributed to the lack of contextual understanding of the problems, unlike PSTs, who can draw upon their prior knowledge and experiences to inform their solutions. In this regard, Wardat et al. (2023b) noted that various factors can influence ChatGPT's accuracy. These include the clarity and detail of the prompts given, the mathematical questions complexity, the relevance and range of ChatGPT's training data, and the specific context and knowledge area involved. Therefore, it is crucial for teachers to verify ChatGPT's responses with additional sources for accuracy and reliability. Enhancing accuracy can also be achieved by using clear, specific prompts, asking questions within ChatGPT's areas of expertise, and cross-verifying results with other sources. Future research should take these factors into account when evaluating ChatGPT's response accuracy.

In relation to the strategies used, ChatGPT mainly relied on algebraic equations and formulas to derive accurate solutions, and its utilization of other problem-solving strategies, such as guess, check, and refine, was not as effective as PSTs and did not consistently produce the correct answers. This finding is not surprising, given that ChatGPT is a language model trained on vast text data, including mathematical formulae and equations. This suggests that while AI models such as ChatGPT can solve mathematical problems, they may not necessarily employ the same problem-solving strategies as human learners. The results from this study also indicate that AI models may require further development to enhance their problem-solving capabilities in a certain context. Using AI systems such as ChatGPT in mathematics education can provide children with a tool for generating solutions to problems quickly and accurately. However, it is essential to ensure that children also develop their own problem-solving strategies and understand the underlying mathematical concepts to assess the validity of the information produced by ChatGPT (UNESCO, 2023). In addition, it has been noticed that ChatGPT may not always interpret a correct solution to a problem accurately, as illustrated in **Figure 10**. Therefore, it has become essential to authenticate the information produced by ChatGPT prior to its utilization. UNESCO (2023) has proposed a flow chart (**Figure 1**) that could assist educators in integrating ChatGPT into their teaching. The flow chart mandates that teachers validate the information generated by ChatGPT before incorporating it into their teaching.

Overall, the study suggests that human problem-solvers have a distinct advantage over AI when it comes to problem-solving tasks that require contextual understanding and creativity. Using problem-solving strategies is essential in helping PSTs understand and solve problems while using algebraic equations and formulas is more effective for ChatGPT. However, there is still much room for improvement in AI problem-solving strategies.

Implication for Teacher Education & Classroom Teaching

The findings of this study have several important implications for teacher education and classroom teaching in primary school mathematics. Firstly, PSTs, including teachers, need to be aware of the potential benefits and limitations of incorporating ChatGPT into primary school mathematics education. They should receive training on effectively integrating and becoming AI literate in such tools into their teaching practices to enhance children's mathematics learning. This approach may include customizing prompts and modifying settings to better match the understanding levels of primary school children. Personalizing learning paths and implementing a feedback system can also refine ChatGPT's utility for each student's unique needs, enhancing its effectiveness as an educational resource in primary education. In addition, teachers need to understand how AI-based tools can support their teaching and how to integrate them into the curriculum in a meaningful way and support guidelines for the safe usage of ChatGPT for mathematics teaching in the primary school context could be developed. Secondly, teachers should emphasize the importance of diverse problem-solving strategies to children. The results of this study suggest that while ChatGPT may be effective in solving mathematical problems, it primarily relies on algebraic equations and formulas. Therefore, teachers should encourage children to use various problem-solving strategies, including drawing diagrams, to develop a deeper understanding of mathematical concepts (Duru et al., 2011; Reys et al., 2021). Thirdly, before introducing AI-based language models like ChatGPT, teachers must first establish a strong foundation of mathematical knowledge and skills among children to verify the accuracy of the information provided (UNESCO, 2023). It is essential that children have a basic understanding of mathematical concepts and problem-solving strategies before they can effectively utilize such tools (Lee et al., 2022). Fourthly, while ChatGPT may effectively solve mathematical problems, it cannot replace the critical thinking and problem-solving skills developed through human instruction. Therefore, teachers might use ChatGPT as a tool to support their teaching and encourage a diverse range of problem-solving strategies.

Limitations & Future Direction

While the findings of this study provide valuable insights into the problem-solving strategies employed by PSTs and ChatGPT, some limitations should be considered. Firstly, the study only used a single mathematical problem, which limits the generalizability of the results. Future studies can explore the problem-solving strategies used by PSTs and AI systems in solving various mathematical problems. Secondly, the study did not investigate the impact of using ChatGPT on children's learning outcomes. Future studies should examine the effectiveness of incorporating AI systems like ChatGPT in improving children's mathematics learning outcomes. Thirdly, the study was limited to only four different prompts and five trials on ChatGPT 3.5, which may have affected the accuracy and diversity of its responses. Additionally, the study's participant pool is relatively small, comprising only 58 PSTs. This limited sample size could potentially affect the generalizability of the findings. Future studies can consider using a larger sample of participants and a wider range of prompts and trials with the latest version of ChatGPT 4.0 to determine the effectiveness of ChatGPT and human participants in problem-solving. Overall, it is important to exercise caution when interpreting the results of this study and to conduct further research to fully explore the potential of ChatGPT as a tool for effective problem-solving in primary school mathematics education.

CONCLUSIONS

The study's findings demonstrated that PSTs utilized a variety of problem-solving strategies, utilizing creativity and adaptability to generate numerous solutions to the problem. On the other hand, ChatGPT relied on a consistent problem-solving strategy based on its training data. While ChatGPT could provide apparently accurate solutions in some cases, it lacked the diversity and flexibility of problem-solving strategies. These results have significant implications for mathematics education, as problem-solving efficiency and accuracy are critical to success in this subject. Although AI systems such as ChatGPT have the potential to assist in problem-solving tasks, they cannot entirely replace human problem-solving strategies. Instead, a combination of AI and human instruction could offer a well-rounded approach to solving mathematical problems, drawing on the strengths of both approaches. Furthermore, these findings emphasize the importance of developing problem-solving skills in primary school children. By encouraging children to explore multiple strategies, teachers can help develop their creativity and flexibility in problem-solving. This supports not only their mathematical thinking but also their cognitive development more broadly.

To sum up, this study highlights the advantages and disadvantages of AI-based problem-solving strategies compared to PSTs' problem-solving approaches. While AI systems like ChatGPT offer efficiency in supporting and assisting problem-solving, the flexibility and creativity of PSTs' problem-solving strategies are still unmatched. These results underscore the importance of fostering problem-solving skills in children and suggest that combining AI and human instruction could offer a comprehensive approach to mathematics education. Future research should investigate the effectiveness of incorporating AI-based language models like ChatGPT into primary school mathematics education and the role of human instruction in supporting problem-solving skills, considering the limitations of this study.

This study further opens several pathways for future research. One key area involves refining AI models, including ChatGPT, to better suit primary school contexts, ensuring they are tailored to the unique needs and learning styles of students across various age groups. Additionally, investigating the most effective methods for integrating AI technologies into teacher education programs presents a valuable direction. This includes exploring how educators can be trained to effectively use AI as a tool for enhancing teaching and learning experiences. Lastly, there is a significant opportunity to examine the impact of AI-based systems on student learning outcomes. This involves assessing not just academic performance but also the influence on students' critical thinking, problem-solving skills, and long-term educational development.

Funding: No funding source is reported for this study.

Ethical statement: The author stated that the University of Southern Queensland Human Research Ethics Committees provided the research study that underpins this publication–approval number ETH2023-0208. Written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by the author.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the author.

REFERENCES

ACARA. (2023). Mathematics: Year 6. *Australian Curriculum, Assessment and Reporting Authority*. https://www.australiancurriculum.edu.au/f-10-curriculum/mathematics/

Böckmann, M., & Schukajlow, S. (2018). Value of pictures in modelling problems from students' perspective. In E. Bergqvist, M. Österholm, M. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (pp. 263-270). PME.

Bourke, S., & Stacey, K. (1988). Assessing problem-solving in mathematics: Some variables related to student performance. *The Australian Educational Researcher, 15*(1), 73-83. https://doi.org/10.1007/BF03219402

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Csíkos, C., & Szitányi, J. (2020). Teachers' pedagogical content knowledge in teaching word problem-solving strategies. *ZDM, 52*(1), 165-178. https://doi.org/10.1007/s11858-019-01115-y

Duru, A., Peker, M., Bozkurt, E., Akgun, L., & Bayrakdar, Z. (2011). Pre-service primary school teachers' preference of the problem-solving strategies for word problems. *Procedia-Social and Behavioral Sciences, 15*, 3463-3468. https://doi.org/10.1016/j.sbspro.2011.04.319

Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., & Berner, J. (2023). Mathematical capabilities of ChatGPT. *arXiv*. https://doi.org/10.48550/arXiv.2301.13867

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem-solving. *Journal of Educational Psychology, 91*(4), 684-689. https://doi.org/10.1037/0022-0663.91.4.684

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

Krawec, J., & Montague, M. (2014). The role of teacher training in cognitive strategy instruction to improve math problem-solving. *Learning Disabilities Research & Practice, 29*(3), 126-134. https://doi.org/10.1111/ldrp.12034

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education, 191*, 104646. https://doi.org/10.1016/j.compedu.2022.104646

Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education, 00*, 1-6. https://doi.org/10.1002/ase.2270

Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based cha bot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development, 70*(5), 1843-1865. https://doi.org/10.1007/s11423-022-10142-8

Mills, K. A., Chandra, V., & Park, J. Y. (2013). The architecture of children's use of language and tools when problem-solving collaboratively with robotics. *The Australian Educational Researcher, 40*, 315-337. https://doi.org/10.1007/s13384-013-0094-z

Mohamed, M. Z. b., Hidayat, R., Suhaizi, N. N. b., Sabri, N. b. M., Mahmud, M. K. H. b., & Baharuddin, S. N. b. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education, 17*(3), em0694. https://doi.org/10.29333/iejme/12132

OpenAI. (2023). *ChatGPT 3.5* [Large language model]. https://chat.openai.com/chat

Pelton, T. & Pelton, L. F. (2023). Adapting ChatGPT to support teacher education in mathematics. In E. Langran, P. Christensen, & J. Sanson (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 1662-1670). Association for the Advancement of Computing in Education.

Qiu, Y., Pan, J., & Ishak, N. A. (2022). Effectiveness of artificial intelligence (AI) in improving pupils' deep learning in primary school mathematics teaching in Fujian Province. *Computational Intelligence and Neuroscience*, *2022*, 1362996. https://doi.org/10.1155/2022/1362996

Ramnarain, U. (2014). Empowering educationally disadvantaged mathematics students through a strategies-based problem-solving approach. *The Australian Educational Researcher, 41*, 43-57. https://doi.org/10.1007/s13384-013-0098-8

Reys, R. E., Lindquist, M. M., Lambdin, D. V., & Smith, N. L. (2021). *Helping children learn mathematics*. Wiley.

Rosé, C. P., McLaughlin, E. A., Liu, R., & Koedinger, K. R. (2019). Explanatory learner models: Why machine learning (alone) is not the answer. *British Journal of Educational Technology, 50*(6), 2943-2958. https://doi.org/10.1111/bjet.12858

Schoenfeld, A. H. (1985). Artificial intelligence and mathematics education: A discussion of Rissland's paper. In E. Silver (Ed.), *Teaching and learning mathematical problem-solving: Multiple research perspectives* (pp. 177-187). Lawrence Erlbaum Associates.

Schoenfeld, A. H. (2016). Learning to think mathematically: Problem-solving, metacognition, and sense making in mathematics. *Journal of Education, 196*(2), 1-38. https://doi.org/10.1177/002205741619600202

Shakarian, P., Koyyalamudi, A., Ngu, N., & Mareedu, L. (2023). An independent evaluation of ChatGPT on mathematical word problems (MWP). *arXiv*. https://doi.org/10.48550/arXiv.2302.13814

Strzelecki, A. (2023). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*. https://doi.org/10.1007/s10755-023-09686-1

Szabo, Z. K., Körtesi, P., Guncaga, J., Szabo, D., & Neag, R. (2020). Examples of problem-solving strategies in mathematics education supporting the sustainability of 21st-century skills. *Sustainability, 12*(23), 10113. https://doi.org/10.3390/su122310113

UNESCO. (2023). ChatGPT and artificial intelligence in higher education: Quick start guide. *United Nations Educational, Scientific and Cultural Organization International Institute for Higher Education in Latin America and the Caribbean*. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

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

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM Mathematics Education, 52*, 1-16. https://doi.org/10.1007/s11858-020-01130-4

Wardat, Y., Alali, R., Jarrah, A. M., & Alzyoudi, M. (2023a). Neutrosophic theory framework for building mathematics teachers capacity in assessment of high school students in the United Arab Emirates. *International Journal of Neutrosophic Science, 21*(1), 33-50. https://doi.org/10.54216/IJNS.210103

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