

Students' mathematical justification abilities in analyzing ChatGPT's answers

Joko Sutrisno AB^{1*}, Elvandri Yogi Pratama¹, Andika Putra R²,
Syaiful Hamzah Nasution³, Latifah Mustofa Lestyanto⁴

¹Department of Mathematics Education, STKIP-PGRI Bandar Lampung, Lampung, Indonesia

²Department of Mathematics Education, Universitas Tamansiswa, West Sumatra, Indonesia

³Department of Mathematics Education, Universitas Negeri Malang, East Java, Indonesia

⁴School of Mathematical and Statistical Sciences, University of Galway, Ireland

*Correspondence: jokosutrisnoab@gmail.com

Received: Sep 16, 2024 | Revised: Jan 27, 2025 | Accepted: Feb 3, 2025 | Published Online: Mar 3, 2025

Abstract

This study investigates students' mathematical justification abilities in analyzing answers obtained from one of the Artificial Intelligence (AI) based services, namely ChatGPT. The subjects of this descriptive qualitative study are 13 fourth-semester students of STKIP-PGRI Bandar Lampung. They were asked to utilize ChatGPT to solve two linear equation system questions adopted from Indonesian university entrance tests in 2019. The students were then requested to analyze and justify the responses generated by ChatGPT. The justifications from students were then analyzed to see their abilities and determine whether the answers from ChatGPT were correct. As many as 61.54% of the respondents could understand the instructions of the two questions and provide reviews for the answers generated by ChatGPT. 38.46% of the respondents did not understand the questions' instructions. The results showed that students may have lacked the necessary ability to analyze the AI-generated answers in depth and explain them in detail. This may be due to weak mastery of the basic concepts of the material and a lack of accuracy. The findings call for paying more attention to students' mathematical justification abilities in the current AI-assisted era.

Keywords:

Artificial intelligence, ChatGPT, Mathematical justification ability

How to Cite:

Sutrisno AB, J., Pratama, E. Y., Putra R, A., Nasution, S. H., & Lestyanto, L. M. (2025). Students' mathematical justification abilities in analyzing ChatGPT's answers. *Infinity Journal*, 14(2), 445-460. <https://doi.org/10.22460/infinity.v14i2.p445-460>

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. INTRODUCTION

Technological advancements have spoiled us in carrying out various daily activities. In today's advanced digital era, the role of Artificial Intelligence (AI) technology plays is growing significantly, as it influences different aspects of human life, one of which is education. As argued by Iqbal et al. (2024), AI has had a significant impact on various industries, including education, by addressing various learning problems through the use of

technology, such as virtual tutors and online learning systems. Thus, the rapid development of numerous AI-based tools, has given a tremendous impact on the world of education. One interesting development in the context of learning is the rise of natural language models such as ChatGPT.

The presence of AI such as ChatGPT has brought significant changes to the world of education, especially in mathematics learning. The ease of access to ChatGPT and its ability to quickly generate mathematical solutions offer an extraordinary potential to improve the effectiveness and efficiency of learning. Cheung et al. (2023) stated that ChatGPT can be used to answer questions, write essays, solve problems, explain complex topics, and even support research. Moreover, with its ability to generate human-like texts, respond to complex questions, and adapt to various contexts, ChatGPT has the potential to revolutionize education and research (Rahman & Watanobe, 2023). The quality of ChatGPT's performance, according to Lo (2023), varies across subject domains, ranging from outstanding (e.g., economics) and satisfactory (e.g., programming) to unsatisfactory (e.g., mathematics). In a similar line, Baidoo-anu and Ansah (2023) maintained that ChatGPT can improve the quality of teaching and learning in various fields, including mathematics. Furthermore, the potential of ChatGPT as an advanced conversational tool lies in its ability to serve as an "object to think with", foster reflective and critical thinking, and facilitate understanding of concepts in STEM, which are all relevant to mathematics research (Vasconcelos & Santos, 2023).

It is important to remember that the integration of ChatGPT in education, especially in mathematics learning, is not without challenges. One of the main challenges is students' heavy dependence on technology that can hinder the development of their critical thinking skills, conceptual understanding, and problem-solving abilities (Holmes et al., 2023; Zawacki-Richter et al., 2019). Furthermore, cheating and plagiarism are some potential negative impacts of using ChatGPT, which educators should be aware of (Wei et al., 2024). In fact, Welding (2023) found that half of the students (51%) in their study agree that using AI tools like ChatGPT for schoolwork is an act of cheating or plagiarism, although about one out of five students still use it.

Being dependent on AI without having a strong understanding of mathematical concepts may lead students to blindly accept AI-generated responses at face value, without being able to evaluate their truth and validity. This is concerning, because it contradicts the core objectives of mathematics education, which emphasize the development of logical reasoning, critical, and creative thinking skills (OECD, 2018). The OECD (2022) emphasizes that as education systems are adapting to the development of digital technology, they need to ensure that students have the necessary skills to succeed in the digital age, such as critical thinking skills, problem-solving, and digital literacy.

Hence, in this context, mathematical justification ability becomes highly crucial. Mathematical justification ability involves providing logical reasoning and valid evidence to support or validate mathematical claims and solutions. This skill entails higher-order thinking processes that require students to analyze information, identify patterns, make generalizations, and build valid arguments (Hershkowitz, 2020).

Furthermore, Hanna and Jahnke (1996) added that mathematical justification not only focuses on proving the truth, but also on explaining the “why” behind a mathematical statement, so that students can “understand the connections between mathematical concepts and develop critical thinking skills.” Mathematical justification can be divided into several types, such as empirical justification (based on observations and examples), deductive justification (based on logical rules), and visual justification (using diagrams or visual representations). In addition to these types, Sukirwan et al. (2020) classified some aspects that affect the justification of four types of arguments built by students, namely, inductive, algebraic, visual, and perceptual. Since the development of mathematical justification ability is a gradual process that requires structured guidance and practice, improving students’ justification abilities in the classroom is challenging for mathematics teachers at all school levels (Sarumaha & Rizkianto, 2022). Santos et al. (2022) stated that teachers should focus their attention on the process of making generalization and justification in their professional development.

The presence of ChatGPT gives a new dimension to the development of mathematical justification ability. Although ChatGPT can help students solve math problems, excessive dependence without adequate conceptual understanding can hinder the justification process and lead to the acceptance of wrong answers without critical evaluation (Kung et al., 2023). Thus, it is important to balance students’ positive perception of ChatGPT as a learning tool with a sufficient understanding about its limitations and potential for misuse (Xu et al., 2024). For instance, students should realize that while ChatGPT has the potential to increase student motivation and engagement in learning mathematics, it has some limitations in solving complex mathematical problems as well as those requiring deep conceptual understanding. Most importantly, ChatGPT cannot replace the role of teachers in guiding students to develop mathematical justification ability.

Due to these reasons, students should be trained not only to run ChatGPT as a tool to help them solve mathematical problems, but also to develop their ability to analyze and evaluate the AI-generated responses. This has been argued by scholars like García-López et al. (2025) who highlighted the need to have a comprehensive understanding of ChatGPT’s capabilities, challenges, limitations, and implications in education and across various fields

Thus, in this context, the topic of this study is also related to other relevant issues in education in the digital age, such as ethics in the use of AI, the digital divide, and the role of teachers in the age of AI. The finding of this study can contribute to the existing literature on the use of AI in mathematics education and provide recommendations for the more effective learning strategies in the digital age.

Based on these problems, this study aims to answer the following research questions: How can students analyze and provide mathematical justification for the answers generated by ChatGPT to linear equation problems?

2. METHOD

This study employs a qualitative approach with a descriptive research design. The qualitative approach was selected to explore and examine how students interpret, analyze, and justify solutions generated by AI, thus dealing with a phenomenon that cannot be

measured solely by numbers. A descriptive research design was used because it aligns with the purpose of this study, that is, to describe students' abilities in analyzing and justifying ChatGPT-generated answers, with no intervention or manipulation of variables.

The participants of this study comprised 13 fourth-semester students enrolled in the Mathematics Education study program at STKIP-PGRI Bandar Lampung. The participant selection was based on the premise that prospective mathematics teachers are required to have sufficient conceptual understanding and justification ability to analyze mathematical solutions generated by both humans and AI tools.

Data were collected through Google Forms in April 2024 using an instrument consisting of two linear equation problems which were adapted from the State University entrance exam question and University of Indonesia entrance exam questions from 2019. The selection of questions from these two sources was based on considerations of the level of difficulty, the required higher-order thinking skills, and the relevance of the questions to the context of prospective mathematics teachers.

The data collection procedure included three stages. First, participants solved the problems by using ChatGPT. They were asked to access ChatGPT and use it to solve the two linear equation problems. Participants were allowed to explore various ways of using ChatGPT to answer the questions. Then, the participants analyzed the responses generated by ChatGPT and provided justification for the responses from ChatGPT. They were asked to explain whether the responses generated by ChatGPT were correct, to indicate the location of errors if any, and to provide reasons or evidence to support their arguments. Third, the data in the form of students' analysis and justification were collected through Google Forms and analyzed by using descriptive qualitative analysis techniques, which included data reduction, data presentation, and drawing conclusions.

Data reduction involved a systematic and careful reading of the participants' responses and identifying key information that aligns with the research objective, namely, assessing students' abilities to analyze and justify ChatGPT-generated responses. Relevant information was then extracted and grouped based on certain categories. The reduced data were then presented in the form of descriptive narratives that were systematic, clear, and easy to understand. To facilitate understanding and interpretation of the data, tables and graphs were used. Finally, conclusions about students' mathematical justification ability in analyzing ChatGPT's answers were drawn. The conclusions were articulated clearly and concisely to answer the research question.

To improve the validity and reliability of the study, we carried out data triangulation by collecting data from various sources, such as participants' written responses from Google Forms and documentation of participant interactions with ChatGPT. Documentation of interactions with ChatGPT were collected in the form of screenshots and screen recordings which contained the participants' dialogues with ChatGPT, including the questions they asked as well as the responses by ChatGPT. This data triangulation aims to obtain a more comprehensive and in-depth picture of students' mathematical justification ability. To enhance the credibility, we also conducted member checking by engaging the participants in a collaborative discussion of our data analysis. This was done to ensure that our interpretations accurately reflected their intentions and experiences. Finally, peer debriefing

was carried out through discussions with colleagues and mathematics education experts to elicit critical input and reflective insights about the research method and findings. These steps were implemented to produce valid and reliable data and findings. To facilitate the analysis of students' evaluations of ChatGPT-generated responses, the participants' names were anonymized and labelled from M1 to M13.

3. RESULTS AND DISCUSSION

3.1. Results

Figure 1 shows the results of students' analysis of ChatGPT-generated answers in response to the provided mathematical questions. As seen from Figure 1, when using ChatGPT to solve the mathematical problems, most participants either directly retype the questions or paraphrase them.

1. If x and y are positive integers that satisfy $4x - 5y = a$ and $8x + 5y = 34$ and $x + a$ is a prime number between 2 and 6, then $x - y = \dots\dots$
2. If $(p^2 - 1)x + y = 0$ and $-2x + (p^2 - 4) + y = 0$ with $x \neq 0$ and $y \neq 0$, the smallest p^2 value that satisfies the system of linear equations is

Figure 1. Questions

It is notable that none of the students ran step-by-step commands. Figure 2 presents some of the commands that students inputted to ChatGPT.

If x and y are positive integers that satisfy $4x - 5y = a$ and $8x + 5y = 34$ and $x + a$ is a prime number between 2 and 6, then $x - y$ is...

Find $x - y$ if $4x - 5y = a$, $8x + 5y = 34$, $x + a$ is a prime number between 2 and 6.

Figure 2. Some of the commands that students inputted to solve question 1

Of the thirteen participants, only eight provided comprehensive analysis of ChatGPT's responses to the given mathematical questions. The remaining five participants struggled to understand the instructions given in the instrument, and merely collected the answers from ChatGPT without critically analyzing them (see Figure 3). This finding indicates that 61.54% of participants could understand the given instructions. However, even among these participants, not all were able to precisely identify the errors in ChatGPT's answers.

Instructions:
 Work on the following problems with the help of **ChatGPT** and then **analyze whether the answer given by ChatGPT is correct or incorrect**. If the answer is correct, explain whether it is in accordance with the concept and your understanding. If the answer is wrong, explain the reason and the location of the error. Answers can be handwritten or typed (each question on a different page) and then collected in PDF format via the following link.
http://gg.gg/Tes_Kemampuan_Berpikir_Matematis_Mahasiswa

Figure 3. Working instructions

Table 1 presents a summary of participants' analysis of ChatGPT-generated answers to the mathematical questions.

Table 1. Recapitulation of respondents' answers

Participants	Analysis
M1	The analysis done by M1 lacked detail. M1 did not elaborate on the location of the error in ChatGPT's answers. ChatGPT's answers for the first question was correct, but the second one was incorrect.
M2	ChatGPT's answer was incorrect, but M2's own answer was correct. However, M2 was unable to point out the error in ChatGPT's responses.
M3	M3 gave limited reviews of the answers from ChatGPT. The answers obtained from ChatGPT were all correct.
M4	The answers that M4 obtained from ChatGPT were all incorrect. M4 was not able to identify the errors in ChatGPT-generated answers.
M5	ChatGPT provided the correct answer, so M5 did not provide much feedback.
M6	M6 provided very limited feedback by comparing the answers from ChatGPT and their own answers, but the attached answer images were unclear or blurry.
M7	M7 did not provide much analysis. For question 1, both ChatGPT's and the respondent's answers were correct. For question 2, while ChatGPT's answer was incorrect and M7's own answer was correct, M7 could not point out the errors in ChatGPT's response.
M8	M8 did not give much analysis. Instead, it appeared that M8 did not understand question 1 because the incorrect answer given by ChatGPT was accepted as correct.
M9	M9 did not understand the instructions of the questions, because M9 gave their own answers, and not reviewing the answers obtained from ChatGPT. However, M9's own answers were correct.
M10	M10 wrote the answers only to question number 1, and did not finish the second one.
M11	M11 did not understand the instructions but gave correct answers on their own.
M12	M12 gave correct answers based in their own work, which shows that they did not understand the instructions.
M13	M13 did not understand the instructions because they gave their own answers. Even then, it was only question number 2 and the answer was wrong.

The results of several analyses conducted by the participants can be seen in [Figure 4](#).

If x and y are positive integer that satisfy $4x - 5y = a$, $8x + 5y = 34$, and $x + a$ is a prime number between 2 and 6, then $x - y = \dots$	
<p>Translation Chat GPT Responses</p> <p>We can use the elimination method to solve this system of equations. First, we will add both equations to eliminate the y.</p> $(4x - 5y) + (8x + 5y) = a + 34$ $12x = a + 34$ $x = \frac{a + 34}{12}$ <p>Now, we will use the obtained value of x to find the value of y.</p> $4\left(\frac{a + 34}{12}\right) - 5y = a$ $\frac{4a + 136}{12} - 5y = 1$ $4a + 136 - 60y = 12a$ $136 - 60y = 8a$ $60y = 136 - 8a$ $y = \frac{136 - 8a}{60}$ <p>Next, we will find the value of $x + a$</p> $x + a = \frac{a + 34}{12} + a = \frac{13a + 34}{12}$	<p>To satisfy the condition that $x + a$ is a prime number between 2 and 6, we need to find the value of a that makes this expression a prime number.</p> $\frac{13a + 34}{12}$ <p>To check if this is a prime number, we need to check if it is divisible by 2, 3, 4, or 5. Since we want it to be a prime number between 2 and 6, we only need to check 2 and 3.</p> $\frac{13a + 34}{12} = 2 \quad \text{or} \quad \frac{13a + 34}{12} = 3$ $(13a + 34) = 24 \quad \text{or} \quad 13a + 34 = 36$ $13a = -10 \quad \text{or} \quad 13a = 2$ $a = -\frac{10}{13} \quad \text{or} \quad a = \frac{2}{13}$ <p>However, since a must be a positive integer, we will only consider $a = 2$.</p> <p>then,</p> $x = \frac{2 + 34}{12} = \frac{36}{12} = 3 \quad \text{and} \quad y = \frac{136 - 8(2)}{60} = \frac{120}{60} = 2$ <p>So,</p> $x - y = 3 - 2 = 1$
<p>Student feedback on ChatGPT's responses</p> <p>In conclusion, I think the answers provided by ChatGPT are quite random. Sometimes they're correct, and sometimes they're wrong. I've tried asking the same questions multiple times and received different answers each time. However, the answer you provided above aligns with my understanding and concept.</p>	

Figure 4. The results of M1's analysis of the answer to question 1

M1 explained that for question 1, several experiments were conducted, and ChatGPT's answers were different from previous responses. M1 stated that the answer presented was correct and aligned with the concept and understanding. For question 2 (see Figure 5), M1 described ChatGPT's answer as tentative and found it difficult to understand because it was excessively lengthy. M1 concluded that ChatGPT-generated responses were inconsistent, as some were correct and others were incorrect.

If x and y are positive integer that satisfy $4x - 5y = a$, $8x + 5y = 34$, and $x + a$ is a prime number between 2 and 6, then $x - y = \dots$	
<p>Translation Chat GPT Responses</p> <p>Given the system of equations:</p> <ol style="list-style-type: none"> $(p^2 - 1)x + y = 0$ $-2x + (p^2 - 4) + y = 0$ <p>We will use the elimination method. From equation (2), we can express y as:</p> $y = 2x - (p^2 - 4)$ <p>Then we substitute y into equation (1):</p> $(p^2 - 1)x + (2x - (p^2 - 4)) = 0$ <p>Now we can simplify this equation:</p> $(p^2 - 1)x + 2x - p^2 + 4 = 0$ $(p^2 + 2)x - p^2 + 4 = 0$ $p^2x + 2x - p^2 + 4 = 0$ <p>Now we want to find the smallest value of p^2 that makes the system of equation have a nontrivial solution, where $x \neq 0$ and $y \neq 0$.</p> <p>To do this, we will find the value of p^2 that makes the determinant of the coefficient of the equations is zero.</p> $\text{Determinant} = (p^2 + 2) \cdot 1 - p^2 = 0$ $p^2 + 2 - p^2 = 0$ $2 = 0$ <p>Apparently, this is an impossible equation. Let try another approach. Let's look at equations (1) and (2) again:</p> <ol style="list-style-type: none"> $(p^2 - 1)x + y = 0$ $-2x + (p^2 - 4) + y = 0$ <p>We can also express x from equation (1):</p> $x = \frac{-y}{p^2 - 1}$ <p>Then we substitute x into equation (2):</p> $-2\left(\frac{-y}{p^2 - 1}\right) + (p^2 - 4) + y = 0$	<p>Now we can simplify this equation:</p> $\frac{2y}{p^2 - 1} + (p^2 - 4) + y = 0$ $\frac{2y + (p^2 - 1)(p^2 - 4)}{p^2 - 1} = 0$ <p>Now we want to find the smallest value of p^2 that makes this equation have no trivial solution, where $y \neq 0$.</p> <p>To do this, we want to ensure that the numerator cannot be simplified to zero:</p> $2y + (p^2 - 1)(p^2 - 4) \neq 0$ $2y + (p^4 - 5p^2 + 4) \neq 0$ $2y + p^4 - 5p^2 + 4 \neq 0$ $p^4 - 5p^2 + 2y + 4 \neq 0$ <p>We want to find the smallest value of p^2 that makes this expression non-zero, with $y \neq 0$.</p> <p>Looking at the coefficients of p^4 and p^2, we want to ensure that the coefficient of p^4 is positive (to keep the expression positive) and the coefficient of p^2 is negative (to prevent the expression from becoming zero). Since we are looking for the smallest value of p^2, we will try $p^2 = 1$:</p> $(1)^4 - 5(1) + 2y + 4 \neq 0$ $1 - 5 + 2y + 4 \neq 0$ $2y \neq 0$ $y \neq 0$ <p>With the value of $y \neq 0$, the smallest value of p^2 that satisfies the system of equation is $p^2 = 1$.</p>
<p>Student feedback on ChatGPT's responses</p> <p>In conclusion, the answer is tentative (could be right or wrong). In my understanding, the answer is too long or too broad to be easily understood.</p>	

Figure 5. The results of M1's analysis of the answer to question 2

The analysis conducted by M1 on ChatGPT-generated answers suggests that M1 had difficulty in analyzing ChatGPT-generated answers. M1 has not yet fully analyzed the answers and struggled to pinpoint the errors when ChatGPT provided incorrect responses. For question 2 (see Figures 6 to 7), it was challenging for M1 to comprehend the answers made by ChatGPT and to identify where the error occurred. This may indicate that M1 has not fully grasped the concept related to question 2.

Translation Chat GPT Responses		Translation answer of M2's																				
<p>To solve this problem, let's take it step by step:</p> <ol style="list-style-type: none"> Solve the system of equations to find the values of x and y. Find the value of 'a' that meets the condition that $x + a$ is a prime number between 2 and 6. Calculate the value of $x - y$. <p>Let's start with the first step: solving the system of equations. The system of equations is:</p> <ol style="list-style-type: none"> $4x + 5y = 34$ $x + a$ is a prime number between 2 and 6. <p>To solve this system, we can try some values of x and y that satisfy the first equation, then check the value of $x + a$ to find a prime number between 2 and 6. Would you like to continue with this step?</p>	<p>Okay, now we have $x = 6$ and $y = 2$. We have to find the value of a that makes $x + a$ is a prime number between 2 and 6.</p> <p>Let's try each value of a from 2 to 6:</p> <ul style="list-style-type: none"> If $a = 2$, then $x + a = 6 + 2 = 8$. 8 is not a prime number. If $a = 3$, then $x + a = 6 + 3 = 9$. 9 is not a prime number. If $a = 4$, then $x + a = 6 + 4 = 10$. 10 is not a prime number. If $a = 5$, then $x + a = 6 + 5 = 11$. 11 is a prime number. <p>So, $a = 5$ fulfills the requirement. Now we have found the values of x, y, and a. We can calculate $x - y$.</p> $x - y = 6 - 2 = 4$ <p>Therefore, the result is 4.</p>	<p>If x and y are positive integers that satisfy $4x - 5y = a$ and $8x + 5y = 34$ and $x + a$ is a prime number between 2 and 6, then $x - y = \dots$</p> <p>Answer :</p> $\begin{array}{l} 8x + 5y = 34 \\ 4x - 5y = a \end{array} \times (-1) \left \begin{array}{l} 8x + 5y = 34 \\ -4x + 5y = -a \end{array} \right. \frac{-}{12x = 34 + a}$ <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;">$x + a = 3$</td> <td style="width: 50%; border: none;">$x + a = 5$</td> </tr> <tr> <td style="border: none;">$a = 3 - x$</td> <td style="border: none;">$a = 5 - x$</td> </tr> <tr> <td style="border: none;">$12x = 34 + a$</td> <td style="border: none;">$12x = 34 + a$</td> </tr> <tr> <td style="border: none;">$12x = 34 + (3 - x)$</td> <td style="border: none;">$12x = 34 + (5 - x)$</td> </tr> <tr> <td style="border: none;">$13x = 37$</td> <td style="border: none;">$13x = 39$</td> </tr> <tr> <td style="border: none;">$x = 2,8$ (Incorrect)</td> <td style="border: none;">$x = 3$ (Integer)</td> </tr> <tr> <td style="border: none;">(Not an Integer)</td> <td style="border: none;">↓</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">$a = 5 - x$</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">$= 5 - 3$</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">$= 2$</td> </tr> </table> $\begin{array}{l} 4x - 5y = a \\ 4(3) - 5y = a \\ 12 - 5y = 2 \\ -5y = 2 - 12 \\ -5y = -10 \\ y = 2 \end{array} \rightarrow x - y = 3 - 2 = 1$	$x + a = 3$	$x + a = 5$	$a = 3 - x$	$a = 5 - x$	$12x = 34 + a$	$12x = 34 + a$	$12x = 34 + (3 - x)$	$12x = 34 + (5 - x)$	$13x = 37$	$13x = 39$	$x = 2,8$ (Incorrect)	$x = 3$ (Integer)	(Not an Integer)	↓		$a = 5 - x$		$= 5 - 3$		$= 2$
$x + a = 3$	$x + a = 5$																					
$a = 3 - x$	$a = 5 - x$																					
$12x = 34 + a$	$12x = 34 + a$																					
$12x = 34 + (3 - x)$	$12x = 34 + (5 - x)$																					
$13x = 37$	$13x = 39$																					
$x = 2,8$ (Incorrect)	$x = 3$ (Integer)																					
(Not an Integer)	↓																					
	$a = 5 - x$																					
	$= 5 - 3$																					
	$= 2$																					

Figure 6. The results of M2's analysis of the answer to question 1

Translation Chat GPT Responses		Translation answer of M2's	
<p>find the smallest value of p^2 that satisfies the system of equations, we need to solve the system of equations:</p> <ol style="list-style-type: none"> $(p^2 - 1)x + y = 0$ $-2x + (p^2 - 4)y = 0$ <p>We can solve this system by finding the value of p^2 that satisfies both equations. Let's try step by step:</p> <ol style="list-style-type: none"> Substitute $y = -(p^2 - 1)x$ from the first equation into the second equation: $-2x + (p^2 - 4) - (p^2 - 1)x = 0$ Simplify the equation: $-2x + (p^2 - 4) - (p^2 - 1)x = 0$ $-2x + (p^2 - 4) - p^2x + x = 0$ $-2x - p^2x + x + p^2 - 4 = 0$ $-x(2 + p^2) + x + p^2 - 4 = 0$ <p>Therefore, the smallest value of p^2 that satisfies the system of linear equations is 0.</p>	<ol style="list-style-type: none"> Factor out x: $x(-2 - p^2 + 1) + p^2 - 4 = 0$ $x(-p^2 - 1) + p^2 - 4 = 0$ $x(-p^2 - 1) = 4 - p^2$ $x = \frac{4 - p^2}{-p^2 - 1}$ Now, since $x \neq 0$, the denominator cannot be zero. So, we can write: $\frac{-p^2 - 1 \neq 0}{p^2 \neq -1}$ <p>From here, we need to find the smallest value of p^2 that satisfies $p^2 \neq -1$. Since $p^2 \neq -1$, we must find the value of p^2 that is greater than -1. The smallest value that satisfies this is 0.</p>	<p>If $(p^2 - 1)x + y = 0$ and $-2x + (p^2 - 4)y = 0$, with $x \neq 0$ and $y \neq 0$, find the smallest possible value of p^2 that satisfies the system of linear equations.</p> <p>Answer:</p> $\begin{array}{l} (p^2 - 1)x + y = 0 \\ y = -(p^2 - 1)x \\ \frac{y}{x} = -(p^2 - 1) \end{array}$ $\begin{array}{l} -2x + (p^2 - 4)y = 0 \\ (p^2 - 4)y = 2x \\ \frac{y}{x} = \frac{2}{p^2 - 4} \end{array}$ $\begin{array}{l} -(p^2 - 1) = \frac{2}{p^2 - 4} \\ -(p^2 - 1)(p^2 - 4) = 2 \end{array}$ <p>Let $m = p^2$</p> $\begin{array}{l} -(m - 1)(m - 4) = -2 \\ -(m^2 - 4m - m + 4) = 2 \\ -(m^2 - 5m + 4) = 2 \\ m^2 - 5m + 4 = \frac{2}{-1} \\ m^2 - 5m + 4 = -2 \\ m^2 - 5m + 6 = 0 \end{array}$	<p>Factor the quadratic equation</p> $\begin{array}{l} ax^2 + bx + c = 0, \text{ if } a = 1 \\ (x + P)(x + Q) = 0 \\ P + Q = b \quad P \cdot Q = c \end{array}$ $\begin{array}{l} m^2 - 5m + 6 = 0 \\ (m - 3)(m - 2) = 0 \\ m - 3 = 0 \quad m - 2 = 0 \\ m = 3 \quad m = 2 \\ p^2 = 3 \quad p^2 = 2 \quad \checkmark \end{array}$ <p>So, the answer from Chat GPT and my answer are different. Chat GPT's first answer uses substitution, while I use factoring and substitution, so Chat GPT's answer is wrong.</p>

Figure 7. The results of M2's analysis of the answer to question 2

On the contrary, M2 noted that ChatGPT's answers differed from their own. M2 explained that for question 1, ChatGPT relied on memorization, while he used factors and substitution to solve the problem. Based on this difference, M2 concluded that the answer by ChatGPT for question 1 was incorrect.

Furthermore, the feedback given by M2 was quite limited. M2 simply stated that ChatGPT's answer was incorrect but did not identify where the error was or provide a detailed review to address the error. Additionally, M2 only reviewed question 1 and did not give any review for question 2. Instead, they attached ChatGPT's responses and his own work without providing further analysis. Notably, there is a difference in the answers to question 2, with ChatGPT's response being incorrect and M2's work being correct.

CONCLUSION

- ❖ For question 1, the answer given by ChatGPT is correct and the concept is the same as the one I used, there are only slight differences in the assumptions and substitution, but in essence they are the same.
- ❖ For question 2, the answer given by ChatGPT is correct and the concept is also the same as the one I use, which is by using substitution, but there are slight differences in the substitution and assumptions, but in essence they are the same, because both use assumptions and substitutions.

Figure 8. The results of M3's analysis of the answers from ChatGPT

Seen from [Figure 8](#), M3's analysis of ChatGPT's answers was presented directly as a conclusion. For question 1, the answer generated by ChatGPT was correct and aligned with the concept they used, though there was a slight difference in modelling and substitution. Similarly, for question 2, M3 found ChatGPT's answer correct and consistent with their own concept, despite minor differences in the use of substitution and modelling. M3 emphasized that both answers were fundamentally similar. The answers and analysis of ChatGPT's responses submitted by M3 arrived at correct final results. However, M3 did not provide a detailed explanation about where the differences occurred.

M4 observed that although ChatGPT could answer general questions, it struggled with complex mathematical questions. It often failed to provide the correct concept or final answer. M4 also explained that ChatGPT sometimes only provided steps for solving the problem without reaching a final answer. In some cases, it gave no answer at all and even started with incorrect steps. Regarding question 1, M4 noted that ChatGPT only outlined the steps and failed to provide a final result. The steps that it presented were quite unclear to those without a strong grasp of the basic concepts of the mathematical problem. As for question 2, M4 stated that the answer given by ChatGPT was incorrect from the beginning, which resulted in an obviously wrong final response.

ChatGPT's answer to question number 1	ChatGPT's answer to question number 2
<p>Let's solve these problems one by one. First, we need to solve the system of equations:</p> $4x + 5y = a$ $8x + 5y = 34$ <p>We can find the values of x and y this way. Then, we will get the value of a from the first equation.</p> <p>By subtracting the second equation from the first, we can eliminate the y variable:</p> $4x + 5y - (8x + 5y) = a - 34$ $4x - 8x = a - 34$ $-4x = a - 34$ $x = \frac{34-a}{4}$ <p>Then, we can substitute the value of x into one of the original equations to get the value of y. For example, we use the second equation:</p> $8x + 5y = 34$ $8\left(\frac{34-a}{4}\right) + 5y = 34$ <p>Now, we can find the value of y and then get the value of a.</p> <p style="text-align: center;">Conclusion:</p> <p>ChatGPT can indeed answer essay questions from general questions but for math questions ChatGPT still cannot answer all complicated math problems with the correct concept and final answer, sometimes only given the steps to work on it or no final answer is given, sometimes there is no answer at all or the steps from the beginning are wrong. So ChatGPT's answer to this question only gave the steps of the process not the final result and the steps were given in a form that could not be understood by people who did not understand the basic concept of the question above.</p>	<p>To solve the system of linear equations, we need to find the value of p^2 that satisfies both equations. let's solve them one by one.</p> <p>First equation:</p> $(p^2 - 1)x + y = 0$ <p>Second equation:</p> $-2x + (p^2 - 4)y = 0$ <p>We will use the elimination method to find the value of p^2.</p> <p>From the first equation, we can express y as:</p> $y = (1 - p^2)x$ <p>Then, we substitute y into the second equation:</p> $-2x + (p^2 - 4)(1 - p^2)x = 0$ <p>Now, we can simplify and find the value of p^2.</p> $-2x + (p^2 - 4 - p^4 + 4p^2)x = 0$ $-2x + (p^2 - p^4)x = 0$ $x(-2 + p^2 - p^4) = 0$ <p>Since we are told that $x \neq 0$, it must be:</p> $-2 + p^2 - p^4 = 0$ $p^4 - p^2 + 2 = 0$ <p>We now have a quadratic equation in p^2. To find the smallest value of p^2 that satisfies this equation, we must solve it.</p> <p>If we call it $z = p^2$, then this equation becomes:</p> $z^2 - z + 2 = 0$ <p>We can use the quadratic formula to find the solution. However, note that the discriminant value is $\Delta = (-1)^2 - 4 \cdot 1 \cdot 2 = 1 - 8 = -7$. Because the discriminant is negative, this equation has no real roots.</p> <p>So, there is no value of p^2 that satisfies the system of equations.</p> <p style="text-align: center;">Conclusion:</p> <p>For this question, the conclusion is the same as number 1, but for number 2, the answer given by ChatGPT was wrong from the beginning and it is clear that the final result is also wrong.</p>

Figure 9. The results of M4's analysis of the answers to questions 1 and 2

Based on the analysis given by M4 (see [Figure 9](#)), it is evident that M4 merely stated that the answer from ChatGPT was incorrect without identifying the error or providing the correct answer. Moreover, M4's analysis of ChatGPT's responses also did not show whether M4 understood and could answer the two mathematical questions. This suggests that M4 did not fully understand the instructions given by the researchers.

ChatGPT's answer to question number 1	ChatGPT's answer to question number 2
<p>Let's solve the system of equations to find the values of x and y. First, let's set up these equations: $4x - 5y = a$; $8x + 5y = 34$ Add the two equations to get rid of y: $4x - 5y + 8x + 5y = a + 34$ $12x = a + 34$ $x = \frac{a+34}{12}$ Now, we can find the value of x from the second equation: $8x + 5y = 34$ $8\left(\frac{a+34}{12}\right) + 5y = 34$ $\frac{2a+68}{3} + 5y = 34$ $2a + 68 + 15y = 102$ $2a + 15y = 34$ $y = \frac{34-2a}{15}$ To find the values that satisfy x and y as positive integers, we will try some possible values of a between 2 and 6 and check if x and y are positive integers. We are given that $x + a$ is a prime number between 2 and 6. Since x is a positive integer, and a is an integer, we can conclude that x should also be in the range of 2 to 6. Therefore, we can try values of a from 0 to 4. For each value of a, we will find the values of x and y and check if x and y are positive integers. Then we will check if $x + a$ is a prime number. After that, we can calculate $x - y$. $x = \frac{2+34}{12} = 3$; $y = \frac{34-2(2)}{15} = 2$ Both x and y are positive integers. Now let's check if $x + a$ is a prime number: $x + a = 3 + 2 = 5$ $x + a$ is a prime number. So, $x - y = 3 - 2 = 1$. So, for $a = 2$; $x - y = 1$ Since we have already found a solution, we don't need to test other values of a. So, $x - y = 1$ when $a = 2$.</p> <p>So, the conclusion I draw from the answer given by chat gbt is very helpful. however, we cannot just believe it, we must first prove the answer given, because it could be true or false. For the question above I have searched several times in the gbt chat and the answers are different, but the answer I took above is the easiest answer to understand and in accordance with my understanding.</p>	<p>To find the smallest p^2 value that satisfies the system of linear equations, we must first ensure that the determinant of the coefficient matrix is equal to zero. In this case, the coefficient matrix is: $\begin{bmatrix} p^2 - 1 & 1 \\ -2 & p^2 - 4 \end{bmatrix}$ The determinant of this matrix must equal zero: $(p^2 - 1)(p^2 - 4) - (-2)(1) = 0$ We solve the equation: $(p^2 - 1)(p^2 - 4) + 2 = 0$ Now we can simplify this expression: $p^4 - 5p^2 + 6 = 0$ We factor this expression: $(p^2 - 2)(p^2 - 3) = 0$ This provides two solutions for $\frac{p^2}{(p^2)} = 2$ dan $p^2 = 3$ However, we want the smallest p^2 value, which is $p^2 = 2$ Jadi, nilai p^2 terkecil yang memenuhi sistem persamaan linear tersebut adalah $p^2 = 2$. Conclusion number 2 is also the same as conclusion number 1 above. This is the answer that matches my understanding and is easy to understand. Apart from the gbt chat answers which are always different for each question.</p>

Figure 10. The results of M5's analysis of the answers to questions 1 and 2

M5 reviewed the answers after conducting several experiments until he found a response from ChatGPT that was easiest to understand and aligned with his understanding (see Figure 10). M5 also noted that while ChatGPT's answers were very helpful, they could not be fully accepted without conducting verification. M5 stressed the importance of validating the responses generated by ChatGPT, as they could be either correct or incorrect. However, M5's analysis only highlighted the answer considered correct, without addressing the previous answers considered incorrect.

Translation answer of M6's	
<p>1. $4x - 5y = a$ $y = \frac{a-4x}{-5}$ $8x + 5\left(\frac{a-4x}{-5}\right) = 34$</p> <p>$8x + 5y = 34$</p> <p>$8x + 4x - a = 34$</p> <p>$12x - a = 34$</p> <p>$x + a = 3$ $x + a = 5$</p> <p>$a = 3 - x$ $a = 5 - x$</p> <p>$12x - (3 - x) = 34$ $12x - (5 - x) = 34$</p> <p>$13x = 37$ $13x = 39$</p> <p>$x = 2,8$ $x = 3$ ✓</p> <p>$a = 5 - x$</p> <p>$= 5 - 3$</p> <p>$= 2$</p> <p>$4(3) - 5y = a$</p> <p>$12 - 5y = 2$</p> <p>$y = 2$ \longrightarrow $x - y = 3 - 2 = 1$</p>	<p>2. $(p^2 - 1)x + y = 0$ $-2x + (p^2 - 4)y = 0$</p> <p>$y = -(p^2 - 1)x$ $(p^2 - 4)y = 2x$</p> <p>$\frac{y}{x} = -(p^2 - 1)$ $\frac{y}{x} = \frac{2}{p^2 - 4}$</p> <p>$-(p^2 - 1) = \frac{2}{p^2 - 4}$</p> <p>$-(p^2 - 1)(p^2 - 4) = 2$</p> <p>Let $n = p^2$</p> <p>$-(n - 1)(n - 4) = -2$</p> <p>$-(n^2 - 5n + 4) = 2$</p> <p>$n^2 - 5n + 4 = \frac{2}{-1}$</p> <p>$n^2 - 5n + 4 = -2$</p> <p>$n^2 - 5n + 6 = 0$</p> <p>$(n - 3)(n - 2) = 0$</p> <p>$n = 3$ $n = 2$</p> <p>$p^2 = 3$ $p^2 = 2$ ✓</p>
<p>Conclusion:</p> <ul style="list-style-type: none"> The answer provided by Chat GPT is correct and matches my answer. The method used is also easy to understand. In my opinion, the answer given also uses the same method that I used 	

Figure 11. The results of M6's analysis of the answers to questions 1 and 2

M6 provided limited analysis as they simply displayed ChatGPT's answers along with their own (see Figure 11). M6 explained that ChatGPT's answers were correct and aligned with M6's own. The method used in ChatGPT's answers was also easy to understand and identical to the one that M6 employed. Unfortunately, M6's analysis was so minimal that it did not reflect M6's level of understanding of ChatGPT's responses.

Similar to M6, M7 provided only ChatGPT's answers alongside their own, which were followed by a brief conclusion. M7 mentioned that for question 1, ChatGPT's response aligned with their own answers in terms of the concept and method used. For question 2, M7 noted that the methods used were nearly identical, but the results differed. However, M7 did not offer a detailed review of the differences between ChatGPT's answers and their own. Although M7's own answers were correct, they could not identify and explain the specific differences or errors in ChatGPT's responses.

M8's analysis highlighted their positive attitude towards the use of ChatGPT. M8 explained that ChatGPT can sometimes be helpful, as it helps him find materials and learn mathematics, especially topics that are not well explained by lecturers or difficult to understand. However, M8 also pointed one of its disadvantages: the questions and answers provided by ChatGPT do not always align with the intended concept of the question. Regarding M8's experience with the two mathematical questions, they initially believed that the answers from ChatGPT were correct. M8's conclusion indicates a lack of understanding of the concept behind question 1. M8 argued that the answer given by ChatGPT was correct, despite the attached response from ChatGPT being incorrect. This clearly demonstrates that M8 was unable to analyze the answers obtained from ChatGPT, even though he acknowledged that ChatGPT's responses are sometimes was not in line with the intended concept.

3.2. Discussion

One crucial skill for mathematics students to master is the ability to justify statements, which is commonly referred to as mathematical justification. According to Fatmanissa et al. (2024), mathematical justification involves presenting arguments that support or challenge mathematical claims using mathematical reasoning and knowledge.

Based on the students' responses, it is clear that most students lack the necessary ability to justify the answers provided by ChatGPT. In general, the answers and explanations from the eight students were not consistent or entirely correct. For instance, some students considered the answers obtained from ChatGPT to be correct, even when those answers were inaccurate.

To our understanding, the weakness in students' mathematical justification abilities can be attributed to their limited understanding of relevant mathematical concepts and procedural skills. This is evident from the fact that some students do not verify the answers provided by ChatGPT and simply accept them as correct. According to Sarumaha and Rizkianto (2022), having a strong conceptual understanding and enhancing mathematical skills are critical for improving students' mathematical justification abilities. Sukirwan et al. (2020) argued that there are three factors that influence students' ability to justify mathematical arguments: their understanding of the claims, the treatment given, and the factual evidence in the arguments. Furthermore, students struggled to identify errors in ChatGPT's responses. We believe this is likely due to a lack of attention to detail, weak procedural skills, and failure to check whether the final answer actually solves the problem.

Students recognized that the explanations and answers generated by chatGPT were not always correct, but they were not fully able to identify the specific steps where ChatGPT's answers went wrong. In some cases, ChatGPT's answers could be trusted, while in others, they were incorrect. This is in line with the study by Rahman and Watanobe (2023), who noted that although ChatGPT is a powerful tool capable of generating impressive responses across various topics, it still has several limitations, such as a lack of logic, potential bias, difficulty with complex reasoning, and an inability to process visuals.

However, as Fijačko et al. (2023) pointed out, ChatGPT not only provides thorough answers to stand-alone questions, but also gives in-depth explanations to support the responses. While the references given by ChatGPT are often general, the rationale behind the answer is usually much more complex. Farhi et al. (2023) also concluded that the performance of ChatGPT varied across different subject domains, with outstanding results in several areas. However, they argued that its performance in mathematics was rather unsatisfactory. Therefore, the introduction of AI tools like ChatGPT in mathematics education is beneficial, but it needs to be accompanied by an awareness of the challenges that it presents (Govender, 2023).

The accuracy and effectiveness of ChatGPT's answers to mathematical questions can vary based on the complexity of the equation, the input data, and the prompts provided to ChatGPT (Wardat et al., 2023). In addition, Ojeda et al. (2023) concluded that ChatGPT is a powerful and impactful tool for teaching and learning in higher education. While ChatGPT offers both advantages and disadvantages, its effectiveness largely depends on how it is utilized. Hence, it is important to carefully consider the ethical, technological, and other

challenges associated with AI use in education and implement appropriate measures to address these challenges and ensure AI is used ethically and effectively (Zhai, 2022).

4. CONCLUSION

In conclusion, it is crucial to give more attention to students' mathematical justification abilities. This is evident from the tendency of students to rely on ChatGPT's answers without thoroughly checking them and their difficulty in identifying errors in the responses generated by the AI tool. We argue that the primary factors contributing to this lack in students' mathematical justification abilities include insufficient understanding of mathematical concepts, weak procedural skills, and a lack of attention to detail when verifying AI-generated answers. Although ChatGPT can certainly serve as a useful tool for learning mathematics, we caution that its limitations, particularly in logic, complex reasoning, and potential bias, must be carefully considered. Therefore, we believe that improving students' mathematical justification abilities should focus on strengthening their conceptual understanding, procedural skills, and critical thinking ability. This will equip them to independently assess and evaluate information, including that generated by AI tools such as ChatGPT.

Declarations

Author Contribution : JSAB: Conceptualization, Resources, Visualization, Writing - original draft, and Writing - review & editing; EYP: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Visualization, and Writing - original draft; APR: Writing - review & editing; SHN: Conceptualization, and Writing - review & editing; LML: Writing - review & editing.

Funding Statement : This research was not funded.

Conflict of Interest : The authors declare no conflict of interest.

Additional Information : Additional information is available for this paper.

REFERENCES

- Baidoo-anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Cheung, B. H. H., Lau, G. K. K., Wong, G. T. C., Lee, E. Y. P., Kulkarni, D., Seow, C. S., Wong, R., & Co, M. T.-H. (2023). ChatGPT versus human in generating medical graduate exam multiple choice questions—A multinational prospective study (Hong Kong S.A.R., Singapore, Ireland, and the United Kingdom). *PLoS One*, 18(8), e0290691. <https://doi.org/10.1371/journal.pone.0290691>
- Farhi, F., Jeljeli, R., Aburezeq, I., Dweikat, F. F., Al-shami, S. A., & Slamene, R. (2023). Analyzing the students' views, concerns, and perceived ethics about chat GPT usage.

Computers and Education: Artificial Intelligence, 5, 100180.
<https://doi.org/10.1016/j.caeai.2023.100180>

- Fatmanissa, N., Siswono, E., Yuli, T., Lukito, A., & Ismail, I. (2024). Utilización de una tarea de toma de decisiones: justificación matemática de los estudiantes en la resolución colaborativa de problemas [Utilizing decision-making task: Students' mathematical justification in collaborative problem solving]. *Avances de investigación en Educación Matemática*, 26, 85-103.
<https://doi.org/10.35763/aiem26.5341>
- Fijačko, N., Gosak, L., Štiglic, G., Picard, C. T., & John Douma, M. (2023). Can ChatGPT pass the life support exams without entering the American heart association course? *Resuscitation*, 185, 109732. <https://doi.org/10.1016/j.resuscitation.2023.109732>
- García-López, I. M., González González, C. S., Ramírez-Montoya, M.-S., & Molina-Espinosa, J.-M. (2025). Challenges of implementing ChatGPT on education: Systematic literature review. *International Journal of Educational Research Open*, 8, 100401. <https://doi.org/10.1016/j.ijedro.2024.100401>
- Govender, R. (2023). The impact of artificial intelligence and the future of ChatGPT for mathematics teaching and learning in schools and higher education. *Pythagoras*, 44(1), a787. <https://doi.org/10.4102/pythagoras.v44i1.787>
- Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 877-908). Springer Netherlands. https://doi.org/10.1007/978-94-009-1465-0_24
- Hershkowitz, R. (2020). Shape and space: Geometry teaching and learning. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 774-779). Springer International Publishing. https://doi.org/10.1007/978-3-030-15789-0_138
- Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. In *Data ethics : building trust : how digital technologies can serve humanity* (pp. 621-653). Globethics Publications. <https://doi.org/10.58863/20.500.12424/4276068>
- Iqbal, M., Khan, N. U., & Imran, M. (2024). The role of artificial intelligence (AI) in transforming educational practices: Opportunities, challenges, and implications. *Qlantic Journal of Social Sciences*, 5(2), 348-359.
<https://doi.org/10.55737/qjss.349319430>
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2), e0000198.
<https://doi.org/10.1371/journal.pdig.0000198>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- OECD. (2018). *Education at a Glance 2018: OECD Indicators*. OECD Publishing.
<https://doi.org/10.1787/eag-2018-en>
- OECD. (2022). *Trends Shaping Education 2022*. OECD Publishing.
<https://doi.org/10.1787/6ae8771a-en>

- Ojeda, A. D., Solano-Barliza, A. D., Alvarez, D. O., & Cárcamo, E. B. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria [Analysis of the impact of artificial intelligence ChatGPT on the teaching and learning processes in university education]. *Formación universitaria*, 16(6), 61-70. <https://doi.org/10.4067/S0718-50062023000600061>
- Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13(9), 5783. <https://doi.org/10.3390/app13095783>
- Santos, L., Mata-Pereira, J., da Ponte, J. P., & Oliveira, H. (2022). Teachers' understanding of generalizing and justifying in a professional development course. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(1), em2067. <https://doi.org/10.29333/ejmste/11488>
- Sarumaha, Y. A., & Rizkianto, I. (2022). Promoting mathematical justification through realistic mathematics education classroom. *Jurnal Pendidikan Matematika (JUPITEK)*, 5(2), 83-94. <https://doi.org/10.30598/jupitekvol5iss2pp83-94>
- Sukirwan, S., Muhtadi, D., Saleh, H., & Warsito, W. (2020). Profile of students' justifications of mathematical argumentation. *Infinity Journal*, 9(2), 197-212. <https://doi.org/10.22460/infinity.v9i2.p197-212>
- Vasconcelos, M. A. R., & Santos, R. P. d. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2296. <https://doi.org/10.29333/ejmste/13313>
- 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), em2286. <https://doi.org/10.29333/ejmste/13272>
- Wei, X., Chu, X., Geng, J., Wang, Y., Wang, P., Wang, H., Wang, C., & Lei, L. (2024). Societal impacts of chatbot and mitigation strategies for negative impacts: A large-scale qualitative survey of ChatGPT users. *Technology in Society*, 77, 102566. <https://doi.org/10.1016/j.techsoc.2024.102566>
- Welding, L. (2023). *Half of college students say using AI on schoolwork is cheating or plagiarism*. BestColleges. Retrieved from <https://www.bestcolleges.com/research/college-students-ai-tools-survey>
- Xu, X., Su, Y., Zhang, Y., Wu, Y., & Xu, X. (2024). Understanding learners' perceptions of ChatGPT: A thematic analysis of peer interviews among undergraduates and postgraduates in China. *Heliyon*, 10(4), e26239. <https://doi.org/10.1016/j.heliyon.2024.e26239>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X. (2022). *ChatGPT: Artificial Intelligence for Education*. Retrieved from <https://doi.org/10.13140/RG.2.2.35971.37920>