



Jurnal Pendidikan Fisika

<https://journal.unismuh.ac.id/index.php/jpf>

DOI: 10.26618/jpf.v13i1.16624



Case Study on ChatGPT's Performance in Assisting Students with Physics Tests

Innal Mafudi^{1,2)}, Heru Kuswanto¹⁾, Jumadi¹⁾, Intan Fatmawati³⁾

¹⁾Physics Education Department, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia

²⁾Physics Education Department, Universitas PGRI Madiun, Madiun, 63118, Indonesia

³⁾Vocational High School of Bodronoyo, Madiun, 63161, Indonesia

*Corresponding author: innalmafudi.2023@student.uny.ac.id

Received: September 02, 2024; Accepted: December 24, 2024; Published: January 20, 2024

Abstract – The rapid development of artificial intelligence (AI), particularly ChatGPT, has sparked interest in its application in education. This study aims to investigate the potential of ChatGPT in helping students understand and solve physics problems, focusing on the Test of Understanding Graphs in Kinematics and the Determining and Interpreting Resistive Electric Circuit Concepts Test. The study involved 25 physics education students who completed these tests independently and with ChatGPT's assistance. The results revealed that students with a strong foundational understanding and reflective abilities interacted more effectively with ChatGPT, leading to improved answers and deeper conceptual understanding. In contrast, students with weaker prior knowledge tended to accept ChatGPT's answers without critical reflection, perpetuating errors. Furthermore, ChatGPT showed limitations in interpreting image-based questions, reading scales, and providing consistent responses to concept-specific queries. These findings suggest that while ChatGPT has the potential to enhance learning, it requires thoughtful integration, particularly in helping students develop critical thinking and problem-solving skills. Teachers are encouraged to use ChatGPT's limitations to design assessments that minimize the risk of cheating and foster deeper understanding. In conclusion, this study underscores the importance of combining AI tools with strong conceptual foundations and active reflection to optimize learning outcomes in physics education. Future research should focus on refining strategies for using AI in education to address its current limitations and enhance its effectiveness in complex learning scenarios.

Keywords: artificial intelligence; chatGPT; physics concept understanding

© 2025 Physics Education Department, Universitas Muhammadiyah Makassar, Indonesia.

I. INTRODUCTION

The rapid development of artificial intelligence (AI) technology in recent years has significantly impacted education, particularly in the domain of physics learning. One such AI platform, ChatGPT, offers substantial potential for enhancing the effectiveness of teaching and learning processes (Polverini & Gregorcic, 2024). Commonly used as a virtual assistant, ChatGPT aids in understanding subject material, provides personalized feedback, and supports learning by

offering relevant and accurate information (Akavova et al., 2023; Hashem et al., 2023). Additionally, AI technologies like ChatGPT have expanded their functions to include data analysis, essay writing, and conducting Socratic-style dialogues on introductory physics topics (Abbas et al., 2023; Wang et al., 2023).

The emergence of ChatGPT as a new educational tool is expected to bring positive effects to physics learning. Interviews with 12 prospective physics teacher candidates revealed that AI platforms are frequently used, with ChatGPT being the most commonly used tool among them. These students employ AI to assist in creating papers, reports, proposals, answering practice questions, and seeking learning resources. However, the answers provided by ChatGPT sometimes require further clarification to meet the desired level of accuracy. This indicates that while ChatGPT has potential, its responses can sometimes be problematic and require additional scrutiny before adoption (Hikmatiar et al., 2024).

In physics education, AI has demonstrated a positive impact, especially in designing learning units and creating assessment rubrics that cater to student needs (Cooper, 2023). AI can also analyze data, providing valuable insights for teachers to make informed decisions about students' learning progress (Kasepalu et al., 2022). However, ChatGPT's performance in solving physics-related problems varies significantly, ranging from providing inaccurate responses comparable to low-performing students, to delivering high-quality answers akin to experts (Kim, 2023). For example, in the context of the Force Concept Inventory (FCI) test, ChatGPT offered correct answers to some questions, while it struggled with more complex concepts (Eaton & Willoughby, 2018; Henderson & Stewart, 2018). This highlights both the promise and the challenges of using AI for educational purposes.

Furthermore, ChatGPT faces difficulties when dealing with non-textual information. Tasks involving images must often be transcribed into text before they can be processed by ChatGPT, presenting a limitation in solving certain types of physics problems (Abbas et al., 2023). The performance of AI platforms like ChatGPT still requires further evaluation, particularly in handling verbal, algebraic, and complex tasks, such as programming or data analysis (Kim, 2023). There are concerns about plagiarism and cheating, where students may use ChatGPT to generate answers without understanding the underlying concepts (Khan et al., 2023; Hoa, 2023). This reliance on AI may also reduce critical thinking and analytical skills among students (Farrokhnia et al., 2023). Therefore, it is essential to develop strategies that ensure the ethical and responsible use of AI in education.

While many studies have explored the advantages and limitations of AI in education, few have specifically addressed ChatGPT's performance in solving physics problems, particularly those involving visual representations such as graphs or diagrams. Additionally, there is a lack of

research on how factors such as prior knowledge and reflective ability influence user interaction with ChatGPT. Thus, this study aims to explore the potential and limitations of ChatGPT in physics education, particularly in the context of helping students' complete physics concept understanding tests. By assessing ChatGPT's performance, the study seeks to identify strategies for integrating AI into physics education while addressing its weaknesses. The study will also provide insights for educators on how to design assessments that minimize the risk of cheating.

II. METHODS

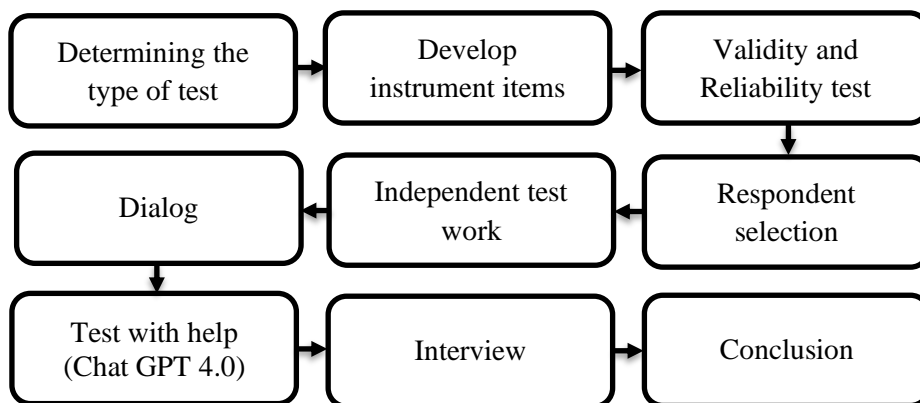
The research method used in this study was descriptive-qualitative, with the case study conducted in October 2024. A total of 25 physics education students participated as respondents, selected using a simple random sampling technique. The test instruments were adapted from the Test of Understanding Graphs in Kinematics (TUG-K) and the Interpreting Resistive Electric Circuit Concepts Test (DIRECT), each consisting of 1 question. The TUG-K test instrument is a diagnostic tool designed to assess proficiency in one-dimensional kinematics concepts (Beichner, 1994; Zavala et al., 2017). This instrument has undergone several versions since its inception in 1994, with the most recent update in 2017. For this study, we adapted the TUG-K 4.0 instrument from 2017, specifically item number 4 from the English version. The DIRECT test instrument is a diagnostic tool used to measure understanding of direct current electrical circuit concepts (Engelhardt & Beichner, 2004). This test comprises 29 questions and has multiple versions. In this study, we adapted the DIRECT 1.0 version, item number 4, English version, as it was more suitable for a qualitative approach. The adapted test instrument consisted of 2 questions, which were then presented as open-ended questions in a paper-and-pencil format. This modification aimed to obtain a work process that would be used as analysis material.

The test instrument was validated through expert judgment by 5 experts, resulting in a V'Aiken score of 0.93, which falls under the valid category. An empirical test was then conducted involving 18 student respondents. The empirical test data were analyzed using Rasch modeling with the help of the Winstep application. The test reliability score, based on the Cronbach Alpha (KR-20) Person Raw Score "Test" Reliability value, was 0.51, which is categorized as low. Additionally, all questions functioned properly for measurement purposes (item fit). The item fit criteria included the following: OUTFIT MNSQ values between $0.5 < \text{MNSQ} < 1.5$, OUTFIT ZSTD values between $-2.0 < \text{ZSTD} < +2.0$, and PT-Mean Corr values between $0.4 < \text{PT-Mean Corr} < 0.85$. The results of the item fit measurement can be seen in Figure 1.

ENTRY	TOTAL	TOTAL		MODEL	INFIT	OUTFIT	PT-MEASURE	EXACT MATCH						
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	Item	
2	9	18	.69	.87	1.00	.1	1.00	.1	.82	.82	66.7	66.7	Q2	
1	11	18	-.69	.87	1.00	.1	1.00	.1	.81	.81	66.7	66.7	Q1	
MEAN	10.0	18.0	.00	.87	1.00	.1	1.00	.1			66.7	66.7		
S.D.	1.0	.0	.69	.00	.00	.0	.00	.0			.0	.0		

Figure 1. Item fit order

The data collection process began with giving the test questions to the 25 respondents, who were instructed to complete them manually within 10 minutes. Next, an exploration was conducted with several purposively selected respondents to engage in an interactive dialogue. The goal was to clarify the results of their work, explore their thinking process, and understand the reasons behind the answers provided. After the dialogue process, the respondents were asked to complete the same test questions with the help of Chat GPT 4.0 within 5 minutes. The final stage involved conducting interviews with the respondents to assess their confidence in the results, comparing their manually completed answers with those generated using Chat GPT 4.0. The complete design flow of this study can be seen in Figure 2.

**Figure 2.** Research flow diagram

III. RESULTS AND DISCUSSION

Investigation of Case Question Number 1

In the first case, all respondents were presented with an open-ended test question adapted from TUG-K, as shown in Figure 3. The question displays a velocity versus time (v vs t) graph, and the respondents are asked to calculate the distance traveled by an elevator with a mass of 1000 kg over 3 seconds.

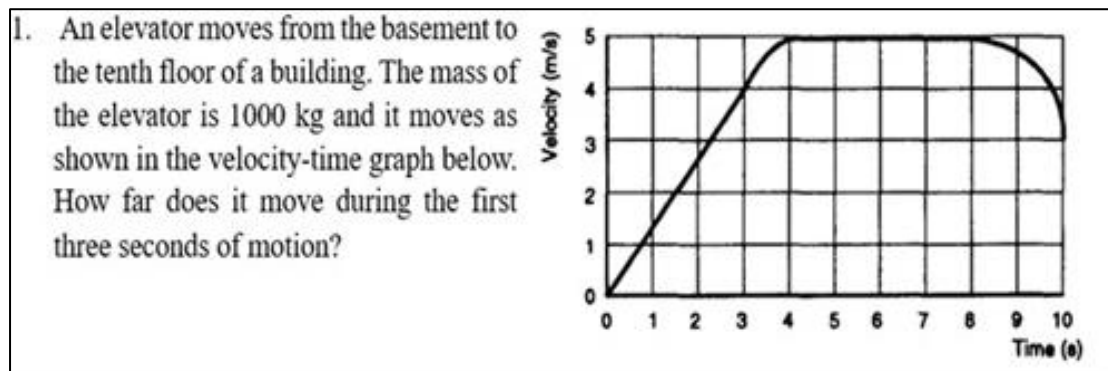


Figure 3. Question number 1 of the TUG-K test

Based on the total answers collected from the respondents, two answers were selected for further analysis because they represent variations in understanding physics concepts and the application of basic concepts in answering questions. In the answers from Respondents A and B, it was evident that both had difficulty understanding the basic concepts needed to solve problems based on velocity-time graphs. Respondent A used the distance calculation approach with the equation $s=Vt$, which is valid only at constant speed. Respondent A did not notice that the graph showed acceleration with a changing speed during the first three seconds. The study's results also revealed that 72% of introductory physics students made mistakes when answering this question (Maries & Singh, 2013). Difficulty distinguishing the meaning of position, velocity, and acceleration versus time graphs has been reported (Beichner, 1994). The failure to interpret kinematic graphs is due to a lack of contextual knowledge in mathematics and physics, which are fundamental for understanding graphic representations (Phage et al., 2017). A similar issue in Indonesia was found with errors in using formulas to solve kinematic problems (Amin et al., 2020). Respondent B also made an error in applying the concept. Respondent B included mass in the calculation, even though mass is irrelevant in this context because it does not affect the distance traveled in the velocity-time graph. Difficulty understanding kinematics is a common issue (Phage, 2018; Suganda et al., 2020; Warsono et al., 2020). Respondents' answers in completing the TUG-K open test questions are presented in Figures 4a and 4b.

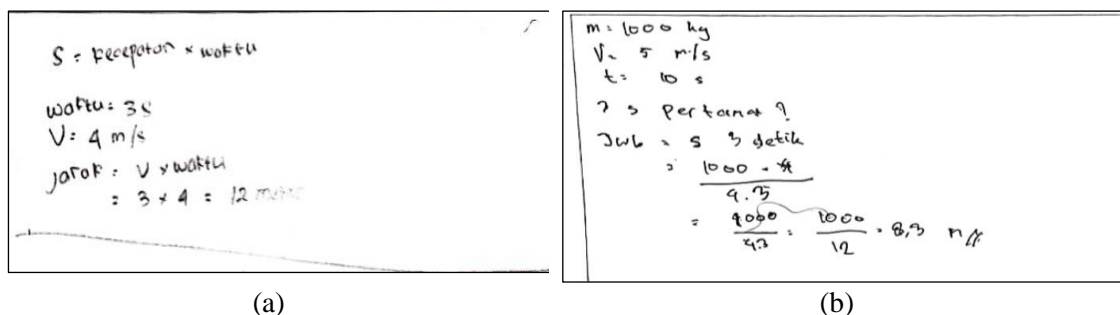


Figure 4. Answers manual work on question number 1 (a) respondent A and (b) respondent B

Respondent A has a partial understanding of the basic concept of calculating distance because they only know the basic formula $s = v \cdot t$ to calculate distance. However, they do not understand that this formula only applies at constant speed and ignore the change in speed shown in the graph. The issue where students often fail to link the concept of acceleration with the graph depicting speed changes was also revealed (Sutopo et al., 2017). On the other hand, Respondent B showed a more fundamental misunderstanding. Respondent B involved mass in calculating distance, which was irrelevant for the velocity-time graph in this context. The case with Respondent B proposing an alternative concept often hindered their understanding (Linuwih, 2013). This finding shows that Respondent B does not understand the relationship between the concepts of mass, speed, and distance in physics. These results are further reinforced by the transcription of the dialogue conducted with Respondents A and B, shown in Table 1.

Table 1. Transcription of the dialogue with respondents discussing question number 1

Respondent A	Respondent B
<p>Dialogue:</p> <p>P: Try to explain why you can get the answer 12 meters.</p> <p>Ra: I worked by looking at the unit of speed, which is (m/s); based on that unit, I determined the formula to answer the question of the distance traveled by the elevator in 3 s. Then, I drew a conclusion based on the graph: everyone travels a distance of 4 meters. So I multiplied $3 \times 4 = 12$ meters.</p> <p>P: So what concept of motion are you applying to this question, uniform linear motion or uniformly accelerated motion?</p> <p>Ra: If based on the formula I use, this is uniform linear motion.</p> <p>P: Does that mean uniform linear motion?</p> <p>Ra: Yes, because it was constant earlier.</p> <p>P: If you look at this graph, is the speed constant or not?</p> <p>Ra: No, if from zero to 3 seconds, it is not constant because the speed increases, so it is not constant.</p> <p>P: If it is like that, should it be uniform linear motion or uniformly accelerated motion?</p> <p>Ra: uniformly accelerated motion.</p>	<p>Dialogue:</p> <p>P: Try to explain where you got the answer 8.3 meters.</p> <p>Rb: I divide the elevator mass of 1000 kg by this with v.s, which is $1000/4.3 = 8.3$ meters</p> <p>P: Where did the equation you used come from?</p> <p>Rb: I looked at the known data in the question.</p> <p>P: Do you know what concept the question uses?</p> <p>Rb: I don't know.</p> <p>P: Have you ever studied the concepts of uniform linear motion and uniformly accelerated motion?</p> <p>Rb: I did during the basic Physics 1 lecture, but I forgot.</p> <p>P: Is mass an influential variable in the concept of kinematics?</p> <p>Rb: I don't know.</p> <p>P: Do you want to try working on it with the help of GPT Chat?</p> <p>Rb: Yes. I will try working on it with the help of GPT Chat.</p>

P: Oh. Does that mean uniformly accelerated motion? Does that mean if you use the equation $s=v.t$, is it correct or not? Or is there another equation that can be applied to solve this problem?

Ra: Wrong

P: Does that mean you applied the concept wrong? In the uniformly accelerated motion case, you should apply the uniform linear motion concept equation.

Rb: Yes, I misapplied the concept when working on the problem.

P: Do you want to fix it by applying the uniformly accelerated motion concept?

Rb: Yes, but the problem is I forgot the equation

P: Okay, please use GPT Chat to help

*Note: P (Researcher), Ra (Respondent A), Rb (Respondent B)

The dialogue shows differences in understanding the concept of kinematics between Respondents A and B and highlights several basic conceptual errors. Both respondents were then advised to use ChatGPT to obtain a structured explanation, as done by Kasepalu et al. (2022) and Wang et al. (2023). The results of the work of the two respondents with the help of ChatGPT are shown in Figure 5.

To determine how far the elevator moves during the first three seconds, we can analyze the velocity-time graph.

The area under the velocity-time graph represents the distance traveled. From 0 to 3 seconds, the graph shows a linear increase in velocity, indicating uniform acceleration.

- The graph shows that the velocity increases linearly from 0 to 5 m/s in 3 seconds.
- Since the section is a triangle with a base of 3 seconds and a height of 5 m/s, we can calculate the area (and thus the distance) as:

$$\text{Distance} = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 3 \text{ s} \times 5 \text{ m/s}$$

Let's calculate this.

The elevator moves 7.5 meters during the first three seconds of motion. [↩]

why are the results different instead of the speed equation $s=vt$ so $4 \times 3 = 12 \text{ m/s}$

The reason why we don't use the simple equation $s = v \times t$ in this case is because the velocity of the elevator is not constant over the first 3 seconds. The equation $s = v \times t$ applies only when the object moves at a constant velocity.

In this scenario:

- The velocity starts at 0 m/s and increases linearly to 5 m/s over the first 3 seconds, which means the elevator is accelerating.
- When velocity changes uniformly like this, the distance traveled is actually the area under the velocity-time graph.

Since the shape under the curve is a triangle, we used the area of a triangle formula:

$$\text{Distance} = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 3 \text{ s} \times 5 \text{ m/s} = 7.5 \text{ m}$$

So, the distance is 7.5 meters, not 12 meters, due to the initial acceleration.

oh yes that's right, but on the graph at $t = 3$ seconds the value of $v = 4 \text{ m/s}$ and apply using the glbb equation

< 2/2 >

You're correct that we can also approach this problem using kinematic equations for uniformly accelerated motion (also known as "GLBB" or "Gerak Lurus Berubah Beraturan" in Indonesian). Given the graph, we see that at $t = 3$ seconds, the velocity v is 4 m/s.

Steps to solve using kinematic equations:

- Determine the acceleration:** Since the velocity increases linearly from 0 to 4 m/s in 3 seconds, we can find the acceleration a using:
$$a = \frac{\Delta v}{\Delta t} = \frac{4 \text{ m/s} - 0 \text{ m/s}}{3 \text{ s}} = \frac{4}{3} \text{ m/s}^2$$
- Use the kinematic equation for distance:** Since the elevator starts from rest, the distance s traveled in time $t = 3$ seconds can be calculated using:
$$s = v_0 t + \frac{1}{2} a t^2$$

where:

- $v_0 = 0 \text{ m/s}$ (initial velocity),
- $a = \frac{4}{3} \text{ m/s}^2$,
- $t = 3 \text{ s}$.

Let's calculate this distance.

It seems like I can't do more advanced data analysis right now. Please try again later.

However, you can calculate it manually as follows:

$$s = v_0 t + \frac{1}{2} a t^2$$

Substituting the values:

$$s = 0 + \frac{1}{2} \cdot \frac{4}{3} \cdot (3)^2 = \frac{1}{2} \cdot \frac{4}{3} \cdot 9 = 6 \text{ meters}$$

So, the elevator moves 6 meters during the first three seconds. [↩]

(a)

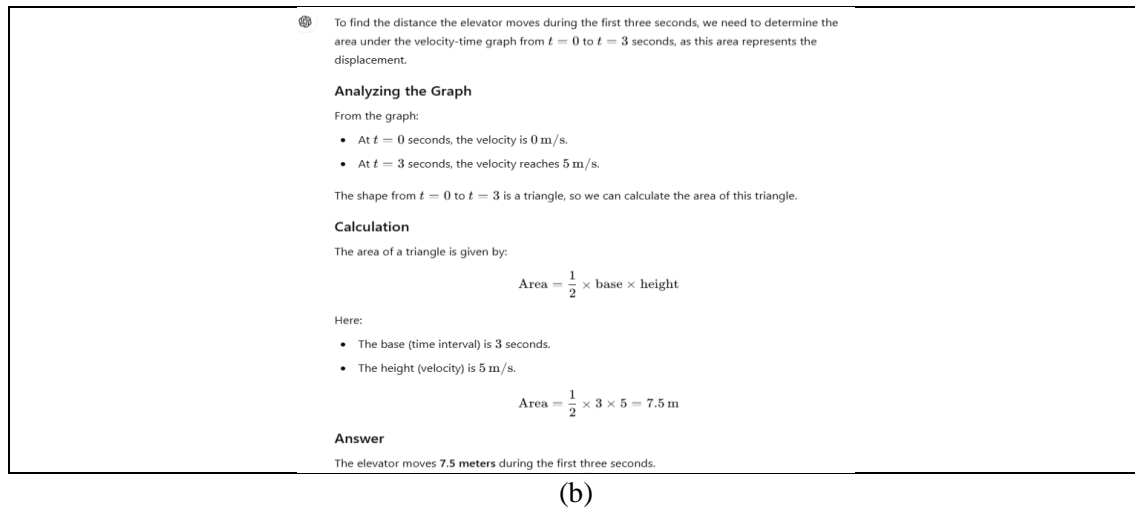


Figure 5. Answers from Chat GPT question number 1 (a) respondent A and (b) respondent B

The difference between Respondent A and Respondent B in using ChatGPT as an aid can be seen in how they understand the concept and interact with the platform. After being assisted by ChatGPT, Respondent A attempted to understand his mistake and even asked for an alternative explanation so that ChatGPT could provide a solution using the uniformly accelerated linear motion equation. This more critical interaction shows that Respondent A actively uses ChatGPT to deepen his understanding of the concept. On the other hand, Respondent B showed a weaker understanding from the outset, including errors in understanding basic kinematic concepts, such as associating mass with irrelevant calculations. When using ChatGPT, Respondent B tended to accept answers without questioning or seeking alternative approaches, so their interactions were limited to understanding the basic steps presented by ChatGPT. As a result, even though ChatGPT assistance was provided, Respondent B's answers remained incorrect. It can be concluded that using ChatGPT assistance to work on problems involving complex concepts is complicated and yields different results (Eaton & Willoughby, 2018; Henderson & Stewart, 2018). Other studies also confirm that initial conceptual knowledge of physics and critical thinking are essential when using ChatGPT (Hidaayatullaah, 2022; Kasepalu et al., 2022). These results were further validated through interviews, as shown in Table 2.

Table 2. Transcription of interviews with respondents for question no. 1

Respondent A	Respondent B
P: Which version of the answer are you sure is correct?	P: Which version of the answer are you sure is correct?
Ra: The answer from the work with the help of GPT Chat.	Rb: Answer from ChatGPT.
P: Do you want to change the results of your work at the beginning?	P: What makes you sure about ChatGPT's answer?
Rb: Yes. I will adjust it to the uniformly accelerated motion	Even though the answer does not use the concept of uniform linear motion or uniformly accelerated motion equations, it instead solves it with the concept of the area of a triangle.

concept based on GPT Chat suggestions.	Rb: I don't know what I should ask ChatGPT. P: Do you want to change the answer to ChatGPT? Rb: Yes, I will change it.
--	--

*Note: P (Researcher), Ra (Respondent A), Rb (Respondent B)

This interview highlights important differences in understanding and response between Respondent A and Respondent B to the assistance provided by ChatGPT. Respondent A showed better understanding and reflective ability, not only accepting the answer from ChatGPT but also recognizing the need to change the concept used from uniform linear motion to uniformly accelerated motion. He actively stated that he would adjust his answer according to ChatGPT's suggestions, indicating that he understood and adopted the new concept to improve his work. This phenomenon demonstrates that reflective ability and active involvement in the learning process are essential for a better understanding of the concept. In contrast, Respondent B tended to accept the answer from ChatGPT without fully understanding it. Although ChatGPT used a triangle area method, which differs from the uniform linear motion or uniformly accelerated motion concept, Respondent B did not explore further or inquire about the reasoning behind the method. As a result, the answers provided by both respondents show that Respondent A changed his answer (correct), while Respondent B's answer remained incorrect. See Figures 6a and 6b.

$s = \frac{1}{2} a t^2$
 $s = \frac{1}{2} \cdot 4 \cdot 3^2$
 $s = 6 \text{ meter}$
 $3s = 4$
 $a = \frac{4}{3} = 1,3$

(a)

$s = \frac{1}{2} a \cdot t$
 $= \frac{1}{2} \cdot 3 \cdot 5$
 $= 7.5 \text{ Meter}$
 Jarak yang ditempuh elevator adalah 7.5 meter

(b)

Figure 6. answers after interaction with ChatGPT (a) Respondent A and (b) Respondent B

Investigation on case question number 2

The next investigative activity involved a series and parallel circuit case adapted from the DIRECT test, as shown in Figure 7. In this question, four circuits with two light bulbs were presented, and respondents were asked to determine which circuit was parallel to the battery.

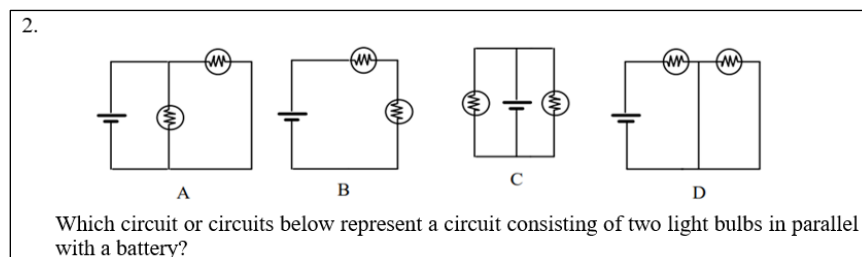


Figure 7. Question number 2 test (DIRECT)

The answers shown in Figure 8 reveal that Respondent A chose circuits A and D as circuits with two light bulbs in parallel with the battery. In contrast, Respondent B only chose circuit A

because the two light bulbs had circuit paths that were not in the same direction. The analysis of the results shows that both Respondents A and B did not fully understand the concept. Other studies have revealed that errors in identifying parallel circuits are often caused by a lack of understanding of basic concepts, errors in reading circuit diagrams, and an inability to consistently apply current and voltage rules (Engelhardt & Beichner, 2004). The lack of understanding of basic concepts, weaknesses in electrical topology knowledge, and the use of irrelevant information in circuit analysis have also been identified as causes of student errors (Ivanjek et al., 2021).



Figure 8. Answers to manual work on question number 2 (a) respondent A and (b) respondent B

In depth analysis was conducted to reveal the respondents' conceptions through a dialogue, as shown in Table 3. The results of the dialogue between the researcher and Respondents A and B showed differences in their levels of understanding of the electric circuit concept. Respondent A initially answered that circuits A and D were parallel. However, after being tested using the thought experiment method by removing one of the light bulbs, Respondent A realized his mistake. He recognized that Circuit D did not meet the parallel criteria. Respondent A then concluded that the correct answer was circuits A and C. Meanwhile, Respondent B answered that only circuit A was parallel because the two light bulbs had paths that were not in the same direction. However, Respondent B did not know other characteristics of parallel circuits, such as the criterion for the independence of the electric current path. After being tested using the thought experiment method, Respondent B maintained his answer and was unable to identify Circuit C as a parallel circuit.

Table 3. Transcription of dialogue with respondents discussing question no. 2

Respondent A	Respondent B
Dialogue: P: In the question, there are four types of circuits consisting of cables, two light bulbs, and one battery. Which circuit do you think is a parallel circuit? Pa: Circuits A and D. P: Why circuits A and D? Pa: As seen from the picture, circuit A is clearly not in one path, so it is clearly a parallel circuit; Circuit B light bulbs are arranged in one path, so it is a series circuit; Circuit C is also the same as B, still in one battery current so it is a series circuit, Circuit D is	Dialogue: P: In question number 2, there are four different types of circuits; from your answer, Circuit A is a parallel circuit because the two light bulbs have circuit paths that are not unidirectional. Pb: Yes, that is right. P: Are there no parallel circuits other than circuit A Pb: None, because circuits B, C, and D light bulbs are connected in a unidirectional circuit path.

<p>parallel because the path between the two light bulbs is cut.</p> <p>P: Are you sure Circuit D is parallel?</p> <p>Pa: God willing</p> <p>P: Are you sure or not?</p> <p>Pa: Sure.</p> <p>P: Let's test it, for example, circuit A, we give the vertical light bulb (bulb) a sign as bulb 1 and the horizontal light bulb we call bulb 2, and the characteristic of a parallel circuit is that if one light bulb is taken, the other light bulb will still be on.</p> <p>Pa: Agreed.</p> <p>P: For example, in circuit A, I take bulb 1. Will bulb two light up?</p> <p>Pa: It lights up because it is still connected.</p> <p>Q: Now, in circuit B, if I remove one of the bulbs, will the other bulbs still light up?</p> <p>Pa: Off, Because in one path (series circuit)</p> <p>P: Now, in circuit C, if I take one of the bulbs, will the other bulbs light up?</p> <p>Pa: On</p> <p>P: Huh.. does that mean circuit C is a parallel circuit?</p> <p>Pa: Yes Parallel.</p> <p>P: Circuit D, if I take one of the bulbs, will the other bulbs light up?</p> <p>Pa: On.</p> <p>P: Sure, it lights up; try checking again. Will the current still be connected to the (-) battery terminal if the bulb is taken?</p> <p>Pa: Oh, yes. the bulb will turn off.</p> <p>P: So your answer was right or wrong?</p> <p>Pa: Correct in circuit A only.</p> <p>P: So the correct answer is?</p> <p>Pa: Circuits A and C</p> <p>P: Do you want to test it with the help of Chat GPT?</p> <p>Pa: Yes. I will do it.</p>	<p>P: Oh, so what are circuits B, C, and D called?</p> <p>Pb: Series Circuit</p> <p>P: All series circuits, huh?</p> <p>Pb: All series circuits.</p> <p>P: What types of circuits are there?</p> <p>Pb: Series and parallel circuits</p> <p>P: Oh.okay. Besides unidirectional and non-unidirectional paths, are there other characteristics that can indicate a series or parallel circuit?</p> <p>Pb: I don't know.</p> <p>P: Let's test it, for example, in circuit A, we give the vertical light bulb (bulb) a sign as bulb 1 and the horizontal light bulb we call bulb 2, for example, if I take bulb one, will bulb 2 still light up?</p> <p>Pb: It lights up. Q: What about series B, C, and D?</p> <p>Pb: The lights will go off.</p> <p>P: Oh, okay. Do you want to try working on it with the help of GPT Chat?</p> <p>Pb: Yes. Okay, I will try.</p>
--	--

*Note: P (Researcher), Ra (Respondent A), Rb (Respondent B)

After the dialogue, both respondents were advised to use ChatGPT to correct their conceptual errors. The goal was to help them obtain more structured guidance. The results of the two respondents' work with ChatGPT are shown in Figure 9.

The question is asking which circuit shows two light bulbs in parallel with a battery.

- Circuit A:** Two light bulbs are connected in parallel with each other and with the battery.
- Circuit B:** Two light bulbs are connected in series with each other, not in parallel.
- Circuit C:** Two light bulbs are connected in parallel with the battery.
- Circuit D:** Two light bulbs are connected in series with each other, not in parallel.

So, the correct answer is **Circuits A and C**. Both circuits A and C represent two light bulbs connected in parallel with a battery.

Why is the circuit c parallel? even though the positive and negative poles of both light bulbs are on one cable

In Circuit C, although it may look like both light bulbs are connected directly across the battery without separate branches, they are actually connected in parallel because each bulb has its own direct path to the battery's terminals.

In a parallel circuit:

- Both components (light bulbs, in this case) have both ends connected to the same two points in the circuit, specifically the positive and negative terminals of the battery.
- This means each bulb operates independently, as the current splits at the junction before each bulb and recombines afterward.

In Circuit C:

- Both light bulbs have one side connected to the positive terminal of the battery and the other side connected to the negative terminal.
- This allows each bulb to receive the full voltage of the battery, characteristic of a parallel connection.

Therefore, Circuit C is indeed a parallel circuit despite the appearance of having only two wires, as the key to parallel circuits is that each component has a direct connection across the battery's terminals.

I usually see in books that the form of a d circuit is parallel

You're correct that, at first glance, Circuit D might look like it could be a parallel circuit because it has two paths branching from the battery. However, in this particular case, Circuit D is actually a **series circuit**, not a parallel one. Here's why:

In a **parallel circuit**, each component should have a direct path to the battery's terminals, allowing current to split into separate branches. This means:

- Each light bulb (or component) has its own, independent path from the positive terminal to the negative terminal.
- The voltage across each component in a parallel circuit is the same as the battery's voltage.

Why Circuit D is Series

In Circuit D:

- The two light bulbs are connected in a single loop, one after the other.
- There is only one path for current to follow, moving through one bulb and then the other before returning to the battery.

So, in Circuit D, the current flows through the first light bulb and then through the second, which is a hallmark of a **series circuit**. The bulbs do not each have independent paths to the battery, as they do in parallel circuits.

Why Circuit C is Parallel

In Circuit C, each light bulb is connected directly across the battery terminals, providing two separate paths for current. Each bulb receives the full voltage of the battery, which matches the structure of a parallel circuit.

In summary:

- Circuit D** is a **series circuit** because the current flows sequentially through each bulb in a single loop.
- Circuit C** is a **parallel circuit** because each bulb is independently connected across the battery, forming two parallel branches.

(a)

The question asks which circuit(s) represent two light bulbs in parallel with a battery. Let's analyze each option:

- Circuit A:** This shows two light bulbs connected in parallel with each other, and the parallel combination is connected to the battery. **This is a correct answer.**
- Circuit B:** This shows two light bulbs connected in series with the battery, not in parallel. **This is not correct.**
- Circuit C:** This circuit has two batteries connected in parallel with each other, with each light bulb connected to a separate battery. **This does not match the description of two bulbs in parallel with a single battery.**
- Circuit D:** This circuit shows two light bulbs in parallel with each other, connected to a single battery. **This is also a correct answer.**

Answer: Circuits A and D represent two light bulbs in parallel with a battery.

(b)

Figure 9. Answers from Chat GPT question number 2 (a) respondent A and (b) respondent B

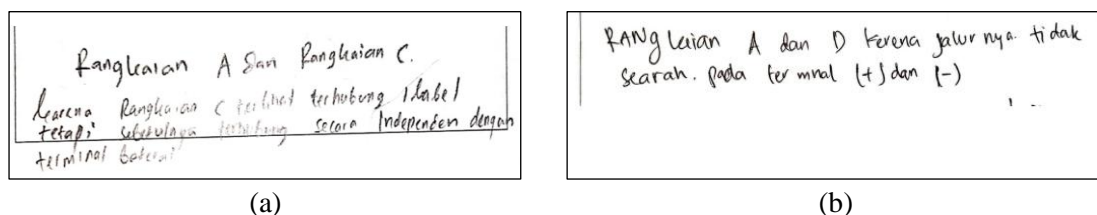
The results of the interaction between Respondents A and B with ChatGPT showed differences in their levels of understanding of parallel and series circuits. Respondent A received a detailed explanation from ChatGPT that Circuits A and C are parallel because each light bulb has an independent current path to the battery terminal. ChatGPT also explained that Circuit D is a series circuit because the current flows sequentially without branching. This active interaction helped Respondent A understand the difference between parallel and series circuits. On the other hand, the interaction between Respondent B and ChatGPT was more passive. Respondent B simply accepted the clarification that the correct answer was circuits A and D as parallel circuits. However, this answer was wrong because Circuit D is a series circuit. The difference in answers provided by ChatGPT to the two respondents confirmed that this technology struggles with solving problems involving images (Remoto, 2023). These results were later confirmed through interviews, as shown in Table 4.

Table 4. Transcription of interviews with respondents for question number 2.

Respondent A	Respondent B
P: Are you sure about the answer given by Chat GPT?	P: Which version of the answer are you sure is correct?
Ra: Yes, I am sure about the answer from the work with the help of Chat GPT.	Rb: Answer from GPT Chat.
P: What makes you sure?	P: What makes you sure?
Rb: Because the answer given by Chat GPT has been confirmed with several questions to confirm the answer is in accordance with the concept that I understand.	Rb: Because GPT Chat also states that circuit A is parallel and in GPT Chat it also says that Circuit D is also a parallel circuit.
P: Will you adjust your answer at the beginning?	P: Have you changed your mind to change your answer to circuit A and D?
Pb: Yes, I will adjust it.	Pb: Yes, I will change it
	P: Why not parallel circuit C?
	Pb: Because circuit C is connected 1 path on the (+) and (-) terminals

*Note: P (Researcher), Ra (Respondent A), Rb (Respondent B)

From the interview, Respondent A stated that he was confident in the answer given by ChatGPT but still showed openness to confirm the answer and align it with the concept he understood. In contrast, Respondent B expressed confidence in ChatGPT's answer, even though there was an error in understanding the concept of parallel circuits, especially in Circuit C. Both respondents then revised their answers according to the results suggested by ChatGPT, as seen in Figure 10. This interaction shows that ChatGPT helps respondents confirm their answers but is not entirely practical, especially if the user does not have a strong conceptual foundation. This highlights the need for further assistance to validate the answers provided by ChatGPT.

**Figure 10.** Answers after interaction with ChatGPT (a) respondent A and (b) respondent B

The results of this study found that students still have difficulty understanding and applying the concepts of kinematics, as well as the concepts of parallel and series electrical circuits. In question number 1, Respondents A and B showed a fundamental difference in their understanding of how to complete the velocity-time graph. Respondent A used an equation that only applies to uniform linear motion, without considering the changes in velocity on the graph, thus failing to apply the concept of acceleration. Respondent B included mass in the calculation, which is irrelevant for determining the distance on the velocity-time graph. In question number 2, both respondents made mistakes in identifying parallel circuits. Respondent A initially stated that Circuit D was parallel but managed to correct his answer after being given the thought experiment

method. In contrast, Respondent B maintained his initial incorrect answer even though he was provided with explanations and thought experiments.

This finding supports the results of previous studies. More than 70% of students failed to understand velocity-time graphs because they had difficulty connecting kinematic concepts with graphical representations (Maries & Singh, 2013). Conceptual errors in understanding the relationship between position, velocity, and acceleration graphs are often caused by a lack of mathematical skills and contextual knowledge in physics (Beichner, 1994). A study conducted by Phage, (2018) added that errors often occur because students cannot integrate their understanding of physics and mathematics effectively. In the context of electrical circuits, it was found that students often confuse parallel and series circuits (Engelhardt & Beichner, 2004). A lack of understanding of basic concepts and weaknesses in reading electrical circuit diagrams are the leading causes of student errors (Ivanjek et al., 2021).

An interesting difference in results was found regarding the effectiveness of using technology such as ChatGPT to support conceptual understanding. Research by Kasepalu et al. (2022) showed that ChatGPT can help students understand physics concepts more deeply if used actively. This aligns with the results from Respondent A, who used ChatGPT to improve his concepts reflectively. In contrast, Respondent B did not show significant development in understanding. The effectiveness of learning technology is highly dependent on students' prior knowledge base and their involvement in the learning process (Henderson & Stewart, 2018; Chiu, et al., 2024). The consistency of answers provided by ChatGPT remains a challenge, especially in the context of physics-based questions involving images, scales, and specific concepts. In this study, it was found that ChatGPT gave different responses to respondents who asked similar questions, depending on the method of interaction and initial input provided. ChatGPT's ability to solve visual-based problems and complex concepts depends heavily on the clarity of user input (Polverini & Gregorcic, 2024; Remoto, 2023). ChatGPT tends to provide diverse answers when faced with problems that require in depth interpretation of the context (Gregorcic & Pendrill, 2023; Henderson & Stewart, 2018). This inconsistency is one of the weaknesses that must be considered when using ChatGPT as a learning aid, so teacher guidance is essential to ensure that the answers received follow the correct concept.

The results of this study have important implications for physics education. First, teachers must integrate learning strategies that emphasize conceptual understanding through multiple representation-based approaches, such as graphs, simulations, and direct experiments. This approach can help students connect physics concepts with visual and mathematical representations. Second, guidance in using technology such as ChatGPT is essential to ensure that students can use this technology critically and reflectively. Third, diagnostic-based formative

evaluations must routinely be carried out to identify and address student misconceptions early on. With this approach, physics education can be more effective in improving students' conceptual understanding and problem-solving skills, while helping them overcome the challenges of learning in the technological era.

IV. CONCLUSION AND SUGGESTION

This study found that the ChatGPT 4.0 has great potential to assist the physics learning process, especially in identifying and correcting conceptual errors. The findings show that respondents with adequate prior knowledge and reflective ability can interact positively with ChatGPT. These interactions have implications for improving their answers and deepening their conceptual understanding. In contrast, respondents who lack prior knowledge tend to accept answers from ChatGPT without critical reflection, leading to repeated errors. Additionally, this study revealed ChatGPT's limitations in solving image-based problems, carefully reading scales, and understanding problems that require the application of specific concepts. These findings highlight the importance of having a strong conceptual foundation and reflective ability to utilize AI-based tools optimally.

As a suggestion, educators are encouraged to utilize the results of this study to design learning experiences that can strengthen students' conceptual understanding before using technologies such as ChatGPT. The use of experimental, simulation, and interactive discussion-based approaches can improve students' reflective ability and their capacity to use AI technology critically. Furthermore, educators can take advantage of ChatGPT's limitations by creating evaluation questions that involve visual representations or other elements that are not easily solved by AI, thus minimizing the potential for cheating and improving students' analytical abilities. Further research is needed to develop optimal learning strategies that leverage the advantages of AI technology while overcoming its weaknesses in physics education. By taking this approach, the use of ChatGPT is expected to support more effective and adaptive learning in the digital era.

REFERENCES

- Abbas, N., Ali, I., Manzoor, R., Hussain, T., & Hussain, M. H. L. (2023). Role of artificial intelligence tools in enhancing students' educational performance at higher levels. *Journal of Artificial Intelligence Machine Learning and Neural Network*, 3(5), 36–49. <https://doi.org/10.55529/jaimlenn.35.36.49>

- Akavova, A., Temirkhanova, Z., & Lorsanova, Z. (2023). Adaptive learning and artificial intelligence in the educational space. *E3s Web of Conferences*, 451, 06011, 1-4. <https://doi.org/10.1051/e3sconf/202345106011>
- Amin, B. D., Sahib, E. P., Harianto, Y. I., Patandean, A. J., Herman, & Sujiono, E. H. (2020). The interpreting ability on science kinematics graphs of senior high school students in South Sulawesi, Indonesia. *Jurnal Pendidikan IPA Indonesia*, 9(2), 179–186. <https://doi.org/10.15294/jpii.v9i2.23349>
- Beichner, R. (1994). Testing student interpretation of kinematic graphs. *American Journal of Physics*, 62, 750-762. <https://doi.org/10.1119/1.17449>
- Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, 32(7), 3240–3256. <https://doi.org/10.1080/10494820.2023.2172044>
- Cooper, G. (2023). Examining science education in ChatGPT: an exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32, 444–452. <https://doi.org/10.1007/s10956-023-10039-y>
- Eaton, P., & Willoughby, S. D. (2018). Confirmatory factor analysis applied to the force concept inventory. *Physical Review Physics Education Research*, 14, 1-11. <https://doi.org/10.1103/physrevphyseducres.14.010124>
- Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, 72, 98–115. <https://doi.org/10.1119/1.1614813>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Gregorcic, B., & Pendrill, A. M. (2023). ChatGPT and the frustrated socrates. *Physics Education*, 58,1-9. <https://doi.org/10.1088/1361-6552/acc299>
- Hashem, R., Ali, N., Zein, F. E., Fidalgo, P., Khurma, O. A. (2023). AI to the rescue: Exploring the potential of chatgpt as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, 19, 1-26. <https://doi.org/10.58459/rptel.2024.19023>
- Henderson, R., & Stewart, J. (2018). Racial and ethnic bias in the force concept inventory. *Conference Proceedings*, 172–175. <https://doi.org/10.1119/perc.2017.pr.038>
- Hidaayatullaah, H. N. (2022). The science literacy profile based on madrasah students' misconceptions on science concepts. *Proceedings of the International Conference on Madrasah Reform*, 111-116. <https://doi.org/10.2991/assehr.k.220104.017>

- Hikmatiar, H., Sya'bania, N., Jayadin., Kasman, R. A., Imranah., Sahlan., & Saputra, S. (2024). The effectiveness of Chatgpt in completing astronomy lectures: Building awareness of its use. *Jurnal Pendidikan Fisika*, 12(2), 121-130. <https://doi.org/10.26618/jpf.v12i2.13587>
- Hoa, N. Q. (2023). AI and plagiarism: Opinion from teachers, administrators and policymakers. *Proceedings of the Asiacall International Conference*, 4, 75–85. <https://doi.org/10.54855/paic.2346>
- Ivanjek, L., Morris, L., Schubatzky, T., Hopf, M., Burde, J. P., Haagen-Schützenhöfer, C., Dopatka, L., Spatz, V., & Wilhelm, T. (2021). Development of a two-tier instrument on simple electric circuits. *Physical Review Physics Education Research*, 17, 1-15. <https://doi.org/10.1103/PhysRevPhysEducRes.17.020123>
- Kasepalu, R., Prieto, L. P., Ley, T., & Chejara, P. (2022). Teacher artificial intelligence-supported pedagogical actions in collaborative learning coregulation: A wizard-of-oz study. *Frontiers in Education*, 7, 1-15. <https://doi.org/10.3389/educ.2022.736194>
- Khan, R. A., Jawaidd, M., Khan, A. R., & Sajjad, M. (2023). ChatGPT - Reshaping medical education and clinical management. *Pakistan Journal of Medical Sciences*, 39(2), 605-607. <https://doi.org/10.12669/pjms.39.2.7653>
- Kim, S.-W. (2023). Change in attitude toward artificial intelligence through experiential learning in artificial intelligence education. *International Journal on Advanced Science Engineering and Information Technology*, 13(5), 1953–1959. <https://doi.org/10.18517/ijaseit.13.5.19039>
- Linuwih, S. (2013). Konsepsi alternatif mahasiswa calon guru fisika tentang gaya-gaya yang bekerja pada balok. *Jurnal Pengajaran MIPA*, 18(1), 69-77. <https://doi.org/10.18269/jpmipa.v18i1.259>
- Maries, A., & Singh, C. (2013). Exploring one aspect of pedagogical content knowledge of teaching assistants using the test of understanding graphs in kinematics. *Physical Review Physics Education Research*, 9, 1-14. <https://doi.org/10.1103/physrevstper.9.020120>
- Phage, I. (2018). Undergraduate physics students' conceptual understanding in the learning of kinematics using a blended approach. *Ijaedu- International E-Journal of Advances in Education*, 4(11), 199–204. <https://doi.org/10.18768/ijaedu.455623>
- Phage, I. B., Lemmer, M., & Hitge, M. (2017). Probing factors influencing students' graph comprehension regarding four operations in kinematics graphs. *African Journal of Research in Mathematics, Science and Technology Education*, 21(2), 200–210. <https://doi.org/10.1080/18117295.2017.1333751>
- Polverini, G., & Gregorcic, B. (2024). Performance of ChatGPT on the test of understanding graphs in kinematics. *Physical Review Physics Education Research*, 20, 1-16. <https://doi.org/10.1103/PhysRevPhysEducRes.20.010109>
- Remoto, J. P. (2023). ChatGPT and other AIs: Personal relief and limitations among mathematics-oriented learners. *Environment and Social Psychology*, 9(1), 1-13. <https://doi.org/10.54517/esp.v9i1.1911>

- Suganda, T., Kusairi, S., Azizah, N., & Parno, P. (2020). The correlation of isomorphic, open-ended, and conventional score on the ability to solve kinematics graph questions. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 6(2), 173–180.
<https://doi.org/10.21009/1.06204>
- Sutopo, S., Parno, P., & Angin, S. L. (2017). Pemahaman mahasiswa tentang multi representasi konsep percepatan. *Jurnal Riset dan Kajian Pendidikan Fisika*, 4(2), 48-53.
<https://doi.org/10.12928/jrkpf.v4i2.6551>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success. *Applied Sciences*, 13(11), 1-15. <https://doi.org/10.3390/app13116716>
- Warsono, W., Nursuhud, P. I., Darma, R. S., Supahar, S., Oktavia, D. A., Setiyadi, A., & Kurniawan, M. A. (2020). Multimedia learning modules (MLMs) based on local wisdom in physics learning to improve student diagram representations in realizing the nature of science. *International Journal of Interactive Mobile Technologies (Ijim)*, 14(6), 148-158.
<https://doi.org/10.3991/ijim.v14i06.11640>
- Zavala, G., Tejeda, S., Barniol, P., & Beichner, R. J. (2017). Modifying the test of understanding graphs in kinematics. *Physical Review Physics Education Research*, 13(2), 20111, 1-16.
<https://doi.org/10.1103/PhysRevPhysEducRes.13.020111>