



ChatGPT in Physics Education: A Content-Based Analysis on Newtonian Force Problems

^{1*} Ni Nyoman Sri Putu Vewawati, ¹ Wahyudi, ¹ Nina Nisrina, ² Muhammad Asy'ari

¹ Physics Education Department, University of Mataram, Mataram, Indonesia

² Physics Education Department, Universitas Pendidikan Mandalika, Mataram, Indonesia

Corresponding author e-mail: veyra@unram.ac.id

Received: January 2025; Revised: March 2025; Published: April 2025

Abstract

The integration of artificial intelligence (AI) in education has significantly transformed learning environments, particularly through the use of large language models (LLMs) such as ChatGPT. While these tools show promise in supporting science and technology education, their effectiveness in solving domain-specific problems, such as Newtonian mechanics, remains under-explored. This study aims to evaluate the capability of ChatGPT in solving essay-type physics problems involving Newton's Laws of Motion, with a specific focus on force analysis. Using a content-based qualitative evaluation method, the research was conducted in three stages: development and validation of conceptual physics problems, submission of these problems to ChatGPT, and assessment of the AI-generated responses by expert reviewers. The problem used in this study required decomposition of forces on an inclined plane under idealized, frictionless conditions. ChatGPT's responses were evaluated across three dimensions: scientific accuracy, logical coherence, and contextual relevance. The findings indicate that while ChatGPT was able to provide structured and numerically accurate responses, it lacked depth in reasoning and failed to explicitly articulate physical assumptions and validation steps, such as analyzing counteracting gravitational forces. These limitations point to the model's partial conceptual understanding and highlight the need for human oversight. The study concludes that ChatGPT holds potential as a supplementary learning aid, particularly for reinforcing procedural knowledge. However, its use must be carefully integrated into instructional contexts that promote critical thinking and conceptual verification. Recommendations are offered for its pedagogical implementation, along with a call for further research into AI's role in physics education.

Keywords: ChatGPT; Newton's Laws of Motion; Physics Education; Artificial Intelligence in Learning; Qualitative Evaluation

How to Cite: Vewawati, N., Wahyudi, W., Nisrina, N., & Asy'ari, M. (2025). ChatGPT in Physics Education: A Content-Based Analysis on Newtonian Force Problems. *Prisma Sains : Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, 13(2), 335-347. doi:<https://doi.org/10.33394/j-ps.v13i2.15824>



<https://doi.org/10.33394/j-ps.v13i2.15824>

Copyright© 2025, Vewawati et al.

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



INTRODUCTION

The integration of artificial intelligence (AI) in the educational sector marks a pivotal shift in how knowledge is delivered and acquired. With the proliferation of machine learning algorithms and their deployment across various educational platforms, there has been a notable transformation in teaching methodologies, administrative operations, and student engagement. Recent studies have identified key domains where AI exerts profound influence, such as in supporting personalized learning models, optimizing administrative functions, and revolutionizing pedagogical strategies to enhance educational equity and quality (Qin & Ao, 2023; Soyombo-Erdene, 2024). In this evolving educational paradigm, AI serves as both an instructional aid and an administrative tool, driving efficiencies that were previously unattainable through conventional educational practices.

The application of AI in personalizing learning trajectories is one of its most celebrated innovations. Through sophisticated data analytics, AI systems adapt content delivery based on student performance, learning pace, and cognitive style, making differentiated instruction more accessible and impactful. This personalization is especially relevant in settings with limited educational resources, where such technology can help bridge gaps in teaching quality and learner engagement (Liulka et al., 2024). Beyond its instructional capacity, AI contributes significantly to streamlining administrative processes within educational institutions. Systems enhanced by AI are now capable of automating grading, tracking student progress, and offering data-driven insights to guide instructional decisions (Bearman et al., 2022; Xing, 2023). These developments, while promising, are accompanied by growing ethical concerns regarding algorithmic transparency, data privacy, and the safeguarding of student rights, necessitating a balanced and conscientious approach to the deployment of AI in educational settings (He, 2023).

As AI technologies evolve, one of the most impactful innovations in educational contexts is the emergence of large language models (LLMs) such as ChatGPT. These models, designed to process and generate human-like text based on vast data corpora, have demonstrated considerable utility in domains that demand complex reasoning and contextual understanding. In science and technology education, LLMs facilitate personalized learning by interacting with students in real-time, offering explanations, and adapting to individual learning styles and comprehension levels (Nurjanah et al., 2024; Madhu et al., 2024). In this regard, ChatGPT has been utilized not only as a supplementary tutor but also as an exploratory learning tool that enhances accessibility and fosters independent inquiry.

However, the capabilities of LLMs like ChatGPT in effectively solving domain-specific problems, particularly in physics, remain under-examined. Given the cognitive complexity required to understand and solve classical mechanics problems—such as those involving Newton's Laws—it is imperative to rigorously evaluate whether ChatGPT can produce answers that are not only syntactically correct but also conceptually accurate. The central research question thus emerges: Can AI models like ChatGPT provide correct and logically consistent responses to problems grounded in Newtonian physics? This inquiry is significant not only for understanding the limitations of AI in content-specific education but also for exploring the potential of integrating LLMs in curricula that require both analytical reasoning and domain expertise.

The proliferation of research into the educational applications of ChatGPT has produced several encouraging findings. Studies such as Kieser et al. (2024) demonstrate that LLMs are capable of performing well across a range of academic disciplines, including physics, highlighting their capacity for general problem-solving and student engagement. However, these studies often lack domain-specific granularity, especially in assessing the efficacy of AI in solving problems rooted in fundamental physical laws. As such, while the initial data may suggest potential, the absence of targeted investigations into classical mechanics indicates a significant research void.

Bitzenbauer (2023) initiated one of the few empirical inquiries into the role of ChatGPT in physics education. His pilot study suggests that while students benefit from ChatGPT's support in conceptual discussions, the reliability of the model's responses to detailed problem-solving exercises in classical physics remains questionable. This finding is echoed in Rasul et al. (2023), who emphasize the importance of accuracy and dependability in deploying AI for higher education. Their work points to the necessity of rigorous evaluation mechanisms, especially in disciplines like physics, where even minor conceptual inaccuracies can lead to fundamentally flawed understandings.

Additionally, a study by Gregorcic and Pendrill (2023) documented inconsistencies and factual inaccuracies in ChatGPT's responses to basic physics questions, highlighting the model's limitations in contexts requiring precise and structured reasoning. These

inconsistencies underline a critical concern: while ChatGPT can generate plausible answers linguistically, it may lack the epistemological framework to accurately handle the logical structure of physics problems. The findings of these studies collectively underscore the importance of domain-specific assessments of LLM performance, particularly in subjects such as Newtonian mechanics where problem-solving requires a deep understanding of interrelated concepts and principles.

Physics education, particularly at the secondary and early tertiary levels, faces a multitude of pedagogical challenges. Newton's Laws of Motion, forming the cornerstone of classical mechanics, are conceptually rich yet frequently misunderstood due to their abstract nature and dependence on both qualitative and quantitative reasoning. Research has shown that student motivation is a critical determinant of physics learning success. When students perceive physics as disconnected from their interests or real-world experiences, they are less likely to engage meaningfully with the content (Bøe & Henriksen, 2013; Hazari et al., 2010). Conversely, context-based instruction that situates Newtonian concepts in relatable scenarios enhances motivation and comprehension (Crouch & Heller, 2014).

Compounding these issues is the disparity in student preparation for physics coursework. Many learners—particularly those from non-technical backgrounds—lack the foundational skills necessary to navigate Newtonian mechanics effectively (Kaczmarek & Greczyło, 2025; Geller et al., 2018). For example, students from life sciences disciplines often find themselves struggling to reconcile the mathematical rigor and conceptual abstraction of physics with their prior educational experiences. This skills gap is further exacerbated by traditional teaching approaches that prioritize rote memorization over inquiry and application (Crouch et al., 2012). These methodologies often fail to cultivate problem-solving competencies, leading to superficial learning and limited retention.

The challenges associated with online and hybrid learning environments have further complicated the delivery of physics education. According to Mercado (2021) and Stanlee & Swanto (2022), the loss of interactivity and peer collaboration in digital settings has led to diminished student engagement and learning outcomes. This makes the case for integrating AI technologies, such as ChatGPT, even stronger—as potential tools to reintroduce interactivity and provide personalized support in remote learning environments. However, these tools must be rigorously evaluated for their efficacy and reliability before they can be recommended as integral components of educational strategy.

Objective, Novelty, and Scope of the Study

In response to these challenges and gaps in the existing literature, the present study aims to evaluate the capacity of ChatGPT to solve problems related to Newton's Laws, specifically problems involving forces. The objective is to assess whether ChatGPT can produce solutions that are accurate, logically consistent, and aligned with the conceptual underpinnings of Newtonian physics. Through this evaluation, the study seeks to determine whether ChatGPT can be reliably used as a supplementary educational tool in physics instruction.

This research is novel in its targeted focus. While previous studies have explored ChatGPT's general educational applications, few have delved into its performance within a defined scientific domain such as Newtonian mechanics. By narrowing the scope to force-related problems within the framework of Newton's Laws, this study offers a more precise and methodologically sound evaluation of ChatGPT's cognitive modeling in a complex conceptual space. Furthermore, by situating this inquiry within the context of the Indonesian educational system, the study contributes region-specific insights that are currently lacking in the global literature.

The scope of the study is deliberately bounded to include open-ended questions related to Newton's Laws at the senior secondary and early undergraduate levels. The focus is on assessing ChatGPT's accuracy, logical coherence, and conceptual fidelity in solving these

problems. The research does not aim to replace conventional teaching methodologies but rather to explore how AI tools can complement and support instructional strategies in physics education. The findings are intended to inform educators, curriculum developers, and policymakers about the potentials and limitations of LLMs in STEM education, contributing to more effective and ethically responsible integration of AI technologies in the classroom.

METHOD

This study adopted a Content-Based Qualitative Evaluation approach (Schilling et al., 2024) to assess the performance of ChatGPT in solving essay-based physics problems specifically related to Newton's Laws of Motion, with a focus on force. This methodological framework is particularly suited for analyzing textual outputs generated by artificial intelligence models, especially when the goal is to evaluate the accuracy, logical coherence, and conceptual relevance of their responses to scientifically grounded prompts. The qualitative evaluative design of this research enables a detailed interpretation of how ChatGPT processes, interprets, and articulates complex physics content without reliance on statistical generalization, making it appropriate for in-depth content analysis. The research flow is presented in Figure 1.

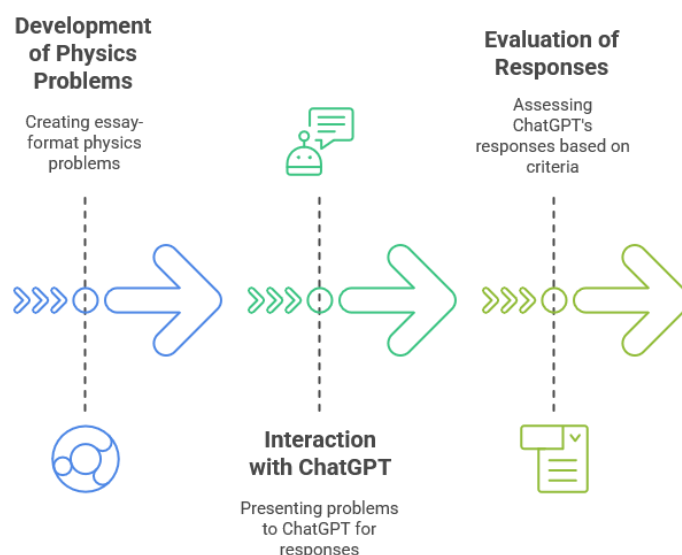


Figure 1. The research flow

The first stage of the research involved the development of physics problems in essay format. These problems were systematically constructed to cover fundamental and applied aspects of Newtonian mechanics, such as net force, normal and frictional forces, and Newton's first, second, and third laws. The problems were designed based on learning objectives from national high school and introductory university physics curricula. The aim was to ensure that each question required conceptual understanding, critical reasoning, and, in some cases, quantitative analysis. To validate the quality and relevance of the questions, a panel of two physics educators with expertise in classical mechanics reviewed the items. Their feedback was used to refine the questions to ensure alignment with instructional goals and cognitive demand levels appropriate for the intended academic level.

In the second stage, the validated questions were submitted to ChatGPT (GPT-4 version) through structured prompts. Each prompt was designed to elicit a detailed and explanatory response from the AI model, mimicking a typical student-AI interaction in a tutoring or problem-solving context. To ensure consistency, all prompts were delivered in a single conversational session without external context or follow-up clarification. The answers

generated were captured verbatim, with no alterations, and were organized for subsequent evaluation. The use of the standard version of ChatGPT, without custom training or optimization, was intended to reflect the real-world application of the model as accessed by general users in educational settings.

The third phase comprised the evaluation of ChatGPT's responses using three principal criteria: (1) scientific accuracy, assessing whether the answers aligned with established physical laws and concepts; (2) logical coherence, evaluating the internal consistency and structure of reasoning within the response; and (3) contextual relevance, analyzing how appropriately the AI interpreted and responded to the specific problem context. The evaluation was conducted independently by two physics education experts, using an assessment rubric designed to capture varying degrees of conceptual understanding and logical development.

Data analysis was conducted qualitatively through content analysis, identifying recurring patterns, strengths, and weaknesses in ChatGPT's responses. Findings are presented narratively and supported with excerpts from the AI's answers, categorized by levels of conceptual comprehension and reasoning structure. This analytical framework facilitates a nuanced understanding of ChatGPT's capabilities and limitations in addressing discipline-specific problems in physics, and offers valuable insights into the pedagogical implications of integrating AI tools into science education.

RESULTS AND DISCUSSION

Development of Physics Problems

In order to evaluate ChatGPT's capability in solving Newtonian mechanics problems involving force components on an inclined plane, a series of essay-type physics problems were constructed and submitted to the model. These problems were designed to assess the model's conceptual understanding and problem-solving accuracy, particularly in applying Newton's Laws of Motion. One illustrative example used in the study is presented in Figure 2, which depicts a cart being pulled up an inclined plane at an angle of 30° , under frictionless conditions. The problem prompt asked ChatGPT to determine the necessary force parallel to the incline when a pulling force of 200 N is applied at a 30° angle with respect to the horizontal, and the cart moves at a constant speed.

The first stage of the research involved the development of physics problems in essay format. The example item in Figure 2 requires integration of Newton's Second Law and vector resolution of forces acting along and perpendicular to an inclined plane. The scenario was chosen to probe the AI's handling of equilibrium conditions and its ability to distinguish between force components. This question evaluates the AI's competence in identifying the gravitational component acting down the incline and in computing the force required to maintain uniform motion under idealized, frictionless conditions.

Solve the following problem using Newton's Laws:

Observe the image. A cart is pulled upwards with a force of 200 N at an angle of 30 degrees with a constant speed. How much force parallel to the inclined plane is needed if frictional effects are ignored?

Provide detailed steps in solving this problem.

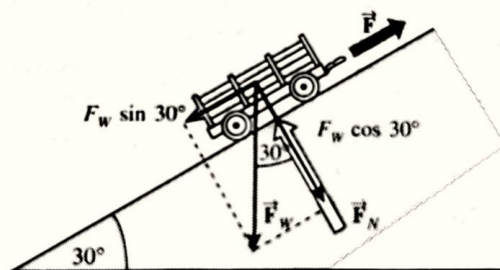


Figure 2. Newtonian mechanics problems involving force components on an inclined plane

To ensure the clarity, scientific validity, and instructional alignment of the essay-type problem presented in Figure 2, a validation process was undertaken involving a panel of two experienced physics educators. These reviewers possess academic and instructional expertise in classical mechanics, particularly in the area of Newtonian dynamics at the senior high school and early undergraduate levels. The objective of this validation was to ascertain whether the problem item accurately represents the physical principles of Newton's Laws, is free from ambiguity, and is appropriate for use in evaluating the conceptual understanding and reasoning ability of AI models such as ChatGPT.

Both reviewers examined the problem by considering three main aspects: the conceptual accuracy of the physical principles involved, the clarity and cognitive load imposed by the problem structure, and its alignment with the intended learning objectives in physics education. Their analysis concluded that the problem is valid across all dimensions and suitable for the intended academic level. The reviewers highlighted that the problem appropriately integrates Newton's Second Law, particularly in contexts involving force resolution on an inclined plane. The inclusion of numerical values and a visual diagram was seen as pedagogically sound, providing sufficient scaffolding for problem-solving without simplifying the conceptual challenge.

With regard to clarity, the reviewers agreed that the wording of the prompt was straightforward and accessible, particularly due to the explicit instruction to "Provide detailed steps in solving this problem." This directive was found to be useful in guiding both students and AI systems to not only produce final answers but also articulate their reasoning processes. The diagram accompanying the problem was deemed effective in visualizing the orientation and components of force, although one reviewer suggested a minor enhancement—namely, adding a note that the mass of the cart is unspecified—to help clarify assumptions that need to be made by the solver.

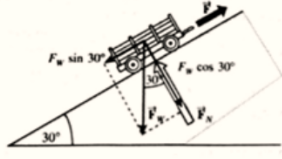
In terms of instructional alignment, the problem was judged to be consistent with standard competencies found in physics curricula. It reflects the type of applied mechanics question typically encountered in higher-level assessments and is suitable for eliciting multi-step reasoning. The reviewers emphasized that the problem not only tests computational skill but also encourages critical interpretation of physical systems and an understanding of equilibrium conditions—key objectives in the study of Newton's Laws.

ChatGPT Response to Structured Prompt

In the second stage of the study, the validated physics problem was submitted to ChatGPT (GPT-4 version) using a structured prompt that mimicked a natural student-AI interaction. This stage was intended to simulate a real-world educational context in which a learner seeks conceptual and procedural assistance from an AI tutor. To maintain the integrity of the interaction, all prompts were delivered in a single session, and no contextual follow-up or clarifying modifications were provided. The responses were recorded exactly as generated—without editing or reinforcement learning—thereby reflecting the model's unassisted reasoning and problem-solving capacity. The specific answer provided by ChatGPT is displayed in Figure 3.

In its response, ChatGPT began by correctly identifying the implications of constant velocity motion, namely that the net force acting along the incline must be zero. This is a sound application of Newton's First Law, indicating an initial awareness of equilibrium conditions. The model then proceeded to decompose the applied force of 200 N into components parallel and perpendicular to the inclined plane, using the correct trigonometric expressions $F_{\parallel (\text{parallel})} = F \cdot \cos(\theta)$ and $F_{\perp (\text{perpendicular})} = F \cdot \sin(\theta)$. Using standard values for cosine and sine at 30° , the model accurately computed: $F_{\parallel} = 200 \cdot \cos(30^\circ) \approx 173.2 \text{ N}$ and $F_{\perp} = 200 \cdot \sin(30^\circ) = 100 \text{ N}$.

ChatGPT 4



Solve the following problem using Newton's Laws:

Observe the image. A cart is pulled upwards with a force of 200 N at an angle of 30 degrees with a constant speed. How much force parallel to the inclined plane is needed if frictional effects are ignored? Provide detailed steps in solving this problem.

To solve the problem of finding the force parallel to the inclined plane when pulling the cart upward with a force of 200 N at a 30° angle to the incline, we use Newton's laws. The key to solving this problem lies in decomposing the force into its components parallel and perpendicular to the incline.

Given:

- $F = 200 \text{ N}$ (total force applied)
- $\theta = 30^\circ$ (angle of the applied force with respect to the inclined plane)
- The cart moves at a constant speed (implying no net acceleration, $a = 0$)

Steps to Solve

- Resolve the Force:** Since the force is applied at an angle to the inclined plane, we need to find its components along and perpendicular to the plane:
 - Parallel Component (F_{\parallel}):** The force component that acts parallel to the incline which is responsible for moving the cart up the incline.
 - Perpendicular Component (F_{\perp}):** The force component that acts perpendicular to the incline, affecting the normal force but not directly contributing to motion along the incline.

The formulas for these components are:

 - $F_{\parallel} = F \cdot \cos(\theta)$
 - $F_{\perp} = F \cdot \sin(\theta)$
- Calculate the Components:**
 - $F_{\parallel} = 200 \cdot \cos(30^\circ)$
 - $F_{\perp} = 200 \cdot \sin(30^\circ)$

Using trigonometric values:

 - $\cos(30^\circ) \approx 0.866$
 - $\sin(30^\circ) = 0.5$

So,

 - $F_{\parallel} = 200 \cdot 0.866 = 173.2 \text{ N}$
 - $F_{\perp} = 200 \cdot 0.5 = 100 \text{ N}$
- Force Required to Move the Cart:** Since the motion is at a constant speed, the net force along the incline must be zero. This means that the component of the gravitational force pulling the cart down the incline plus any friction (if present) must be balanced by F_{\parallel} . As the problem specifies to ignore friction and the cart is moving at a constant speed, the force needed parallel to the inclined plane is simply F_{\parallel} .

Conclusion

The force parallel to the inclined plane required to pull the cart upwards at a constant speed (ignoring friction) is **173.2 N**.

Message ChatGPT

Figure 3. The specific answer provided by ChatGPT

These calculations were correctly executed and numerically valid. The response also provided brief definitions of what each force component represents in the physical system: the parallel component responsible for moving the cart, and the perpendicular component influencing the normal force. Such inclusion indicates an attempt at pedagogical completeness and reinforces its applicability in instructional contexts.

However, while the response was procedurally correct, it lacked deeper conceptual elaboration that would typically be expected in expert problem-solving discourse. Notably, the model did not mention or compute the gravitational force component acting along the plane,

which is critical to verifying that the applied force component truly balances the system. Instead, ChatGPT inferred—correctly in principle—that because motion is constant and friction is neglected, the resolved component F_{\parallel} (parallel) is the effective force needed to overcome the component of gravity along the incline. This conclusion is stated without completing the logical feedback loop that would involve showing the equivalence between this value and $m \cdot g \cdot \sin(\theta)$.

Another limitation is the model's assumptive reasoning. While it implicitly recognizes the absence of friction and relies on idealized conditions, these assumptions are not stated explicitly. In scientific problem-solving, clearly articulating boundary conditions and constraints is essential, both for clarity and for building a sound argument. The absence of this reflective commentary, though minor in routine exercises, becomes more critical in contexts where learners are expected to justify each step and assumption.

Despite these limitations, the response in Figure 3 demonstrates ChatGPT's strong syntactic structure and general competence in executing standard computational techniques within Newtonian mechanics. Its ability to parse the prompt, apply appropriate equations, and perform the relevant mathematical steps confirms its utility as a support tool in physics education. The clarity of its language and the organization of steps also make it accessible to learners seeking procedural guidance.

ChatGPT's answer to the problem in Figure 3 reveals a competent but not fully comprehensive understanding of the physical system. It successfully identifies and computes the key components necessary to arrive at a plausible answer, yet stops short of validating the result through complete force analysis. This underscores a broader finding of the study: ChatGPT performs reliably in executing isolated procedures but requires oversight when engaged in conceptual synthesis or critical reasoning tasks. As such, it holds value as a supplementary instructional aid, particularly for reinforcing procedural knowledge, but it should be integrated thoughtfully alongside guided instruction to ensure conceptual completeness.

Evaluation of ChatGPT's Responses

The third and final stage of this research involved a critical evaluation of ChatGPT's responses to the validated essay-type problem related to Newton's Laws of Motion. This stage aimed to assess the quality, depth, and educational value of the AI-generated solution along three key dimensions: scientific accuracy, logical coherence, and contextual relevance. The analysis focused on the output presented in Figure 3, which was based on the force problem previously illustrated in Figure 2. The purpose of this evaluative stage was to determine whether ChatGPT's explanations reflect the cognitive and conceptual demands typically required in upper secondary or early university-level physics.

Two physics educators independently reviewed the AI's response using a structured qualitative rubric. This rubric was aligned with instructional standards and research-based indicators for evaluating scientific reasoning in classical mechanics. The evaluators brought with them extensive experience in physics pedagogy and assessment and applied their expertise to identify both the strengths and deficiencies in the AI's performance. Their findings offer insights into how ChatGPT functions in an educational context when applied to conceptually layered problem-solving tasks.

ChatGPT's response demonstrated moderate scientific accuracy. It correctly interpreted the core condition of the problem—that motion at a constant speed implies a net force of zero along the incline. Based on this, the model proceeded to resolve the applied force into components acting parallel and perpendicular to the inclined plane. These numerical computations were technically correct, aligning with standard methods used in physics classrooms. As in the findings of Nurjanah et al. (2024) and Madhu et al. (2024), ChatGPT

showed competence in translating known physical quantities into structured steps, reflecting procedural fluency.

However, the model did not acknowledge or calculate the resisting force due to gravity, which is a fundamental element in analyzing force balance on an inclined plane. The absence of this comparison represents a conceptual gap. As noted by Bitzenbauer (2023), this tendency to omit counterbalancing forces is a recurrent issue when LLMs are tasked with physics problems that require system-level reasoning. While ChatGPT concluded that the component of the applied force was sufficient, it did so without verifying whether it matched the opposing force acting on the cart, which is necessary to confirm equilibrium.

The model also did not address the effect of the vertical component of the force on the surface interaction—namely, the normal force. Although this omission does not directly influence the main objective of the problem due to the frictionless condition, it reduces the completeness of the explanation. As highlighted by Gregorcic and Pendrill (2023), incomplete articulation of secondary forces can lead to misunderstandings, especially among learners still developing conceptual schemas in mechanics.

Logical Coherence

From a structural standpoint, the response followed a clear and systematic order. It began with a general statement of the problem, proceeded to resolve the force into components, and concluded with a final answer. This mimics the structured reasoning expected in a formal educational context and aligns with findings that LLMs can effectively simulate instructional discourse (Qin & Ao, 2023; Jayavardhini, 2024).

However, logical coherence goes beyond organization—it includes the internal linkage of ideas and whether conclusions logically follow from preceding steps. In this regard, ChatGPT's reasoning was only partially coherent. Although the model correctly stated the condition of zero net force, it did not explicitly relate this condition back to its resolved force values. The response lacked a validation step that would confirm whether the conditions for balanced forces had been met.

As noted by Rasul et al. (2023), this type of superficial coherence is common in AI-generated educational content. The steps appear connected, but the underlying logic is not rigorously developed. In the present study, this gap manifested in the form of implicit assumptions—such as assuming the sufficiency of the calculated force component—without justification. Effective scientific reasoning requires not only procedural correctness but also the ability to justify assumptions, link cause and effect, and consider alternative interpretations (Eysenbach, 2023; Altınay et al., 2024).

Contextual Relevance

ChatGPT interpreted the context of the problem accurately. It identified the scenario, recognized the significance of the incline, and applied Newton's laws appropriately. These actions reflect the model's ability to align its output with contextual cues and task requirements, which has been previously described as one of its pedagogical strengths (Adeleye et al., 2024; Simon et al., 2024). In that sense, ChatGPT demonstrated a functional level of contextual relevance.

Nonetheless, its response did not discuss the boundary conditions or data limitations that could influence the solution. For instance, the model did not comment on the absence of the cart's mass or reflect on how this omission constrained the calculation of the opposing force due to gravity. Moreover, the frictionless assumption, although implied by the problem statement, was not reinforced in the response. Such oversights reduce the epistemic transparency of the solution—a critical component of scientific communication (Макаренко et al., 2024; He, 2023).

Contextual understanding in physics requires more than interpreting numerical values; it involves recognizing which variables are present, which are missing, and how these elements interact within a defined physical system. The failure to explicitly engage with these contextual layers reveals that ChatGPT's reasoning, while operationally accurate, does not fully meet the explanatory standards expected in physics education.

Pedagogical Implications

The evaluators concluded that ChatGPT is most useful as a supportive tool for reinforcing core concepts and providing initial guidance in problem-solving. Its ability to present structured, linguistically coherent explanations can benefit students who need step-by-step assistance or are revisiting familiar content (Zhai et al., 2023). This is especially relevant in environments with limited access to expert instruction, where AI can serve as an accessible reference point (Liulka et al., 2024; Xing, 2023). However, ChatGPT's limitations necessitate a pedagogical framework that encourages critical evaluation of its outputs. Educators should prompt students not only to follow the AI's steps but to question its assumptions, verify its logic, and reflect on alternative solutions. This aligns with the reflective learning strategies proposed by Crouch and Heller (2014), in which learners deepen their understanding by identifying and correcting flaws in reasoning. The role of ChatGPT should thus be positioned within a hybrid model of instruction, where AI augments but does not replace the teacher's role in fostering deep conceptual learning. As Erbaşı et al. (2023) emphasize, educational innovation must be guided by principles that balance technological affordances with human-centered pedagogy.

The final evaluative stage of this study found that ChatGPT exhibits strong potential as an instructional aid, particularly in tasks that involve procedural reasoning and basic conceptual application. Its response to the Newtonian problem was structured and partially accurate but lacked the analytical completeness and inferential depth required in higher-level physics education. These findings support previous literature indicating that while AI can support learning, it should be implemented alongside critical thinking practices and with active educator oversight (Bearman et al., 2022; Baskara, 2023).

CONCLUSION

This study set out to explore whether ChatGPT can provide scientifically accurate and logically consistent responses to problems involving Newton's Laws of Motion, specifically those related to the concept of force. Based on a multi-stage evaluation process that included problem construction, AI interaction, and expert review, the findings reveal that while ChatGPT demonstrates competence in basic procedural tasks—such as identifying motion conditions and decomposing forces—it falls short in delivering fully reasoned and contextually complete solutions. The model tended to omit key conceptual elements, such as counteracting gravitational components and the articulation of physical assumptions, which are essential for validating equilibrium in Newtonian systems. These outcomes affirm the initial research concern: ChatGPT can produce responses that appear correct at a surface level, but these responses require further scrutiny to be pedagogically reliable.

In alignment with the study's objective, the results underscore the potential of ChatGPT as a supplementary educational tool rather than a primary instructional resource. Its ability to support foundational learning, particularly in under-resourced or self-directed learning environments, is promising—yet its application must be critically framed and supported by educator oversight. When integrated with reflective teaching practices and guided verification, ChatGPT can serve as a catalyst for deeper student engagement and conceptual development in physics. However, to achieve instructional integrity, especially in subjects like classical mechanics that demand precision and layered reasoning, the deployment of AI in education must remain grounded in human expertise and pedagogical judgment.

RECOMMENDATION

Based on the findings of this study, it is recommended that educators and curriculum developers integrate ChatGPT as a complementary tool in physics instruction, particularly for supporting problem-solving practice and reinforcing conceptual understanding. However, its use should be coupled with guided pedagogical frameworks that encourage students to critically evaluate the AI's outputs, verify conceptual accuracy, and articulate underlying assumptions. Training for both teachers and students on how to effectively engage with AI-generated responses is essential to maximize educational benefits while mitigating the risk of misconceptions. Further research is also encouraged to explore the performance of ChatGPT across varied physics topics and student populations, including longitudinal studies that assess its impact on learning outcomes.

ACKNOWLEDGMENT

The authors extend their deepest gratitude to the research team whose dedication and collaborative efforts were instrumental in the successful completion of this study. Special thanks are directed to the team members responsible for designing the problem items, conducting the AI interaction sessions, and carrying out the evaluation and analysis processes with rigor and integrity. Their commitment to academic quality and innovation in physics education significantly contributed to the depth and credibility of this research.

REFERENCES

- Adeleye, O., Eden, C., & Adeniyi, I. (2024). Innovative teaching methodologies in the era of artificial intelligence: A review of inclusive educational practices. *World Journal of Advanced Engineering Technology and Sciences*, 11(2), 69–79. <https://doi.org/10.30574/wjaets.2024.11.2.0091>
- Altınay, Z., Altınay, F., Dağlı, G., Shadiev, R., & Othman, A. (2024). Factors influencing AI learning motivation and personalisation among pre-service teachers in higher education. *MIER Journal of Educational Studies, Trends & Practices*, 14(2), 462–481. <https://doi.org/10.52634/mier/2024/v14/i2/2714>
- Baskara, F. (2023). Chatbots and flipped learning: Enhancing student engagement and learning outcomes through personalised support and collaboration. *International Journal of Recent Educational Research*, 4(2), 223–238. <https://doi.org/10.46245/ijorer.v4i2.331>
- Bearman, M., Ryan, J., & Ajjawi, R. (2022). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*, 86(2), 369–385. <https://doi.org/10.1007/s10734-022-00937-2>
- Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, 15(3), ep430. <https://doi.org/10.30935/cedtech/13176>
- Bøe, M., & Henriksen, E. (2013). Love it or leave it: Norwegian students' motivations and expectations for postcompulsory physics. *Science Education*, 97(4), 550–573. <https://doi.org/10.1002/sce.21068>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/access.2020.2988510>
- Crouch, C., & Heller, K. (2014). Introductory physics in biological context: An approach to improve introductory physics for life science students. *American Journal of Physics*, 82(5), 378–386. <https://doi.org/10.1119/1.4870079>
- Crouch, C., Heller, K., Rebello, N., Engelhardt, P., & Singh, C. (2012). Teaching physics to life science students—Examining the role of biological context. *American Institute of Physics Conference Proceedings*. <https://doi.org/10.1063/1.3680019>
- Erbaşı, Z., Tural, B., & Çoşkun, İ. (2023). The role and potential of artificial intelligence and gamification in education: The example of Vakıf Participation Bank. *Orclever*

- Proceedings of Research and Development*, 3(1), 243–254. <https://doi.org/10.56038/oprd.v3i1.329>
- Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: A conversation with ChatGPT and a call for papers. *JMIR Medical Education*, 9, e46885. <https://doi.org/10.2196/46885>
- Geller, B., Turpen, C., & Crouch, C. (2018). Sources of student engagement in introductory physics for life sciences. *Physical Review Physics Education Research*, 14(1), 010118. <https://doi.org/10.1103/physrevphyseducre.14.010118>
- Gregorcic, B., & Pendrill, A. (2023). ChatGPT and the frustrated Socrates. *Physics Education*, 58(3), 035021. <https://doi.org/10.1088/1361-6552/acc299>
- Hazari, Z., Sonnert, G., Sadler, P., & Shanahan, M. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47(8), 978–1003. <https://doi.org/10.1002/tea.20363>
- He, J. (2023). Analysis of the application of artificial intelligence in education and teaching. *Advances in Educational Technology and Psychology*, 7(2). <https://doi.org/10.23977/aetp.2023.070210>
- Jayavardhini, P. (2024). AI study partner: Development of an LLM and Gen AI-enhanced study assistant tool. *International Journal of Scientific Research in Engineering and Management*, 8(3), 1–5. <https://doi.org/10.55041/ijssrem29626>
- Kaczmarek, M., & Greczyło, T. (2025). Challenges in designing basic physics course for undergraduate biologists—Investigating difficulties with learning physics. *Journal of Physics: Conference Series*, 2950(1), 012013. <https://doi.org/10.1088/1742-6596/2950/1/012013>
- Kieser, F., Tschisgale, P., Rauh, S., Bai, X., Maus, H., Petersen, S., ... & Wulff, P. (2024). David vs. Goliath: Comparing conventional machine learning and a large language model for assessing students' concept use in a physics problem. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1408817>
- Liulka, V., Savenkova, O., & Dedukhno, A. (2024). The peculiarities of using artificial intelligence in teaching foreign languages in higher education institutions in Ukraine. *Humanities Science Current Issues*, 2(73), 195–201. <https://doi.org/10.24919/2308-4863/73-2-30>
- Madhu, N., Latha, P., & Savitha, N. (2024). Revolutionizing education: Harnessing AI for personalized learning pathways and student success. *International Journal for Multidisciplinary Research*, 6(5). <https://doi.org/10.36948/ijfmr.2024.v06i05.28371>
- Макаренко, О., Борисенко, О., Горюхивська, Т., Козуб, В., & Яременко, Д. (2024). Embracing artificial intelligence in education: Shaping the learning path for future professionals. *Multidisciplinary Science Journal*, 6, 2024ss0720. <https://doi.org/10.31893/multiscience.2024ss0720>
- Mercado, J. (2021). Resilient mechanism in learning modern physics: A grounded theory. *Research Square*. <https://doi.org/10.21203/rs.3.rs-690636/v1>
- Nurjanah, A., Salsabila, I., Azzahra, A., Rahayu, R., & Marlina, N. (2024). Artificial intelligence (AI) usage in today's teaching and learning process: A review. *Syntax Idea*, 6(3), 1517–1523. <https://doi.org/10.46799/syntax-idea.v6i3.3126>
- Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 135–153. <https://doi.org/10.2478/jolace-2019-0025>
- Qin, Q., & Ao, L. (2023). A new world of open space: Reflections and perspectives of Chinese scholars on the digital transformation of education and future education research. *Modern Comparative Education and Technology*. <https://doi.org/10.55162/mcet.05.153>

- Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F., Ladeira, W., ... & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching*, 6(1). <https://doi.org/10.37074/jalt.2023.6.1.29>
- Schilling, Y., Hillebrand, L., & Kuckuck, M. (2024). A Qualitative-Content-Analytical Approach to the Quality of Primary Students' Questions: Testing a Competence Level Model and Exploring Selected Influencing Factors. *Education Sciences*, 14(9), 1003. <https://doi.org/10.3390/educsci14091003>
- Simon, S., Coelho, R., Marfisi-Schottman, I., & Pea, R. (2024). Generative AI tools in an undergraduate computer science program. *International Conference on Learning Sciences*. <https://doi.org/10.22318/icls2024.271910>
- Soyombo-Erdene, U. (2024). Research on the application of artificial intelligence in second language teaching. *International Journal of New Developments in Education*, 6(4). <https://doi.org/10.25236/ijnde.2024.060406>
- Stanlee, T., & Swanto, S. (2022). Professional development for ESL educators during COVID-19 pandemic era at Sabah's tertiary institutions, Malaysia. *EAI Conference Proceedings*. <https://doi.org/10.4108/eai.14-8-2021.2317649>
- Xing, C. (2023). Research on the application of artificial intelligence empowered education management. *Journal of Artificial Intelligence Practice*, 6(6). <https://doi.org/10.23977/jaip.2023.060602>
- Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 180–210. <https://doi.org/10.14742/ajet.7492>
- Zhai, G., Guo, K., & Li, S. (2023). Research on innovation of ideological and political education in universities based on artificial intelligence. In *Proceedings of the 2023 International Conference on Education and Social Sciences* (pp. 27–32). https://doi.org/10.2991/978-2-38476-158-6_5