JURNAL PEMBELAJARAN FISIKA
http://jurnal.fkip.unila.ac.id/index.php/JPF
VOL. 14 (1), JUNE 2026, 64 – 76 || P-ISSN 2302-0105, E-ISSN 2684-9828
LITERATURE REVIEW ARTICLE

ChatGPT in Physics Education: A Systematic Literature Review and Bibliometric Mapping of Pedagogical Roles, Challenges, and Implications

Dhita Thivani Linch Hutabarat*, Sintia Sintia, Heru Kuswanto

Master of Physics Education Study Program, Yogyakarta State University, Yogyakarta, Indonesia

*Corresponding Author, E-mail: dhitathivani.2025@student.uny.ac.id

Article History

Received:
14 April 2026

Accepted:
1 July 2026

Published:
8 Agust 2026

Keywords

ChatGPT; Physics education; Artificial intelligence in education; Large language model

Abstract

The rapid advancement of Artificial Intelligence (AI), particularly Large Language Models (LLMs) such as ChatGPT, has accelerated its integration into education. However, evidence regarding its pedagogical value and limitations in physics education remains fragmented, especially concerning conceptual accuracy, discipline-specific hallucination risks, and teacher roles in AI-supported learning. This study provides a physics-specific synthesis of ChatGPT implementation in physics education by examining its contributions to conceptual learning, assessment, instructional design, and teacher development. A Systematic Literature Review (SLR) following the PRISMA protocol was conducted using the Scopus database in the Physics and Astronomy subject category. The search employed the keywords "ChatGPT", "Artificial Intelligence", and "Physics Education", yielding 35 records. Following screening and eligibility assessment, 15 studies were retained for bibliometric mapping, while seven core studies directly investigating ChatGPT implementation in physics learning were selected for qualitative synthesis. Bibliometric analysis using VOS-viewer identified three thematic clusters, AI capabilities, physics implementation, and pedagogical integration. The findings indicate that ChatGPT supports conceptual discussions, experimental data analysis, lesson plan development, and teacher pedagogical training. Automated assessment accuracy reached approximately 70% after domain-specific fine-tuning, whereas conceptual inaccuracies remained evident, with reported error rates ranging from 15% to 25% in precision physics contexts. The review demonstrates that ChatGPT should not function as an autonomous physics tutor but should be embedded within teacher-validated instructional designs that explicitly evaluate conceptual accuracy, reasoning transparency, and student verification processes in physics learning.

DOI: http://dx.doi.org/10.23960/jpf.v14i1.29

INTRODUCTION

Significant transformations in contemporary education are driven by the rapid advancement of digital technologies, particularly Artificial Intelligence (AI). As computational systems capable of simulating aspects of human reasoning, language processing, and problem-solving, AI technologies have increasingly evolved into adaptive learning tools that can respond to the diverse needs of learners and support personalized educational experiences (Jang et al., 2024). These developments align closely with the demands of twenty-first-century education, which emphasize digital literacy, critical thinking, collaboration, and independent learning skills.

Among recent AI innovations, Large Language Models (LLMs) such as ChatGPT have attracted considerable attention because of their ability to generate contextual responses, explain concepts, and engage in interactive dialogue. In educational settings, ChatGPT has gradually emerged as a virtual tutor, discussion partner, and instructional assistant capable of supporting learning activities across disciplines. Its ability to provide immediate feedback and natural language explanations offers opportunities to enrich learning experiences, particularly in science education where conceptual understanding and reasoning processes are central components of learning (Brown et al., 2020)

Physics represents a particularly interesting domain for examining the educational use of ChatGPT because learning physics requires not only factual knowledge but also conceptual reasoning, mathematical derivation, graphical interpretation, and the ability to connect multiple representations of scientific phenomena. Students frequently encounter difficulties in understanding abstract concepts and translating mathematical expressions into physical meaning. Theoretically, ChatGPT has the potential to support constructivist learning approaches and Socratic dialogue practices that encourage active conceptual exploration and reflective reasoning among learners (Elson & et al., 2025).

Existing studies have demonstrated several promising applications of ChatGPT in physics education. Previous investigations have reported its use at secondary and higher education levels, across topics such as mechanics, kinematics, and experimental analysis, and for educational functions including problem-solving support, automated assessment, lesson planning, and teacher professional development. Research designs have varied from case studies and qualitative explorations to experimental classroom implementations, while reported outcomes include improvements in engagement, conceptual discussion, and assessment efficiency (Jang et al., 2024; Gregorcic et al., 2024).

Despite these encouraging findings, the current body of knowledge remains fragmented and several important questions remain unresolved. Empirical evidence regarding actual improvements in physics learning outcomes remains limited, particularly evidence derived from controlled or longitudinal studies capable of evaluating conceptual retention and transfer of learning. Much of the existing literature focuses on feasibility and perceptions rather than measurable educational effectiveness.

Then, a disciplinary gap persists because most reviews on generative AI in education examine broad educational applications without adequately addressing the unique characteristics of physics learning. Physics education requires high levels of conceptual precision, mathematical consistency, representational competence, and scientific reasoning. Consequently, issues such as hallucinations, incorrect derivations, contradictory explanations, and misconceptions generated by AI systems may have more serious implications in physics than in many other disciplines (Park & Park, 2024). Challenges associated with technological dependency and academic misuse further reinforce the need for discipline-specific evaluation frameworks (Gregorcic et al., 2024).

In other that, a synthesis gap can be identified in the absence of reviews that integrate evidence regarding student learning, automated assessment, teacher professional development, teacher–AI collaboration, and conceptual accuracy within a single physics-specific analytical framework. Unlike broad reviews of generative AI in education, physics education requires an integrated understanding of how AI influences conceptual reasoning, problem-solving processes, assessment practices, and instructional decision-making simultaneously. Therefore, existing literature has not yet provided a comprehensive framework capable of connecting these dimensions within the aspect of physics pedagogy (Hasan & et al., 2024). Unlike broader AI-in-education reviews, this study specifically integrates evidence on conceptual reasoning, problem-solving, automated assessment, teacher–AI collaboration, and model-generated scientific errors to provide a physics-specific framework for future instructional design and research development (Gregorcic & Pendrill, 2023) .

METHOD

Review Design and Reporting Standard

This study employed a Systematic Literature Review (SLR) design to synthesize evidence regarding the implementation of ChatGPT in physics education. The review process was conducted in accordance with the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting guideline to ensure transparency and consistency in reporting the review procedures and findings (Polverini & Gregorcic, 2024). In this study, PRISMA functioned as a reporting framework rather than as a research method or data collection instrument.

Search Procedure

The literature search was conducted in the Scopus database because of its broad coverage of peer-reviewed international publications and strong indexing in science education research. Searches were limited to the Physics and Astronomy subject category and included publications available up to December 2025. The final search was completed in January 2026. The search strategy employed the following Boolean query:

("ChatGPT" OR "Generative AI" OR "Large Language Model" OR "LLM") AND ("Physics Education" OR "Physics Learning" OR "Physics Teaching")

The search initially identified 35 records relevant to the topic of AI implementation in physics education.

Screening Procedure and Eligibility Criteria

The screening process consisted of identification, screening, eligibility assessment, and inclusion stages following the PRISMA 2020 workflow. Inclusion criteria consisted of:

1) peer-reviewed journal articles;
2) publications written in English;
3) studies focusing on the application of ChatGPT or LLM technologies in physics education; and
4) studies reporting pedagogical implications, learning outcomes, assessment practices, or teacher-related applications.

Exclusion criteria included conference abstracts, editorials, opinion papers, duplicate records, and studies focusing on AI applications outside educational physics contexts, such as medicine or engineering.

Following the screening process, 15 studies met the eligibility criteria and were included in the bibliometric mapping stage. Among these, seven core studies directly investigated the implementation of ChatGPT in physics learning and were therefore selected for in-depth qualitative synthesis (Rahman & Watanobe, 2023) .

Bibliometric Mapping and Data Collection

Bibliometric mapping was conducted using VOS-viewer software to identify patterns of relationships among keywords and emerging research themes (Redish, 2003). Keyword co-occurrence was used as the unit of analysis using full counting methods. Due to the relatively small number of eligible studies, low occurrence thresholds were applied to maximize network visibility and thematic coverage.

The authors acknowledge that bibliometric indicators such as link strength, centrality, and cluster density derived from a limited number of publications should be interpreted cautiously. Consequently, the bibliometric results are intended to provide exploratory thematic insights rather than statistically robust representations of the entire research field.

Qualitative Synthesis and Data Analysis

The seven core studies were analyzed using narrative qualitative synthesis techniques covering study objectives, research design, educational level, physics topics, AI models, learning outcomes, pedagogical functions, reported benefits, and identified risks (Russell & Norvig, 2021). This analytical framework enabled comparisons across studies and facilitated the identification of implementation patterns, opportunities, and limitations associated with ChatGPT integration in physics education.

The synthesis aimed to develop a physics-specific understanding of how generative AI influences conceptual reasoning, problem-solving activities, automated assessment, teacher professional development, and teacher–AI collaboration within physics learning environments.

PRISMA Flow Diagram

The PRISMA flow diagram presented in Figure 1 was prepared using standard academic English terminology, including Identification, Screening, Eligibility, and Included Studies, in accordance with international reporting practices for systematic reviews.

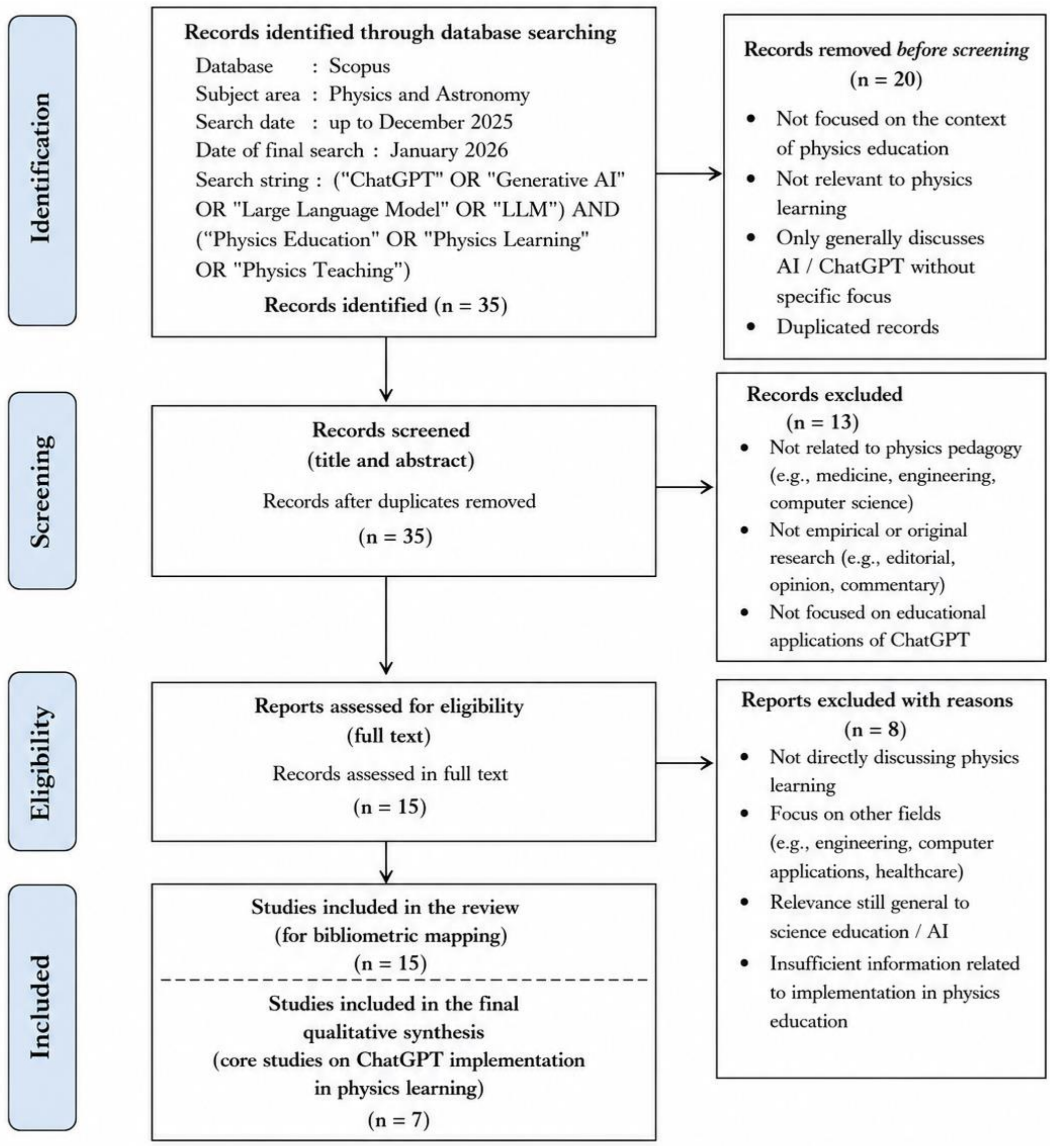

Figure 1. PRISMA Flowchart

RESULTS AND DISCUSSION

Selection and Characteristics of Included Studies

The PRISMA 2020 review procedure identified 35 records from the Scopus database within the Physics and Astronomy subject category. Following title and abstract screening, 15 studies met the eligibility criteria for bibliometric mapping. Subsequently, full-text assessment identified seven studies that directly examined the implementation of ChatGPT in physics education and were therefore included in the qualitative synthesis.

The relatively small number of eligible studies reflects the emerging nature of ChatGPT research in physics education rather than limitations of the search strategy. While research on generative AI in education has expanded rapidly since the public release of ChatGPT, studies focusing specifically on physics pedagogy remain limited. This situation is understandable because the integration of generative AI into physics classrooms is still in an exploratory phase, with most studies seeking to evaluate feasibility, opportunities, and potential risks rather than long-term effectiveness.

Another notable characteristic is the heterogeneity of the included studies. The reviewed articles involved diverse educational levels ranging from secondary school students to pre-service teachers and undergraduate students. Similarly, the educational functions of ChatGPT varied considerably, including conceptual explanation, problem-solving assistance, automated assessment, lesson planning, and teacher professional development. This diversity suggests that ChatGPT is not being implemented for a single educational purpose but rather as a flexible instructional tool that can support multiple dimensions of physics learning.

However, the evidence base remains methodologically immature. Most studies employed qualitative designs, exploratory case studies, or proof-of-concept approaches. Controlled experiments, longitudinal interventions, and randomized controlled trials remain absent from the current literature. Consequently, findings should be interpreted as preliminary evidence rather than definitive proof of educational effectiveness.

Table 1. Characteristics of Included Studies

Author	Educational Level	Physics Topic	Educational Function	Research Design	AI Model	Main Findings	Risk Category
Elson & et al. (2025)	Undergraduate	General Physics	Problem-solving support	Experimental	ChatGPT	Improved engagement and problem-solving support	Academic misuse
Jang et al. (2024)	High school	Simple harmonic motion	Conceptual learning	Case study	ChatGPT	Assisted conceptual understanding	Conceptual inaccuracies
Lee (2025)	Secondary and higher education	Assessment	Automated scoring	Fine-tuning study	Fine-tuned ChatGPT	Improved scoring consistency	Evaluation bias
Gregorcic et al. (2024)	Pre-service teachers	Physics pedagogy	Socratic dialogue training	Proof-of-concept	ChatGPT	Supported reflective dialogue	Conceptual inconsistency
Park & Park (2024)	Pre-service teachers	Lesson planning	Teacher development	Qualitative	ChatGPT	Improved lesson planning structure	Overdependence
Polverini & Gregorcic (2024)	Higher education	Conceptual reasoning	Conceptual explanation	Conceptual analysis	ChatGPT	Prompt quality influenced responses	Hallucination risk
Gregorcic & Pendrill (2023)	General physics	Kinematics	Conceptual dialogue	Dialogue analysis	ChatGPT	Contradictory explanations detected	Scientific error

Publisher: Physics Education Study Program, Faculty of Teacher Training and Education, Universitas Lampung in Collaboration with the Physical Society of Indonesia (PSI)

Forms of ChatGPT Implementation in Physics Education

The reviewed studies reveal four major forms of ChatGPT implementation in physics education: conceptual learning support, problem-solving assistance, automated assessment, and teacher professional development. These findings indicate that the role of ChatGPT extends beyond content delivery and increasingly involves instructional, evaluative, and pedagogical functions.

Table 2. Classification of Educational Functions

Educational Function	Studies	Main Benefits	Main Risks
Conceptual learning support	Jang et al. (2024); Polverini & Gregorcic (2024)	Additional explanations and conceptual discussion	Misconceptions and hallucinations
Problem-solving assistance	Elson & et al. (2025)	Scaffolding and immediate feedback	Academic dishonesty
Automated assessment	Lee (2025)	Faster feedback and scoring consistency	Evaluation bias
Teacher development	Gregorcic et al. (2024); Park & Park (2024)	Lesson planning and reflective teaching	Overdependence on AI

1) Conceptual Learning Support

Several studies identified ChatGPT as a potentially useful tool for supporting conceptual discussions in physics classrooms. Unlike traditional search engines that provide fragmented information, ChatGPT allows students to engage in iterative dialogue, request clarification, and obtain alternative explanations for difficult concepts. Such interactions align with constructivist learning perspectives, where knowledge is actively constructed through interaction and reflection rather than passive reception.

The usefulness of ChatGPT appears particularly evident when students encounter abstract concepts that require multiple explanations before meaningful understanding is achieved. Jang et al. (2024) reported that ChatGPT facilitated conceptual discussions in simple harmonic motion by providing explanations that students perceived as more accessible than textbook descriptions. Nevertheless, conceptual support does not necessarily guarantee conceptual correctness. Because LLMs generate responses probabilistically rather than through scientific reasoning processes, fluent explanations may conceal inaccuracies or misconceptions. Consequently, conceptual learning support should always be accompanied by teacher validation and student verification practices.

2) Problem-Solving Assistance

Problem-solving emerged as another major implementation domain. Elson & et al. (2025) reported that ChatGPT assisted students in decomposing complex problems into smaller procedural steps, thereby supporting analytical thinking and reducing initial cognitive barriers.

This finding can be interpreted through the perspective of self-regulated learning theory, where immediate feedback and scaffolding help learners monitor their own understanding and identify gaps in knowledge. ChatGPT may therefore function as an auxiliary learning partner that encourages independent study outside formal classroom environments.

However, the same capability also introduces risks of academic misuse. Students may become dependent on AI-generated solutions without critically evaluating the reasoning process involved. Such behavior may produce an illusion of understanding in which learners become familiar with solution procedures without developing genuine conceptual mastery.

3) Automated Assessment

Automated assessment represents one of the most promising applications identified in the literature. Lee (2025) demonstrated that domain-specific fine-tuning improved the consistency of short-answer scoring and increased alignment with human evaluation practices. The implications for physics education are considerable because assessment activities often require substantial teacher workload, particularly in large classes. AI-supported assessment systems may provide rapid formative feedback and reduce administrative burdens while allowing educators to focus more on instructional interactions.

Despite these advantages, concerns regarding transparency and fairness remain unresolved. Current LLMs frequently operate as black-box systems, making it difficult for educators to understand how evaluation decisions are generated. Consequently, automated assessment should be viewed as a supplementary rather than replacement tool for teacher judgement.

4) Teacher Professional Development

The reviewed studies also demonstrated the growing role of ChatGPT in teacher professional development. Applications included lesson planning support, generation of instructional activities, and simulation of Socratic dialogue for teacher training.

Gregorcic et al. (2024) showed that interactions with ChatGPT enabled pre-service teachers to practice questioning strategies and reflective pedagogical discussions. Similarly, Park & Park (2024) found that AI collaboration improved lesson plan organization and instructional sequencing.

These findings suggest that AI may contribute not only to student learning but also to teacher learning. However, excessive dependence on AI-generated instructional materials may reduce opportunities for pedagogical creativity and professional judgement if not accompanied by critical evaluation.

Emerging Research Themes from Bibliometric Mapping

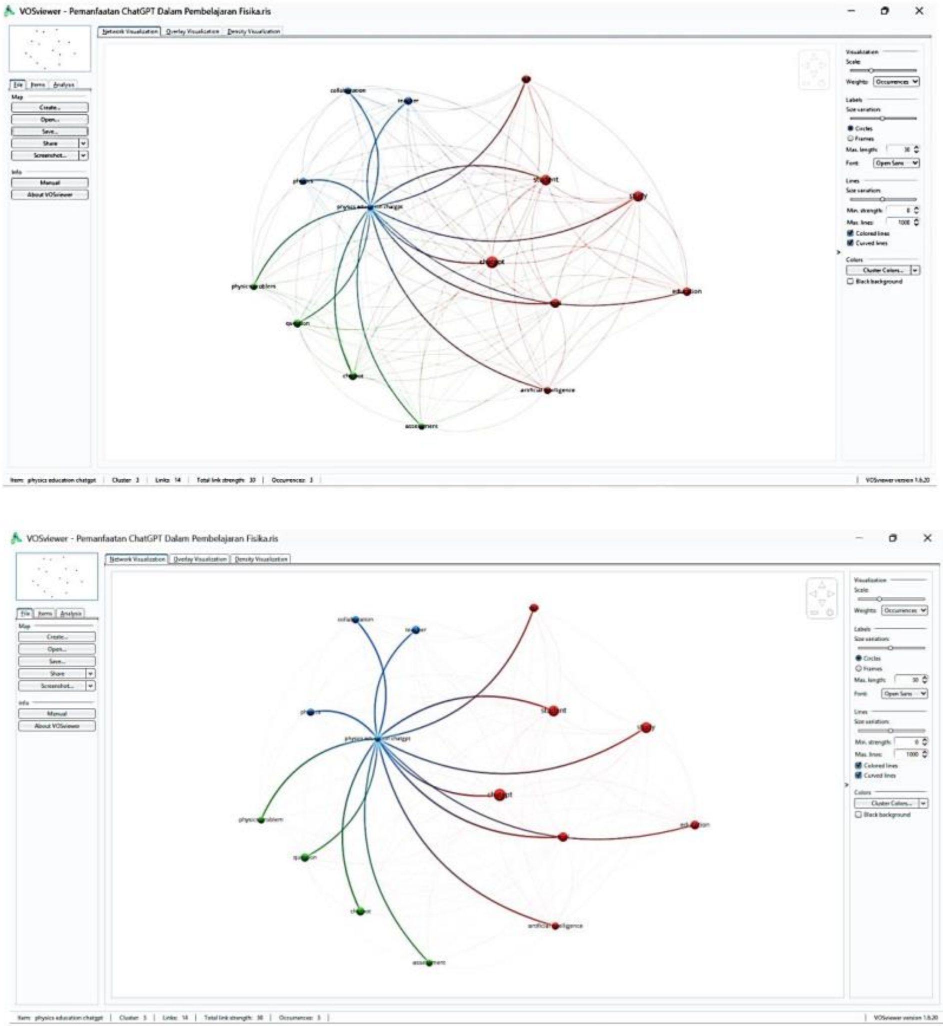

Figure 2. Visualization of Term Relationship Network in ChatGPT and Physics Education Research

Bibliometric mapping identified three thematic clusters associated with AI capabilities, practical implementation, and pedagogical collaboration.

Table 3. Bibliometric Indicators

Indicator	Value
Number of keywords	15
Number of clusters	3
Number of links	105
Total link strength	339

Rather than representing isolated research themes, these clusters illustrate the multidimensional character of ChatGPT integration in physics education. Research attention appears to move simultaneously toward technological capability, classroom application, and teacher involvement. The emergence of a pedagogical cluster is particularly noteworthy because it indicates that current research no longer views AI solely as an instructional technology but increasingly as a collaborative component within educational ecosystems.

Nevertheless, bibliometric indicators should be interpreted cautiously. Because the analysis was conducted on only fifteen eligible studies and seven implementation studies, measures such as centrality, density, and total link strength primarily provide exploratory insights rather than statistically robust representations of the field.

Toward a Teacher-Validated Framework for ChatGPT Integration in Physics Education

One of the most important findings of this review concerns conceptual accuracy. Contrary to optimistic narratives surrounding generative AI in education, the reviewed studies consistently identified conceptual errors as a major challenge in physics applications. While ChatGPT demonstrated considerable potential in supporting conceptual discussion, lesson planning, and formative feedback, the available evidence suggests that its performance varies substantially across educational functions and physics domains.

Physics differs from many educational disciplines because successful learning requires the simultaneous integration of mathematical derivation, symbolic reasoning, graphical interpretation, vector analysis, and multiple representations of physical phenomena. These requirements create particular challenges for probabilistic language models that generate responses based on statistical token prediction rather than formal scientific reasoning or symbolic computation (Polverini & Gregorcic, 2024). Consequently, responses that appear linguistically coherent may not necessarily be scientifically valid.

Gregorcic & Pendrill (2023) illustrated this limitation through contradictory explanations related to vertical acceleration in kinematics problems. Although the generated responses appeared plausible and internally consistent at the linguistic level, the underlying physical reasoning contained inconsistencies that could reinforce misconceptions among learners rather than correct them. Similar concerns were reported by Polverini & Gregorcic (2024), who argued that response quality in physics depends heavily on prompt formulation and that inappropriate prompting strategies may increase hallucination risks in precision-oriented scientific domains.

These findings can be interpreted through the perspective of conceptual change theory, which emphasizes the importance of replacing misconceptions through cognitive conflict, reflection, and evidence-based reasoning rather than through passive information reception. Incorrect AI-generated explanations may instead strengthen alternative conceptions and delay conceptual restructuring processes. Furthermore, cognitive load theory suggests that scientifically inaccurate explanations increase extraneous cognitive load and divert learners' attention away from meaningful schema construction and problem-solving processes (Luckin et al., 2016). Consequently, the educational risks associated with hallucinations may be greater in physics than in disciplines that rely less heavily on mathematical precision and multi-representational reasoning.

A consistent finding across the reviewed studies is that ChatGPT performs most effectively when embedded within teacher-mediated instructional environments rather than functioning as an autonomous tutor (Gregorcic et al., 2024; Park & Park, 2024). The evidence suggests that teachers continue to perform several essential functions, including validating conceptual accuracy, designing effective prompts, identifying hallucinations, contextualizing explanations, and facilitating metacognitive reflection. These responsibilities position teachers as critical mediators between AI-generated information and student understanding.

These finding challenges narratives portraying generative AI as a substitute for educators. Instead, the reviewed studies support a teacher-in-the-loop model in which AI functions as a collaborative instructional partner rather than a replacement for pedagogical expertise (Gregorcic et al., 2024). Such an approach appears particularly relevant in physics education, where conceptual precision, scientific reasoning, and mathematical consistency remain central learning objectives.

The present review also highlights several limitations that constrain current conclusions regarding the effectiveness of ChatGPT in physics education. First, the available evidence remains limited and reflects the emerging nature of the field. Second, the reviewed studies employed heterogeneous research designs, educational settings, and outcome measures, preventing formal meta-analysis or effect-size synthesis. Third, quantitative findings reported in individual studies should not be interpreted as pooled estimates for physics education as a whole because reported assessment accuracy and conceptual error rates originated from specific contexts and measurement procedures.

These limitations suggest several priorities for future research. Longitudinal interventions and controlled experimental studies are needed to examine conceptual retention and transfer effects associated with AI-assisted learning. Comparative investigations involving different LLM versions and prompting strategies may also clarify the relationship between model architecture and conceptual accuracy in physics problem-solving. Furthermore, future work should focus on developing teacher validation protocols, prompt engineering guidelines, and AI literacy frameworks specifically designed for physics education contexts.

The findings of this review suggest that ChatGPT should not be implemented as an autonomous physics tutor. Instead, its educational value appears greatest when embedded within teacher-validated instructional designs that explicitly evaluate conceptual accuracy, reasoning transparency, and student verification practices. Such a framework may provide a safer and more pedagogically sound pathway for integrating generative AI into physics learning environments while preserving the central role of teachers in supporting conceptual understanding and scientific reasoning.

CONCLUSION

This systematic review demonstrates that ChatGPT has emerged as a promising supporting tool in physics education through its roles in conceptual discussion, problem-solving assistance, automated assessment, and teacher professional development.

However, the findings consistently indicate that physics presents unique challenges for generative AI because effective learning requires mathematical derivation, symbolic reasoning, graphical interpretation, and multiple representations of physical phenomena. Several reviewed studies reported conceptual inconsistencies and hallucinations, particularly in precision-dependent topics such as kinematics and conceptual reasoning. Consequently, ChatGPT should not be implemented as an autonomous physics tutor but rather as a collaborative instructional tool operating within teacher-validated learning environments that prioritize conceptual accuracy, reasoning transparency, and student verification practices.

Future research should prioritize longitudinal and controlled studies to evaluate the long-term effects of ChatGPT-assisted learning on conceptual understanding and scientific reasoning in physics education. Furthermore, the development of physics-specific prompt protocols represents an important research direction to reduce hallucination risks and improve response reliability. Such protocols may include explicit statement of assumptions, step-by-step derivations, multi-representational explanations, and verification prompts that encourage students to critically evaluate AI-generated responses rather than accept them uncritically. These approaches may provide a safer and more pedagogically effective pathway for integrating generative AI into physics learning environments.

REFERENCE

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. 34th Conference on Neural Information Processing Systems (NeurIPS 2020), 33, 1877–1901. https://doi.org/10.48550/arXiv.2005.14165

Elson, L., Gaughan, L., Hopton, B., Lloyd-Brown, S., Briggs, C., Gharamti, M., & Moriarty, P. (2025). ‘ChatGPT did my homework’: lessons learnt from embedding a chatbot in undergraduate coursework. Physics Education, 60(2). https://doi.org/10.1088/1361-6552/adb363

Gregorcic, B., & Pendrill, A.-M. (2023). ChatGPT and the frustrated Socrates. Physics Education, 58(3). https://doi.org/10.1088/1361-6552/acc299

Gregorcic, B., Polverini, D., & Sarlah, M. (2024). ChatGPT as a tool for honing teachers’ Socratic dialogue skills. Physics Education, 59(045005). https://doi.org/10.1088/1361-6552/ad3d21

Hasan, N., Polin, J. A., Ahmmed, M. R., Sakib, M. M., Jahin, M. F., & Rahman, M. M. (2024). A novel approach to analyzing the impact of AI, ChatGPT, and chatbot on education using machine learning algorithms. Bulletin of Electrical Engineering and Informatics, 13(4), 2951–2958. https://doi.org/10.11591/eei.v13i4.7158

Jang, M., Kahng, Y. H., Lee, S. D., Lim, O. Y., & Park, C. J. (2024). Case study of high school physics class development utilizing ChatGPT. New Physics: Sae Mulli, 74, 779–787. https://doi.org/10.3938/NPSM.74.779

LeeJ. (2025). Research on automated scoring for short constructed-response questions through fine-tuning ChatGPT. New Physics: Sae Mulli, 75, 381–393. https://doi.org/10.3938/NPSM.75.381

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. In London: Pearson. Pearson. https://static.googleusercontent.com/media/edu.google.com/en//pdfs/Intelligence-Unleashed-Publication.pdf

Park, H. K., & Park, J. (2024). Lesson plan development for pre-service elementary school teachers through collaboration with AI: Focused on ChatGPT. New Physics: Sae Mulli, 74, 215–226. https://doi.org/10.3938/NPSM.74.215

Polverini, D., & Gregorcic, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. European Journal of Physics, 45(2). https://doi.org/10.1088/1361-6404/ad1420

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. Applied Sciences, 13(9), 5783. https://doi.org/10.3390/app13095783

Redish, E. F. (2003). Teaching physics with the physics suite. In Hoboken, NJ : John Wiley & Son. John Wiley \& Sons. https://www.academia.edu/112737022/Teaching_Physics_with_the_Physics_Suite

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A modern approach. Pearson Education. http://lib.ysu.am/disciplines_bk/efdd4d1d4c2087fe1cbe03d9ced67f34.pdf