STEM Education
STEM Education, 6(4): 539–583
DOI: 10.3934/steme.2026023
Received: 04 February 2026
Revised: 11 May 2026
Accepted: 15 May 2026
Published: 04 June 2026

https://www.aimspress.com/journal/steme

Research article

Artificial intelligence in physics education: A systematic review of content coverage, implementation models, learning impact, and pedagogical challenges

Roziqin Roziqin[1], Achmad Samsudin[1,*], Duden Saepuzaman[1], Haslinda Nawang Sari[2] and Mimin Iryanti[1]

[1] Program of Physics Education, Faculty of Mathematics and Natural Sciences Education, Universitas Pendidikan Indonesia, 40154 Bandung, Indonesia; roziqin@upi.edu, achmadsamsudin@upi.edu, dsaepuzaman@upi.edu, mien_iryanti@upi.edu

[2] Department of Chemistry Education, Universitas Negeri Yogyakarta, 55281 Yogyakarta, Indonesia; haslindanawang.2025@student.uny.ac.id

* **Correspondence:** Email: achmadsamsudin@upi.edu.

Academic Editor: Feng-Kuang Chiang

Abstract: Artificial intelligence (AI) has made a fast entry into the world of physics education research, but the empirical literature is still fragmented according to content domain, pedagogical function, and evaluation criteria. Leveraging disciplinary epistemic practices, instructional scaffolding theory, and self-regulated learning, this is a systematic literature review of 44 empirical studies published from 2021 to 2025 to examine the implementation of AI in physics education, where it is concentrated, the learning functions it serves, and the epistemic risks associated with its use. Following the guidelines of PRISMA, the review combines frequency analysis and thematic synthesis in order to search for patterns beyond isolated outcomes across studies. The results show that there is a high concentration of AI applications in the areas of mechanics, thermodynamics, and modern physics, which are characterized by machine-interpretable problem structures and high demands for abstraction. AI is mainly applied in conversational tutoring, automated feedback, and personalization systems, and the benefits have been shown to be focused on short-term learning performance, engagement, and instructional efficiency. However, common themes among the contexts are observed, such as epistemic unreliability, overreliance on AI outputs, lack of AI literacy, and fragility of methodology, which directly pose a threat to disciplinary reasoning practices that are central to learning physics. Far from reflecting pedagogical transformation, the current use of AI appears from the synthesized evidence to be instrumented instructional support, frequently divorced from explicit theory-driven learning design. This review introduces a theory-informed, integrative, and explanatory framework for relationships between characteristics of physics content, AI pedagogical functions, and epistemic risks to provide a basis for future research to go beyond tool-centered evaluations, toward sustainable, discipline-sensitive AI integration in physics education.

Keywords: artificial intelligence, physics education, systematic literature review, learning outcomes, AI-based instruction, pedagogical challenges

1. Introduction

Digital transformation has become a preeminent demand of modern higher education, leading to more bespoke, flexible, and learner-centered pedagogies. The European Digital Education Action Plan 2021–2027 sets out the necessity of using digital technologies in order to create collaborative and creative learning environments [1,2]. Within this broader landscape, a new concept of the influential factor behind pedagogical evolution has emerged in the form of artificial intelligence (AI) and generative systems such as ChatGPT that are capable of generating near-human-like content and strengthening instructional processes [3–5]. In terms of STEM education, AI-based tools such as intelligent tutoring systems, adaptive platforms, automated assessments, and virtual laboratories have shown great potential in increasing learner engagement, personalization, and instructional efficacy [6,7]. Recent systematic reviews also report a marked increase in AI-in-education publications since 2022, suggesting that the rise of generative AI has accelerated both experimentation and scholarly attention in digital learning environments [8,9].

These advancements are particularly relevant against the background of persistent constraints in conventional physics education. The abstract and mathematically demanding nature of physics often results in heavy cognitive load, thus limiting motivation and leading to superficial understanding [10–12]. Empirical research has proven that active and collaborative forms of learning provide better learning outcomes than purely lecture-based approaches, if they are used strategically [13–15]. Accordingly, AI-supported tools that enable immediate feedback, adaptive explanations, and interactive problem-solving provide exciting opportunities to address persistent pedagogical deficits in learning physics [16].

Recent literature indicates that AI use in physics education is diversifying across classroom discussion, problem-solving support, assessment, and laboratory activities. For example, ChatGPT-based activities have been piloted in quantum physics classrooms [17], student and AI responses to physics problems have been compared to examine sensemaking and mechanistic reasoning [18], and generative AI has begun to be explored as a laboratory partner in school physics settings [19]. While AI-supported tools may enhance critical thinking, engagement, and task completion, concerns remain regarding incorrect or misleading outputs, particularly in conceptually complex areas of physics [20,21]. These early studies therefore suggest both growing pedagogical potential and the need for more systematic evidence on where AI is used, how it is implemented, and what limitations accompany its use in physics learning.

Despite the increasing adoption rate, the implementation of AI for the educational sector of physics is still fragmented. AI tools are used in a variety of ways, as a problem-solving assistant, content producer, assessment aid, and adaptive learner, but there is significant heterogeneity in the institutional policy, infrastructural provision, and faculty preparedness [22]. To make sense of this diversity, recent review works suggest that AI integration should be examined not only in terms of technology adoption but also in relation to pedagogy, task transformation, and disciplinary content. In particular, implementation-oriented reviews have used Laurillard's Conversational Framework and the SAMR model and have proposed a GenAI-TPACK lens to examine how AI aligns with teaching goals, reshapes learning tasks, and fits subject-specific knowledge in higher education [23]. Nonetheless, empirical evidence identifying which physics topics benefit most from AI, through what implementation models, and under what instructional conditions, remains limited [24–26].

Furthermore, persistent technical and non-technical impediments such as lack of AI literacy, infrastructural deficits, ethical apprehensions, and the danger of excessive dependency continue to pose challenges to efficacious deployment in a range of educational settings and milieus [27]. These impediments highlight that effective AI integration depends on technological readiness, pedagogical readiness, AI literacy, and ethical governance rather than tool availability alone.

In light of these opportunities and challenges, a systematic synthesis of recent empirical evidence is needed. Existing reviews have examined AI in education broadly [8], AI in science education more generally [26], and generative AI in science education at the subject-cluster level [28]. Recent implementation-oriented reviews have also synthesized GenAI case studies in higher education and proposed integrative lenses such as Laurillard, SAMR, and GenAI-TPACK [23]. However, these reviews do not specifically map how AI is implemented across physics content domains, pedagogical functions, reported outcomes, and contextual challenges in empirical studies published during the recent generative-AI surge. Accordingly, this review addresses a more specific gap by synthesizing empirical studies published between 2021 and 2025 and focusing on four dimensions: physics content coverage, implementation models, reported outcomes, and pedagogical challenges. This gap is what motivates the following research questions:

1. Which physics content domains are represented in empirical studies of AI use in physics education?
2. What pedagogical roles and implementation patterns of AI are reported in physics education?
3. What student, teacher, and instructional outcomes are reported in empirical studies of AI use?
4. What methodological, pedagogical, and contextual challenges are reported?

2. Literature review

2.1. The theory of artificial intelligence in education

Artificial intelligence (AI) has become an important strategic component in modern educational practice, mainly because of its ability to support personalized learning through adaptive content, teaching approaches, and feedback mechanisms that can be aligned with learners' needs, progress, and performance data [29–31]. Contemporary AI-based systems, such as highly sophisticated intelligent tutoring systems and virtual assistant technologies, are designed to track learning paths and patterns of mistakes and thereby produce automated assessments and provide immediate context-sensitive feedback. These capabilities promote greater student engagement and greater efficiency of learning [32–34].

Nevertheless, a substantial body of empirical scholarship highlights major pedagogical dangers that emerge with the use of AI without explicit instructional scaffolding. An excessive dependence on AI may destroy students' initiative, the originality of their work, and their capacity for autonomous critical thinking. Moreover, unstructured usage is frequently seen as favoring superficial learning and ignoring core disciplinary concepts, particularly if AI is viewed as an answer-generating substitute rather than as an aid for cognitive support [35–37]. Consequently, there is a strong consensus that AI should act as an augmentative tool to complement educators and learners and therefore be implemented within explicit pedagogical and ethical frameworks [38–41]. Recent implementation-oriented reviews further suggest that AI integration should be analyzed through frameworks such as Laurillard's Conversational Framework, SAMR, and GenAI-TPACK so that technological affordances remain aligned with pedagogy, task design, and disciplinary content [23]. Conceptual studies in related educational fields have also described AI-supported learning as an iterative process that integrates data collection, dynamic modeling, instructional feedback, and personalized interventions [42]. Although these studies focus on physical education rather than physics education, they remain conceptually relevant for understanding how AI can support the design of responsive learning.

2.2. Use of artificial intelligence in the education of physics

In the area of physics education, AI is widely used to address the field's inherent abstraction and conceptual complexity. Through AI-driven simulations and interactive environments, learners can dynamically explore physical systems, manipulate key variables, and observe immediate consequences, thereby supporting conceptual understanding and expanding learning opportunities beyond the limitations imposed by traditional laboratory settings [43,44]. This role is particularly consistent with constructivist and inquiry-oriented physics learning, in which students refine understanding by testing representations, receiving feedback, and revising misconceptions through interaction with tasks and tools. In addition to offering adaptive learning pathways that can customize explanations to the capability of the learner and enable personalized repetition, AI-enabled AR/VR environments make traditionally intangible subjects, such as wave phenomena and principles of optics, more tangible and accessible to learners [32,33,45].

However, existing research still tends to focus on particular content domains, most commonly mechanics, thermodynamics, electromagnetism, and modern physics, and does not yet provide a systematic mapping of the integration of AI throughout the curriculum of physics or a detailed exploration of the implications for the roles, workload, and pedagogical adaptations of educators [46–50]. Moreover, empirical data describing how AI moderates learners' cognitive processes, reasoning strategies, and interaction patterns in physics classrooms are still sparse. The lack of comprehensive metrics points to the need for focused research to not only examine the effectiveness of interventions using AI-based approaches but also to examine which instructional contexts, phases of physics instruction, and learner demographics benefit the most [51]. Accordingly, a more targeted synthesis is needed to identify which physics content domains have actually been addressed in empirical AI studies and how those patterns vary across implementations.

3. Method

The current study employed a systematic literature review (SLR) design and was reported in accordance with the PRISMA 2020 guidelines, thus ensuring transparency and traceability in the process of reviewing literature [52]. Methodologically, this study is best characterized as a systematic literature review with mapping elements. Unlike a scoping review, which is generally used to examine the breadth of a broad field, or an evidence and gap map, which primarily visualizes where evidence exists, the present review applied explicit eligibility criteria and structured coding to synthesize empirical studies in relation to four predefined research questions [8,53]. The review consisted of three major stages: (1) systematic literature searching, (2) article selection according to explicit eligibility criteria, and (3) data extraction and thematic coding of included studies. All the screening and coding procedures were performed by a single reviewer with domain knowledge in physics education and educational technology. The review followed a predefined protocol that was aligned with the research questions and that included consistent inclusion, exclusion, and coding criteria at all stages of the review. This procedural framework was developed in order to ensure that the review captured only empirical studies on the use of artificial intelligence in physics education within the given publication period. The process of study identification, screening, and inclusion is summarized in the PRISMA flow diagram, which describes the number of records that were kept or excluded at each stage. Because study selection and coding were conducted by a single reviewer, supervisory consultation was used to clarify ambiguities; however, this should still be acknowledged as a methodological limitation because independent multi-reviewer screening is generally recommended to reduce erroneous exclusion of eligible studies [54].

3.1. Literature search

A systematic literature search was conducted using six major academic databases relevant to education and educational technology: Scopus, ScienceDirect, IEEE Xplore, ERIC, Springer Nature, and Taylor & Francis. The search query focused on empirical research on the use of AI in physics education, using the core search words "physics education" and "AI", combined through the Boolean operator AND. More specifically, the search strategy was anchored in the combination “physics education” AND “AI”, with database-specific field adjustments to maximize relevance across platforms. In line with PRISMA-S, the review reports the databases searched, record counts retrieved from each source, and the search limits applied [52]. Search fields were carefully customized within each database in order to maximize relevance. Consequently, the search yielded a total of 4928 records, consisting of 4385 records from Scopus, 118 records from ScienceDirect, 28 records from IEEE Xplore, 52 records from ERIC, 269 records from Springer Nature, and 76 records from Taylor & Francis. To ensure the collection of the latest views on AI-assisted learning in physics, the publication window was limited to the period 2021–2025. Only journal articles and conference papers published in English were considered for further screening.

3.2. Literature selection

The article selection process followed a three-step screening protocol, in line with the rules of evidence in PRISMA, as shown in Figure 1. The inclusion criteria covered publication year, keyword relevance, language, document type, empirical design, classroom or instructional implementation, and relevance to the review questions. The exclusion criteria removed reviews, bibliometric studies, conceptual papers, non-English publications, and AI development studies without actual educational use.

• At the **first stage of retrieval**, there were 4928 records identified. Following the removal of duplicate references (n = 196) using Zotero, 4732 unique articles advanced to the abstract screening phase.
• During the **abstract screening** process, the titles and abstracts of the studies were evaluated using the Rayyan AI platform. Articles were excluded when they were not relevant to AI or physics education (n = 3846), when they were in other publication formats (n = 89), when they were published in languages other than English (n = 96), and when they were outside the predetermined publication range (n = 82). This process culminated in 619 articles that were considered eligible for full-text review.
• In the **full-text assessment stage,** further exclusions were carried out based on topic mismatch (n = 501), unavailability of full texts (n = 27), non-empirical research designs (n = 6), AI development studies without classroom implementation (n = 40), and failure to address the review questions (n = 1). Overall, 44 empirical studies met all eligibility criteria and were thus included in the final analysis.

This staged procedure was intended to balance breadth and relevance. At the same time, the reliance on single-reviewer screening should be interpreted cautiously, because more than one reviewer is commonly recommended in systematic review selection procedures [54].

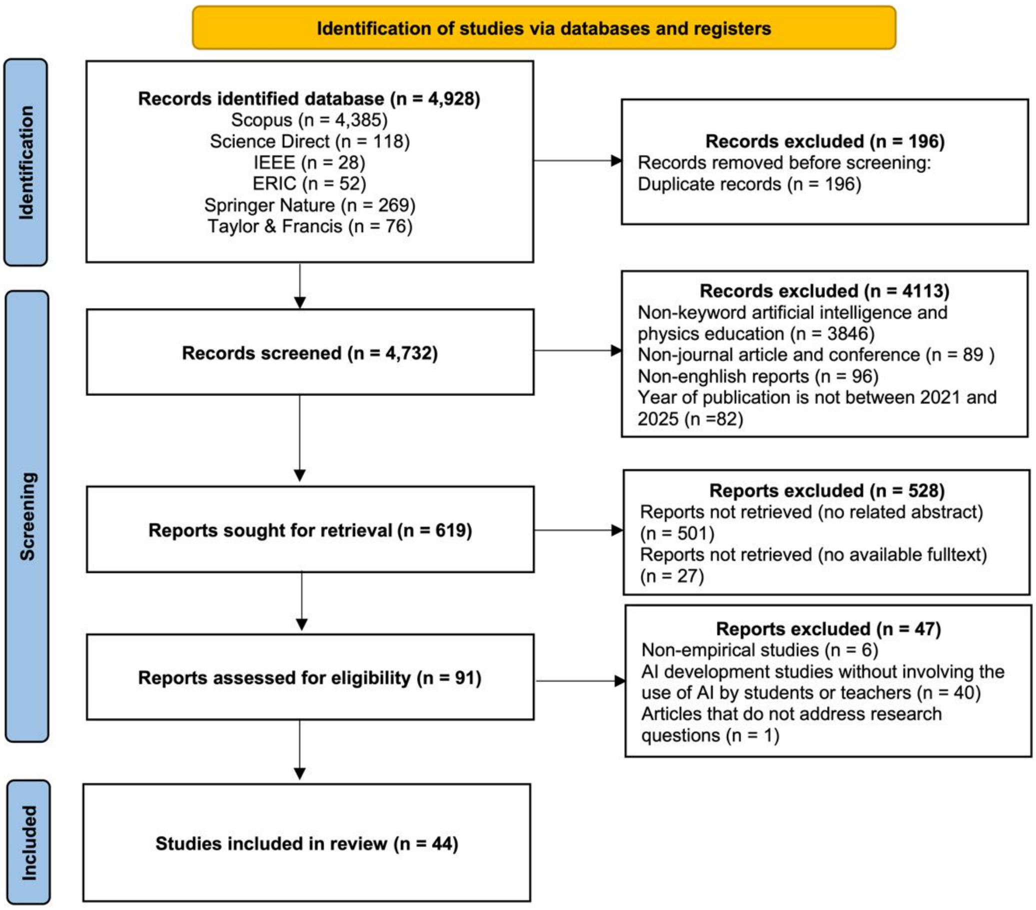


Figure 1. PRISMA 2020 flow diagram for study identification and selection.

3.3. Literature coding

In this review, the chosen studies were carefully coded in light of the four predetermined research questions (RQ1–RQ4). RQ1 focused on the content domains of physics that relate to artificial intelligence; RQ2 addressed the most common ways of deploying AI; RQ3 addressed the pedagogical implications of using AI by both learners and instructors; and RQ4 addressed the barriers that are faced when implementing AI. In addition to coding studies according to these four themes, descriptive study characteristics were also extracted, including publication year, country or region, sample size, participant type, research design, and type of AI used. This enabled the review to group the evidence not only by topic but also by year, geography, and AI category. The eligibility criteria used for study selection are presented in Table 1.

Each article was analyzed in its entirety and coded deductively according to established categories, but with the openness to emerging new themes. The physics curriculum content (RQ1) was categorized based on curriculum-aligned topic categories. AI applications (RQ2) were listed in functional groups of applications that include conversational agents, assessment and feedback mechanisms, personalization applications, simulations, learning analytics, multimodal interfaces, and AI literacy initiatives. The learning impacts that were determined (RQ3) included academic performance, motivation, personalization, teacher roles, and potential adverse effects. Implementation challenges (RQ4) were outlined in terms of dimensions such as accuracy, ethics, infrastructure, AI literacy, overreliance, and methodology.

To strengthen the synthesis, methodological notes were also recorded for each study, especially sample size, research design, duration of intervention, and whether the study used a control or comparison group [55]. Particular attention was given to studies with weak designs, such as one-shot implementations, descriptive reports, or interventions without controls, because such features limit the strength of causal interpretation and the comparability of findings across studies [43,56]

Consistent application of uniform coding procedures was applied throughout the corpus by one reviewer. Ambiguities that emerged in the coding guidelines were subsequently clarified through consultation with the supervisor. The findings of the synthesis were analyzed using a combination of frequency analysis and thematic coding, and then presented in tabular form to provide insight into dominant trends and emerging patterns in the literature.

Table 1. Eligibility criteria.

Code	Criterion	Inclusion	Exclusion
1	Year of publication	2021–2025	<2021
2	Keyword relevancy	Contains terms related to AI and physics education	Not related to AI or physics education
3	Language	English	Non-English
4	Document type	Journal articles, conference papers	Book chapters, books, theses
5	Screening criteria (title and abstract)	Relevant to AI use in physics education; primary study	Literature reviews, SLRs, bibliometric studies, or irrelevant topics
6	Research design	Empirical studies	Non-empirical studies
7	Development stage	AI implemented in instructional or learning contexts involving students or teachers	AI development without classroom use
8	Relevance to research questions	Addresses at least one of RQ1–RQ4	Does not address the review questions

4. Result

4.1. Descriptive

4.1.1. Characteristics of included studies

To provide an overview of the general characteristics of the selected literature, the detailed descriptive and methodological profile of the included studies is presented in Appendix A (Table 2).

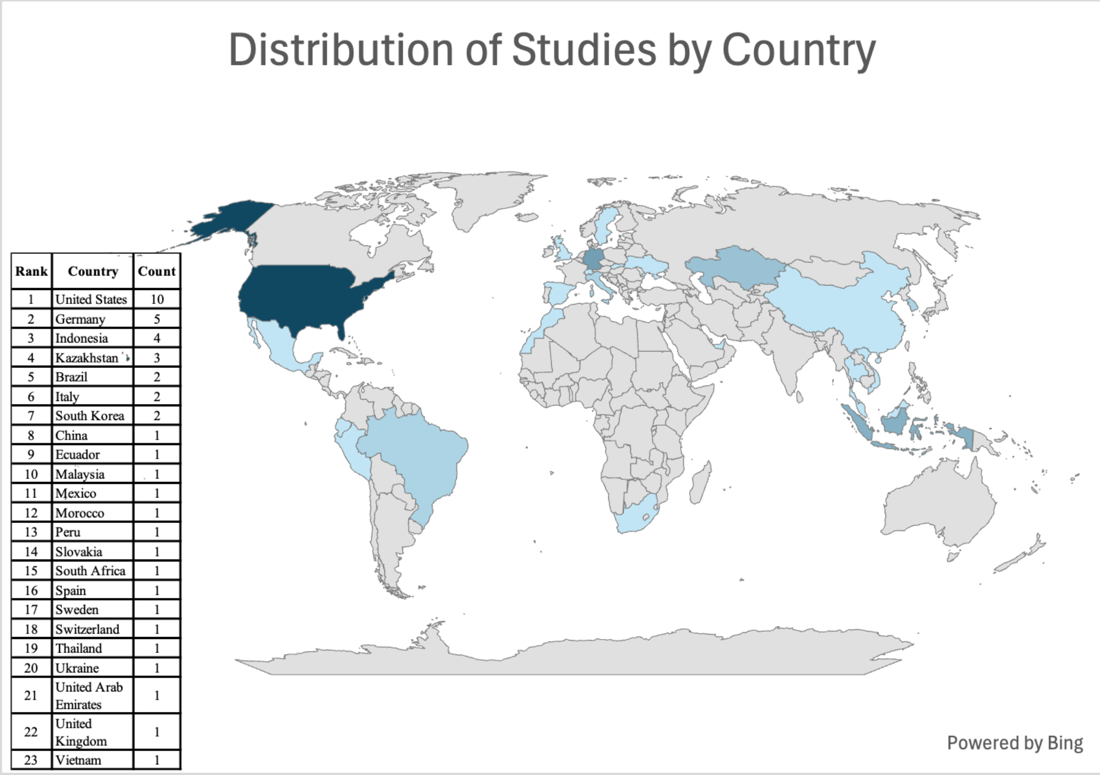
Appendix A (Table 2) shows that the corpus is dominated by journal articles, spans multiple countries, and is strongly concentrated in generative-AI applications, especially ChatGPT and other large language models. This pattern suggests that AI in physics education is entering a more established empirical phase; however, publication format alone should not be taken as a proxy for methodological strength. Moreover, the literature review suggests that recent work is increasingly oriented toward classroom use, feedback, assessment, and instructional support rather than purely technical experimentation [17,70,80].

Importantly, the same table reveals a very unbalanced state of research orientation: studies explicitly framed around AI tools and systems are much more numerous than studies fundamentally grounded in the disciplinary domain of physics education. This pattern implies that a significant percentage of the literature takes a technology-centric perspective, rendering visible the AI tools and functionalities over the physics-specific learning problems, such as conceptual misconceptions, representational competence, or disciplinary reasoning. As such, the field of AI often becomes a generic instructional adjunct in physics pedagogy, rather than one that is integrative. In other words, the corpus also suggests that many studies foreground tool functionality more clearly than explicit alignment with physics-specific pedagogical goals. From a constructivist or inquiry-oriented perspective, this matters because the educational value of AI depends not only on tool availability but also on how well the tool supports conceptual change, guided feedback, scientific reasoning, and meaningful learner interaction. Recent implementation-oriented reviews similarly argue that AI integration should be interpreted through pedagogical frameworks rather than through technology adoption alone [23].

At the same time, the methodological profile of the corpus appears uneven. These descriptive patterns should be interpreted cautiously because much of the corpus appears exploratory, context-bound, or methodologically limited. Where studies rely on one-group interventions, self-reported perceptions, or designs without control or comparison groups, claims about effectiveness should be interpreted cautiously because such designs provide weaker support for causal inference. This also affects external validity, since findings from specific institutions, cohorts, or instructional settings may not generalize straightforwardly to other populations or contexts. Accordingly, the expanded version of Appendix A (Table 2) should include participant type, sample size, and research design so that readers can distinguish exploratory evidence from stronger comparative evidence [64].

4.1.2. Geographical distribution of included studies

To identify dominant regions and underrepresented contexts, the geographical distribution of the included studies is presented below.



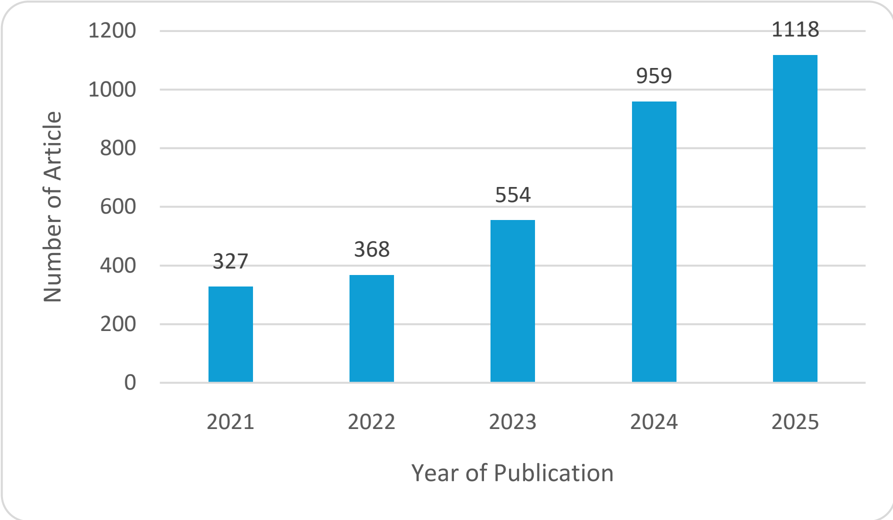
Rank	Country	Count
1	United States	10
2	Germany	5
3	Indonesia	4
4	Kazakhstan	3
5	Brazil	2
6	Italy	2
7	South Korea	2
8	China	1
9	Ecuador	1
10	Malaysia	1
11	Mexico	1
12	Morocco	1
13	Peru	1
14	Slovakia	1
15	South Africa	1
16	Spain	1
17	Sweden	1
18	Switzerland	1
19	Thailand	1
20	Ukraine	1
21	United Arab Emirates	1
22	United Kingdom	1
23	Vietnam	1

Figure 2. Distribution of studies by country.

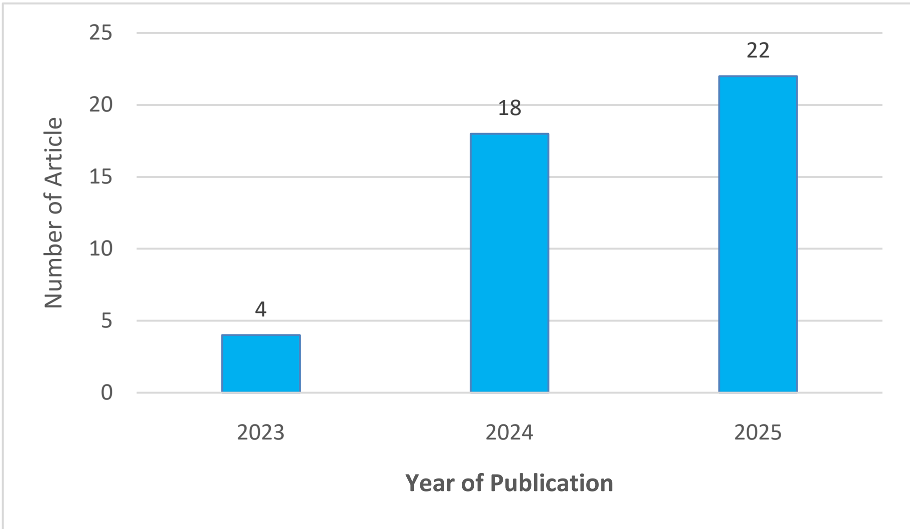
As indicated by Figure 2, the studies are geographically distributed in multiple areas, with a high representation from Europe, Asia, and North America. This reflects a growing international interest in AI-supported physics education, but not a balanced global evidence base. However, this broader spread should not be equated with representational balance, because the evidence still appears concentrated in relatively well-resourced or digitally advancing contexts. Nevertheless, the concentration of studies in such contexts may restrict our understanding of AI implementation in under-resourced settings, different curricular systems, or institutions with lower technological readiness [57,73]. This issue is important not only descriptively but also methodologically, because external validity depends on whether findings can travel across populations, settings, and instructional conditions [41].

4.1.3. Publication trends by year

To illustrate the temporal development of AI research in physics education, publication trends by year are presented in Figure 3.



(a)



(b)

Figure 3. (a) Yearly distribution of retrieved records across databases before screening. (b) Yearly publication trends of empirical AI studies in physics education.

Figure 3a presents the yearly distribution of records retrieved before screening, whereas Figure 3b presents the yearly distribution of the 44 included empirical studies. No eligible empirical studies published in 2021 or 2022 met the final inclusion criteria. Such an increase is consistent with the release of the much-needed generative AI technologies, specifically the large language models, and represents a shift from exploratory curiosity to mainstream scholarly interest. Within the context of physics education, this is a growing trend for AI-enhanced instruction, assessment, feedback, and laboratory experiences.

4.2. RQ1. Which physics content domains are represented in empirical studies of AI use in physics education?

RQ1 identifies the distribution of physics content domains in which AI has been implemented in educational settings and examines how domain-specific epistemic characteristics may shape where AI is more readily implemented. This analysis is intended not just to describe topical coverage but also to identify dominant concentrations and areas of unexplored territory that may represent opportunities for research. The distribution of the physics content covered in the reviewed studies is summarized in Table 3.

Table 3. Distribution of included studies by physics content domain, sub-topic, frequency, and study code.

Physics content domain	Sub-topic	Frequency	Study code
Mechanics	Newton's Laws (I, II, III); collision problems (1D); conservation of momentum; force and motion; kinematics (rolling motion, projectile motion, translational kinematics); translational and rotational dynamics; pendulum (including Wilberforce); Hooke's law and springs; work; mechanics energy; center of mass; moment of inertia; rigid body dynamics; classical mechanics; Gravitron assignment; introductory mechanics course; mechanics and its laboratory.	17	A1; A18; A19; A21; A22; A23; A25; A26; A30; A14; A15; A35; A36; A42; A43; A10; A12
Thermodynamics and heat (including climate physics/cryogenics)	Thermodynamics; calorimetry; P-V cycles; ideal gas laws; isothermal and adiabatic processes; heat transfer (conduction, convection, radiation); thermal conductivity; 2D/3D temperature distribution; 2D plate models; energy and heat; temperature; thermodynamic equilibrium; thermodynamics in the lab; climate physics (greenhouse effect, global warming) as context; cryogenics.	11	A3; A4; A5; A7; A9; A21; A22; A26; A27; A28; A38
Waves and sound (including seismic)	Mechanics waves; standing waves; wave phase; acoustic interference and diffraction; wave properties (wavelength, frequency, amplitude); sound; speed of sound; seismic waves (seismic activity); oscillations and waves in the laboratory.	7	A4; A8; A24; A25; A32; A42; A43
Optics and electromagnetic waves	Light waves; optics; optical interference and diffraction; radio (part of the EM spectrum); electromagnetic waves; electromagnetic spectrum.	6	A4; A5; A8; A32; A39; A42
Electricity and magnetism (electromagnetism)	Electricity; electric current; electric circuit; E and M (electricity and magnetism); electromagnetic field; electromagnetism (general); electrodynamics; electromagnetic problems; electromagnetism in the laboratory.	8	A5; A8; A14; A22; A23; A34; A20; A42
Modern physics (quantum, relativity, atomic/nuclear/particle, solid state, plasma)	Quantum cryptography; quantum mechanics/physics; wave-particle duality; quantum entanglement; eigenvalues; EM spectrum (quantum context); general relativity theory; relativity; curved spacetime and gravity; relativistic concepts; modern physics; atomic, nuclear and particle physics; radioactivity; plasma physics; solid state physics.	11	A2; A5; A17; A13; A31; A34; A22; A41; A39; A44; A42
Astronomy and astrophysics	Astronomy and astrophysics; Earth and space systems (from an astronomical perspective); gravitational wave astronomy.	2	A5; A41

As shown in Table 3, applications of AI in physics education show a systematic focus on particular areas and not a general coverage of the curriculum. Mechanics emerges as the clearest concentration, suggesting that current AI integration is strongly oriented toward foundational topics that lend themselves to structured quantitative reasoning and step-by-step problem solving. The epistemic structure of the mechanics, in which student work is usually externalized in the form of equations, numerical answers, and procedures of solution, makes it especially amenable to the kind of dominant affordances that AI offers in the form of automated feedback, intelligent tutoring, and AI-supported problem generation. From a constructivist and guided-inquiry perspective, this alignment is pedagogically plausible because mechanics tasks often generate explicit learner traces that enable iterative feedback, error diagnosis, and scaffolded problem solving. This helps to understand why topics in mechanics, such as kinematics, Newton's laws, collisions, and energy conservation, keep appearing in the empirical corpus over and over again [21,70,71].

Thermodynamics/heat and modern physics form a second prominent cluster, but their visibility appears to reflect two somewhat different patterns of epistemic and pedagogical alignment. In thermodynamics, AI-supported learning is often linked to context-based and data-rich learning environments (e.g., heat transfer, temperature distributions, and climate-related contexts) in which AI can support interpretation, modeling, and context-based reasoning instead of just procedural checking of answers [57,58]. By contrast, in modern physics, the pattern points more strongly to the use of AI as cognitive and representational scaffolding for abstract, mathematically intensive, and conceptually demanding topics such as quantum physics and relativity. In the language of contemporary physics, the high frequency indicates that the increasing use of AI as cognitive scaffolding is used to support student understanding of abstract and mathematically intense concepts (e.g., quantum and relativity), in which conventional instruction often runs up against entrenched conceptual barriers. Thus, Table 3 suggests that AI adoption is not only concentrated in "easy to automate" domains but also in domains where the demands for representational and explanatory support are high [17,44]. This also indicates that the current corpus is shaped by the types of AI being used: many recent studies rely on LLM-based feedback, ChatGPT-supported tutoring, or multimodal visualization tools rather than traditional rule-based systems alone.

In contrast, electricity and magnetism, waves and sound, and optics/electromagnetic waves appear only moderately represented despite their central role in the physics curriculum. This pattern may reflect structural underrepresentation shaped by the epistemic demands of these domains. Topics like waves, optics, electromagnetism, etc., often demand intensive multi-representational coordination (e.g., phase, interference, field/diagram reasoning, and translation between conceptual and mathematical representations), which is perhaps less directly provided by AI implementations that are dominated by text-based tutoring and feedback. In other words, the present evidence base appears better aligned with domains that produce linear, verbalizable, or procedural traces than with domains that depend heavily on spatial, field-based, or diagrammatic reasoning. As a result, these domains are less commonly found in the empirical literature, even though they are pedagogically important.

Astronomy and astrophysics remain at the margins of the corpus, suggesting that AI integration has been concentrated mainly on classroom-scale physics topics. Rather than being a reflection of low relevance, such scarcity might be a manifestation of the greater design complexity for astronomy learning, in which observational inference, large-scale phenomena, and visualization-centered reasoning are often involved. It may also reflect the relative scarcity of AI implementations specifically designed for content areas that depend on large-scale data interpretation and observational modeling rather than immediate stepwise feedback.

These domain patterns should be interpreted cautiously, given the exploratory and context-bound character of much of the corpus [17,19,21,72]. In addition, where studies rely on one-group interventions or designs without control or comparison groups, conclusions about the comparative effectiveness of AI across content domains remain limited. These features also constrain external validity, because findings from specific institutions, learner groups, or technological settings may not transfer straightforwardly to other educational contexts.

Overall, the results of RQ1 reveal that AI in physics education is currently focused on areas that either (a) generate machine-interpretable traces of learning amenable to procedural tutoring and assessment (e.g., mechanics) or (b) create high demand for scaffolding of abstract reasoning (e.g., modern physics). By contrast, domains requiring complex multi-representational coordination or observational inference remain comparatively underrepresented. Taken together, these findings suggest that current AI integration in physics education is shaped not only by technological availability but also by the degree to which specific content domains align with the pedagogical goals and representational affordances that current AI tools can realistically support.

4.3. RQ2. What pedagogical roles and implementation patterns of AI are reported in physics education?

RQ2 investigates how AI is applied in physics education by identifying dominant patterns of implementation and the pedagogical orientations that they reflect. Rather than treating AI adoption merely as a technological trend, this analysis emphasizes what instructional functions receive attention in empirical work and which functions still remain underexplored. The classification of AI use across the reviewed studies is summarized in Table 4.

As shown in Table 4, eight major categories of AI utilization were identified. The clearest concentration is in conversation-based support, especially chatbots, conversational assistants, and tutor-like systems. This dominance implies that interactive dialogue has emerged as the most accessible and widely tested pathway for integrating AI into physics learning. In many cases, AI is positioned as a real-time cognitive support tool giving explanations, procedural problem-solving guidance, and support for students as they work through complex physics tasks. This pattern aligns with the popularity of LLM-based tutoring in physics learning environments, especially in topics that contain a significant amount of mathematics and are well-structured in terms of procedures [26,66]. Notably, more specialized forms of conversation, such as laboratory or experiment assistants (2) and metacognitive dialogue systems (1), are still rare, suggesting that conversational use focuses predominantly on the provision of assistance in the immediate situation rather than on sustained reflection, monitoring, and epistemic control during learning.

The next most prominent categories are *personalization and recommendations* (18 studies) and *AI assessment and feedback* (16 studies). These categories respond to long-standing problems of physics education, including differences in students' prior knowledge and the practical difficulty of giving timely and individual feedback. Personalized learning systems usually adapt learning pathways or content difficulty, whereas automated grading and formative feedback tools are intended to enhance the efficiency of the instructional processes as well as diagnostic precision [58,80]. However, many of these applications are framed with limited explicit grounding in physics-specific learning theories, such as conceptual change, learning progressions, or representational competence, while prioritizing measurable gains in speed, accuracy, or task completion. This creates a risk that "successful AI use" comes to be defined narrowly in terms of improved performance indicators rather than deeper disciplinary reasoning.

Table 4. Classification of AI use in physics education by main pedagogical function, derivative form, citations, and unique study coverage.

Basic form of usage	Derived usage forms (with frequency per derivative)	Study code	Article frequency (unique per main topic)
Conversation assistant and tutor	Chatbots/conversational assistants (26); smart tutors/virtual tutors (7); lab assistant/experiment assistant (2); metacognitive dialogue with AI (1).	A22; A17; A35; A26; A2; A25; A44; A39; A12; A19; A38; A32; A30; A24; A14; A9; A43; A36; A11; A7; A31; A21; A33; A27; A8; A6; A16	27
AI assessment and feedback	Automated grading system/AI-assisted grading (9); automated/real-time/formative feedback (8); LLM-based grading assistant (3).	A22; A2; A5; A19; A38; A15; A29; A20; A9; A36; A4; A31; A18; A34; A27; A8	16
Personalization and recommendations	Personalized learning/personalized learning systems (16); recommendation systems and automatic LO labeling (7); diagnostic analysis of misconceptions (1).	A22; A40; A35; A32; A3; A16; A36; A10; A23; A41; A28; A4; A31; A37; A34; A27; A8; A42	18
Simulations and virtual experiments	Artificial intelligence-based simulations/virtual experiments (8); mathematical modeling and discovery of equations with genetic algorithms/symbolic regression (1).	A17; A42; A23; A28; A31; A37; A34; A8	8
Content and knowledge generation	Content generation (text, questions, explanations, concept items, arguments, reading materials) (10); image generation AI (1); AI as a source of knowledge/problem-solving copilot (3).	A13; A26; A5; A38; A32; A29; A14; A44; A7; A18; A34; A8	12
Learning analytics and data analysis	Learning analytics/data analysis (11); plagiarism detection and academic integrity monitoring (1).	A22; A40; A25; A15; A29; A20; A14; A9; A10; A41; A7; A28; A34	13
Multimodal interface (text-voice-image)	Image recognition/OCR/object detection (5); speech-to-text and text-to-speech (3).	A5; A9; A36; A41; A21; A27; A44	7
AI literacy and perception studies	AI literacy tools and AI literacy training (2); AI perception/acceptance studies (4).	A22; A5; A25; A29; A1; A31	6

At a moderate level of representation, the categories were *content and knowledge generation* (12 studies) and *learning analytics and data analysis* (13 studies). Within this grouping, there is a tendency for AI to function either as a productive agent (generating explanations, questions, and learning materials) or as an analytical agent (extracting patterns from learning data). A notable point in Table 4 is that, despite widespread concern about AI misuse in education, only one subtype within this category directly addresses academic integrity monitoring. This implies that there has been a faster development of the literature in terms of the adoption of AI for instructional support than in the establishment of systematic protections for the evaluation of validity and ethical use of AI in education, a disparity of growing importance when generative AI is becoming more embedded in student learning practices [17,67].

Less frequently explored areas, *simulations and virtual experiments* (8 studies), *multimodal interfaces* (7 studies), and *AI literacy and perception studies* (6 studies) remain promising yet relatively underdeveloped areas. This pattern is especially salient in physics education because physics learning is strongly model-based, relies heavily on experimentation, and requires multi-representational coordination. AI-driven simulations and virtual experiments may offer scalable alternatives to resource-intensive or complicated lab setups, but these usually require substantially more design work, including validated models, careful task constraints, and meaningful integration with inquiry-based learning [69]. Similarly, multimodal AI systems, including those that integrate OCR/object detection and speech interfaces, have the potential to support learning in physics, where diagrams, graphs, and visual representations are central to disciplinary thinking. However, these are relatively infrequent in occurrence, which means that multimodal support is a peripheral rather than a mainstream implementation strategy [59]. Finally, the few AI literacy and perception studies suggest that current scholarship is focused on short-term instructional functionality and not on longer-term readiness, teacher capacity, and issues of sustainable integration that are essential to the responsible and consistent use of AI in physics classrooms [21].

Overall, AI use in physics education is concentrated in conversational support, personalization, and assessment, whereas metacognitive support, multimodal scaffolding, and simulation-based inquiry remain comparatively underexplored. These patterns suggest that this is a rapidly emerging field that is experimenting with AI as a learning support tool, but is still in the process of building more theory-grounded and more discipline-sensitive designs that match the epistemic practices that are the core of physics learning.

4.4. RQ3. What student, teacher, and instructional outcomes are reported in empirical studies of AI use?

RQ3 investigates the effects of AI implementation in physics education on some of the most important user groups, such as students, teachers, and the broader learning process. This assessment is necessary in order to evaluate not only the potential benefits of learning but also the pedagogical risks that may come with the routine use of AI in physics. Based on the synthesis of the selected studies, five major categories of AI impact were identified, including both reported benefits and critical challenges. The distribution of these outcomes is summarized in Table 5.

As Table 5 indicates, the most commonly reported impact concerns teacher and researcher efficiency (14 studies). Across these studies, AI is applied to support workload-intensive activities in teaching and research in physics, such as automated task grading, learning objective labeling, lesson and question generation, data coding, and AI-enabled teaching assistance [21,62]. This pattern suggests that current AI adoption is often framed in the context of instructional efficiency and professional support as opposed to wholesale pedagogical redesign [82,88,89]. Practically speaking, the role of AI is being positioned in the literature as a way to reduce time spent on routine instructional labor, with the potential of freeing up teacher attention to facilitation, discussion, and higher-order sense making that is central to physics learning.

Table 5. Reported learning, instructional, and user-related outcomes of AI use in physics education.

Outcome category	Sub-category	Users (teachers/students)	Study code	Frequency
Student learning outcomes and skills	AI is used as chatbots/LLMs, adaptive learning modules, AI-generated physics content, student record analysis, or diagnostic systems to improve students' test scores, concept understanding, modeling, problem solving, and scientific skills.	Students	A2; A5; A12; A16; A17; A19; A23; A28; A35; A22; A44; A27	12
Student motivation, involvement, and attitude	AI is used in the form of a chatbot, as an AI tutor, integrated as ChatGPT in course activities, or in the form of AI-based media that was designed specifically to measure its effects on motivation, situational interest, positive affect, and students' attitudes toward physics and AI.	Students	A7; A11; A13; A31; A32; A39; A29; A33; A24	9
Self-guided learning and individualization	AI is used as a personalization engine (machine learning algorithms, adaptive digital modules) to identify at-risk students, provide culturally relevant interventions, adjust material levels/pace, and provide instant feedback that supports independent learning.	Students	A3; A4	2
Efficiency, role, and development of teachers/researchers	AI is used as an automatic grader, atomic learning objective labeler, research data coder, question and lesson plan generator, intelligent teaching assistant (ITAS, AI tutor, PhysicsAssistant), and AI training/literacy tool, thereby reducing workload and increasing the professional capacity of teachers/researchers.	Teachers and researchers (main), students as recipients of the impact	A1; A6; A9; A10; A15; A18; A38; A20; A36; A37; A41; A34; A42; A22	14
Risks, limitations, and negative consequences	AI (ChatGPT, AI, cheat sites, competition copilots, lab partners) is actually used by students/teachers, and its negative impacts have been analyzed: dependence on instant answers, decline in critical thinking and metacognition, misconceptions, reduced human interaction, cheating, and overtrust in incorrect answers.	Students and teachers	A1; A14; A21; A24; A25; A26; A30; A39; A43	9

In terms of student-centered outcomes, positive effects are most apparent in student learning outcomes and skills (12 studies) and motivation, engagement, and attitudes (9 studies). The reviewed studies often conclude that conversational support, AI-assisted feedback, AI-generated learning resources, and AI-supported laboratories or simulations are associated with improvements in conceptual understanding, problem-solving performance, and task engagement

[43,44,65,68,71,73,75]. Affective outcomes are also an area of frequent discussion, especially when AI is implemented in the form of a tutor-like presence, a structured classroom tool, or an activity in a course that alters the way that students approach tasks in physics [17,64,79,80,81,87]. Overall, these studies suggest that much of the existing impact evidence is focused on near-term cognitive and affective indicators, such as performance, perceived understanding, interest, and attitudes, rather than more distant-term disciplinary outcomes such as sustained epistemic agency and scientific reasoning practices.

However, self-directed learning and contextual personalization are the least represented category (2 studies) despite the more general presence of AI systems that technically make adaptation possible. In Table 5, very few studies frame personalization as an explicit means of promoting learner autonomy, for example, by identifying at-risk students, accommodating pace and level, or incorporating culturally responsive interventions [57,58]. This is important to consider because personalization can easily be "system-driven" (optimizing the delivery of content) without necessarily being "learner-driven" (assisting in planning, monitoring, and self-regulation). This pattern suggests that, within the reviewed corpus, autonomy-oriented personalization is more often treated as a technical capability than as a deliberate pedagogical goal.

Crucially, the literature also reports risks, limitations, and negative outcomes, which show that these concerns are not marginal. The concerns that remain are based on overreliance on AI-generated responses, reduced metacognitive engagement, misconceptions from flawed outputs, instances of cheating, violations of academic integrity, and excessive trust in seemingly plausible reasoning [19,20,41,66,74,78,90]. These findings complicate overly optimistic accounts of AI integration. Even when AI increases efficiency or short-term performance, the literature regularly states the potential that learners outsource verification of their work, reduce productive struggle, and embrace fallacious explanations of physics in the absence of guided and imposed constraints.

Overall, the empirical corpus reports stronger evidence for short-term gains in learning, engagement, and instructional efficiency than for long-term development of disciplinary reasoning, and these benefits remain accompanied by non-trivial epistemic and pedagogical risks. As such, future empirical efforts should think about AI not as an auxiliary technology that helps people perform their tasks better but as an instructional component that is part of the curriculum and is explicitly scaffolded with metacognitive supports, verification protocols, and AI literacy training, particularly for educators responsible for managing the everyday dynamics of classroom practice [21,56].

4.5. RQ4. What methodological, pedagogical, and contextual challenges are reported?

RQ4 examines the methodological, pedagogical, and contextual challenges associated with the use of artificial intelligence in physics education. The reviewed studies reveal 10 major categories of challenge, indicating that barriers to AI adoption are multifaceted and extend well beyond purely technical issues. These difficulties include epistemic reliability, pedagogical integrity, human factors such as literacy and professional identity, infrastructural readiness, and the strength of the empirical basis used to justify adoption. The full categorization is presented in Table 6.

Table 6. Reported methodological, pedagogical, and contextual challenges of AI use in physics education.

Challenge category	Description of challenge	Affected users	Study code	Frequency
Accuracy, conceptual errors, and AI output hallucinations	AI responses are often inaccurate, contain hallucinations, incoherent reasoning, or only provide a superficial understanding of physics concepts.	Students who use AI as a tutor/source of answers, and teachers/lecturers who use AI for questions and feedback.	A1; A5; A6; A7; A12; A13; A14; A18; A19; A20; A21; A25; A26; A27; A24; A35; A36; A38; A39; A42	20
Excessive dependence and decline in critical thinking	Students rely too much on ChatGPT/AI as an "answer machine" and accept the output without verification, resulting in a decline in independent problem-solving and critical thinking skills.	High school students/physics students who use AI to complete assignments and conceptual problems.	A1; A7; A12; A20; A21; A22; A27; A31; A32; A34; A42; A43; A33	13
Low AI literacy, technological anxiety, and training needs	Teachers and students do not yet understand how AI works, its potential, and its risks; there is technological anxiety, a lack of confidence, and a strong need for training and AI literacy.	Physics teachers (in-service and pre-service), prospective teachers, and students who must critically assess the quality of AI outputs.	A8; A22; A3; A4; A27; A28; A29; A32; A31; A34; A37; A41; A44; A17	14
Limitations in infrastructure, access, and cost	The integration of AI requires a stable internet connection, compatible devices, and ICT infrastructure; many schools/institutions in developing regions do not yet have adequate support for this.	Public schools in developing regions, universities with limited IT budgets, and institutions with low resources.	A3; A4; A5; A27; A8; A32; A31; A34; A36; A37; A42	11
Ethical issues, data privacy, academic integrity, and bias	Challenges related to plagiarism and academic dishonesty, personal data protection, lack of model transparency, and algorithmic bias that can perpetuate injustice.	Students (risk of plagiarism and misuse of AI), lecturers/teachers (academic integrity), and institutions/government (ethical policies and data protection).	A1; A5; A7; A20; A22; A27; A31; A32; A34; A36; A41; A38; A42	13
Feedback scalability and teacher workload	Providing personalized written feedback in large classes is very time-consuming; AI helps, but it still requires curation and verification by teachers, so the workload does not disappear.	Physics lecturers/teachers who teach large classes and teaching assistants who handle assessment and feedback.	A18; A19; A9; A11; A36	5
Limitations in AI capabilities (complex physics, multimodal, local languages)	AI has difficulty handling advanced conceptual physics, mathematics, visual/lab data, and non-English languages; it is also limited in generating new problems or creative solution variations.	Students studying advanced physics topics, laboratory practicums, and users in local languages (e.g., Korean, Thai).	A1; A7; A14; A18; A19; A20; A21; A35; A36; A34; A44; A23; A40	13
Limitations of the	Many studies have small sample sizes, short	Physics education	A2; A3; A4; A10; A11;	14

methodological framework and empirical evidence	durations, highly specific contexts, and do not yet provide a strong methodological framework for evaluating the effectiveness of AI in physics education.	researchers, curriculum developers, and policymakers who want to adopt AI in a systematic and evidence-based manner.	A12; A29; A30; A31; A32; A22; A16; A17; A40	
Changes in the role and professional identity of teachers	The integration of AI triggers role identity conflicts, concerns about losing professional authority, and pressure to adapt to the new role of teachers as facilitators who collaborate with AI.	High school physics teachers and lecturers are expected to utilize AI in teaching and experiments.	A1; A8; A22; A37; A34; A42	6
Technical skills: programming, prompting, and evaluating AI responses	Teachers and students often lack the skills to program/set up AI systems, write effective prompts, and critically evaluate the quality of AI responses.	Physics teachers/lecturers who develop AI activities or systems, and students who interact directly with chatbots or other AI tools.	A8; A16; A15; A28; A41; A44	6

As highlighted in Table 6, the most prominent challenge concerns inaccurate outputs, conceptual errors, and hallucinations. The stakes are high enough in the case of physics because modest conceptual lapses (the misstatement of the laws of forces, the misuse of signs in dynamics, or an erroneous understanding of the causal factors of phenomena, for example) can become petrified into misconceptions when expressed with seeming confidence. Across the reviewed studies, AI-generated explanations may appear coherent while remaining conceptually impoverished or incorrect, which becomes pedagogically dangerous when students treat textual fluency as a proxy for epistemic authority [21,66]. From the perspective of physics teaching, the problem is not limited to wrong answers; it also includes failures in dimensional analysis, boundary condition reasoning, model use, and consistency across. Closely related to this is excessive dependence on AI and the resulting erosion of critical thinking. Table 6 suggests that when AI is used as an "answer machine", students may avoid the use of verification, modeling, and sense-making processes that are central to physics education, especially when the value of a task is built on the construction and evaluation of reasoning, rather than a simple, numerical answer [61]. This problem is intensified by the issue of accuracy: overreliance on systems prone to hallucination weakens students' metacognitive oversight while increasing exposure to plausibly presented misconceptions [66].

Beyond epistemic reliability, Table 6 also brings human-factor barriers into the foreground. The reviewed literature points to low AI literacy, technological apprehension, training deficits, and gaps in practical skills such as prompting, output evaluation, and basic system setup. Collectively, the evidence is suggestive of the fact that the successful use of AI tools in physics education is not dependent merely on the fact that they are accessible in realistic contexts but on the proficiency of users who can critically interrogate AI outputs and embed them in pedagogical practice without abdicating epistemic authority [26,27]. In the physics classroom, which is characterized by high representational demands and a need to evaluate the quality of given explanations with regard to formal constraints, AI literacy needs to be a set of discipline-specific evaluative practices, including the ability to identify validity, justify claims, and balance consistency between equations, graphs, and physical reasoning.

Structural constraints are also significant. The literature also shows that AI integration is uneven across contexts, particularly in developing regions and low-resource institutions, where reliable internet access, device availability, and ICT infrastructure remain limited [57,58]. At the same time, ethical issues involving privacy, academic integrity, and bias highlight institutional concerns that cannot be resolved through classroom enthusiasm alone. The juxtaposition of easy text generation and assessment pressure makes academic integrity a rather tangible, as opposed to theoretical, dilemma, especially in physics courses, in which homework-style problems abound, and might be outsourced without rigorous task redesign [66]. These ethics-related barriers point to the need for explicit governance and deliberate assessment design, rather than relying only on pedagogical "warnings" to students.

Another recurring issue concerns the present limits of AI capability in physics-specific contexts. The literature reports difficulties with advanced mathematics, multimodal reasoning involving diagrams and laboratory data, and the handling of local languages, all of which are especially consequential in physics, where visual representation, measurement, and uncertainty reasoning are integral to disciplinary practice. When considered alongside concerns about feedback scalability and teacher workload, these limitations suggest that the perceived applicability of AI is often overstated. Although AI can serve to relieve educators from some routine labor, it does not relieve educators from their role in curation and validation, and alignment of feedback to curricular goals, especially where correctness cannot be understood simply from the plausibility of appearance [62].

A further challenge lies in the methodological fragility of the evidence base itself. Many interventions are short-term, context-specific, or based on small samples, which limits the strength of conclusions about sustained learning gains or long-term risks [44,91,92]. This limitation is not merely academic. Without more rigorous designs, institutions must make adoption decisions under uncertainty, while educators face pressure to "embrace AI" without clear guidance about its benefits, risks, or the most effective safeguards.

Overall, the challenge profile summarized in Table 6 shows that successful AI use in physics education depends as much on pedagogical and institutional readiness as on technical capability. The most persistent barriers (epistemic unreliability, overdependence, literacy gaps, integrity concerns, and methodological weakness) are deeply intertwined and imply that improvements to the tools alone are unlikely to result in progress. Future work would therefore be well served by longitudinal and theoretically grounded studies that treat verification practices, AI literacy, and assessment integrity as core design requirements, while also addressing infrastructural limitations and the changing professional role of physics teachers in AI-mediated learning environments [21,26].

5. Discussion, future research direction, conclusions, limitations

5.1. Discussion

The present review indicates that the recent expansion of AI in physics education should be interpreted with some caution. While the number of empirical studies has increased markedly after 2023, this growth does not necessarily indicate that the field has already reached pedagogical or methodological maturity. Rather, the literature seems to reflect a period of rapid experimentation in which the adoption of AI has progressed faster than the development of strong learning designs, clear theoretical grounding, and robust comparative evidence [17,66,71].

When the findings across RQ1–RQ4 are considered together, a more coherent picture emerges. AI adoption in physics education is not evenly distributed across the curriculum. It is concentrated most strongly in mechanics, thermodynamics, and modern physics, which share features that are relatively compatible with current AI tools, such as symbolic problem solving, step-by-step explanation, and high levels of abstraction that often challenge learners in conventional instruction [21,44,68]. By contrast, optics, waves, and astronomy appear less frequently in the reviewed studies, despite their importance in school and university physics. This imbalance is revealing. It suggests that the current spread of AI in physics has been shaped not only by educational need but also by the extent to which a topic fits the present strengths of AI systems. Topics that require stronger visual interpretation, spatial thinking, and coordination across multiple representations still seem less visible in current implementations [27,43,59].

A similar pattern appears in the dominant forms of implementation. Across the reviewed studies, AI is used most often as a conversational tutor, a feedback provider, an assessment aid, or a personalization tool. On the one hand, these functions respond to long-standing practical difficulties in physics teaching, such as limited teacher time, large classes, and substantial variation in students' prior knowledge [44,70,80]. On the other hand, it is important to note that many of these uses remain focused on instructional convenience and immediate support. In many cases, AI is added to explain, answer, or assist more quickly, but is not yet integrated into learning designs that deliberately strengthen the kinds of thinking central to physics, such as connecting equations to concepts, comparing representations, checking assumptions, and judging whether an answer makes physical sense [20,60,62,71].

This point becomes clearer when the findings are interpreted from an educational perspective. From a constructivist point of view, learning physics involves more than receiving correct answers. Students need to build understanding actively by connecting new ideas with prior conceptions, testing those ideas against tasks and evidence, and revising them when inconsistencies appear. From a social constructivist perspective, this process is supported by dialogue, guided participation, and carefully structured feedback. In this sense, AI can be educationally valuable when it functions as a scaffold that helps students explain, reflect, compare, and revise. However, when AI is used mainly as a shortcut to produce answers, it may reduce the effort needed for checking, interpretation, and conceptual reconstruction. This distinction is particularly important in physics, where understanding is demonstrated not only by obtaining a result but also by justifying it, relating it to representations, and explaining it in terms of physical principles [44,65,87].

The reported outcomes should therefore be read carefully. A considerable number of studies describe positive short-term effects on performance, motivation, engagement, and instructional efficiency [64,65,75,83]. These findings are meaningful and suggest that AI can reduce barriers to support, make feedback more immediate, and increase responsiveness in some learning situations. However, much less attention is given to whether these gains are accompanied by stronger conceptual understanding, more stable reasoning, or greater ability to evaluate explanations independently. The literature, therefore, seems more conclusive about improvement in task performance than about lasting improvement in understanding. This is an important distinction, because fluent AI-supported responses may appear convincing even when learners themselves have not yet developed a secure grasp of the underlying physics [20,60,77].

The challenges identified across the reviewed studies make this concern more explicit. Inaccurate answers, misleading explanations, overreliance, weak AI literacy, academic integrity concerns, and unequal infrastructure are not merely side issues. They are closely connected to the very features that make AI attractive, namely speed, accessibility, and fluency of response [21,27,66,90]. In physics, these risks are particularly significant because students need to learn how to check whether a solution is reasonable, whether a sign convention is correct, whether a diagram and an equation tell the same story, and whether an explanation is scientifically sound. If AI use is not accompanied by these checking practices, then the technology may help students complete tasks more quickly while weakening the habits of careful reasoning that physics education is supposed to develop [19,20,41,90].

A further issue concerns the methodological profile of the field itself. A substantial proportion of the reviewed studies are pilot studies, case studies, exploratory implementations, or perception-based reports. In many cases, there is no control group, no comparison condition, a limited duration, and little attention to delayed outcomes. Under such conditions, it becomes difficult to determine whether the reported benefits are produced by AI itself, by novelty effects, by increased teacher attention, or simply by more time spent on tasks. This weakens the strength of causal claims. It also limits external validity, because many studies come from relatively well-resourced institutions or specific local settings. As a result, the present evidence base is promising, but still not strong enough to support broad conclusions about what forms of AI integration work best across different students, topics, and teaching conditions [61,66,72,89].

Overall, these findings suggest that the next stage of research in this area should move beyond tool-centered adoption. The central issue is not simply whether AI can be used in physics education, but how it can be integrated in ways that remain closely connected to the aims of learning physics. For that reason, the present review proposes a framework for AI integration in physics education that places equal emphasis on the nature of the physics content, the intended learning goals, support for students' thinking, and responsible design and evaluation. Based on the synthesis of the reviewed studies, a suggested framework for AI integration in physics education can be seen in Figure 4.

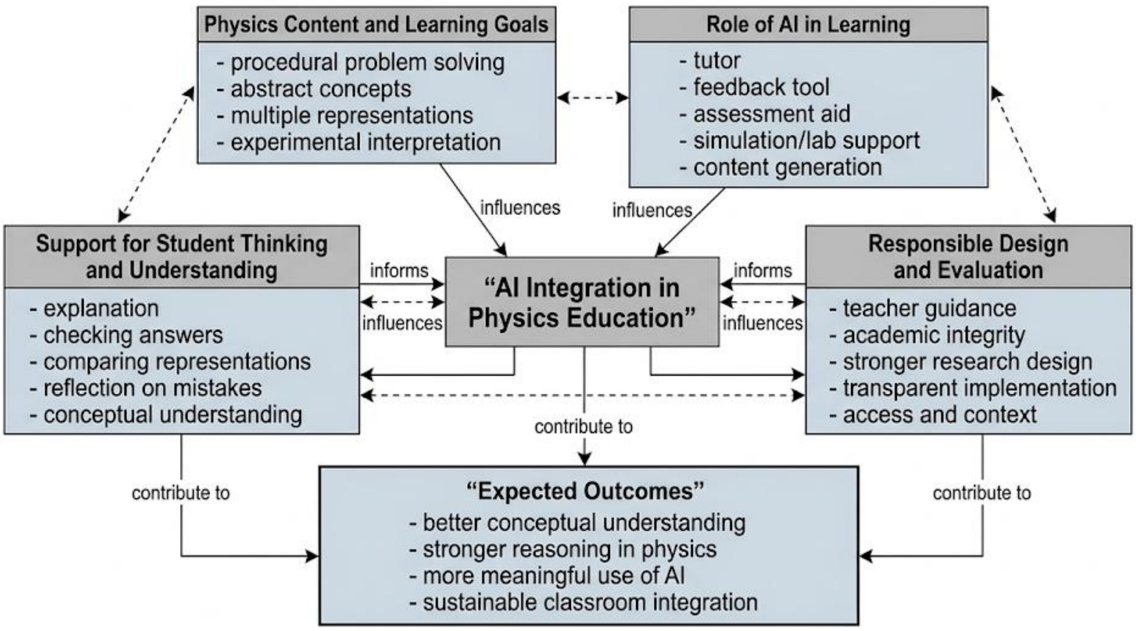


Figure 4. Framework for AI integration in physics education.

5.2. Limitations

The present review has several limitations that should be taken into account when interpreting its findings. First, while 44 empirical studies were included, most used short-term interventions with relatively small sample sizes, which limits the generalizability of findings and the conclusions regarding their influence on long-term learning [61,84]. As such, the existing synthesized evidence is largely based on immediate or short-term impacts rather than sustained changes in disciplinary understanding or changes in learning practices. In addition, the heterogeneity of study designs and contexts further limits the strength of cross-study comparisons.

Second, research examining the impact of AI on higher-order outcomes from a cognitive perspective, such as scientific reasoning, self-regulation, students' understanding, and reasoning processes, is limited. Most of the studies focus on performance-based indicators and affective measures, which leaves deeper cognitive and dimensions of understanding and reasoning not fully explored [58,80]. Furthermore, differences in how learning outcomes are defined and measured across studies make it difficult to establish consistent comparisons.

Third, the reviewed literature is geographically biased toward contexts with relatively advanced digital infrastructure, resulting in an underrepresentation of rural, low-resource, and developing educational settings, thus limiting the transferability of the results to a global context [60,64]. This imbalance may also reflect differences in research capacity and publication access rather than actual differences in educational need.

Finally, this review did not systematically disaggregate the effects of different AI technologies (such as machine-learning systems, computer-vision applications, or multimodal AI) and different physics content areas. As a result, the conclusions drawn in this study should be interpreted as reflecting general trends of AI use in physics education and not technology-specific or content-specific causal impacts [41,88]. In addition, the classification and coding of studies involve a degree of interpretive judgment, which may introduce some level of subjectivity despite efforts to maintain consistency.

Moreover, this review is limited by its search strategy, including the selection of databases, keywords, and inclusion criteria, which may have excluded relevant studies not indexed in the selected sources. Potential publication bias and the focus on English-language articles may also affect the overall representation of findings.

5.3. Future research direction

The findings of this review point to several directions for future research. First, more studies are needed that investigate long-term learning rather than short-term performance only. This includes delayed posttests and follow-up designs that can show whether gains in understanding remain over time. Second, stronger comparative research designs are necessary, especially studies with control groups or clear comparison conditions, so that claims about effectiveness are not based mainly on descriptive or self-reported findings. Third, future work should pay greater attention to underrepresented areas such as optics, waves, and astronomy, where the demands on representation and interpretation may differ from more procedural topics. Fourth, more research is needed on how AI can be integrated into inquiry, modeling, and laboratory learning without weakening student independence. Fifth, future studies should pay closer attention to how AI affects reasoning, conceptual understanding, and students' ability to judge whether a scientific explanation is sound [19,20,41,43,60,64,83,89].

More broadly, future research should move beyond asking whether AI works in physics education and instead ask what kind of AI, used for what purpose, under what teaching conditions, and for which learners, can support meaningful understanding of physics. Only then can the field move from rapid experimentation toward more stable and useful knowledge for teaching practice.

In this sense, AI clearly has potential in physics education, but that potential remains conditional. It will not be fulfilled by speed, fluency, or automation alone. Its value depends on whether it can be used in ways that strengthen explanation, understanding, checking, and careful reasoning in physics. Without such alignment, AI may improve efficiency while remaining educationally shallow. With stronger alignment, it may become a meaningful part of sustainable and discipline-sensitive physics teaching.

5.4. Conclusions

This review has shown that AI is becoming an increasingly visible part of physics education research, especially in the recent period of rapid growth in generative AI use. However, overall findings suggest that the field is still at an early stage of consolidation. The existing literature provides encouraging evidence that AI can support learning performance, engagement, feedback, and instructional efficiency in selected contexts. At the same time, the reviewed studies also indicate that current uses of AI remain concentrated in a limited range of physics topics and are often shaped more by the present strengths of the technology than by a comprehensive reconstruction of physics pedagogy [17,43,44,66,70,83].

A central conclusion of this review is that the educational value of AI in physics does not lie in the technology alone but in the quality of its pedagogical integration. AI appears most promising when it is used to support explanation, feedback, guided inquiry, and structured interaction with physics tasks. Yet the literature also shows that such promise remains conditional. If AI is used mainly to provide fast answers, complete tasks, or simulate understanding without deeper checking, then its benefits may remain superficial. In physics education, this is particularly important because meaningful learning depends on more than correct output. It requires students to connect concepts, equations, and representations, to examine assumptions, and to judge whether a solution is scientifically and physically reasonable [20,60,78,90].

The review also makes it clear that reported benefits should not be interpreted uncritically. Much of the current evidence is strongest for short-term and relatively visible outcomes, whereas stronger evidence concerning conceptual understanding, reasoning, transfer, and long-term learning remains limited. Likewise, the risks identified in the literature, including inaccurate outputs, overreliance, weak AI literacy, academic integrity concerns, and unequal infrastructure, suggest that AI use in physics education cannot be reduced to a matter of technical access or innovation alone. Rather, its successful use depends on deliberate design, teacher mediation, and the cultivation of student habits of checking, reflection, and responsible use [20,21,27,56,90].

Overall, the findings of this review suggest that AI should be understood not as a stand-alone solution for physics education but as one component within a broader instructional design. The field would benefit from moving beyond tool-centered experimentation toward more carefully aligned uses of AI that take into account the nature of the physics content, the intended learning goals, the support provided for student thinking, and the need for responsible implementation. Under such conditions, AI has the potential to become a meaningful and sustainable part of physics teaching and learning rather than merely an instrument for efficiency or convenience [21,27,44,80].

Author contributions

Roziqin: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing - original draft; Writing - review & editing. Achmad Samsudin: Conceptualization; Methodology; Supervision; Writing - review & editing. Duden Saepuzaman: Conceptualization; Methodology; Supervision; Writing - review & editing. Haslinda Nawang Sari: Formal analysis; Visualization; Writing - original draft; Writing - review & editing. Mimin Iryanti: Supervision; Writing - review & editing.

Use of Generative-AI tools declaration

During the preparation of this work, the authors used ChatGPT to improve English grammar, language clarity, and overall flow, and Rayyan.ai to support literature screening, organisation, and management of the systematic review process. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgements

The authors would like to thank LPDP (Lembaga Pengelola Dana Pendidikan), which is part of the Ministry of Finance of the Republic of Indonesia, for its assistance in funding the master's program so that this article could be published. The scholarship award numbers are as follows: Roziqin, 2025071142903536; Haslinda Nawang Sari, 2025061143002653.

Conflict of interest

The authors declare that they have no conflicts of interest.

Ethics declaration

The author declared that no ethics approval is required for the study.

References

1. Moreira, J.A., Nunes, C.S. and Casanova, D., Digital Competence of Higher Education Teachers at a Distance Learning University in Portugal. *Computers*, 2023, 12(9): 169. https://doi.org/10.3390/computers12090169
2. Outeda, C.C., European Education Area and Digital Education Action Plan (2021–2027): One More Step Towards the Europeanisation of Education Policy. In *E-Governance in the European Union: Strategies, Tools, and Implementation,* 2024, 187–206. https://doi.org/10.1007/978-3-031-56045-3_13
3. Jauhiainen, J.S. and Garagorry Guerra, A., Generative AI and education: dynamic personalization of pupils' school learning material with ChatGPT. *Front. Educ. (Lausanne)*, 2024, 9: 1288723.

https://doi.org/10.3389/feduc.2024.1288723
4. Lindsay E.D., Zhang, M., Johri, A. and Bjerva, J., The Responsible Development of Automated Student Feedback with Generative AI. *IEEE Global Engineering Education Conference (Educon)*, 2025, 1–10. https://doi.org/10.1109/EDUCON62633.2025.11016572
5. Chen, J., Mokmin, N.A.M. and Su, H., Integrating generative artificial intelligence into design and art course: Effects on student achievement, motivation, and self-efficacy. *Innovations in Education and Teaching International*, 2025, 62(5): 1431–1446. https://doi.org/10.1080/14703297.2025.2503857
6. Avci, H., Lunn, S.J. and Hazari, Z., Exploring STEM educators' perspectives on the integration of AI-enabled technologies in teaching and learning. *Computers and Education Open*, 2025, 100304. https://doi.org/10.1016/j.caeo.2025.100304
7. Huwer, J., Thyssen, C., Becker-Genschow, S., von Kotzebue, L., Finger, A., Kremser, E., et al., Competencies for teaching with and about artificial intelligence in the natural sciences – DiKoLAN AI. *Computers and Education Open*, 2025, 100303. https://doi.org/10.1016/j.caeo.2025.100303
8. Garzón, J., Patiño, E. and Marulanda, C., Systematic Review of Artificial Intelligence in Education: Trends, Benefits, and Challenges. *Multimodal Technologies and Interaction*, 2025, 9(8): 84. https://doi.org/10.3390/mti9080084
9. Wang, F., Zhou, X., Li, K., Cheung, A.C.K. and Tian, M., The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. *Interactive Learning Environments*, 2025, 33(7): 4633–4652. https://doi.org/10.1080/10494820.2025.2470319
10. Gupta, U. and Zheng, R.Z., Cognitive Load in Solving Mathematics Problems: Validating the Role of Motivation and the Interaction Among Prior Knowledge, Worked Examples, and Task Difficulty. *European Journal of Stem Education*, 2020, 5(1): 5. https://doi.org/10.20897/ejsteme/9252
11. Bouchée, T., de Putter-Smits, L., Thurlings, M. and Pepin, B., Towards a better understanding of conceptual difficulties in introductory quantum physics courses. *Stud. Sci. Educ.*, 2022, 58(2): 183–202. https://doi.org/10.1080/03057267.2021.1963579
12. Oon, P.T. and Subramaniam, R., On the declining interest in physics among students-from the perspective of teachers. *Int. J. Sci. Educ.*, 2011, 33(5): 727–746. https://doi.org/10.1080/09500693.2010.500338
13. Asem, E.K. and Rajwa, B., Impact of combination of short lecture and group discussion on the learning of physiology by nonmajor undergraduates. *Adv. Physiol. Educ.*, 2023, 47(1): 1–12. https://doi.org/10.1152/advan.00022.2022
14. Awwad, F., Enhancing Electronics Courses Education: Active Learning Strategies for Undergraduate Engineering Students. *International Journal of Engineering Pedagogy*, 2025, 15(2): 42–73. https://doi.org/10.3991/ijep.v15i2.51739
15. Margolin, J., Ba, H., Friedman, L.B., Swanlund, A., Dhillon, S. and Liu, F., Examining the impact of a play-based middle school physics program. *Journal of Research on Technology in Education*, 2021, 53(2): 125–139. https://doi.org/10.1080/15391523.2020.1754973
16. Hartley, K., Hayak, M. and Ko, U.H., Artificial Intelligence Supporting Independent Student Learning: An Evaluative Case Study of ChatGPT and Learning to Code. *Educ. Sci. (Basel)*, 2024,

14(2): 120. https://doi.org/10.3390/educsci14020120
17. Bitzenbauer, P., ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemp. Educ. Technol.*, 2023, 15(3): ep430. https://doi.org/10.30935/cedtech/13176
18. Sirnoorkar, A., Zollman, D., Laverty, J.T., Magana, A.J., Rebello, N.S. and Bryan, L.A., Student and AI responses to physics problems examined through the lenses of sensemaking and mechanistic reasoning. *Computers and Education: Artificial Intelligence*, 2024, 7: 100318. https://doi.org/10.1016/j.caeai.2024.100318
19. Kilde-Westberg, S., Johansson, A. and Enger, J., Generative AI as a lab partner: a case study. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(2): 020119. https://doi.org/10.1103/ggy1-3kjk
20. Dahlkemper, M.N., Lahme, S.Z. and Klein, P., How do physics students evaluate artificial intelligence responses on comprehension questions A study on the perceived scientific accuracy and linguistic quality of ChatGPT. *Phys. Rev. Phys. Educ. Res.*, 2023, 19(1): 010142. https://doi.org/10.1103/PhysRevPhysEducRes.19.010142
21. Jang, H. and Choi, H., A Double-Edged Sword: Physics Educators' Perspectives on Utilizing ChatGPT and Its Future in Classrooms. *J. Sci. Educ. Technol.*, 2025, 34(2): 267–283. https://doi.org/10.1007/s10956-024-10173-1
22. Durgungoz, A. and Durgungoz, F.C., Exploring effortless AI-generated gamified quizzes in an online special education module: evaluating question quality, student engagement, and its potential to identify at-risk students. *Educ. Inf. Technol. (Dordr)*, 2025, 30(17): 25335–25357. https://doi.org/10.1007/s10639-025-13765-5
23. Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., et al., Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education Artificial Intelligence*, 2025, 8: 100407. https://doi.org/10.1016/j.caeai.2025.100407
24. Benczik, I.J. and Toldy, E.H., AI-aided formative assessment in physics education. *EMI. Educ. Media Int.*, 2025, 62(4): 450–458. https://doi.org/10.1080/09523987.2025.2588530
25. Mok, R., Akhtar, F., Clare, L., Li, C., Ida, J., Ross, L., et al., Using large language models for grading in education: an applied test for physics. *Phys. Educ.*, 2025, 60(3): 035006. https://doi.org/10.1088/1361-6552/adb92b
26. Abdulayeva, A., Zhanatbekova, N., Andasbayev, Y., Khaimuldanov, Y. and Zhiyembayev, Z., The Role of Artificial Intelligence (AI) in Personalised Physics Education. *Jurnal Pendidikan IPA Indonesia*, 2025, 14(3). https://doi.org/10.15294/jpii.v14i3.29752
27. Yehya, F., ElSayary, A., Al Murshidi, G. and Al Zaabi, A., Artificial intelligence integration and teachers' self-efficacy in physics classrooms. *Eurasia Journal of Mathematics, Science and Technology Education*, 2025, 21(8): em2679. https://doi.org/10.29333/ejmste/16660
28. Gunbatar, S.A., Kiran, B.E., Boz, Y. and Oztay, E.S., A systematic review of green and sustainable chemistry training research with pedagogical content knowledge framework: current trends and future directions. *Chemistry Education Research and Practice*, 2025, 26: 34–52. https://doi.org/10.1039/D4RP00166D
29. Fortuna, A., Prasetya, F., Samala, A.D., Rawas, S., Criollo-C, S., Kaya, D., et al., Artificial intelligence in personalized learning: A global systematic review of current advancements and shaping future opportunities. *Social Sciences and Humanities Open*, 2025, 12: 102114. https://doi.org/10.1016/j.ssaho.2025.102114
30. Bouguettaya, S., Pupo, F., Chen, M. and Fortino, G., A Meta-Survey of Generative AI in Education:

Trends, Challenges, and Research Directions. *Big Data and Cognitive Computing*, 2025, 9(9): 237. https://doi.org/10.3390/bdcc9090237

31. Abrar, M., Aboraya, W., Khaliq, R.A., Subramanian, K.P., Al Husaini, Y. and Al Hussaini, M., AI-Powered Learning Pathways: Personalized Learning and Dynamic Assessments. *International Journal of Advanced Computer Science and Applications*, 2025, 16(1): 454–462. https://doi.org/10.14569/IJACSA.2025.0160145
32. Rakhmetov, M., Sadvakassova, A., Saltanova, G., Kuanbayeva, B. and Zhusupkalieva, G., Evaluation of an AI-Based Feedback System for Enhancing Self-Regulated Learning in Digital Education Platforms. *Electronic Journal of E-Learning*, 2025, 23(4): 126–141. https://doi.org/10.34190/ejel.23.4.4150
33. Naseer, F. and Khawaja, S., Mitigating Conceptual Learning Gaps in Mixed-Ability Classrooms: A Learning Analytics-Based Evaluation of AI-Driven Adaptive Feedback for Struggling Learners. *Applied Sciences (Switzerland)*, 2025, 15(8): 4473. https://doi.org/10.3390/app15084473
34. Slepankova, M., Kilianova, K., Kockova, P., Kostolanyova, K., Kotyrba, M. and Habiballa, H., Student Perceptions and Preferences in Personalized AI-driven Learning. *Acta Informatica Pragensia*, 2025, 14(2): 261–271. https://doi.org/10.18267/j.aip.278
35. Abubakar, S., Jeilani, A. and Yusuf, M., The role of over-reliance on AI in the negative consequences of student learning: The moderating effects of ethical concerns and institutional policies. *Cogent Education*, 2025, 12(1): 2591503. https://doi.org/10.1080/2331186X.2025.2591503
36. Zhai, C., Wibowo, S. and Li, L.D., The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, 2024, 11(1): 28. https://doi.org/10.1186/s40561-024-00316-7
37. Khodadad, D., ChatGPT in engineering education: a breakthrough or a challenge?. *Phys. Educ.*, 2025, 60(4): 045006. https://doi.org/10.1088/1361-6552/add073
38. Mills, E., Mizouri, A. and Peach, A., Prompting Better Feedback: A Study of Custom GPT for Formative Assessment in Undergraduate Physics. *Educ. Sci. (Basel)*, 2025, 15(8): 1058. https://doi.org/10.3390/educsci15081058
39. Saqr, M. and López-Pernas, S., Why explainable AI may not be enough: predictions and mispredictions in decision making in education. *Smart Learning Environments*, 2024, 11(1): 52. https://doi.org/10.1186/s40561-024-00343-4
40. Chkarka, F. and Fatmi, H., Comparative Examination of Master's Students' and Faculty Members' Maintenance of Academic Integrity in the Age of AI. *J. Acad. Ethics*, 2026, 24(1): 1–23. https://doi.org/10.1007/s10805-025-09684-2
41. Tong, D., Jin, B., Tao, Y., Ren, H., Atiquil Islam, A.Y.M. and Bao, L., Exploring the role of human-AI collaboration in solving scientific problems. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(1): 010149. https://doi.org/10.1103/PhysRevPhysEducRes.21.010149
42. Cui, B., Jiao, W., Gui, S., Li, Y. and Fang, Q., Innovating physical education with artificial intelligence: a potential approach. *Frontiers in Psychology,* 2025, 16: 1490966. https://doi.org/10.3389/fpsyg.2025.1490966
43. de Souza, M.G., Won, M., Treagust, D. and Serrano, A., Visualising relativity: assessing high school students' understanding of complex physics concepts through AI-generated images. *Phys. Educ.*, 2024, 59(2): 025018. https://doi.org/10.1088/1361-6552/ad1e71

44. Coban, A., Dzsotjan, D., Küchemann, S., Durst, J., Kuhn, J. and Hoyer, C., AI support meets AR visualization for Alice and Bob: personalized learning based on individual ChatGPT feedback in an AR quantum cryptography experiment for physics lab courses. *EPJ Quantum Technol.*, 2025, 12(1): 15. https://doi.org/10.1140/epjqt/s40507-025-00310-z
45. Wang, D. and Huang, X., Transforming education through artificial intelligence and immersive technologies: enhancing learning experiences. *Interactive Learning Environments*, 2025, 33(7): 4546–4565. https://doi.org/10.1080/10494820.2025.2465451
46. Jiao, L., Song, X., You, C., Liu, X., Li, L., Chen, P., et al., AI meets physics: a comprehensive survey. *Artif. Intell. Rev.*, 2024, 57(9): 256. https://doi.org/10.1007/s10462-024-10874-4
47. Raut, P., Panda, D.K. and Goyal, A.K., A Comprehensive Review on Next-Generation Modeling and Optimization for Semiconductor Devices. *IEEE Access*, 2025, 13: 123724–123742. https://doi.org/10.1109/ACCESS.2025.3587721
48. Jin, L., Su, H., Guo, S., Song, J. and Yang, Q., A Review of AI-Based Computational Modelling Studies of Electromagnetic Fields | 基于 AI 的电磁场计算建模研究综述. *Diangong Jishu Xuebao Transactions of China Electrotechnical Society*, 2025, 40(10): 3013–3029. Doi: 10.19595/j.cnki.1000-6753.tces.240877
49. Zhang, W., Kou, J. and Liu, Y., Prospect of artificial intelligence empowered fluid mechanics | 智能赋能流体力学展望. *Hangkong Xuebao Acta Aeronautica Et Astronautica Sinica*, 2021, 42(4). Doi: 10.7527/S1000-6893.2020.24689
50. von Gillern, S., Wright, W. and Winger, K., A Survey of Teachers’ Perspectives on Benefits and Concerns of Artificial Intelligence in Literacy Education. *Lit. Res. Instr.*, 2025, 1–24. https://doi.org/10.1080/19388071.2025.2557795
51. Garcia-Suarez, D., Guzmán, B.R.G., Alamillo, C.A.H. and De Cote, E.P.M.L.M., AI-Driven Personalized Learning Profiles to Enhance Student Performance in Basic Physics: A Pilot Study. In *IEEE Global Engineering Education Conference (Educon)*, 2025, 1–5. https://doi.org/10.1109/EDUCON62633.2025.11016294
52. Rethlefsen, M.L., Kirtley, S., Waffenschmidt, S., Ayala, A.P., Moher, D., Page, M.J., et al., PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Journal of the Medical Library Association*, 2021, 109(2): 174–200. https://doi.org/10.5195/jmla.2021.962
53. Al-Kamzari, F. and Alias, N., A systematic literature review of artificial intelligence (AI) in secondary school physics: applications, benefits, and challenges. *Interactive Learning Environments*, 2025, 34(2): 676–693. https://doi.org/10.1080/10494820.2025.2508323
54. Waffenschmidt, S., Knelangen, M., Sieben, W., Bühn, S. and Pieper, D., Single screening versus conventional double screening for study selection in systematic reviews: A methodological systematic review. *BMC Med. Res. Methodol.*, 2019, 19(1): 132. https://doi.org/10.1186/s12874-019-0782-0
55. Fabregues, S., Sainz, M., Romano, M.J., Escalante-Barrios, E.L., Younas, A. and Lopez-Perez, B.S., Use of mixed methods research in intervention studies to increase young people’s interest in STEM: A systematic methodological review. *Front. Psychol.*, 2023, 13: 956300. https://doi.org/10.3389/fpsyg.2022.956300
56. Abdulayeva, A., Zhanatbekova, N., Andasbayev, Y. and Boribekova, F., Fostering AI literacy in pre-service physics teachers: inputs from training and co-variables. *Front. Educ. (Lausanne)*, 2025, 10: 1505420. https://doi.org/10.3389/feduc.2025.1505420

57. Jufrida, J., Kurniawan, W., Furqon, M., Anwar, K., Falah, H.S. and Riantoni, C., AI-Driven Ethnoscience Learning: Enhancing Physics Education Through Malay Cultural Insights. *Journal of Information Technology Education: Innovations in Practice*, 2025, 24: 013. https://doi.org/10.28945/5520
58. Simbolon, M., Pongkondek, J.J., Henukh, A. and Ronchintaniawati, D., AI-Driven Sociocultural Interactive Digital Module for Papua: Advancing Educational Technology to Sustainable Development Goals. *International Journal of Learning, Teaching and Educational Research*, 2025, 24(2): 543–559. https://doi.org/10.26803/ijlter.24.2.27
59. Dmitrenko, N., Shkola, I. and Shkola, O., AI-Powered Tools for Teaching ESP to Pre-Service Physics Teachers. *Analele Universitatii Ovidius Constanta, Seria Filologie*, 2025, XXXVI.
60. Zollman,D., Sirnoorkar, A. and Laverty, J.T., Analyzing AI and student responses through the lens of sensemaking and mechanistic reasoning. In *Physics Education Research Conference Proceedings*, 2023, 415–420. https://doi.org/10.1119/perc.2023.pr.Zollman
61. Robledo-Rella, V. and Toh, B.Y., Artificial Intelligence in Physics Courses to Support Active Learning. In *ACM International Conference Proceeding Series*, 2024, 68–75. https://doi.org/10.1145/3678610.3678631
62. Kortemeyer, G. and Nöhl, J., Assessing confidence in AI-Assisted grading of physics exams through psychometrics: An exploratory study. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(1): 010136. https://doi.org/10.1103/PhysRevPhysEducRes.21.010136
63. Liu, N., Sonkar, S., Basu Mallick, D., Baraniuk, R. and Chen, Z., Atomic Learning Objectives and LLMs Labeling: A High-Resolution Approach for Physics Education. In *15th International Conference on Learning Analytics and Knowledge, LAK 2025*, 2025, 620–630. https://doi.org/10.1145/3706468.3706550
64. Lademann, J., Henze, J. and Becker-Genschow, S., Augmenting learning environments using AI custom chatbots: Effects on learning performance, cognitive load, and affective variables. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(1): 010147. https://doi.org/10.1103/PhysRevPhysEducRes.21.010147
65. Duy, H.T., Ngoc, C.T. and Hai, N.T., Building and using chatbots in the process of self-studying physics to improve the quality of learners' knowledge. *International Journal of Education and Practice*, 2024, 12(4): 1165–1185. https://doi.org/10.18488/61.v12i4.3859
66. Kortemeyer, G. and Bauer, W., Cheat sites and artificial intelligence usage in online introductory physics courses: What is the extent and what effect does it have on assessments. *Phys. Rev. Phys. Educ. Res.*, 2024, 20(1): 010145. https://doi.org/10.1103/PhysRevPhysEducRes.20.010145
67. Fussell, R.K., Flynn, M., Damle, A., Fox, M.F. and Holmes, N.G., Comparing large language models for supervised analysis of students' lab notes. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(1): 010128. https://doi.org/10.1103/PhysRevPhysEducRes.21.010128
68. Kurniawan, W., Sutrisno, M., Marzal, J. and Anwar, K., Construction of an Intelligent Teacher Assistant System Using the TPACK Framework and Machine Learning to Diagnose Work and Energy Misconceptions. *International Journal of Information and Education Technology*, 2025, 15(5): 1084–1102. https://doi.org/10.18178/ijiet.2025.15.5.2312
69. dos Anjos, J.R., de Souza, M.G., Prodanov, T.S. and de Andrade Neto, A.S., Disruptive Education: Integrating ChatGPT into an Active Methodology for Teaching Sciences and Mathematics. *Acta Sci. (Canoas)*, 2024, 26(1): 334–369.

70. Wan, T. and Chen, Z., Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning. *Phys. Rev. Phys. Educ. Res.*, 2024, 20(1): 010152. https://doi.org/10.1103/PhysRevPhysEducRes.20.010152
71. El-Adawy, S., Macdonagh, A. and Abdelhafez, M., Exploring Large Language Models as Formative Feedback Tools in Physics. In *Physics Education Research Conference Proceedings*, 2024, 126–131. https://doi.org/10.1119/perc.2024.pr.El-Adawy
72. Jere, S., Bessong, R., Mpeta, M. and Litshani, N.F., Exploring Pre-Service Teachers' Perceptions of ChatGPT Integration into Physical Sciences Teaching: A Case Study at a Rural South African University. *International Journal of Learning, Teaching and Educational Research*, 2024, 23(11): 464–486. https://doi.org/10.26803/ijlter.23.11.24
73. Llorella, F., Antonio Cebrián, J., Corbi, A. and María Pérez, A., Fostering scientific methods in simulations through symbolic regressions. *Phys. Educ.*, 2024, 59(4): 045010. https://doi.org/10.1088/1361-6552/ad3cad
74. Borovský, D., Hanč, J. and Hančová, M., Innovative approaches to high school physics competitions: Harnessing the power of AI and open science. In *Journal of Physics: Conference Series*, 2024, 2715(1): 012011. https://doi.org/10.1088/1742-6596/2715/1/012011
75. Wibowo, F.C., Hasbey, H., Aziz, T.A., Darman, D.R., Kusuma, A.F.A.K. and Bunyami, M.A.H., Integrating artificial intelligence chatbot climate change with physics digital modules for student learning assistants. *Multidisciplinary Science Journal*, 2025, 7(12): 583. https://doi.org/10.31893/multiscience.2025583
76. Shektibayev, N., Turekulova, A., Alimbekova, G., Ospanbekov, Y., Auyelbekov, O. and Yessentayev, K., Integration of Artificial Intelligence into Solving Physical Problems to Improve Pedagogical Competence. *Eurasian Journal of Educational Research*, 2024, 2024(113): 210–226.
77. Her, P. and Hamerski, P., Investigating student perceptions of creativity and generative ai in computational physics. In *Physics Education Research Conference Proceedings*, 2024, 181–186. https://doi.org/10.1119/perc.2024.pr.Her
78. Jang, H., Less is more: A multimodal exploration of how limited AI shapes team physics problem solving. *New Physics: Sae Mulli*, 2025, 75(5): 449–467. https://doi.org/10.3938/NPSM.75.449
79. Tafhi, O., Khattabi, K.E., Lamrabet, M., Ouzennou, F. and Chikhaoui, A., Mapping Physical Moroccan Sciences Student's Perceptions of AI: A Case Study on Generative AI (ChatGPT). *International Journal of Information and Education Technology*, 2025, 15(1): 175–186. https://doi.org/10.18178/ijiet.2025.15.1.2230
80. Guo, F., Li, T. and Cunningham, C.J.L., One year in the classroom with ChatGPT: empirical insights and transformative impacts. *Front. Educ. (Lausanne).*, 2025, 10: 1574477. https://doi.org/10.3389/feduc.2025.1574477
81. Torre, M., Physics education in the age of ChatGPT: The art of asking questions. In *Nuovo Cimento della Societa Italiana di Fisica*, 2025. Doi: 10.1393/ncc/i2025-25258-y
82. Wattanakasiwich, P., Kaewkhong, K. and Katwibun, D., Physics instructors' acceptance and implementation of generative AI. *Phys. Rev. Phys. Educ. Res.*, 2025, 21(1): 010155. https://doi.org/10.1103/r2fn-kdy4
83. Beltozar-Clemente, S. and Díaz-Vega, E., Physics XP: Integration of ChatGPT and Gamification Improve Academic Performance and Motivation in Physics 1 Course. *International Journal of*

Engineering Pedagogy, 2024, 14(6): 82–92. https://doi.org/10.3991/ijep.v14i6.47127
84. Latif, E., Parasuraman, R. and Zhai, X., PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations. In *IEEE International Workshop on Robot and Human Communication, RO-MAN*, 2024, 864–871. https://doi.org/10.1109/RO-MAN60168.2024.10731312
85. Tian, M. and Jiang, Q., Role adaptation of Maysian secondary school physics teachers to AI-assisted experimental teaching: Integrating social psychological challenges with. *Environment and Social Psychology*, 2025, 10(8). https://doi.org/10.59429/esp.v10i8.3906
86. El-Adawy, S., Liao, I., Lad, V., Abdelhafez, M. and Dourmashkin, P., Streamlining Physics Problem Generation to Support Physics Teachers in Using Generative Artificial Intelligence. *Phys. Teach.*, 2024, 62(7): 595–598. https://doi.org/10.1119/5.0201458
87. Ding, L., Li, T., Jiang, S. and Gapud, A., Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education*, 2023, 20(1): 63. https://doi.org/10.1186/s41239-023-00434-1
88. Piccione, A., Massa, A.A., Ruggiero, M.L., Serio, M., Rinaudo, M., Marocchi, D., et al., Training teachers on new topics and new tools in Physics education. In *Journal of Physics: Conference Series*, 2024, 2693(1): 012010. https://doi.org/10.1088/1742-6596/2693/1/012010
89. Guerrero-Zambrano, M., Sanchez-Alvarado, L., Valarezo-Chamba, B. and Lamilla-Rubio, E., Transforming Physics Teacher Training Through ChatGPT: A Study on Usability and Impact. *Educ. Sci. (Basel)*, 2025, 15(7): 887. https://doi.org/10.3390/educsci15070887
90. Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K.E., Lukowicz, P., Kuhn, J., et al., Unreflected Acceptance - Investigating the Negative Consequences of ChatGPT-Assisted Problem Solving in Physics Education. In *Frontiers in Artificial Intelligence and Applications*, 2024, 199–212. https://doi.org/10.3233/FAIA240195
91. McNeish, D., Challenging Conventional Wisdom for Multivariate Statistical Models With Small Samples. *Rev. Educ. Res.*, 2017, 87(6): 1117–1151. https://doi.org/10.3102/0034654317727727
92. Tipton, E., Hallberg, K., Hedges, L.V. and Chan, W., Implications of Small Samples for Generalization: Adjustments and Rules of Thumb. *Eval. Rev.*, 2017, 41(5): 472–505. https://doi.org/10.1177/0193841X16655665

Author's biography

Roziqin is a master's student affiliated with the Department of Physics Education, Faculty of Mathematics and Natural Sciences Education, Universitas Pendidikan Indonesia, Bandung, Indonesia. His academic interests focus on physics education and the development of innovative learning approaches and instructional media.

Dr. Achmad Samsudin is an Associate Professor at Universitas Pendidikan Indonesia and currently serves as Head of the Bachelor and Master of Physics Education Study Program. His research focuses on physics education, conceptual change, misconceptions, cognitive psychology in physics learning, and science education. He has an extensive publication record and is widely recognized for his contributions to physics education research in Indonesia.

Dr. Duden Saepuzaman is a senior lecturer in the Physics Education Program at Universitas Pendidikan Indonesia. He holds a doctoral degree and conducts research in physics education, especially in assessment, misconceptions, meta-analysis, and earth physics/geophysics. His academic work reflects sustained engagement in improving physics learning and assessment practices in higher education and school contexts.

Haslinda Nawang Sari is a master's student affiliated with the Department of Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia. Her academic interests focus on chemistry education and the development of innovative learning resources, particularly in green chemistry, sustainable development, and instructional media.

Dr. Mimin Iryanti is a physics lecturer at Universitas Pendidikan Indonesia with an academic profile in Earth sciences and environmental studies. Her ORCID record lists her employment at Universitas Pendidikan Indonesia since 2002, and her public scholar profile highlights research activity in Earth sciences and the environment. She is also publicly listed in association with the Physics Study Program at UPI.

Appendix A

Table 2. Descriptive and methodological profile of included empirical studies (n = 44).

Title	Year	Country/region	Publication type	Participant group	Sample size (n)	Research design	AI type	Reference	Study code
A Double-Edged Sword: Physics Educators' Perspectives on Utilizing ChatGPT and Its Future in Classrooms	2025	Korea	Article	Physics educators (secondary and higher education)	10	Qualitative (consensual qualitative research; semi-structured interviews)	Generative AI (ChatGPT; NLP-based chatbot)	[21]	1
AI support meets AR visualization for Alice and Bob: personalized learning	2025	Germany	Article	University students (physics laboratory courses)	21 groups	Mixed-method (quasi-experimental crossover design; eye-tracking analysis)	Generative AI (GPT-4/ChatGPT-based feedback system)	[44]	2

based on individual ChatGPT feedback in an AR quantum cryptography experiment for physics lab courses									
AI-Driven Ethnoscience Learning: Enhancing Physics Education Through Malay Cultural Insights	2025	Indonesia	Article	Secondary school students	150 (141 valid responses)	Mixed-method (educational data mining; machine learning prediction)	Machine learning (Random Forest; Naïve Bayes)	[57]	3
AI-Driven Sociocultural Interactive Digital Module for Papua: Advancing Educational Technology to Sustainable Development Goals	2025	Indonesia	Article	Middle school students and science teachers	200 students; 40 teachers	Mixed-method (ethnographic approach and quasi-experimental pretest–posttest control group design)	AI-based adaptive learning system (personalized learning and real-time feedback)	[58]	4
AI-Powered Tools For Teaching Esp To Pre-Service Physics Teachers	2025	Ukraine	Article	Pre-service physics teachers (university students)	20	Mixed-method case study (qualitative data supported by pretest–posttest analysis)	AI-powered language learning tools (speech recognition, AI-assisted content generation, AI-guided discussion	[59]	5

							tools, and AI image generation)		
Analyzing AI and student responses through the lens of sensemaking and mechanistic reasoning	2023	United States	Conference	Undergraduate physics students (introductory level)	8 students and 8 AI-generated responses	Qualitative comparative study (discourse analysis using sensemaking and mechanistic reasoning frameworks)	Generative AI (ChatGPT; NLP-based language model)	[60]	6
Artificial Intelligence in Physics Courses to Support Active Learning	2024	Mexico	Conference	Undergraduate engineering students (freshman level)	91	Quasi-experimental study (comparative analysis of AI-assisted and traditional instruction with survey data)	Generative AI (LLMs: ChatGPT, Copilot; integrated with symbolic computation tools such as WolframAlpha)	[61]	7
Artificial intelligence integration and teachers' self-efficacy in physics classrooms	2025	United Arab Emirates (UAE)	Article	Secondary school physics teachers	15	Qualitative study (interpretivist approach; semi-structured interviews with thematic analysis)	AI in education (adaptive learning systems, AI tools such as ChatGPT, simulations, and intelligent tutoring tools)	[27]	8
Assessing confidence in AI-assisted grading of physics exams through psychometrics: An	2025	Switzerland	Article	Undergraduate engineering students (thermodynamics exam)	252 (from 434 total; opt-in sample)	Exploratory empirical study; mixed-method (AI experimentation and psychometric modeling: CTT and IRT and statistical evaluation)	Generative AI (LLMs: GPT-4o for OCR, GPT-4-Turbo for grading) and psychometric models (IRT, Bayesian	[62]	9

exploratory study							uncertainty estimation)		
Atomic Learning Objectives and LLMs Labeling: A High-Resolution Approach for Physics Education	2025	United States	Conference	Undergraduate physics problem datasets (introductory physics course and OpenStax textbook); expert annotator	131 questions; 194 learning objectives across 9 chapters	Empirical experimental study; comparative evaluation of LLMs with human expert labeling; mixed-method (quantitative metrics and qualitative error analysis)	Large language models (GPT-4o, GPT-3.5, LLaMA-70B, LLaMA-8B) for automated learning objective labeling	[63]	10
Augmenting learning environments using AI custom chatbots: Effects on learning performance, cognitive load, and affective variables	2025	Germany	Article	Sixth-grade secondary school students	214 students	Randomized controlled trial (RCT); between-group experimental design (EG vs. CG); quantitative analysis (Mann–Whitney U, effect size)	AI custom chatbot (GPT-4-based, RAG-style augmentation for generating learning explanations)	[64]	11
Building and using chatbots in the process of self-studying physics to improve the quality of learners' knowledge	2024	Vietnam	Article	Grade 10 high school students	100 students (50 CG, 50 EG)	Quasi-experimental (control vs. experimental group); mixed methods (test and questionnaire)	Teacher-built chatbot (scripted and keyword-based, platform: Messnow, Messenger-integrated)	[65]	12
ChatGPT in physics education: A pilot study on	2023	Germany	Article	Secondary school students (Grade 12 physics	53	Pilot study; one-group pretest–posttest design	Generative AI (ChatGPT; large language model)	[17]	13

easy-to-implement activities				classes)					
Cheat sites and artificial intelligence usage in online introductory physics courses: What is the extent and what effect does it have on assessments?	2024.	United States	Article	Undergraduate students (introductory physics courses for scientists and engineers)	221	Mixed-method (survey and statistical analysis, clustering, regression)	Generative AI (LLMs such as ChatGPT) and online problem-solving platforms	[66]	14
Comparing large language models for supervised analysis of students' lab notes	2025	United Kingdom	Article	Undergraduate physics students (introductory experimental physics laboratory courses)	873 lab notes (58,369 sentences); 205 manually coded subset	Quantitative experimental study with supervised machine learning comparison (bag-of-words, BERT, LLaMA, few-shot LLM)	Large language models (BERT, LLaMA) and traditional NLP (bag-of-words)	[67]	15
Construction of an Intelligent Teacher Assistant System Using the TPACK Framework and Machine Learning to Diagnose Work and Energy Misconceptions	2025	Indonesia	Article	High school students and physics teachers	150 students; 30 teachers (main study); initial pilot: 50 students, 10 teachers	Research and Development (Design Thinking: Empathize-Define-Ideate-Prototype-Test) with experimental validation and expert evaluation	Machine learning (supervised learning, SVM-based classification) integrated with TPACK framework	[68]	16
Disruptive	202	Brazil	Article	High	Not	Qualitative	Generative AI	[69]	17

Education: Integrating ChatGPT into an Active Methodology for Teaching Sciences and Mathematics	4			school students	reported	educational intervention/design-based classroom study	(ChatGPT; large language model)		
Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning	2024	United States	Article	Undergraduate physics students; student researchers; physics instructors	Students: 85 (responses); Stage I: 20; Stage II: 16; Stage III: 65; Raters: 4 students, 4 instructors	Mixed-method exploratory study (three-stage design: prompt development, student evaluation, instructor evaluation)	Generative AI (GPT-3.5, few-shot prompting, prompt engineering)	[70]	18
Exploring Large Language Models as Formative Feedback Tools in Physics	2024	United States	Conference	Undergraduate students (introductory calculus-based mechanics course) and physics instructors	179 student responses (Question 1: 103; Question 2: 76)	Mixed-method study (quantitative scoring comparison and qualitative thematic analysis of feedback)	Generative AI (ChatGPT-3.5; prompt-engineered formative feedback system)	[71]	19
Exploring Pre-Service Teachers' Perceptions of ChatGPT Integration into Physical Sciences Teaching: A	2024	South Africa	Article	Pre-service physical sciences teachers (final-year Bachelor of Education Honors students)	11 participants	Qualitative case study (semi-structured interviews; thematic analysis; TPACK framework)	Generative AI (ChatGPT; large language model for educational support)	[72]	20

Case Study at a Rural South African University									
Exploring the role of human-AI collaboration in solving scientific problems	2025	China	Article	First-year high school students	120 (final analyzed: HHC = 60; HAI = 55)	Quasi-experimental comparative study (pretest–posttest design; HAI vs. HHC; statistical analysis including t-tests; dialogue analysis; CSAT self-assessment)	Generative AI (ChatGPT-4o; conversational large language model with multimodal capabilities)	[41]	21
Fostering AI literacy in pre-service physics teachers: inputs from training and co-variables	2025	Kazakhstan	Article	Pre-service physics teachers (undergraduate students across 9 universities)	136 (intervention = 59; control = 77)	Quasi-experimental (pretest–posttest control group design; mediation/path analysis; Bayesian RM ANCOVA)	Generative AI tools (e.g., ChatGPT, Gemini, Eduaide.ai, MagicSchool.ai)	[56]	22
Fostering scientific methods in simulations through symbolic regressions	2024	Spain	Article	K-12 students	360	Design-based educational technology study with field testing of an AI-enhanced simulation framework	Grey-box AI/machine learning (genetic algorithms and symbolic regression; dimensional analysis-enhanced custom framework; PySR used in some implementations)	[73]	23
Generative AI as a lab partner: a case study	2024	Sweden	Article	High school physics students	19 (7 groups)	Qualitative exploratory case study using video observations,	Generative AI (ChatGPT, GPT-3.5)	[19]	24

				(Years 2–3; Physics 1–3 courses)		interviews, chat logs, and inductive narrative analysis framed by variation theory			
How do physics students evaluate artificial intelligence responses on comprehension questions? A study on the perceived scientific accuracy and linguistic quality of ChatGPT	2023	Germany	Article	First- and second-year university physics students	102 (initial), 94 fully analyzed participants (+3 partial)	Quantitative survey with experimental manipulation; Likert-scale ratings; repeated-measures ANOVA, ANCOVA, and exploratory factor analysis	Generative AI (ChatGPT, GPT-3.5)	[20]	25
Innovative approaches to high school physics competitions: Harnessing the power of AI and open science	2024	Slovakia	Conference	High school students preparing for physics competitions and their tutors	Not reported	Exploratory sequential mixed-methods study (QUAL for quan), including literature review and pilot case study	Generative AI chatbots and open science tools (ChatGPT/GPT-3.5 and GPT-4, Bard, Claude, Vicuna, Jupyter AI, SageMath)	[74]	26
Integrating artificial intelligence chatbot climate change with physics digital modules for student	2025	Indonesia	Article	Undergraduate students (environmental physics course; student learning assistants)	124 (experimental = 62; control = 62)	Experimental research (control vs. experimental group; descriptive statistical analysis of questionnaire responses)	AI Chatbot (natural language processing–based conversational agent; integration with ChatGPT module and	[75]	27

learning assistants							CONVAI; text-to-speech features)		
Integration of Artificial Intelligence into Solving Physical Problems to Improve Pedagogical Competence	2024	Kazakhstan	Article	Mixed participants (teachers, students, and researchers in physics and AI)	97 (control vs. experimental groups)	Quasi-experimental design (control vs. experimental group; quantitative and qualitative analysis; survey and performance comparison)	Neural networks; machine learning; physics-informed neural networks (PINNs); deep learning	[76]	28
Investigating student perceptions of creativity and generative ai in computational physics	2024	United States	Conference	Upper-division undergraduate physics students (computational physics course)	6 participants	Qualitative interpretive study (narrative essays; thematic coding using Four C Model of creativity)	Generative AI (e.g., ChatGPT; large language models)	[77]	29
Less is more: A multimodal exploration of how limited AI shapes team physics problem solving	2025	South Korea	Article	Undergraduate students (multidisciplinary backgrounds: physics-related, engineering, AI, software)	14 participants	Randomized controlled experimental study with multimodal analysis (performance data, discourse analysis, chatbot logs, video/audio recordings)	Generative AI (ChatGPT-based physics chatbots; large language models with pedagogical customization)	[78]	30
Mapping Physical Moroccan Sciences Student's Perceptions of AI: A Case Study on Generative AI	2025	Morocco	Article	Undergraduate, Master’s, and final-year high school students in physical sciences	300 valid responses (from 350 collected)	Quantitative, descriptive, and exploratory survey using a structured questionnaire (Google Forms) with statistical analysis (Excel)	Generative AI (ChatGPT; NLP-based large language model using Transformer architecture and RLHF)	[79]	31

(ChatGPT)									
One year in the classroom with ChatGPT: empirical insights and transformative impacts	2025	United States	Article	Undergraduate and graduate students in psychology (statistics) and physics courses	169 students	Year-long empirical experimental study (two semesters) combining classroom interventions (coding and conceptual learning activities) with post-activity survey (Likert-scale) and subgroup comparative analysis	Generative AI (ChatGPT; GPT-based large language model using NLP and Transformer architecture)	[80]	32
Physics education in the age of ChatGPT: The art of asking questions	2025	Italy	Article	High school students (ages 16–17)	108 students	Classroom-based educational experiment (8 instructional hours; qualitative pedagogical intervention with observational insights)	Generative AI (ChatGPT; large language model based on NLP and Transformer architecture)	[81]	33
Physics instructors' acceptance and implementation of generative AI	2025	Thailand	Article	High school and university physics instructors	320 instructors (206 GenAI users; 114 nonusers)	Mixed-method study (online survey, structural equation modeling, descriptive statistics, and qualitative content analysis)	Generative AI/Large Language Models (e.g., ChatGPT, Gemini, Claude)	[82]	34
Physics XP: Integration of ChatGPT and Gamification to Improve Academic	2024	Peru	Article	Undergraduate engineering students enrolled in Physics 1 course	188 students (98 experimental; 90 control)	Quasi-experimental study (control vs. experimental group; statistical comparison of exam scores; Likert-scale	Generative AI (ChatGPT; NLP-based large language model) integrated with gamification	[83]	35

Performance and Motivation in Physics 1 Course						motivation survey)	tools		
PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations	2024	United States	Conference	K-12 students (8th-grade physics students)	10 students	Experimental user study with human expert evaluation (Bloom's taxonomy-based scoring; comparative analysis with GPT-4)	Multimodal AI (Large Language Model: GPT-3.5-turbo; Computer Vision: YOLOv8; Speech Recognition and Text-to-Speech)	[84]	36
Role adaptation of Maysian secondary school physics teachers to AI-assisted experimental teaching: Integrating social psychological challenges with	2025	Malaysia	Article	Secondary school physics teachers	420 survey participants; 30 interview participants; 6 focus groups; 12 classroom observation cases	Explanatory sequential mixed-methods design (questionnaire survey, in-depth interviews, focus groups, and classroom observations)	AI-assisted experimental teaching systems (including virtual experiments, digital recording systems, and AI-supported instructional tools)	[85]	37
Streamlining Physics Problem Generation to Support Physics Teachers in Using Generative Artificial	2024	United States	Article	Physics teachers and graduate teaching assistants in a classical mechanics course	Not reported (two graduate teaching assistants involved in workflow developm	Design-based study/methodological exploration (workflow development and demonstration using ChatGPT)	Generative AI (Large Language Model: ChatGPT 3.5; NLP-based)	[86]	38

Intelligence					ent)				
Students' perceptions of using ChatGPT in a physics class as a virtual tutor	2023	United States	Article	Undergraduate students enrolled in an introductory physics course	40 students	Mixed-method classroom-based study (physics make-up exam activity, survey, trust clustering analysis, MANOVA/ANOVA, and drawing-based qualitative analysis)	Generative AI (ChatGPT; large language model based on machine learning/NLP)	[87]	39
The Role Of Artificial Intelligence (AI) In Personalised Physics Education	2025	Kazakhstan	Article	Secondary school students (Grade 10)	58 students (31 experimental, 27 control)	Mixed-methods experimental study (pedagogical experiment with control vs. experimental groups, statistical analysis using t-test and ANOVA, thematic qualitative analysis)	AI-driven adaptive learning system (machine learning-based personalized learning platform with content adaptation, predictive analytics, and intelligent tutoring features)	[26]	40
Training teachers on new topics and new tools in Physics education	2024	Italy	Conference	Secondary school physics teachers	Approximately 100 registered participants (≈30 per session)	Descriptive project-based study (training program implementation with workshops, seminars, and participant feedback evaluation)	Mixed AI applications (Machine Learning for image classification, AI-based data analysis tools, virtual assistants)	[88]	41
Transforming Physics Teacher Training	2025	Ecuador	Article	Pre-service Physics teachers	24 participants	Quantitative descriptive-comparative study (pre-/post-interven	Generative AI (ChatGPT; NLP-based large language	[89]	42

Through ChatGPT: A Study on Usability and Impact						tion Likert-scale surveys following a workshop)	model)		
Unreflected Acceptance Investigating the Negative Consequences of ChatGPT-Assisted Problem Solving in Physics Education	2024	Germany	Conference	Undergraduate STEM students (physics and non-physics background s)	39 participants (ChatGPT group: 27; Search engine group: 12)	Mixed-methods experimental study (between-subject design comparing ChatGPT vs. search engine; includes pretest, main task, interaction logs, questionnaires, and exit interviews)	Generative AI (ChatGPT 3.5; large language model)	[90]	43
Visualising relativity: assessing high school students' understanding of complex physics concepts through AI-generated images	2024	Brazil	Article	High school students (Year 12)	10 students	Qualitative case study with mixed elements (pre-/post-test, AI-based activity, interviews using report-aloud protocol, and artifact analysis of prompts and images)	Generative AI (AI image generation via chatbot; Bing AI image generator using NLP-based prompting)	[43]	44