



Received: 26-08-2025
Accepted: 06-10-2025

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

ChatGPT and DeepSeek in Physics Education: A Narrative and Thematic Literature Review of Pedagogical Implications

Konstantinos T Kotsis

Lab of Physics Education and Teaching, Department of Primary Education, University of Ioannina, Greece

DOI: <https://doi.org/10.62225/2583049X.2025.5.5.5074>

Corresponding Author: Konstantinos T Kotsis

Abstract

This paper presents a narrative and thematic literature review examining the pedagogical implications of integrating two leading large language models, ChatGPT and DeepSeek, into physics education. It investigates how their distinct architectures, interaction styles, and affordances influence conceptual understanding, problem-solving practices, and instructional design. ChatGPT's dialogic and adaptive nature aligns with constructivist and inquiry-based approaches, fostering metacognitive engagement and conceptual change through scaffolded explanations, analogies, and multimodal representations. Conversely, DeepSeek emphasizes computational efficiency, precision, and iterative refinement, making it particularly effective in structured problem-solving contexts and high-stakes learning environments. By synthesizing

recent empirical and conceptual studies, the review identifies complementary pedagogical roles for these tools. It emphasizes the importance of strategic teacher mediation to avoid cognitive passivity and ensure meaningful learning. Ethical and practical challenges, including academic integrity, data privacy, algorithmic bias, and teacher professional development, are critically examined. The paper concludes by outlining future research directions for integrating generative AI in physics education, emphasizing the need for theoretically grounded, classroom-based studies that address both pedagogical opportunities and epistemic implications. This review provides a comprehensive foundation for understanding how AI can reshape physics teaching and learning in a rapidly evolving educational landscape.

Keywords: ChatGPT, DeepSeek, Physics Education, Artificial Intelligence, Pedagogy

1. Introduction

The swift incorporation of artificial intelligence (AI) into educational settings has generated increasing interest in comprehending both its theoretical foundations and its practical implementations within particular disciplinary frameworks. Although recent academic research has explored the conceptual frameworks, pedagogical opportunities, and ethical implications of AI in education ^[1, 2], broader systematic reviews have shown that much of the existing research focuses predominantly on technological aspects, with limited attention to the pedagogical role of educators ^[3]. There is, therefore, a pressing necessity for empirical studies that assess the actual efficacy and classroom influence of particular AI platforms. Recent research has begun to address this need by proposing concrete lesson designs and practical strategies for AI integration in physics education ^[4]. This study conducts a comparative analysis of two leading AI-driven conversational agents, ChatGPT and DeepSeek, within the realm of physics education, building on the theoretical and pedagogical foundations laid in a prior companion paper and extending recent comparative perspectives in the field ^[5, 6]. This emphasis facilitates the examination of their unique affordances, constraints, and prospective functions in enhancing physics education and learning, aligning with recent discussions on the broader transformation of physics teaching in the era of AI integration ^[7].

Physics education provides a distinctive platform for integrating AI, due to the field's dual focus on conceptual comprehension and problem-solving abilities ^[8, 9]. Educators often face the challenge of making abstract concepts comprehensible while also fostering analytical thinking and empirical reasoning. AI tools like ChatGPT and DeepSeek can aid in these tasks by providing context-specific explanations, facilitating dialogue-driven inquiries, and generating examples or problem sets tailored to various levels of student proficiency. Nonetheless, variations in their training data, interaction methods, and pedagogical alignment may affect their efficacy in the classroom.

This study intends to investigate these differences via a systematic, classroom-focused comparative analysis. This research

aims to assess the performance of ChatGPT and DeepSeek across diverse instructional contexts to discern patterns in accuracy, conceptual clarity, adaptability to student requirements, and their ability to facilitate inquiry-based and problem-solving methodologies in physics.

The paper prioritizes evidence from specific educational contexts over generic AI capabilities, aiming to generate actionable insights for educators and policymakers. This approach enhances the theoretical insights presented in the previous paper by integrating empirical evidence and practical implications, thereby providing a more comprehensive understanding of the role of AI in modern physics education.

This paper prioritizes evidence from specific educational contexts over generic AI capabilities, aiming to generate actionable insights for educators and policymakers. This approach enhances the theoretical insights presented in the previous paper by integrating empirical evidence and practical implications, thereby providing a more comprehensive understanding of the role of AI in modern physics education.

Accordingly, this study is a comparative empirical research article situated at the intersection of physics education and the integration of artificial intelligence. It is grounded in classroom-focused evidence. It systematically evaluates the distinct affordances of ChatGPT and DeepSeek in supporting both conceptual understanding and structured problem-solving. Its scope and methodological orientation align with the characteristics of original research papers in science and physics education journals, offering insights that are simultaneously pedagogical, empirical, and ethically attuned.

2. Methodology

This article adopts a narrative and thematic literature review methodology, aiming to synthesize emerging research on the pedagogical integration of ChatGPT and DeepSeek in physics education. Rather than conducting a systematic review with predefined coding protocols, the objective here is to develop a coherent and critically informed account of how these AI systems are discussed and positioned in recent scholarly work, emphasizing conceptual themes, pedagogical frameworks, and ethical considerations. This approach aligns with established traditions in science education research, where narrative reviews serve to map evolving fields, identify conceptual tensions, and articulate directions for further inquiry^[10, 11].

The literature search was conducted between January and August 2025 across major academic databases, including Scopus, Web of Science, ERIC, and Google Scholar. The search terms combined concepts related to artificial intelligence and large language models (e.g., "ChatGPT," "DeepSeek," "large language models," "artificial intelligence in education") with physics education keywords (e.g., "physics learning," "problem solving," "conceptual understanding," "pedagogical integration," "inquiry-based science education"). Additional sources were retrieved through backward citation tracking of key review articles and forward citation alerts, ensuring coverage of both foundational theoretical texts and the most recent empirical and conceptual discussions.

Inclusion criteria focused on peer-reviewed articles, conference proceedings, and pre-prints published from 2019 onwards, reflecting the period during which large language

models began to have a discernible impact on educational practices. Studies were selected if they addressed the use of ChatGPT, DeepSeek, or similar generative AI tools within physics or science education contexts, demonstrating clear pedagogical relevance. Conceptual papers, empirical classroom studies, and policy-focused analyses were all included to provide a comprehensive thematic landscape. Exclusion criteria were applied to publications that focused exclusively on technical model development without educational implications, studies addressing AI in unrelated STEM fields without transferable pedagogical insights, and non-academic commentaries lacking analytical grounding.

Thematic synthesis was employed to organize the literature around central pedagogical and ethical dimensions. Initially, studies were reviewed to identify recurrent themes related to AI-supported conceptual understanding, problem-solving strategies, teacher-AI interaction, assessment practices, ethical concerns, and cross-cultural or contextual considerations in physics education. These themes were iteratively refined through multiple readings and comparative analysis of studies, allowing for the development of an interpretive framework that connects the affordances of ChatGPT and DeepSeek with key debates in contemporary physics education research. Particular attention was paid to contrasting the dialogic, adaptive uses of ChatGPT with the more structured, precision-oriented uses of DeepSeek to map their complementary and divergent roles across educational settings.

This narrative methodology provides a structured yet flexible basis for examining the emerging body of literature. It enables the articulation of thematic patterns and pedagogical implications without reducing the diversity of studies to rigid categories. Moreover, it allows for critical reflection on the conceptual and methodological gaps in current research, highlighting areas where future empirical studies could provide more robust evidence to support the integration of AI in physics classrooms.

3. Key Features and Affordances of ChatGPT and DeepSeek in Physics Education

The emergence of large language models (LLMs) like ChatGPT and DeepSeek in education has created new opportunities for improving instructional design, personalizing learning, and broadening the range of formative feedback. Although both platforms possess the core ability to produce human-like text in response to user prompts, their architecture, design philosophy, and training methodologies result in unique advantages that warrant further analysis within the realm of physics education.

ChatGPT, created by OpenAI, utilizes the Generative Pre-trained Transformer architecture, refined through reinforcement learning from human feedback to enhance contextual relevance, coherence, and alignment with user intent^[12]. This combination enables ChatGPT to respond in a manner that frequently aligns with conversational teaching methods, making it especially effective for scaffolding student comprehension through incremental explanations. In physics classrooms, ChatGPT can facilitate Socratic questioning, create analogies for intricate phenomena, and tailor explanations to diverse levels of prior knowledge. It can also generate structured instructional resources, such as experiment worksheets for teachers, supporting the design of inquiry-based learning activities^[13]. Moreover, recent work has demonstrated ChatGPT's capacity to support the

planning and implementation of hands-on physics experiments in primary school, enhancing experiential learning opportunities and teacher preparation ^[14]. The ability to generate various representations, such as verbal descriptions, symbolic equations, and conceptual diagrams, enhances multimodal learning, which has been demonstrated to promote deeper conceptual understanding in science education ^[15, 16].

DeepSeek, although an advanced LLM, has adopted a distinct design approach, prioritizing computational efficiency, enhanced access to multilingual training datasets, and an emphasis on precision in problem-solving ^[17]. This orientation is well-suited for computational physics tasks and structured problem-solving exercises, where precision and succinct reasoning are essential ^[18]. Moreover, DeepSeek's architecture incorporates mechanisms for iterative refinement, allowing it to revisit and modify prior responses based on user feedback with greater flexibility than certain other LLMs ^[19, 20]. This feature corresponds with inquiry-based learning methodologies in physics, wherein hypothesis modification and iterative experimentation are essential to the educational process ^[21]. The capabilities of these tools are influenced by their interactional style. ChatGPT excels in fostering a collaborative learning environment, emulating a tutor-like presence that engages students in prolonged exploration dialogue. This dialogic nature may prompt students to express their reasoning, a practice linked to enhanced conceptual understanding and meta-cognitive growth ^[22]. Conversely, DeepSeek can provide concise, targeted responses, which are especially beneficial in exam preparation contexts or when students need swift elucidation of procedural steps.

Moreover, both platforms can facilitate differentiated instruction in physics by adjusting the complexity of language, depth of explanation, and the use of technical terminology according to learner requirements. This adaptability corresponds with inclusive pedagogical principles and facilitates the creation of learning experiences that accommodate diverse student populations, including those with differing levels of science capital ^[23]. Empirical studies in laboratory contexts, such as the use of Arduino-based setups for thermal equilibrium experiments, have demonstrated how hands-on and technology-enhanced approaches can support conceptual understanding and differentiated learning pathways ^[24]. The synergistic potential of ChatGPT and DeepSeek indicates that, when strategically integrated, they may enhance each other, ChatGPT facilitating conceptual exploration and narrative-driven elucidation. At the same time, DeepSeek guarantees accuracy and efficiency in structured problem-solving scenarios.

4. Comparative Pedagogical Analysis

A direct pedagogical comparison of ChatGPT and DeepSeek in physics education highlights significant differences in their instructional capabilities, the depth and style of their responses, and their potential compatibility with various teaching methodologies. Although both systems utilize extensive corpora and sophisticated transformer-based architectures, their differing optimization objectives and interactional characteristics place them uniquely within the educational framework.

ChatGPT's ability for dialogic interaction renders it

especially appropriate for constructivist and socio-cultural methodologies in science education. In a constructivist classroom, students actively construct new knowledge through engagement with peers, educators, and instructional resources, frequently by integrating prior understandings with new information ^[25]. ChatGPT's adaptive and context-sensitive responses enable it to function as a "conversational partner," facilitating learners' conceptual refinement and enhancing metacognitive awareness. The model facilitates student reasoning by posing clarifying questions, suggesting analogies, or presenting counterexamples, thereby creating cognitive conflict that can stimulate conceptual change ^[26, 27, 28].

Conversely, DeepSeek exhibits pedagogical advantages in contexts where accuracy, procedural transparency, and temporal efficiency are paramount. In the context of physics problem-solving, particularly within cognitive apprenticeship frameworks ^[29], DeepSeek's succinct and precise presentation can replicate the function of an expert demonstrator. The model's iterative refinement mechanism facilitates a feedback loop akin to teacher-guided correction in conventional education, enhanced by immediate response times. This trait is especially beneficial in high-stakes situations, such as exam preparation, where students often prioritize efficiency over exploratory discussion.

From the standpoint of formative assessment, ChatGPT provides a diverse array of formative feedback modalities, encompassing explanatory elaboration, suggestive questioning, and multiple-solution pathways. These attributes correspond with Black and Wiliam's ^[30] assertion that formative assessment is most efficacious when it fosters self-regulation and ownership of learning. Conversely, DeepSeek offers concise feedback that is readily actionable, proving effective in situations where learners require specific corrections or swift validation of procedural steps instead of extensive conceptual analysis.

In the context of differentiated instruction, both tools present potential, albeit with distinct strengths. ChatGPT excels at adjusting complexity and establishing interdisciplinary connections, such as correlating kinematics with historical scientific experiments, thereby enhancing the educational experience for students with significant scientific capital or interest. DeepSeek, however, excels at deconstructing intricate procedural issues into a series of manageable steps, rendering it accessible for learners who may find cognitive load challenging in problem-solving tasks ^[31].

The incorporation of these tools within the same instructional setting may produce synergistic advantages. A physics instructor may utilize ChatGPT for initial conceptual explorations at the beginning of a unit and subsequently employ DeepSeek to reinforce procedural proficiency. This integrated method corresponds with studies on adaptive expertise, which highlight the capacity to engage in exploratory and efficiency-oriented problem-solving strategies alternately ^[32]. The tools are not competitive; rather, they function as pedagogically distinct yet complementary resources, a conclusion that resonates with recent work on reciprocal AI evaluation in educational contexts ^[6].

5. Discussion

The reviewed literature reveals a rapidly evolving yet conceptually fragmented field, in which the integration of generative AI tools, such as ChatGPT and DeepSeek, into

physics education is being theorized, piloted, and critiqued from multiple perspectives. Across studies, four interrelated thematic domains emerge: the role of AI in fostering conceptual understanding, its contribution to structured problem solving, its implications for teacher-AI interaction and instructional design, and the broader pedagogical and ethical tensions that arise when these tools are embedded in classroom practice.

A first key theme concerns the use of generative AI to support conceptual understanding in physics. Narrative evidence from recent studies suggests that ChatGPT's dialogic interaction style facilitates students' ability to articulate, refine, and negotiate their reasoning processes, aligning with established constructivist and sociocultural approaches to science learning [25, 22]. Recent curriculum-focused work has also highlighted how ChatGPT can be effectively integrated into inquiry-based science instruction at the primary level to support conceptual engagement and structured exploration [33]. Through adaptive questioning, analogical explanations, and multimodal representations, ChatGPT can scaffold learners' engagement with abstract physical concepts, promoting metacognitive reflection and conceptual change [16, 28].

These features are particularly relevant in physics education, where misconceptions often persist despite direct instruction, and dialogic interventions have been shown to catalyze more robust knowledge restructuring.

However, several authors caution that ChatGPT's conversational fluency may create an illusion of understanding if students are not guided to evaluate AI-generated explanations critically, underlining the need for teacher mediation [1, 34].

A second thematic area relates to structured problem solving. DeepSeek is frequently portrayed as complementing ChatGPT by emphasizing precision, iterative refinement, and computational efficiency in tasks such as equation manipulation, algorithmic reasoning, and stepwise solution of physics problems [20, 17]. This orientation aligns with cognitive apprenticeship models, in which expert demonstration and scaffolded feedback are central to the development of procedural skills [29]. In the reviewed literature, DeepSeek is often portrayed as an efficient assistant for tasks that require high degrees of accuracy and speed, such as exam preparation or targeted practice with quantitative problems. At the same time, several studies highlight the danger of students becoming over-reliant on procedural outputs without fully internalizing the underlying principles, echoing broader concerns about AI-induced cognitive offloading [35]. Consequently, successful pedagogical integration is frequently framed as a complementary orchestration, with ChatGPT supporting exploratory conceptual engagement and DeepSeek ensuring procedural rigor.

The third theme focuses on teacher-AI interaction and instructional design. Rather than displacing the teacher, recent literature emphasizes the reconfiguration of the teacher's role as a designer, mediator, and critical interlocutor in AI-supported classrooms [36, 37]. Studies in physics education suggest that teachers who strategically alternate between dialogic engagement with ChatGPT and structured exercises mediated by DeepSeek can tailor instruction to diverse learner profiles and learning objectives [28]. Similar findings have been observed in studies employing online physics labs, which demonstrated that

well-designed digital learning environments can enhance both students' self-efficacy and conceptual understanding of complex topics [38]. This aligns with research on differentiated instruction and adaptive expertise, which highlights the value of switching between exploratory and efficiency-oriented strategies to foster both deep understanding and flexible problem-solving capacities [32]. Moreover, emerging evidence suggests that teacher professional development is a crucial factor: educators with higher AI literacy are better equipped to integrate these tools into inquiry-based pedagogies while maintaining epistemic control over the learning process [39].

Finally, the literature consistently highlights pedagogical and ethical tensions that arise from the integration of AI. Concerns regarding academic integrity, cognitive dependency, data privacy, and algorithmic bias appear repeatedly across studies [40, 34]. These issues are particularly acute in physics education because the subject's precision-oriented epistemology can be undermined by uncritical reliance on AI-generated outputs. Several authors argue for the development of transparent guidelines, institutional policies, and classroom practices that frame AI tools as partners in inquiry rather than authoritative sources of knowledge [41, 42]. Furthermore, cross-cultural analyses highlight that most current studies are situated in English-speaking or Western contexts, raising questions about linguistic bias, cultural inclusivity, and the global relevance of AI-assisted physics education [43, 44].

Taken together, these thematic domains underscore that the pedagogical potential of ChatGPT and DeepSeek lies not in their independent capabilities but in their strategic orchestration within carefully designed instructional environments. ChatGPT's strength in fostering conceptual dialogue can counterbalance the procedural efficiency of DeepSeek, but both require thoughtful mediation by educators to avoid cognitive passivity and ethical pitfalls. The literature suggests a future research agenda that moves beyond exploratory case studies to sustained, theoretically grounded investigations of how these tools reshape physics teaching and learning. This includes examining their impact on conceptual change processes, epistemic practices, and teacher professional development across diverse educational and cultural contexts. By synthesizing these thematic strands, this review highlights both the promise and complexity of integrating large language models into physics education, offering a framework for future empirical and theoretical work in the field.

6. Ethical and Practical Implications

The integration of ChatGPT and DeepSeek into physics education raises complex ethical and practical considerations that extend well beyond questions of technological efficiency. These issues intersect with longstanding debates in science education regarding epistemic authority, learner autonomy, equity, and the role of teachers as mediators of knowledge. The literature increasingly emphasizes that the educational value of AI depends not only on its technical affordance but also on the pedagogical and ethical frameworks within which it is situated [1, 39].

A central concern pertains to academic integrity in physics learning environments. ChatGPT and DeepSeek can generate highly coherent, contextually appropriate solutions to conceptual and quantitative problems, which can tempt

students to substitute AI outputs for their own reasoning. Unlike traditional plagiarism, AI-generated content is often original in form, making detection more difficult and necessitating new institutional and pedagogical strategies [34]. This challenge is particularly acute in physics, where problem-solving tasks are central to learning and assessment. Without appropriate framing, students may bypass the cognitive processes essential for conceptual understanding, leading to superficial forms of engagement. Consequently, many authors advocate for explicit guidelines that define legitimate and illegitimate uses of AI in academic contexts, along with assessment practices that emphasize process over product [42].

Closely linked to this is the issue of cognitive dependency. Overreliance on AI-generated explanations or step-by-step solutions may undermine students' ability to engage in independent reasoning and metacognitive regulation. Research in cognitive psychology has long shown that extensive externalization of cognitive processes can impair knowledge consolidation and problem-solving resilience [35]. In physics education, where conceptual change requires active engagement with evidence, models, and representations, the uncritical use of AI risks reinforcing algorithmic reasoning without fostering deep understanding. Mitigating this risk requires teachers to design learning activities that position AI as a support for inquiry, not a substitute for reasoning. Reflective tasks, comparative evaluations of AI-generated and student-generated solutions, and dialogic scaffolding have been proposed as strategies to counter cognitive offloading while preserving AI's affordances for exploration and feedback [22, 28].

Data privacy and algorithmic bias constitute further critical domains. Both ChatGPT and DeepSeek rely on the collection and processing of user input data, raising questions about data security, anonymization, and the potential use of educational interactions for model retraining [39]. These concerns are amplified in school contexts governed by strict legal frameworks such as the General Data Protection Regulation (GDPR) and the Family Educational Rights and Privacy Act (FERPA). Additionally, biases embedded in training datasets can shape the pedagogical character of AI outputs, privileging certain linguistic, cultural, or epistemological perspectives while marginalizing others [40]. In physics education, where examples, measurement systems, and historical narratives vary significantly across contexts, these biases can lead to the inadvertent privileging of Western-centric or Anglophone frameworks. Addressing these challenges requires not only technological solutions, such as bias mitigation and explainability tools, but also curricular adaptations that foster critical AI literacy among both teachers and students.

Finally, the literature underscores practical implementation challenges related to infrastructure, teacher training, and institutional policy. Access to stable internet, compatible devices, and updated software remains uneven across regions, potentially exacerbating existing educational inequalities [44]. Institutional contexts such as the Greek Experimental Model Schools have demonstrated how structured frameworks can support the implementation of innovative curricula and pedagogical practices, highlighting their potential role in integrating emerging technologies effectively [45]. Furthermore, many teachers lack the

necessary professional development to integrate AI tools in pedagogically meaningful ways. Evidence from studies on preservice teacher preparation in physics education reveals significant gaps in readiness for implementing digital and distance learning approaches, underscoring the need for targeted training programs [46]. Studies indicate that educators with higher levels of AI literacy are more likely to design inquiry-oriented, ethically grounded uses of ChatGPT and DeepSeek, whereas those with limited training tend to rely on them in instrumental or superficial ways [36, 37]. Institutional frameworks that combine infrastructural support with sustained professional learning opportunities are therefore essential for equitable and effective implementation.

Taken together, these ethical and practical considerations underscore that the trans-formative potential of ChatGPT and DeepSeek in physics education is inextricably linked to the conditions under which they are used. Effective integration requires not only clear policies and technological safeguards but also pedagogical intentionality, teacher agency, and a critical understanding of the epistemic implications of AI.

By foregrounding these issues, the literature calls for a more reflexive and ethically attuned approach to the adoption of AI tools in science education, ensuring that technological innovation reinforces rather than undermines the goals of meaningful, inclusive, and epistemically responsible physics learning.

7. Conclusions

The emergence of large language models such as ChatGPT and DeepSeek represents a significant development in the landscape of physics education. Their distinct affordances offer complementary opportunities for supporting both conceptual understanding and structured problem solving, while simultaneously raising critical pedagogical and ethical questions. ChatGPT's dialogic, adaptive style lends itself to inquiry-based and constructivist approaches, encouraging learners to articulate and refine their reasoning. DeepSeek's precision and iterative problem-solving capabilities, on the other hand, align with the structured demands of quantitative reasoning and procedural fluency in physics.

This literature review has shown that the educational value of these tools depends less on their intrinsic capabilities and more on how they are embedded within thoughtful instructional designs. Their effective integration requires teachers who can strategically orchestrate dialogic and procedural modes of engagement, ensuring that AI tools act as mediators of learning rather than substitutes for student reasoning. It also requires institutional frameworks that address issues of academic integrity, cognitive dependency, privacy, bias, and equitable access.

The field is at a formative stage, with research to date highlighting both promising practices and significant gaps. Future studies need to examine, in sustained and context-sensitive ways, how these tools shape conceptual change processes, teacher roles, classroom discourse, and learners' epistemic agency across diverse cultural and institutional settings. By approaching AI not merely as a technological innovation but as a pedagogical and ethical challenge, physics education can develop models of integration that enhance, rather than diminish, the intellectual rigor and inclusivity of science learning.

8. References

- Holmes W, Porayska-Pomsta K, Holstein K, Sutherland E, Baker T, Buckingham Shum S. Ethics of AI in Education: Towards a community-wide framework. *Int J Artif Intell Educ*. 2022; 32:17.
- Luckin R. *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. UCL Institute of Education Press, 2018.
- Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education -- Where are the educators? *Int J Educ Technol High Educ*. 2019; 16:39.
- Vakarou G, Stylos G, Kotsis KT. AI for enhancing physics education: Practical tools and lesson plans. *Int J Sci Math Technol Learn*. 2024; 31(2):159-176.
- Kotsis KT. ChatGPT Develops Physics Experiment Worksheets for Primary Education Teachers. *Eur J Educ Stud*. 2024; 11(5):1-20.
- Kotsis KT. ChatGPT and DeepSeek Evaluate One Another for Science Education. *EIKI J Eff Teach Methods*. 2025; 3(1):98-102.
- Kotsis KT. From Chalkboard to Chatbot: The Future of Physics Education through Artificial Intelligence Integration. *EIKI J Eff Teach Methods*. 2025; 3(2):74-79.
- Docktor JL, Mestre JP. Synthesis of discipline-based education research in physics. *Phys Rev Spec Top - Phys Educ Res*. 2014; 10(2):020119.
- Duit R, Treagust DF. Conceptual change: A powerful framework for improving science teaching and learning. *Int J Sci Educ*. 2003; 25(6):671-688.
- Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *J Chiropr Med*. 2006; 5(3):101-117.
- Grant MJ, Booth A. A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Info Libr J*. 2009; 26(2):91-108.
- OpenAI. GPT-4 technical report [Technical report]. arXiv, 2023.
- Kotsis KT. ChatGPT as Teacher Assistant for Physics Teaching. *EIKI J Eff Teach Methods*. 2024; 2(4):18-27.
- Kotsis KT. ChatGPT in teaching physics hands-on experiments in primary school. *Eur J Educ Stud*. 2024; 11(10):126-143.
- Ainsworth S. DeFT: A conceptual framework for considering learning with multiple representations. *Learn Instr*. 2006; 16(3):183-198.
- Treagust DF, Tsui CY, editors. *Multiple representations in biological education (Models and Modeling in Science Education, Vol. 7)*. Dordrecht: Springer, 2013.
- Madupati B, Jonnalagadda AK, Madupathi S. A Technical Comparison of ChatGPT and DeepSeek: Architecture, Efficiency, and Performance. *Int J Glob Innov Solut*, 2025.
- Manik MMH. ChatGPT vs. DeepSeek: A comparative study on AI-based code generation. arXiv preprint. 2025, arXiv:2502.18467.
- Madaan A, Tandon N, Gupta P, Hallinan S, Gao L, Wiegrefe S, *et al*. Self-refine: Iterative refinement with self-feedback. *Adv Neural Inf Process Syst*. 2023; 36:46534-46594.
- Jiang Q, Gao Z, Karniadakis GE. DeepSeek vs. ChatGPT vs. Claude: A comparative study for scientific computing and scientific machine learning tasks. *Theor Appl Mech Lett*. 2025; 15(3):100583.
- Kotsis KT. The Significance of Experiments in Inquiry-based Science Teaching. *Eur J Educ Pedagogy*. 2024; 5(2):86-92.
- Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educ Psychol*. 2014; 49(4):219-243.
- Archer L, Dawson E, DeWitt J, Seakins A, Wong B. "Science capital": A conceptual, methodological, and empirical argument for extending Bourdieusian notions of capital beyond the arts. *J Res Sci Teach*. 2015; 52(7):922-948.
- Stylos G, Georgopoulos K, Nousis V, Kotsis KT. Using Arduino in Introductory Thermal Energy Experiments, The Case of Thermal Equilibrium. *Phys Teach*. 2024; 62(8):669-672.
- Driver R, Asoko H, Leach J, Mortimer E, Scott P. Constructing scientific knowledge in the classroom. *Educ Res*. 1994; 23(7):5-12.
- Posner GJ, Strike KA, Hewson PW, Gertzog WA. Accommodation of a scientific conception: Toward a theory of conceptual change. *Sci Educ*. 1982; 66(2):211-227.
- Kotsis KT. Correcting Students' Misconceptions in Physics Using Experiments Designed by ChatGPT. *Eur J Contemp Educ E-Learning*. 2024; 2(2):83-100.
- Kotsis KT. Comparing ChatGPT and DeepSeek in Addressing Misconceptions about Physics Concepts. *Eur J Contemp Educ E-Learning*. 2025; 3(2):191-206.
- Collins A, Brown JS, Holum A. Cognitive apprenticeship: Making thinking visible. *Am Educ*. 1991; 15(3):6-11.
- Black P, Wiliam D. Developing the theory of formative assessment. *Educ Assess Eval Account*. 2009; 21(1):5-31.
- Sweller J. Cognitive load during problem solving: Effects on learning. *Cogn Sci*. 1988; 12(2):257-285.
- Hatano G, Inagaki K. Two courses of expertise. In: Stevenson H, Azuma H, Hakuta K, editors. *Child development and education in Japan*. W. H. Freeman, 1986, 262-272.
- Kotsis KT. Integrating ChatGPT into the inquiry-based science curriculum for primary education. *Eur J Educ Pedagogy*. 2024; 5(6):28-34.
- Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innov Educ Teach Int*. 2024; 61(2):228-239.
- Sparrow B, Liu J, Wegner DM. Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*. 2011; 333(6043):776-778.
- Luckin R, Holmes W, Griffiths M, Forcier LB. *Intelligence Unleashed: An Argument for AI in Education*. Pearson Education, 2016.
- Holmes W, Bialik M, Fadel C. *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston, MA: Center for Curriculum Redesign, 2019.
- Stylos G, Christonasis A, Georgopoulos K, Kotsis KT. The impact of an online physics lab on university students' self-efficacy and understanding of thermal concepts during COVID-19 pandemic. *J Math Sci Teach*. 2023; 3(2):em049.
- Regan PM, Jesse J. Ethical challenges of edtech, big

- data, and personalized learning: Twenty-first century student sorting and tracking. *Ethics Inf Technol.* 2019; 21(3):167-179.
40. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, 2021, 610-623.
 41. Nature Editorial. Tools such as ChatGPT threaten transparent science: Here are our ground rules for their use. *Nature.* 2023; 613(7945):612.
 42. Bhavsar D, Marušić A, Ng JY. Policies on artificial intelligence chatbots among academic publishers. *Res Integr Peer Rev.* 2025; 10:1.
 43. Tsiouri E, Tsihouridis C, Kotsis KT. The science curricula for ages 11-12 across the European Union: A comparative analysis. *Eurasian J Sci Environ Educ.* 2024; 4(2):39-46.
 44. UNESCO. *AI and education: Guidance for policy-makers*. Paris, France: UNESCO, 2021.
 45. Kotsis KT, Tsiouri E. The Important Role of the Greek Experimental Model Schools for the New Curricula in Education. *Eur J Contemp Educ E-Learning.* 2024; 2(6):44-60.
 46. Kotsis KT, Nikou G, Stylos G. The Readiness of Preservice Teachers for the Distance Education of Physics in Greece. *Eur J Educ Pedagogy.* 2023; 4(3):26-30.