Research in Science Education
https://doi.org/10.1007/s11165-024-10176-3

Exploring the Impact of Artificial Intelligence in Teaching and Learning of Science: A Systematic Review of Empirical Research

Firas Almasri[1,2,3,4]

Accepted: 7 June 2024
© The Author(s) 2024

Abstract

The use of Artificial Intelligence (AI) in education is transforming various dimensions of the education system, such as instructional practices, assessment strategies, and administrative processes. It also plays an active role in the progression of science education. This systematic review attempts to render an inherent understanding of the evidence-based interaction between AI and science education. Specifically, this study offers a consolidated analysis of AI's impact on students' learning outcomes, contexts of its adoption, students' and teachers' perceptions about its use, and the challenges of its use within science education. The present study followed the PRISMA guidelines to review empirical papers published from 2014 to 2023. In total, 74 records met the eligibility for this systematic study. Previous research provides evidence of AI integration into a variety of fields in physical and natural sciences in many countries across the globe. The results revealed that AI-powered tools are integrated into science education to achieve various pedagogical benefits, including enhancing the learning environment, creating quizzes, assessing students' work, and predicting their academic performance. The findings from this paper have implications for teachers, educational administrators, and policymakers.

Keywords AI · Artificial intelligence · Science education · Engineering students · STEM learning · Science learning

✉ Firas Almasri
firas.almasri@torontomu.ca

[1] College of Engineering and Sciences, International University of Science and Technology in Kuwait, Ardiya, Kuwait

[2] Department of Physics, Toronto Metropolitan University, Toronto, ON M5B 2K3, Canada

[3] Centre for Education Studies, University of Warwick, Coventry, United Kingdom

[4] Institute for Biomedical Engineering, Science and Technology (iBEST), A Partnership Between Toronto Metropolitan University and St. Michael's Hospital, Toronto, ON M5B 1T8, Canada

Published online: 27 June 2024

Introduction

Artificial Intelligence (AI) is a broad field encompassing various technologies that have been developed over the past 50 years to enable machines to perform tasks traditionally requiring human intelligence, such as perceiving, reasoning, learning, and interacting (Ergen, 2019). However, recent advancements in generative AI (GenAI), particularly models like ChatGPT, have brought unprecedented attention to AI's transformative potential across multiple industries (Hong et al., 2022; Lucci et al., 2022). Unlike predictive (pre-generative) AI which focuses predictions and decision making through a variety of machine learning and modelling techniques, Generative AI specializes in creating new content, such as text, images, and codes by using models of deep learning (Dai, 2023; Tang & Nichols, 2024). This distinction is essential to understand the breadth of AI applications in education.

Artificial intelligence in education (AIEd) is an evolving interdisciplinary arena incorporating AI technologies to renovate and enhance teaching and learning environments. Particularly, the application of AI in science teaching and learning is becoming more popular, even as interest in AI's effects on general education is growing (Chiu et al., 2023; Gonzalez et al., 2017). More specifically, machine learning, a specific artificial intelligence technology, has been applied to automatically evaluate scientific models used in the education sector. Zhai et al. (2022) employed machine learning techniques to assess the quality of these models after gathering student responses to activities. Their research demonstrates how artificial intelligence can be used to automate assessment procedures and provide students with timely and detailed feedback on their work in the area of science education (Zhai, C Haudek, Zhai et al., 2020a, b, 2022). Similarly, Popenici and Kerr (2017) conducted a study to investigate the impact of AI on the teaching-learning process in higher education settings. Their study focused on how intelligent technologies are affecting student learning and traditional teaching approaches in education. Their research presents valuable insights into the incorporation of AI within science education contexts.

Zawacki-Richter et al. (2019), in their systematic assessment of AI applications in higher education, focused on the vital role that teachers can play in this domain. Their results suggest how important it is to explore and understand the needs and perceptions of teachers when integrating these technologies into teaching-learning settings. Likewise, Xu and Ouyang (2022) employed a systematic literature review method to identify and summarize research studies and classify the roles of AI in the educational system. Their findings advocate the use of AI is within the education environment to support its role in three ways: (1) AI as a new subject, (2) AI as an immediate mediator, and AI as a complementary aid to impact the teacher-learner, learner-self, and learner-learner relationships.

Though artificial intelligence has flourished in numerous domains within the education system, a comprehensive analysis of its role, advantages, and challenges in science education must be further explored through empirical investigations. This knowledge gap might prompt teachers, policymakers, and educational administrators to base decisions on patchy as well as limited information, lacking potential opportunities to enhance science teaching and learning with the help of AI. To fill this gap, the present paper provides a systematic review that comprehensively examines and consolidates AI's impact on science education, evidenced by the empirical publications published from 2014 to 2023. While GenAI represents a significant leap in AI capabilities, this review considers the full spectrum of AI technologies, including both pre-GenAI and GenAI developments. In this way, we attempt

to provide a holistic perspective on the current landscape, aiding stakeholders in leveraging AI's potential while also considering its challenges and ethical implications for educational domains. The overarching objective of this review is to provide insights that could guide future research endeavors and advocate for evidence-based practices to enrich science education through the effective utilization of artificial intelligence.

Research Background

Overview of Science Education

The goal of science education is not only to teach scientific knowledge but also to develop a scientifically literate populace capable of engaging in scientific reasoning and decision-making (Almasri, 2021; Grinnell, 2021). This aligns with the "Science for All" movement, emphasizing the importance of science education for all students, not just those pursuing careers in science (Almasri et al., 2022; Mansour, 2009). Students' scientific literacy and critical thinking abilities are developed through the teaching and learning of scientific theories, procedures, and experiments in science education (Alharbi et al., 2022; Liu & Pásztor, 2022; Mogea, 2022; Zulyusri et al., 2023).

The nature of science education extends beyond content-based instruction to include student-centered activities and the development of scientific literacy for citizenship (Almasri et al., 2021; Irez, 2006a, b; Kolstø, 2001). National development hinges significantly on robust contributions from the scientific community, driving economic growth and propelling the overall advancement of a nation (Hewapathirana & Almasri, 2022; Kola, 2013). The "Call to Action" for science education highlights an compelling necessity to improve educational approaches and make them consistent with the demands of the 21st century (Holme, 2021; Ibáñez & Delgado-Kloos, 2018). It is essential for developing students' foundational knowledge, intriguing their curiosity, and getting them ready for STEM careers as per the contemporary world's needs. Through AI incorporation, science education can be made more interesting, approachable, and pertinent for students of all ages and backgrounds by emphasizing experiential learning.

Prospects of Incorporating AI in Science Learning

With AI technology's continuous evolution and popularity, the possibilities for its application in science education are promising but not without challenges. AI has the capability to transform the way science is taught and learned. One of the most compelling applications of AI in science education is its ability to simulate scientific experiments and provide virtual laboratory experiences to science learners. This ensures that students can practice and develop their scientific skills in a safe and controlled environment, potentially saving expenses and offering new opportunities for exploring scientific concepts that may not be feasible in traditional laboratory settings (Wahyono et al., 2019). However, these virtual experiences may lack the tactile and hands-on aspects of interaction with the physical world (Tang & Cooper, 2024), which are crucial for certain types of learning.

By leveraging AI, educators can also move away from traditional, one-size-fits-all approaches to education and instead provide personalized and interactive learning experi-

ences for students. AI-powered algorithms can go beyond simply providing recommendations and assessments to conducting deep analyses of students' learning patterns, allowing for highly personalized learning experiences (Zhai et al., 2021; Zhai et al., 2020a, b). However, the effectiveness of these personalized learning systems depends heavily on the quality and representativeness of the data they are trained on, which can sometimes introduce biases and perpetuate existing inequities.

In addition, students can also benefit from immediate feedback and adaptive learning pathways, ensuring that they are able to address any misconceptions or gaps in their understanding of scientific phenomena (Mavroudi et al., 2018). AI can also help science educators track and monitor students' progress more effectively, allowing for targeted interventions and support where necessary. Moreover, the use of AI can enable the development of interactive and immersive learning environments, making science education more engaging and accessible to students with diverse learning styles and needs. As AI continues to advance, the potential for its integration into science learning is likely to grow, presenting exciting opportunities to transform and elevate the science education experiences for students at all levels.

Anticipated Benefits of AI Implementation in Science Teaching

Artificial intelligence (AI) has numerous benefits in science education, profoundly affecting teaching and learning for science subjects. AI programs can study how students learn and change the material to fit each student's needs, skills, and the way they learn. This way of creating educational material helps students learn better and faster. It lets them go at their own speed and in a way that matches how they like to learn (Zawacki-Richter et al., 2019). Also, AI-powered data analysis can help science teachers understand how well their students are doing in specific scientific subjects and where they might need extra help.

Another significant advantage is the improvement of exploratory learning through virtual labs and reenactments. AI-powered instruments have the potential to recreate complex logical tests, which may be illogical or hazardous to conduct in a conventional classroom setting. These virtual situations offer hands-on learning encounters and permit understudies to try distinctive scenarios, improving their understanding of scientific concepts (Ibáñez et al., 2018). This approach was not as supportive of extending understudy engagement but too valuable to democratize access to high-quality science instruction. AI devices can interface understudies and teachers over diverse geographies, empowering the trade of logical thoughts and cultivating a worldwide point of view on logical issues. This interconnecting also permits integrating differing datasets into the educational modules, uncovering understudies to real-world logical challenges and datasets (Holmes et al., 2023).

Ethical Considerations of AI Integration in Education

Even though artificial intelligence (AI) in education has bright futures, important ethical issues need to be resolved when integrating AI in the classroom. Many researchers have stressed considering the ethical implications and the need for character education in the era of AI (Burton et al., 2017; Cathrin & Wikandaru, 2023). The lack of critical reflection on the pedagogical and ethical implications and the risks of implementing AI applications in

higher education underscores the need for a comprehensive ethical framework (Bozkurt et al., 2021).

Additionally, the integration of AI into educational settings presents new ethical obligations for teachers, necessitating a revaluation of ethical frameworks and responsibilities (Adams et al., 2022). Moreover, the introduction of ethics courses in academic training and capacity building of AI development actors can facilitate the integration of ethical values and the development of responsible AI (Kiemde & Kora, 2022). The ethical implications of AI in education extend beyond technical considerations to encompass broader societal impacts, such as privacy protection and social justice (Hermansyah et al., 2023).

Educators and students must understand, evaluate, and familiarize themselves with the uses of generative AI tools and consider their potential impacts on academic integrity. This involves recognizing when and how AI is used, assessing the reliability and validity of AI-generated outputs, and understanding the ethical and social implications of AI applications (Akgun & Greenhow, 2021). Moreover, the application of AI in education brings questions about educational equity and access to the fore. Systemic biases in AI algorithms and data can perpetuate inequities, making it crucial to address these biases effectively (Adams et al., 2022).

Purpose of the Study

The primary goal of the present study is to highlight the potential benefits as well as any disparities that might result from the widespread use of AI in science subjects. These discrepancies could be related to gaps in infrastructure, preparedness in the region, or accessibility. Eventually, the present investigation aims to furnish all relevant stakeholders, i.e., educators, learners, policymakers, and curriculum designers—with a deeper understanding of the current interaction between AI and science education.

The following research questions are addressed in this systematic review:

(1) Impact on Learning Outcomes: How do AI tools impact student learning outcomes and engagement in science education?
(2) Contexts of AI Adoption: What are the potential disparities in the uptake of AI tools within science education, considering differences among countries, educational levels, and subject areas?
(3) Student and Teacher Perceptions: What are the perceptions and attitudes of students and educators towards the use of AI tools in science education?
(4) Pedagogical Challenges: What are the identified challenges associated with using AI in science education?

Methodology

The authors worked diligently to explore how artificial intelligence contributes to science education thoroughly. We followed a structured process suggested by the widely used review methodology called Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Page et al., 2021). The review approach of the current study comprised various

stages – defining the study's purpose along with specific research questions, formulating a protocol, an extensive literature search, a systematic screening process, extracting pertinent data, and synthesizing the findings. The sections below specifically mention how each of these steps was carried out for this study.

Search Strategy

We accessed a range of prominent digital repositories and databases to search the relevant literature. Particularly, IEEE Xplore, Springer, Tylor and Francis, ERIC (U.S. Dept. of Education), Science Direct, and Wiley were targeted to search the relevant literature. We also used Google Scholar and Google to make sure that we didn't miss any important information. We used advanced search features to limit our search results to papers published between 2014 and 2023, ensuring that our search was focused and up to date (Piasecki et al., 2018).

We utilized a smart search strategy along with a range of search terms and operators to accomplish this. Our search strategy used a combination of key terms such as "artificial intelligence", "AI", "generative AI", "ChatGPT," "machine learning", "robotics", "intelligent system," and "expert system" paired with descriptors like "science education," "science learning," or simply "science". These combinations, along with their possible variations, were systematically applied to search within the papers' titles, keywords, and abstracts. This search strategy was created with the aim to identify and consider a broad range of empirical work relating to the use of artificial intelligence in the teaching and learning of science.

Eligibility Criteria

Describing clear eligibility (inclusion and exclusion) criteria allows for setting boundaries for a systematic literature review. These criteria were aimed at creating a structured framework that facilitates the inclusion of studies meeting essential prerequisites while excluding those that don't align with our research objectives. The inclusion criteria are as follows:

- The paper must have employed empirical methods, such as quantitative, qualitative, or/and mixed methods, warranting a rigorous data collection and analysis approach.
- The paper should have conducted research in an educational setting, encircling primary, middle, secondary, or higher education, emphasizing the applicability of the findings in educational environments.
- A pivotal criterion necessitates the use of artificial intelligence in the study. This AI practice should have been applied to the teaching-learning process, and empirical data collected and integrated into the study.
- Studies should be related to a science-related content area, spanning courses like chemistry, physics, biology, engineering, health sciences, or other related disciplines, ensuring applicability to the research topic.
- The timeframe specified for publication years, from 2014 to 2023, targets to capture relevant studies within the past decade, ensuring the examination of recent developments in AI-based learning.

Our exclusion criteria were as follows:

- Excluding studies that are not empirical in nature, such as theoretical papers, reviews, editorials, or opinion pieces, to maintain the focus on empirical research.
- Studies written in languages other than English.
- Studies that did not explicitly mention the AI use within a learning context.
- Excluding studies that are solely available in abstract form and lack full-length publications.

The Screening Process

In the months of November and December 2023, we went on a thorough hunt for the required information. We started by searching through loads of databases and found 5,121 articles. After getting rid of duplicates, checking publication dates and titles, and looking at abstracts to see if they met the eligibility criteria for the present study, we ended up with 128 articles. From there, we excluded 41 studies because they didn't really dive into science education. That left us with 87 articles that we pored over super carefully. We made sure they fit our criteria and answered our research questions before diving into them. From the pool of these 61 articles, ten (13) studies were identified as lacking clear empirical evidence regarding the use of artificial intelligence and were subsequently excluded. This process resulted in a final dataset of 74 articles that were included in the systematic review. See Table 1 for the list of studies included in our review. Figure 1 demonstrates a quick preview of the search strategy and the screening process.

Coding and Analysis

We used a mix of qualitative and quantitative content analysis techniques to synthesize the findings of the empirical papers. To ensure inter-rater reliability in relation to the quality of article coding procedures, a small random sample consisting of 20 selected articles was independently coded by multiple raters. The calculated reliability level exceeded 92%, signifying a high degree of agreement across coding categories. We conducted a comprehensive examination of the studies from various perspectives. Firstly, we analyzed the characteristics of the data set, including the country where the studies were conducted, the journal name, the content area, and the educational level.

Findings

In this comprehensive review of the literature, we carefully evaluated seventy-four (74) empirical studies that deal with the incorporation of AI into science education. Numerous research approaches, such as mixed, qualitative, and quantitative approaches, were used in these studies. Examining the publication dates of the included papers revealed that they were dispersed over the review study's 10-year focal period (2014 to 2023). The year 2023, with twenty-seven (27) papers, led the way, demonstrating researchers' strong interest in the most recent research on the application of artificial intelligence in science education. This was followed by ten (10) studies in 2022, eight (08) studies in 2021, and nine (09) studies in year 2020. For more information on the year-wise publication, see Fig. 2.

Table 1 Articles included in the systematic review

S#	Author(s)	Year	Country
1	Aldabe & Maritxalar	2014	Spain
2	Blikstein et al.	2014	USA
3	Hooshya et al.	2015	Malysia
4	Figueiredo et al.	2015	Portugal
5	Kinnebrew et al.	2016	USA
6	Hooshya et al.	2017	Malysia
7	Ledesma & García	2017	Mexico
8	Klebanov et al.	2017	USA
9	Avelino et al.	2017	USA
10	Gonzalez et al.	2017	USA
11	Jaipal-Jamani & Angeli	2018	Canada
12	Balakrishnan	2018	Malysia
13	García-Gorrostieta et al.	2018	Mexico
14	Anderson et al.	2018	USA
15	Hossain et al.	2018	USA
16	Mao et al.	2018	USA
17	Azcona et al.	2019	Ireland
18	Buenaño-Fernández et al.	2019	Spain
19	Lee et al.	2019	USA
20	Gerard et al.	2019	USA
21	Deo et al.	2020	Australia
22	Mirchi et al.	2020	Canada
23	Nesra et al.	2020	China
24	Hellings & Haelermans	2020	Netherland
25	Forsström & Afdal	2020	Norway
26	Mehboob et al.	2020	Pakistan
27	Koć-Januchta et al.	2020	Sweden
28	Yannier et al.	2020	USA
29	VanLehn et al.	2020	USA
30	Çınar et al.	2021	Turkey
31	Lamb et al.	2021	USA
32	Nawaz et al.	2021	Australia
33	Ferrarelli & Iocchi	2021	Italy
34	Topal et al.	2021	Turkey
35	Maestrales et al.	2021	USA
36	Shute et al.	2021	USA
37	Bertolini et al.	2021	USA
38	Jiao et al.	2022	China
39	AlKanaan	2022	Saudi Arabia
40	Lee et al.	2022	Taiwan
41	Su	2022	Taiwan
42	Celik & Ergin	2022	Turkey
43	Antonenko & Abramowitz	2022	USA
44	Zhai et al.	2022	USA
45	Bosman et al.	2022	USA
46	Alqathani et al.	2022	USA
47	Schulz et al.	2022	Germany
48	Cooper	2023	Australia

Table 1 (continued)

S#	Author(s)	Year	Country
49	Elkhodr et al.	2023	Australia
50	Rodrigues et al.	2023	Brazil
51	Kochmar et al.	2023	Canada
52	Bitzenbauer	2023	Germany
53	Dahlkempe et al.	2023	Germany
54	Kieser et al.,	2023	Germany
55	Küchemann et al.	2023	Germany
56	Wulff et al.	2023	Germany
57	Nasution	2023	Indonesia
58	Ariely et al.	2023	Israel
59	Daher et al.	2023	Israel
60	Park & Kwon	2023	Korea
61	Boels et al.	2023	Netherland
62	Nja et al.	2023	Nigeria
63	Syed et al.	2023	Saudi Arabia
64	Weeger & Idestam-Almquist	2023	Sweden
65	Körpeoğlu & Yildiz	2023	Turkey
66	Darayseh	2023	UAE
67	Alneyadi & Wardat	2023	UAE
68	Darayseh	2023	UAE
69	Ottenbreit-Leftwich et al.	2023	USA
70	Leitner et al.	2023	USA
71	Cochran et al.	2023	USA
72	Chen et al.	2023	USA
73	Shao et al.	2023	USA
74	Nguyen et al.	2023	Vietnam

The review process of the present study involved the consolidation of findings pertaining to four distinct research questions, each of which is presented separately in the following sections.

RQ1: Impact on Outcomes Comparison

The first research question of the current study specifically addressed the primary intention of this systematic research i.e., analyzing the reported impact of AI-enhanced learning on students' learning outcomes in science education. The empirical papers reviewed showed that artificial intelligence has been used within science education for a variety of purposes, such as engaging students in the learning process with a strong sense of motivation and interest (Balakrishnan, 2018), generating tests of science subjects (Aldabe & Maritxalar, 2014; Nasution & Education, 2023), scoring and providing personalized feedback on students' assignments (Azcona et al., 2019; Maestrales et al., 2021; Mirchi et al., 2020), and predicting student performance (Blikstein et al., 2014; Buenaño-Fernández et al., 2019; Jiao et al., 2022a, b).

AI-based tools were found to have a positive influence on student' learning outcomes in science-related courses. The experimental group that was exposed to AI integration in their learning environments exhibited significantly higher scores in their academic tests com-

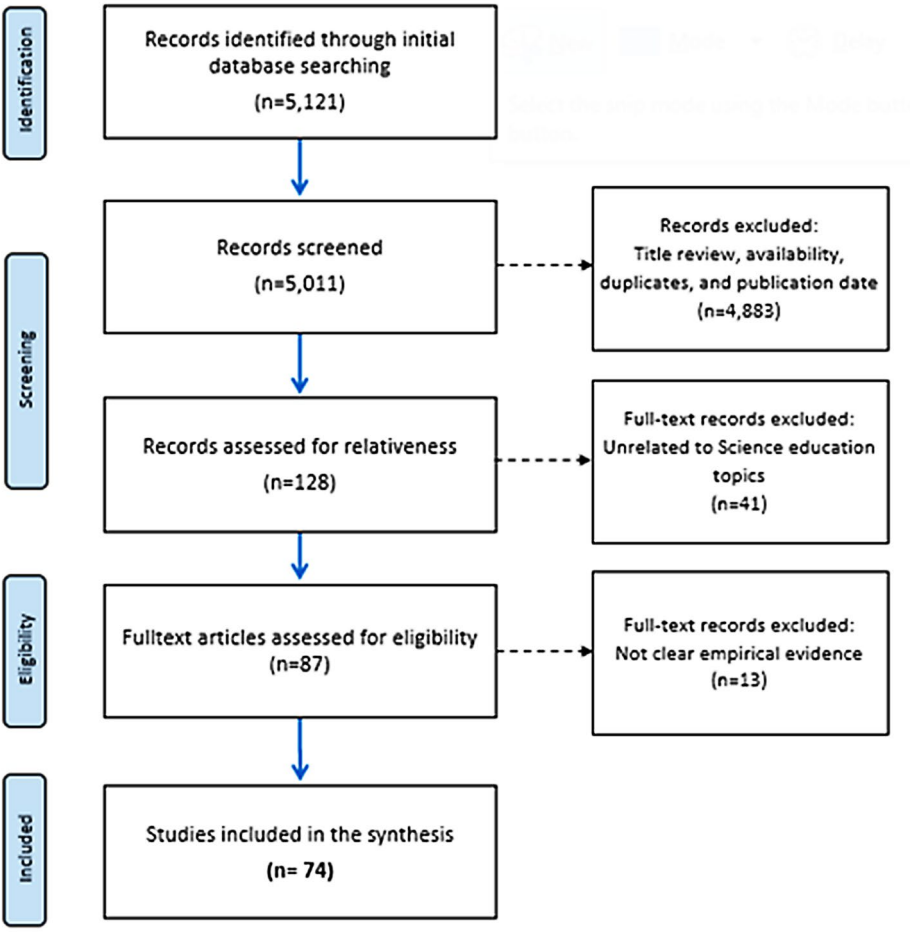


Fig. 1 PRISMA review process

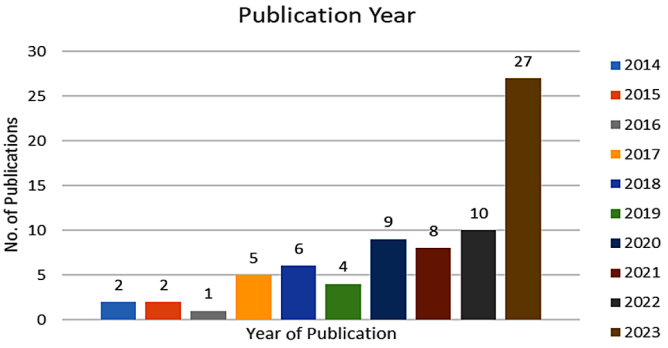


Fig. 2 Year-wise publications

pared to the control group who experienced traditional learning environments (Alneyadi & Wardat, 2023; Koć-Januchta et al., 2020). Ledesma and García (2017) and Lamb et al. (2021) highlighted AI's capacity to identify complex concepts and enhance problem-solving skills significantly in subjects (Lamb et al., 2021; Ledesma & García, 2017). Ferrarelli and Iocchi (2021), Cochran et al. (2023), and Figueiredo and Paixão (2015) showcased how AI is helpful in fostering improved subject understanding and heightened motivation among students, particularly in physics and chemistry (Ferrarelli & Iocchi, 2021; Figueiredo et al., 2016).

Lee et al. (2022) argue that AI-based tools such as chatbots can help students become cognitively more active in the learning process(Lee et al., 2022). Likewise, Azcona (2019) suggests that personalized learning facilitated by AI can help reduce the gap between lower- and higher-performing students. Moreover, AI-powered education can empower students to predict their learning outcomes and strategically regulate their learning behavior (Buenaño-Fernández et al., 2019).

The effectiveness of different AI models varied across studies. Nguyen et al. (2023) highlighted the performance disparities among AI models like Google Bard, ChatGPT, and Bing Chat in addressing biology problems for Vietnamese students (Nguyen et al., 2023). While chatbots positively influenced online learning experiences, their impact on academic achievement remained variable (Almasri, 2022a; Deveci Topal et al., 2021). In essence, these findings underscore the potential of AI to augment science education by enhancing student understanding, motivation, and engagement. However, they also underscore the importance of addressing challenges related to AI's adaptability to subject matter and context and the need for continued exploration into AI's comparative impact on academic achievement vis-à-vis traditional teaching methods in science education. Daher et al. (2023) pointed to AI's limitations in comprehending specific subject matter, which could impact its effectiveness in aiding student learning. Cooper (2023) emphasized the need for educators to critically evaluate and adapt AI-generated resources to suit diverse teaching contexts.

RQ2: Contexts of AI Adoption

In our second research question, we aimed to explore the potential disparities in the uptake of AI tools within science education, considering differences among countries, educational levels, and subject areas. The results disclosed that artificial intelligence has been incorporated in a variety of subject areas within science education, including physical and natural sciences. The studies reviewed were highly dominated by investigations that did not specify any particular domain of science ($n=15$, 20.30%), but they preferred to use "Science" as the subject area in their papers. Next in line, was the subject of Physics with the second-highest number of papers ($n=10$, 13.50%). The list was continued by Biology and Programming with nine ($n=9$, 12.16%) and eight ($n=8$, 10.81%) papers, respectively. The subjects of Mathematics and Engineering occupied about 16% (with 06 papers each) of the total papers. Out of 74 studies, only five (05) studies were conducted to investigate the use of AI for AI education. The subjects of Computers/technology were focused on in four papers. Lastly, only one paper was centered around the use of artificial intelligence in Statistics and Earth Science. Figure 3 provides a summary of the content areas that were the focus of the papers included in our review.

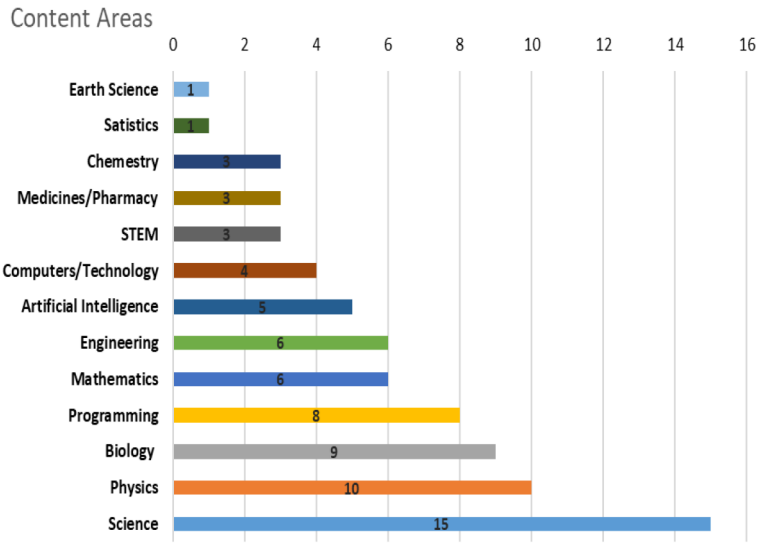


Fig. 3 Studies distributed around subject areas within science education

While examining the various educational levels that benefited from the integration of artificial intelligence in some manner, we found that nearly half of the studies ($n=35$, 47%) belonged to undergraduate level, followed by high schools ($n=15$, 20%) and middle schools ($n=7$, 10%) respectively. Out of the total 74 papers, about 8% of the studies ($n=6$) were conducted in secondary school contexts. Likewise, 8% of the studies involved multiple levels of educational settings. In contrast, three of the studies (about 4%) were conducted in elementary school. Only 2% of the papers belonged to the college level, and only one study was conducted at the postgraduate and college levels. Figure 4 provides a quick distribution of the students in various educational contexts.

Similarly, country-wise categorization of the papers exposed that about 38% of the studies ($n=25$) were conducted in the context of the United States. Germany ranked second in the list with six studies (8%). This was followed by four studies (5.4%) carried out in Turkey and Australia. UAE and Malaysia followed in the race, each with three papers. Eight countries, including Sweden, China, Mexico, Saudi Arabia, Spain, the Netherlands, Israel, and Taiwan, contributed about 21.6% of the total papers, each with two studies. The rest of the papers ($n=10$, 13.51%) were written in the context of 10 different countries across the globe (see Fig. 5 for details).

RQ3: Student and Teacher Perceptions

With our third research question, we attempted to explore science teachers' and students' perceptions regarding the integration of AI. The studies revealed multifaceted perspectives on the integration of AI in science education among both students and teachers. The effectiveness of AI tools in augmenting learning experiences garnered students' attention. Students showcased increased engagement and improved subject understanding through AI-based interventions, indicating positive perceptions of AI's efficacy in enhancing learning

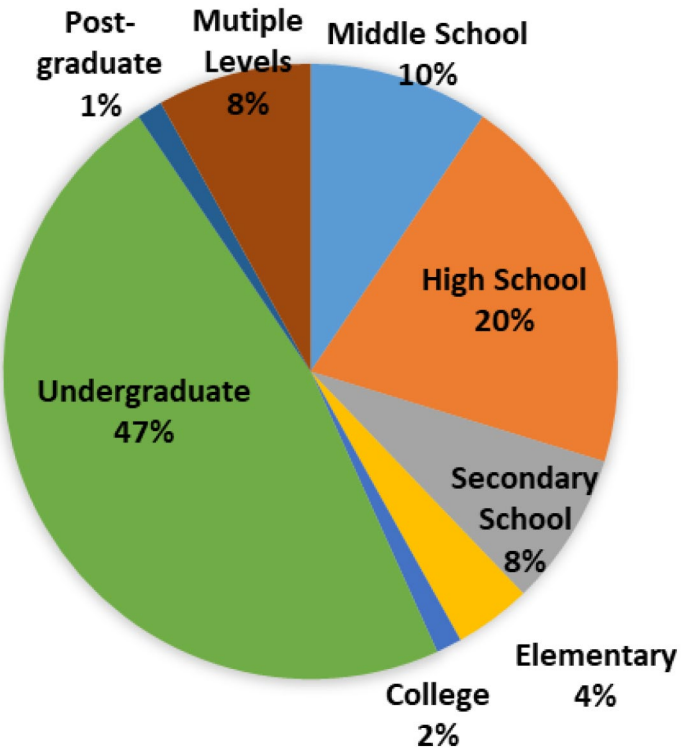


Fig. 4 Studies distributed across various educational levels

outcomes (Ferrarelli & Iocchi, 2021; Ledesma & García, 2017). For example, Bitzenbauer (2023) found that ChatGPT's use in Physics classrooms favorably influenced students' perceptions in Germany. Avelino et al. (2017) echoed this sentiment for undergraduate students in the United States.

Students reported their increased interest in science courses when AI was integrated into the learning environments. Students particularly admired the AI's power to provide prediction and personalized feedback (Azcona et al., 2019). According to Elkhodr et al. (2023), science students perceive AI-based tools as useful and enjoyable learning resources, while most students showed a willingness to use them in the future.

Our analysis suggests that science teachers hold a high level of acceptance and positive attitudes toward AI's utilization in the classroom. Teachers welcome its use with positive correlations to self-efficacy, ease of use, and behavioral intentions (Al Darayseh, 2023). They perceive this technology as the need of the hour to boost student engagement (Almasri, 2022b; Nersa, 2020). Empirical papers included in the current study exposed fluctuating degrees of comfort and adaptability among educators and students in incorporating AI into their teaching and learning processes. Al Darayseh (2023) noted that science teachers exhibited favorable attitudes toward AI's integration, possibly due to the perceived reduced effort in its utilization and their confidence in their essential skills to incorporate AI effectively.

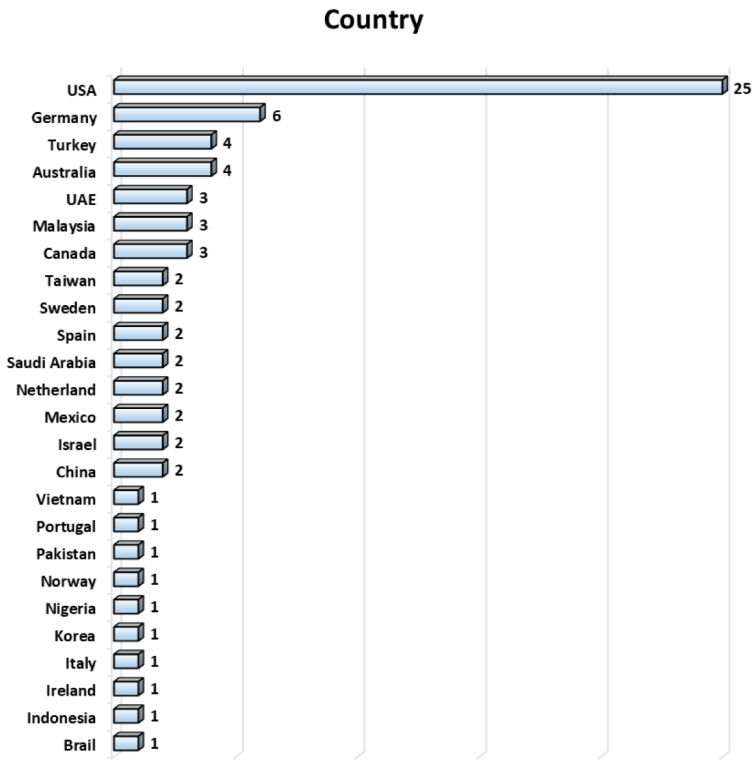


Fig. 5 Country of research context

There are several factors that influence teachers' intentions and behavior regarding the use of AI, including self-esteem, expected benefits, ease of utilization, and their overall attitude toward AI applications. Teachers' favorable disposition towards AI use is also due to their perception of reduced effort in its utilization.(Nja et al., 2023). Overall, teachers consider AI tools like ChatGPT to be helpful in designing science units, rubrics, and quizzes (Cooper, 2023). Yet, challenges associated with AI integration could influence students' and teachers' perceptions of AI's reliability and accuracy in supporting educational goals, posing potential barriers to widespread acceptance and utilization.

RQ4: Pedagogical Challenges

Our analysis uncovered several challenges associated with the integration of AI in terms of complexities and limitations of its use within this particular domain of the education system. One prevalent challenge revolved around AI's capability to comprehend and effectively address specific subject matter. Daher et al. (2023) highlighted instances where AI, like ChatGPT, encountered difficulties in understanding complex concepts in chemistry. They argue that the information provided by AI tools such as ChatGPT is limited because it depends on the data it was taught with. It might not have access to the latest or most complete knowledge in a particular domain.

Adaptability and contextual relevance emerged as significant concerns regarding the use of AI within science teaching. Cooper (2023) stressed that teachers critically evaluate AI-based resources and adapt them to their teaching contexts. He suggested that a one-size-fits-all approach might not suffice in accommodating the intricacies of varied educational environments. Another challenge pertained to the effectiveness and performance variability of different AI models. Nguyen et al. (2023) showcased the varying performance levels of different AI models, indicating disparities in their ability to address specific subject-related challenges. This variability in performance, as seen in different studies, implies the need for thorough evaluation and selection of appropriate AI tools tailored to the needs of specific subject areas. Furthermore, ethical considerations and limitations in AI's current capabilities were notable concerns. Kieser et al. (2023) raised ethical issues regarding students using AI to fabricate data for class assignments. Addressing these challenges requires a nuanced approach that acknowledges the potential and constraints of AI while striving to optimize its role in enhancing science education effectively.

Discussions

The primary objective of this review was to investigate the interaction between artificial intelligence and science education. Our study uncovered a diverse landscape of AI usage within science education. Our results suggested that integrating AI tools in science education consistently improves students' academic performance. This was evident in higher test scores and a better understanding of complex concepts compared to those in traditional learning environments (Alneyadi & Wardat, 2023; Koć-Januchta et al., 2020; Siddaway et al., 2019).

Literature suggests that integrating artificial intelligence into the teaching-learning process facilitates understanding complex scientific topics (Lamb et al., 2021; Ledesma & García, 2017). It also helps develop problem-solving skills considerably, leading to a better understanding of subjects, particularly in fields like physics and chemistry. Furthermore, it was revealed that science teachers use AI-driven tools to engage students effectively and foster their motivation and interest in science-related subjects (Balakrishnan, 2018). Personalized learning through AI tools helps bridge performance gaps between lower and higher-performing students (Azcona et al., 2019), contributing to a more equitable learning environment. AI-generated personalized feedback also contributed to students' increased engagement in the learning process (Azcona et al., 2019; Maestrales et al., 2021; Mirchi et al., 2020).

The current systematic review suggests that the distribution of studies within various subject areas in science education showcases a dominant focus on science in general, followed by physics, biology, programming, and other specific science subjects. Some specific domains, like earth science and statistics, received comparatively the least attention in the reviewed literature.

The distribution of research papers across countries demonstrates certain disparities. The United States had a significantly higher number of studies compared to other nations. Germany ranked second on the list. Turkey and Australia followed, while UAE, Malaysia, and Canada contributed with a moderate number of studies. Several countries had minimal representation, with a diverse spread across multiple nations. Concentration of studies

in certain countries like the United States and Germany might suggest varying levels of research infrastructure or prioritization of AI in education compared to other nations with fewer studies. This could potentially lead to disparities in the implementation and impact of AI tools in science education among different regions globally.

Our analysis found that students exhibit increased engagement and interest in science courses when AI tools are integrated into learning environments. This heightened interest is attributed to AI's ability to provide predictions and personalized feedback (Jiao et al., 2022b), making learning more engaging and enjoyable (Hewapathirana & Almasri, 2022). Students perceive AI-based tools as useful and beneficial for their learning experiences. They acknowledge AI's effectiveness in improving subject understanding and express a willingness to continue using such tools in the future (Elkhodr et al., 2023).

Similar to students, science teachers also demonstrate positive attitudes and acceptance of AI tools in the classroom, correlating with perceived benefits in student engagement and their own teaching efficacy. Teachers view AI integration as a means to enhance student engagement, with some perceiving it as a way to reduce effort while teaching, leading to increased confidence in utilizing AI effectively (Al Darayseh, 2023). Specifically, teachers perceive ChatGPT a valuable resource for designing science units, rubrics, quizzes, and teaching aids, offering convenience and potential enhancement to their teaching methodologies.

While AI showed promise in improving learning outcomes, there are challenges related to its adaptability to subject matter and context. Some studies pointed out limitations in comprehending specific subjects, potentially impacting the effectiveness of AI in aiding student learning. Previous research suggests that AI tools like ChatGPT face difficulties in comprehending and addressing complex concepts in specific subject areas, as seen in instances within chemistry (Daher et al., 2023). The dependency on the data it was trained with limits its access to the latest or most comprehensive knowledge in particular domains. A uniform approach might not adequately cater to the complexities and nuances of varied educational environments, emphasizing the need for adaptable solutions (Cooper, 2023). Addressing these challenges requires a balanced approach that acknowledges AI's potential and constraints in science education. Thus, teachers are advised to critically evaluate AI-generated resources and tailor them to diverse teaching contexts.

Our research provides important implications for teacher preparation and in-service professional development regarding AI in our society and implementing AI tools and processes in K-12 education (Antonenko & Abramowitz, 2023). As a whole, integrating artificial intelligence positively enhances the process and outcome of science education. However, there are certain limitations and challenges associated with its use. Providing training and support to educators to effectively utilize AI tools can enhance their confidence and capabilities in integrating these technologies into teaching practices. Moreover, establishing clear ethical guidelines and frameworks for the responsible use of AI in education can mitigate the risk of misuse and ensure ethical practices among students and educators.

Limitations

Some of the inherent limitations of this research review are discussed in this section. First, just like with other reviews, the search terms and strategies determine which research papers are included. Although a thorough and methodologically rigorous search was the goal, using

different search terms might have turned up more articles that could have been included in the review. Furthermore, a few particular research databases were searched in order to find pertinent empirical literature for inclusion in this research review. An alternative methodological strategy would have involved restricting the search for research to a predetermined list of scholarly, peer-reviewed journals. A smaller sample of literature for inclusion may have occurred due to this strategy. However, greater control over validity, reliability, and credibility during the search and inclusion processes was sought to the best level. Lastly, we may have missed some grey literature, such as dissertations and conference proceedings, that was not indexed in the databases/repositories that we used.

Conclusion

This systematic review examined the impact, perceptions, and challenges associated with the integration of Artificial Intelligence (AI) in the teaching and learning of science. Our analysis uncovered a landscape rich in prospective benefits and challenges. The usage of AI in science education steadily established positive impacts on student learning outcomes. It encourages participation in the educational process, enhances comprehension of the subject, and boosts motivation in the students. Both students and teachers showed positive views of AI's effectiveness and ease of use. Both acknowledged its potential to boost learning experiences. Nevertheless, issues arose from AI's limited ability to understand particular subject matter, its inability to adjust to various educational contexts, and the variation in performance between various AI models. Ethical considerations regarding responsible use also appeared to be an important concern. Addressing these challenges demands a careful approach that considers thorough evaluation and adaptation to diverse contexts. Educators and policymakers should navigate these complexities to join the potential of AI in science education while ensuring ethical practices and maximizing its impact on students' learning journey worldwide.

Acknowledgements None.

Funding None.

Declarations

Consent to Publish I consent to publish identifiable details within the text in the Research in Science Education Journal.

Conflict of Interest The authors reported no potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adams, C., Pente, P., Lemermeyer, G., Turville, J., & Rockwell, G. (2022). Artificial intelligence and teachers' new ethical obligations. *The International Review of Information Ethics*, *31*(1). https://doi.org/10.29173/irie483.

Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2*, 431–440. https://doi.org/10.1007/s43681-021-00096-7.

Aldabe, I., & Maritxalar, M. (2014). Semantic similarity measures for the generation of science tests in basque. *IEEE Transactions on Learning Technologies*, *7*(4), 375–387. https://doi.org/10.1109/TLT.2014.2355831.

Al Darayseh, A. (2023). Acceptance of artificial intelligence in teaching science: Science teachers' perspective. *Artificial Intelligence*, *4*, 100132. https://doi.org/10.1016/j.caeai.2023.100132.

Alharbi, S. M., Elfeky, A. I., & Ahmed, E. S. (2022). The effect of e-collaborative learning environment on development of critical thinking and higher order thinking skills. *Journal of Positive School Psychology*, *6*(6), 6848–6854.

Almasri, F. (2021). *Collaborative learning in science education: effects of student's gender attitudes and achievement in science education* University of Warwick, Retrieved from http://webcat.warwick.ac.uk/record=b3853208.

Almasri, F. (2022a). The impact of e-learning, gender-groupings and learning pedagogies in biology undergraduate female and male students' attitudes and achievement. *Education and Information Technologies*, *27*, 8329–8380. https://doi.org/10.1007/s10639-022-10967-z.

Almasri, F. (2022b). Simulations to teach science subjects: Connections among students' engagement, self-confidence, satisfaction, and learning styles. *Education and Information Technologies*, *27*(5), 7161–7181. https://doi.org/10.1007/s10639-022-10940-w.

Almasri, F., Hewapathirana, G. I., Alhashem, F., Daniel, C. E., & Lee, N. (2022). The effect of gender composition and pedagogical approach on major and non-major undergraduates biology students' achievement. *Interactive Learning Environments*, 1–33. https://doi.org/10.1080/10494820.2022.2066138.

Almasri, F., Hewapathirana, G. I., Ghaddar, F., Lee, N., & Ibrahim, B. (2021). Measuring attitudes towards biology major and non-major: Effect of students' gender, group composition, and learning environment. *Plos One*, *16*(5), e0251453. https://doi.org/10.1371/journal.pone.0251453.

Alneyadi, S., & Wardat, Y. J. C. E. T. (2023). ChatGPT: Revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools. *Contemporary Educational Technology*, *15*(4), ep448. https://doi.org/10.30935/cedtech/13417.

Antonenko, P., & Abramowitz, B. (2023). In-service teachers'(mis) conceptions of artificial intelligence in K-12 science education. *Journal of Research on Technology in Education*, *55*(1), 64–78. https://doi.org/10.1080/15391523.2022.2119450.

Avelino, F., Wittmayer, J. M., Kemp, R., & Haxeltine, A. (2017). Game-changers and transformative social innovation. *Ecology and Society, 22*(4). https://www.jstor.org/stable/26798984

Azcona, D., Hsiao, I. H., & Smeaton, A. F. (2019). Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints. *User Modeling and User-Adapted Interaction*, *29*, 759–788. https://doi.org/10.1007/s11257-019-09234-7.

Balakrishnan, B. (2018). Motivating engineering students learning via monitoring in personalized learning environment with tagging system. *Computer Applications in Engineering Education*, *26*(3), 700–710. https://doi.org/10.1002/cae.21924.

Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *Journal of the Learning Sciences*, *23*(4), 561–599. https://doi.org/10.1080/10508406.2014.954750.

Bozkurt, A., Karadeniz, A., Baneres, D., Rodríguez, M. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability*, *13*(2), 800. https://doi.org/10.3390/su13020800.

Buenaño-Fernández, D., Gil, D., & Luján-Mora, S. (2019). Application of machine learning in predicting performance for computer engineering students: A case study. *Sustainability*, *11*(10), 2833. https://doi.org/10.3390/su11102833.

Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T. (2017). Ethical considerations in artificial intelligence courses. *AI Magazine*, *38*(2), 22–34. https://doi.org/10.1609/aimag.v38i2.2731.

Cathrin, S., & Wikandaru, R. (2023). The future of character education in the era of artificial intelligence. *Humanika Kajian Ilmiah Mata Kuliah Umum*, *23*(1). https://doi.org/10.21831/hum.v23i1.59741.

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, 100118. https://doi.org/10.1016/j.caeai.2022.100118.

Cochran, K., Cohn, C., Rouet, J. F., & Hastings, P. (2023, June). Improving automated evaluation of student text responses using gpt-3.5 for text data augmentation. In *International conference on artificial intelligence in education* (pp. 217–228). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36272-9_18

Cooper, G. (2023). Examining science education in chatgpt: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y.

Daher, W., Diab, H., & Rayan, A. (2023). Artificial intelligence generative tools and conceptual knowledge in problem solving in Chemistry. *Information*, *14*(7), 409. https://doi.org/10.3390/info14070409.

Dai, Y. (2023). Negotiation of epistemological understandings and teaching practices between primary teachers and scientists about artificial intelligence in professional development. *Research in Science Education*, *53*(3), 577–591. https://doi.org/10.1007/s11165-022-10072-8.

Deveci Topal, A., Eren, D., C., & Kolburan Geçer, A. (2021). Chatbot application in a 5th grade science course. *Education and Information Technologies*, *26*(5), 6241–6265. https://doi.org/10.1007/s10639-021-10627-8.

Elkhodr, M., Gide, E., Wu, R., & Darwish, O. (2023). ICT students' perceptions towards ChatGPT: An experimental reflective lab analysis. *STEM Education*, *3*(2), 70–88. https://doi.org/10.3934/steme.2023006.

Ferrarelli, P., & Iocchi, L. (2021). Learning newtonian physics through programming robot experiments. *Technology Knowledge and Learning*, *26*(4), 789–824.https://doi.org/10.1007/s10758-021-09508-3

Figueiredo, M., Esteves, L., Neves, J., & Vicente, H. (2016). A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation. *Chemistry Education Research and Practice*, *17*(1), 156–171. https://doi.org/10.1039/C5RP00144G.

Figueiredo, M., & Paixão, F. (2015). O Papel da Observação na Compreensão da Natureza do Conhecimento Científico: Uma Proposta Didática Baseada nas Reações de Combustão. In *Proceedings do Encontro Nacional de Educação em Ciências – ENEC 2015 - Ciência com Cultura* (pp. 205–209). Lisboa.

Gonzalez, A. J., Hollister, J. R., DeMara, R. F., Leigh, J., Lanman, B., Lee, S. Y., & Wilder, B. (2017). AI in informal science education: Bringing turing back to life to perform the turing test. *International Journal of Artificial Intelligence in Education*, *27*, 353–384. https://doi.org/10.1007/s40593-017-0144-1.

Grinnell, F. (2021). Scientific Inquiry, pluralism and complementarity. *Scientific Inquiry Pluralism and Complementarity*. https://doi.org/10.31235/osf.io/gejwv.

Hermansyah, M., Najib, A., Farida, A., Sacipto, R., & Rintyarna, B. S. (2023). Artificial intelligence and ethics: Building an artificial intelligence system that ensures privacy and social justice. *International Journal of Science and Society*, *5*(1), 154–168. https://doi.org/10.54783/ijsoc.v5i1.644.

Hewapathirana, G., & Almasri, F. (2022). Active learning compared with lecture-based pedagogies in gender and socio-cultural context-specific major and non-major biology classes. *Handbook of research on active Learning and Student Engagement in Higher Education* (pp. 293–319). IGI Global.

Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial intelligence in education*. Globethics.

Holme, T. A. (2021). *Considering the call to action for science education* (Vol. 98, pp. 2739–2740). ACS.

Hong, F., Dou, W., & Chen, S. (2022). *Research on the impact of artificial intelligence on government public service quality*. Paper presented at the 2022 2nd International Conference on Public Management and Intelligent Society (PMIS 2022).

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, *123*, 109–123. https://doi.org/10.1016/j.compedu.2018.05.002.

Irez, S. (2006a). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education*, *90*(6), 1113–1143. https://doi.org/10.1002/sce.20156.

Irez, S. J. S. E. (2006b). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *90*(6), 1113–1143.

Jiao, P., Ouyang, F., Zhang, Q., & Alavi, A. H. (2022a). Artificial intelligence-enabled prediction model of student academic performance in online engineering education. *Artificial Intelligence Review*, *55*(8), 6321–6344. https://doi.org/10.1007/s10462-022-10155-y.

Jiao, P., Ouyang, F., Zhang, Q., & Alavi, A. H. (2022b). Artificial intelligence-enabled prediction model of student academic performance in online engineering education. *Artificial Intelligence Review*, *55*(8), 6321–6344.

Kiemde, S. M. A., & Kora, A. D. (2022). Towards an ethics of AI in Africa: Rule of education. *AI and Ethics*, *2*(1), 35–40. https://doi.org/10.1007/s43681-021-00106-8.

Kieser, F., Wulff, P., Kuhn, J., & Küchemann, S. (2023). Educational data augmentation in physics education research using ChatGPT. *Physical Review Physics Education Research, 19*(2), 020150.

Koć-Januchta, M. M., Schönborn, K. J., Tibell, L. A., Chaudhri, V. K., & Heller, H. C. (2020). Engaging with biology by asking questions: Investigating students' interaction and learning with an artificial intelligence-enriched textbook. *Journal of Educational Computing Research*, *58*(6), 1190–1224.

Kok-Sing, T., & Nichols, K. (Eds.). (2024). Collection: Artificial intelligence in science education. *Research in Science Education*. https://link.springer.com/collections/dbihehajcd.

Kola, A. J. (2013). Importance of science education to national development and problems militating against its development. *American Journal of Educational Research*, *1*(7), 225–229. https://doi.org/10.12691/education-1-7-2.

Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education*, *85*(3), 291–310. https://doi.org/10.1002/sce.1011.

Lamb, R., Hand, B., & Kavner, A. (2021). Computational modeling of the effects of the science writing heuristic on student critical thinking in science using machine learning. *Journal of Science Education and Technology*, *30*, 283–297. https://doi.org/10.1007/s10956-020-09871-3.

Ledesma, E. F. R., & García, J. J. G. (2017). Selection of mathematical problems in accordance with student's learning style. *International Journal of Advanced Computer Science Applications*, *8*(3). https://doi.org/10.14569/IJACSA.2017.080316.

Lee, H. S., Pallant, A., Pryputniewicz, S., Lord, T., Mulholland, M., & Liu, O. L. (2019). Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty. *Science Education*, *103*(3), 590–622.

Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based cha bot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development*, *70*(5), 1843–1865. https://doi.org/10.1007/s11423-022-10142-8.

Liu, Y., & Pásztor, A. (2022). Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis. *Thinking Skills and Creativity*, *45*, 101069. https://doi.org/10.1016/j.tsc.2022.101069.

Lucci, S., Kopec, D., & Musa, S. M. (2022). *Artificial intelligence in the 21st century: Mercury learning and information*.

Maestrales, S., Zhai, X., Touitou, I., Baker, Q., Schneider, B., & Krajcik, J. (2021). Using machine learning to score multi-dimensional assessments of chemistry and physics. *Journal of Science Education and Technology*, *30*, 239–254. https://doi.org/10.1007/s10956-020-09895-9.

Mansour, N. (2009). Science-technology-society (STS) a new paradigm in science education. *Bulletin of Science Technology & Society*, *29*(4), 287–297. https://doi.org/10.1177/0270467609336307.

Mavroudi, A., Giannakos, M., & Krogstie, J. (2018). Supporting adaptive learning pathways through the use of learning analytics: Developments, challenges and future opportunities. *Interactive Learning Environments, 26*(2), 206–220. https://doi.org/10.1080/10494820.2017.1292531

Mirchi, N., Bissonnette, V., Yilmaz, R., Ledwos, N., Winkler-Schwartz, A., & Del Maestro, R. F. J. P. (2020). The virtual operative assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *Plos One*, *15*(2), e0229596. https://doi.org/10.1371/journal.pone.0229596.

Mogea, T. (2022). Students' critical thinking ability in English teaching and learning. *Jurnal Pendidikan Dan Sastra Inggris*, *2*(3), 157–171.

Nasution & Education. (2023). Using artificial intelligence to create biology multiple choice questions for higher education. *Agricultural and Environmental Education*, *2*(1).

Nersa, A. (2020). *The coverage of children Rights in Amhara television children's program*

Nguyen, P., Trương, H., Nguyen, P., Bruneau, P., Cao, L., & Wang, J. (2023). Evaluation of Google Bard on Vietnamese High School Biology Examination.

Nja, C. O., Idiege, K. J., Uwe, U. E., Meremikwu, A. N., Ekon, E. E., Erim, C. M., & Umalili, B. (2023). Adoption of artificial intelligence in science teaching: From the vantage point of the African science teachers. *Smart Learning Environments*, *10*(1), 42. https://doi.org/10.1186/s40561-023-00261-x.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. https://doi.org/10.1136/bmj.n71.

Piasecki, J., Waligora, M., & Dranseika, V. (2018). Google search as an additional source in systematic reviews. *Science and Engineering Ethics*, *24*, 809–810.

Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 1–13. https://doi.org/10.1186/s41039-017-0062-8.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*, 747–770. https://doi.org/10.1146/annurev-psych-010418-102803.

Tang, K. S., & Cooper, G. (2024). The role of materiality in an era of generative artificial intelligence. *Science & Education*, 1–16. https://doi.org/10.1007/s11191-024-00508-0.

Wahyono, I. D., Fadlika, I., Asfani, K., Putranto, H., Hammad, J., & Sunarti. (2019). New adaptive intelligence method for personalized adaptive laboratories. In *2019 International conference on electrical, electronics and information engineering (ICEEIE)* (pp. 196–200). https://doi.org/10.1109/ICEEIE47180.2019.8981477

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0.

Zhai, X., Haudek, C., Shi, K., Nehm, L. H., R., & Urban-Lurain, M. (2020a). From substitution to redefinition: A framework of machine learning-based science assessment. *Journal of Research in Science Teaching*, *57*(9), 1430–1459. https://doi.org/10.1002/tea.21658.

Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. *Journal of Research in Science Teaching*, *59*(10), 1765–1794. https://doi.org/10.1002/tea.21773.

Zhai, X., Shi, L., & Nehm, R. H. (2021). A Meta-analysis of machine learning-based Science assessments: Factors impacting machine-human score agreements. *Journal of Science Education and Technology*, *30*(3), 361–379. https://doi.org/10.1007/s10956-020-09875-z.

Zhai, X., Yin, Y., Pellegrino, J. W., Haudek, K. C., & Shi, L. (2020b). Applying machine learning in science assessment: A systematic review. *Studies in Science Education*, *56*(1), 111–151. https://doi.org/10.1080/03057267.2020.1735757.

Zulyusri, Z., Elfira, I., Lufri, L., & Santosa, T. A. (2023). Literature study: Utilization of the PjBL model in science education to improve creativity and critical thinking skills. *Jurnal Penelitian Pendidikan IPA*, *9*(1), 133–143. https://doi.org/10.29303/jppipa.v9i1.2555.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.