Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 20, Issue 1, June 2026, pp. 1-33. ISSN: 1307-6086

Review Article

# Artificial Intelligence (AI) Research Trends in Science Education: A Critical Review of Ethical Concerns

**Ulku Seher BUDAK[1], Kemal AKOĞLU[2]**

[1] Boğaziçi University, Faculty of Education, Türkiye, ulku.budak@std.bogazici.edu.tr, https://orcid.org/0000-0002-4047-9920

[2] Boğaziçi University, Faculty of Education, Türkiye, kemal.akoglu@bogazici.edu.tr, https://orcid.org/0000-0002-5688-1316

Received : 21.09.2025 Accepted : 21.01.2026

Doi: https://doi.org/10.17522/balikesirnef.1788147

---

*Abstract* – Artificial intelligence (AI) is the science and engineering of intelligent machines to achieve goals that improve humanity and advance the common good. This systematic literature review aimed to examine how AI in science education was addressed in 39 empirical and peer-reviewed studies in science education indexed in the Web of Science database from first mention until the end of 2024, and to address ethical concerns. Content analysis was used to identify trends in AI research in science education. This review revealed that research on AI in science education has increased significantly in recent years. USA and Korea were found to be the origin of most of the research. This study demonstrated that very few of the studies involved AI, such as ChatGPT, as a participant, and most of them involved humans as participants. On the other hand, most studies focused on K-12 students, and the fewest focused on university students and educators. The development of AI raises social and ethical concerns. Nearly half of the reviewed articles did not address AI-related ethical concerns. The most common ethical concerns addressed were bias, privacy, and accuracy. This study can provide important evidence-based support for future studies of AI in science education.

*Keywords*: Artificial intelligence, science education, systematic review, ethics.

---

Corresponding author: Ülkü Seher BUDAK, ulku.budak@std.bogazici.edu.tr

## Introduction

Artificial intelligence (AI) is a constantly evolving tool used to analyse data and recognize and classify patterns in data sets, such as audio and images (Ref Lu et al., 2018; Taani & Alabidi, 2025). AI has increasingly influenced various fields, including education (Lee et al., 2019; Zhai et al., 2021). It aims to develop educational technologies by transforming the education and training process (Hwang et al., 2020; Xu & Ouyang, 2022). Given its growing role in education, AI has also become particularly important in science teaching and learning (Aksit, 2025). AI can support teachers by providing students with feedback and help develop students' 21$^{st}$-century skills through personalized instruction (Erduran, 2023; Ramnarain et al., 2025). According to Kamalov et al. (2023), AI can serve as an effective instructional strategy by teaching strategies in adjusting the content and duration of instruction based on students' individual needs.

According to a recent bibliometric analysis, AI in science education is an emerging research topic (Budak & Yerdelen Damar, 2025). Several reviews have examined AI in education more broadly, for example, reviews covering 2010–2020 (Zhai et al., 2021), 2000–2020 (Ng et al., 2023), and studies focused on Generative AI from 2022–2024 (Yusuf et al., 2024). Additional reviews have considered specific areas such as AI chatbots in education (Labadze et al., 2023) and AI literacy (Casal-Otero et al., 2023; Ng et al., 2023). However, literature reviews specifically addressing AI in education remain limited (Heeg & Avraamidou, 2023). For example, Heeg and Avraamidou (2023) reviewed AI applications in science education between 2010 and 2021, while Jia et al. (2024) examined the impact of AI technologies on learners in primary and secondary schools. Despite these contributions, there is still a need to examine how AI is positioned in science education, especially at the primary and secondary levels (Jia et al., 2024). Recent research has also highlighted important social and ethical implications associated with the development and use of AI in educational contexts (e.g., Zhai et al., 2021). Therefore, it is essential to consider ethical aspects when examining the use of AI in science education (Jia et al., 2024). However, existing reviews, such as Jia et al. (2024), do not sufficiently address the positioning of AI within science education or critically consider ethical concerns, such as data privacy, algorithmic bias, misinformation, and issues related to student confidentiality, at these levels.

The significance of this study lies in contributing to the literature by documenting both how AI is being addressed in science education and the ethical issues associated with its use. Thus, the dual purpose of this review is to conduct a systematic review of empirical, peer-

reviewed studies of AI in science education indexed in the Web of Science (WoS) database from first mention through the end of 2024, and to address ethical concerns, in particular from a pragmatic perspective. Accordingly, this review study sought to answer the following research questions:

1) How did research studies address AI in science education research?
2) How did research studies address ethical concerns in AI for science education research?

## Literature Review

### AI in Education

The boundaries of "AI in education, AI education, and AI literacy" are unclear, and there is no consensus on a common definition of these terms (Ng et al., 2024). Even AI itself lacks a universally accepted definition. McCarthy (2007), considered the father of AI, first used the term AI in the 1950s and defined it as "the science and engineering of making intelligent machines that can achieve goals as humans do" (p. 2).

As the term AI has evolved tremendously since its first coinage, there is no consensus on its definition, as researchers define it according to their own needs and applications (Heeg & Avraamidou, 2023). In this review article, the researchers' stance aligns with the European Commission's High-Level Expert Group on Artificial Intelligence's definition of AI. "AI is not an end in itself, but rather a promising means to increase human flourishing, thereby enhancing individual and societal well-being and the common good, as well as bringing progress and innovation" (HLEG, 2019, p. 5). We have chosen this definition because we argue that AI can also enhance social welfare and the common good, from a pragmatic perspective, while taking into account ethical concerns and issues in its use.

As Dignum (2021) stated, if AI is to revolutionize education and bring about a digital transformation, it is necessary to be aware of the characteristics that distinguish it from other pedagogical approaches. Mittal et al. (2024) stated that AI applications encourage students to learn and increase their curiosity compared to traditional education. Therefore, integrating AI into education aims to make the learning process more meaningful (Limna et al., 2022). In addition to studies on topics as diverse as the particle nature of matter (e.g., Deveci Topal et al., 2021), argumentation (e.g., Zhai et al., 2023), and socio-scientific issues (e.g., Karahan, 2023), the use of AI in science education has been studied in terms of skills and attitudes, such as computational thinking skills (e.g., Huang & Qiao, 2024) and STEM attitudes (e.g.,

Göktepe Körpeoğlu & Göktepe Yıldız, 2024). Due to the increasing research interest in AI, there is a need to identify research trends, especially in K-12 education (Yusuf et al., 2024). Yusuf et al. (2024) stated that most of the articles they compiled as part of their research raised ethical issues. Therefore, the motivation for this review article is to identify general research trends in K-12 education and the ethical issues identified in the studies.

**AI-Related Ethical Concerns**

While AI contributes to society by advancing social development and human welfare, it raises many fundamental ethical issues (Huang et al., 2022; Yampolskiy, 2019). Therefore, addressing the ethical issues of AI should be a primary goal (Huang et al., 2022). Ethics is a complex and convoluted concept and has been defined in many ways (Siau & Wang, 2020). For example, according to Resnik (2011), ethics or morality is defined as the examination of actions in accordance with certain standards to determine their rightness or wrongness. Blanchard and Peale (2011) use the Socratic method for ethical review by asking three famous questions: Is it legal? Is it balanced? And, how will it make me feel about myself? Thus, it is unique in that it uses the Socratic method to evaluate different views in depth and assess whether behaviours or actions are ethical (Resnik, 2011).

AI ethics, or machine ethics, concerns evaluating machines with respect to ethical issues (Anderson & Anderson, 2006). AI ethics concerns the ethical appropriateness of AI attitudes and behaviours toward humans or other machines (Anderson & Anderson, 2007). Heilinger (2022) listed the main issues arising from the use of AI as the limits of AI's responsibility, privacy, bias, autonomy, social injustice, and discrimination. In a recent literature review on the integration of AI in schooling, Karan and Angadi (2024) identified privacy, autonomy, and bias risks, as well as risks to accuracy and functionality, and to social knowledge and skill-building, as highlighted by Heilinger (2022).

One of the biggest threats to AI use is the invasion of privacy, as individuals' information is collected without their consent (Stockman & Nottingham, 2022). Another threat is bias, which includes various biases such as gender and racism. Bias is dangerous because it can amplify social biases, and many AI biases have been reported (Huang et al., 2022). For example, in their review of algorithmic bias in education, Baker and Hawn (2022) emphasize that there are many “unknown unknowns” related to bias, and that certain groups of students are affected by it and may not even be aware of it. Another commonly cited risk associated with AI is the problem of inaccurate results due to faulty algorithms. Even with

adequate machine learning, it may produce incorrect or unclear definitions when the correct data is fed into the machine (Huang et al., 2022).

The accountability of AI is another concern. AI should explain its decisions to users in a reasonable way and account for the moral values and norms it considers when making those decisions (Dignum, 2021). The complexity of the algorithms in the inner workings of machine learning, which can also be described as a black box, makes it difficult for individuals to understand AI and raises questions about its transparency. Because transparency and trust are important for people to trust AI, transparency is considered an ethical concern (Siau & Wang, 2018, 2020).

As Heilinger (2022) points out, it is important to pay closer attention to the ethical issues surrounding AI to use it for the public good and increase social welfare. As the authors of this paper, we are in favour of using AI in a way that is beneficial to society, balancing pragmatic benefits with ethical concerns. In the information society we live in, the ethical stance should be considered collectively rather than individually. In other words, ethics should not be reduced to specific norms and principles, but should critically address the social and contextual dimension of ethics (Sabourin Laflamme & Bruneault, 2022). From a pragmatic perspective, ethics should involve a Socratic discussion of values and actions that takes social groups as a whole as the main context for action (Sabourin Laflamme & Bruneault, 2022). When faced with disagreements over concepts, values, or behaviours, they should not be seen as obstacles but rather as matters that can be agreed upon for the benefit of all (Singer, 2010).

## Method

### Research Design

A qualitative research method was used in this review study. Qualitative research aims to provide in-depth insight into a particular phenomenon of interest (Gay et al., 2012). This study used qualitative research methods to provide in-depth insights into how AI is positioned in science education research. Specifically, this study used content analysis, an approach to categorising text to reveal general trends and frequencies (Vaismoradi et al., 2013). Content analysis requires a rigorous process, and in new research areas that are not widely known, it is particularly well-suited to revealing common themes in texts (Green & Thorogood, 2004).

## Data Collection

The Web of Science Core Collection (WoS) database was purposefully selected as the data source for this systematic review to investigate how AI is positioned in science education research and how ethical concerns are addressed. The reason for selecting the WoS database for this study is that it is not only a catalogue but also provides more selective coverage of journals than other databases and offers highly developed metadata (Birkle et al., 2020; Singh et al., 2021). In this study, the data collection process was based on the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) framework (Moher et al., 2009). PRISMA is recommended for use in systematic reviews and meta-analyses to ensure a rigorous process. PRISMA is a flow diagram that assesses the eligibility of articles that will serve as sources for systematic reviews by defining a set of inclusion and exclusion criteria (Moher et al., 2010). In this review article, the included articles were determined in three stages following the PRISMA flow diagram: identification, screening, and eligibility. Figure 1 shows the selection process of articles on artificial intelligence (AI) in science education using the modified PRISMA flow diagram.

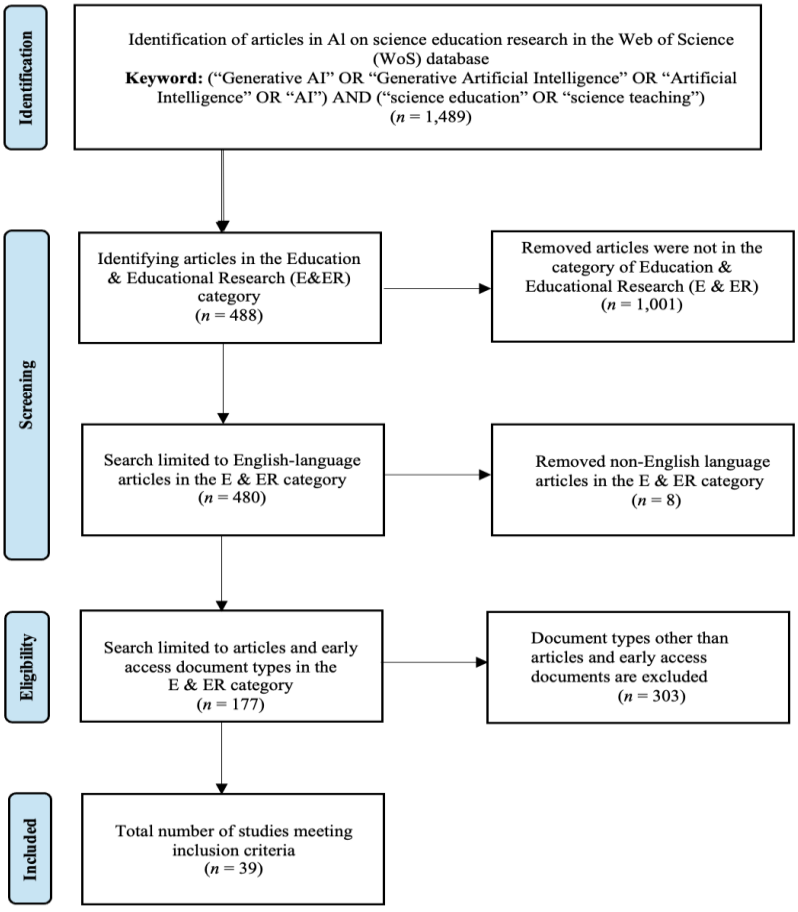



**Figure 1** Selection of Articles on AI in Science Education Published from the Date AI First Mention through the End of 2024 Using the Modified PRISMA Flow Diagram

### *Identification*

The researchers agreed to use the following keywords to investigate how AI is positioned in science education comprehensively.

("Generative AI" OR "Generative Artificial Intelligence" OR "Artificial Intelligence" OR "AI") AND ("science education" OR "science teaching")

Researchers searched the WoS database using these keywords in the 'all fields' category. The reason for searching keywords in 'all fields' categories is to document all research without any field restriction by using a single query. In addition, as the researchers' institute provides open access to the WoS database, the articles were not filtered for open access; rather, non-open-access articles were included in the review for inclusiveness. The database search was completed on 31 December 2024, with all studies published up to and including this date being eligible for inclusion. The rigorous data collection process is illustrated in Figure 1, step by step, to enable other researchers to apply it systematically. The identified keywords were searched in the WoS database across all fields, without time filtering, and 1,489 studies were returned.

### *Screening*

Of the 1,489 studies listed, only the Education & Educational Research (E&ER) category was filtered from the WoS categories to identify articles related to education. As a result of this filtering process, the number of studies was reduced to 488. The English-language articles were then filtered, and 480 studies were identified. 8 studies written in languages other than English were excluded because the researchers were not proficient in them.

### *Eligibility*

As this study aims to conduct a systematic review of empirical, peer-reviewed studies on AI in science education indexed in the WoS database, 480 studies were filtered to include articles and early access, and all other document types were excluded. As a result of this filtering process, the number of articles was reduced to 177. As shown in Table 1, the researchers established inclusion and exclusion criteria to narrow the 177 articles to those eligible for this review.

**Table 1** Inclusion and Exclusion Criteria Set by Researchers to Meet Eligibility

| Inclusion criteria | Exclusion criteria |
|---|---|
| Articles that are related to teaching or learning are included. | Articles that are not related to teaching or learning are excluded. |
| Articles that are related to science education are included. | Articles that are related to disciplines other than science education are excluded. |
| Empirical articles are included. | Non-empirical articles (e.g., commentaries, review articles, and book chapters) are excluded. |

The filtering process was completed by December 31, 2024. The researchers screened the articles according to the inclusion and exclusion criteria. As a result, 39 empirical articles written in English and indexed in the WoS database that met the inclusion criteria were included in this review for in-depth analysis. A list of all articles included in this review can be found in Appendix A.

## Data Analysis

The researchers hyperlinked the articles in an open-access Excel spreadsheet in Google Drive, which enabled collaborative organisation of the selected articles. The selected 39 articles on AI in science education were analysed using content analysis. The coding scheme was developed using a hybrid approach combining deductive and inductive coding. First, deductive codes were generated from the research questions and the relevant literature. Then, additional inductive codes emerged through repeated reading of the studies. Each article was then coded and consistently categorised under the headings of year, country, field of study, type of participants, research designs, data sources, AI-related skills, and AI-related ethical concerns addressed. Themes were generated by clustering related codes and refined through iterative comparison to ensure conceptual clarity and consistency. To validate the coding structure, an independent science education expert coded 20% of the included studies, in addition to the full coding conducted by the first author. Interrater reliability was calculated using Cohen's kappa ($\kappa = .96$), which indicates almost perfect agreement (Landis & Koch, 1977). The first author coded the remaining studies, and any discrepancies between coders were discussed and resolved through consensus. Finally, frequency calculations were then performed to identify trends in each category, and the results were visualised.

## Results

### Research Article Trends by Year

Figure 2 shows the trends in the number of empirical and peer-reviewed articles on AI in science education analysed in this study, published from the first mention in the WoS database until 2024.

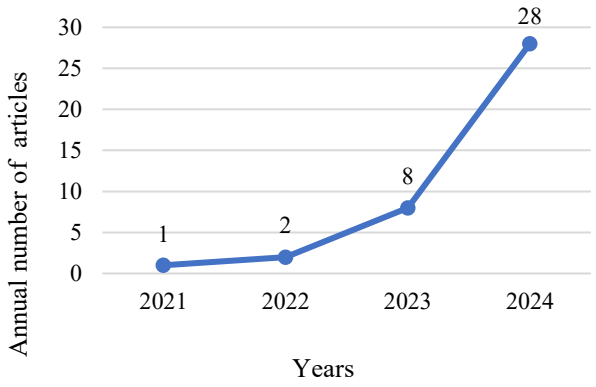



**Figure 2** The Annual Number of Articles in AI on Science Education Research Published from the First Mention to the End of 2024

As shown in Figure 2, artificial intelligence (AI) research in science education was first identified in the WoS database in 2021, and only one study (2.60%) was conducted in that year. In 2022, 2 studies (5.1%) were published, while in 2023, 8 studies (20.50%) were published. The number of articles peaked in 2024. In 2024, 28 (71.80%) research studies were conducted, more than three times the number in 2023. Thus, it is reported that AI research in science education has become an increasing trend between 2021-2024. This sharp rise, particularly after 2022, suggests that the rapid development and wider adoption of AI tools, especially generative AI applications, have intensified research interest in science education. The pattern may also reflect a global emphasis on digital transformation in education, which has encouraged researchers to investigate AI-related opportunities and challenges in science teaching more actively.

### Research Article Trends by Countries

Figure 3 shows the trends by country of origin for AI in the 39 science education research articles included in this review article. As shown in Figure 3, the 39 articles originate from 12 different countries. The origin of most research is the USA (10 articles, 25.64%), Korea (8 articles, 20.51%), China (6 articles, 15.38%), and Türkiye (5 articles, 12.82%). The

fewest research origin was found in African countries, with 3 articles (7.69%); Australia and Taiwan, with 2 articles each (5.13% each); and the United Arab Emirates, Saudi Arabia, Jordan, Cyprus, and Brazil, with 1 article each (2.56% each). Only 2 articles presented data from both the USA and Korea (Lee et al., 2024; Oh & Lee, 2025). The dominance of the USA and Korea may reflect their strong technological infrastructure, national investments in AI innovation, and education policies that actively promote digital transformation. In contrast, the limited number of studies from other regions suggests unequal access to AI resources and varying policy emphases, which may affect the global capacity for AI research in science education.

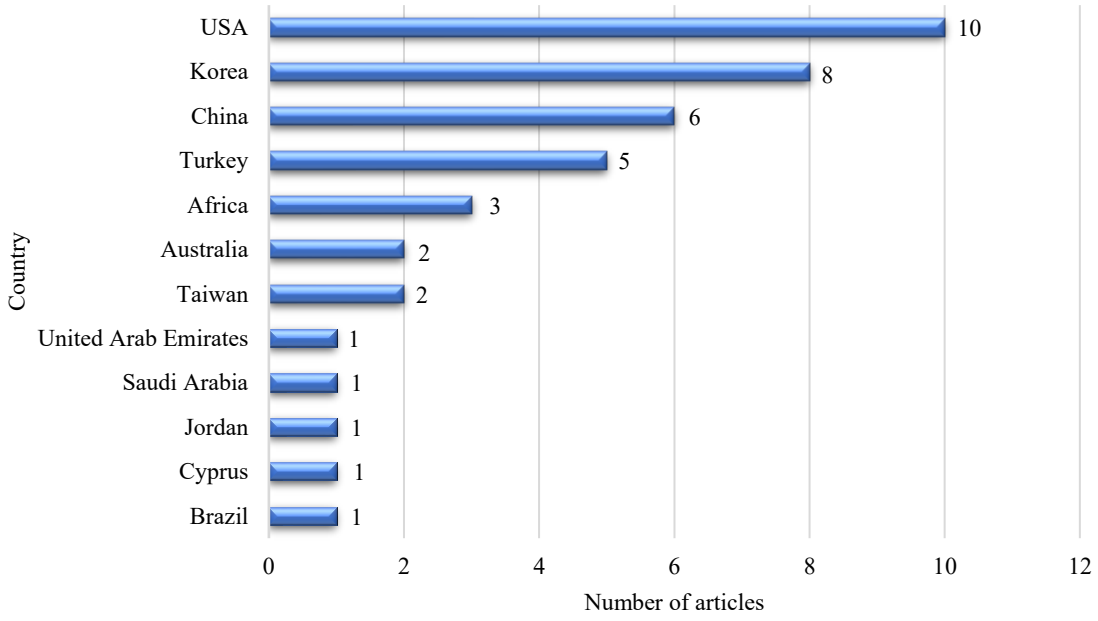



**Figure 3** Trends in AI on Science Education Research Articles by Country of Origin

## Study Areas by Trend

Table 2 shows the frequency of studies on AI in science education according to the areas of study and the references of the related studies. As can be seen in Table 2, most of the studies on AI in science education were from the area of science education (87.18%), while the other studies were from the areas of physics education (7.69%), chemistry education (7.69%), biology education (5.13%), STEM education (5.13%) and mathematics education (2.56%). Only 2 studies reported data from more than one field. One paper reported data from science education, physics education, chemistry education, and biology education (Antonenko & Abramowitz, 2023), while one paper reported data from science education, physics

education, and chemistry education (Lee et al., 2023a). This distribution suggests that AI research in science education is predominantly focused on general science education rather than domain-specific areas. The limited number of studies in physics, chemistry, biology, and STEM education may reflect the slower integration of AI into these subject-specific curricula and teacher education programs, as well as methodological challenges in applying AI tools in laboratory-based or conceptually complex domains.

**Table 2** Frequency of Use of AI in Science Education in the Study Areas

| Area of articles | Total number of articles | References |
|---|---|---|
| Science education | 34 | Antonenko & Abramowitz, 2023; Adeyele & Ramnarain, 2024; AlKanaan, 2022; Alshorman, 2024; Chang et al., 2023; Chen & Liu, 2024; Cheung et al., 2024; Cooper, 2023; Cooper & Tang, 2024; Garofalo & Farenga, 2025; Huang et al., 2025; Jho & Ha, 2024; Karahan, 2023; Göktepe Körpeoglu & Göktepe Yildiz, 2024; Lee et al., 2023; Lee et al., 2023; Lee et al., 2024; Lee & Zhai, 2024; Li & Ironsi, 2025; Monteiro et al., 2024; Ng et al., 2024; Nguyen & Hayward, 2025; Nja et al., 2023; Oh & Lee, 2025; Okulu & Muslu, 2024; Ramnarain at al., 2025; Taani & Alabidi, 2025; Deveci Topal et al., 2021; Wei et al., 2024; Zhai et al., 2022; Zhai et al., 2023; Zhai et al., 2025; Zhang et al., 2025 |
| Physics education | 3 | Antonenko & Abramowitz, 2023; Jang & Choi, 2025; Lee et al., 2023 |
| Chemistry education | 3 | Antonenko & Abramowitz, 2023; Lee et al., 2023; Martin et al., 2024 |
| Biology education | 2 | Antonenko & Abramowitz, 2023; Behrens et al., 2025 |
| STEM education | 2 | Huang & Qiao, 2024; Uyanik Aktulun et al., 2024 |
| Mathematics education | 1 | Taani & Alabidi, 2025 |

## Participant Types

The 39 studies of AI in science education included in this review were analysed as human (teacher or student) or non-human (AI) in terms of participants. Of the 39 studies, 34 studies (87.18%) were conducted with humans, and 4 studies (10.26%) were conducted with AI. One study (2.56%) included both human and AI participants (Zhang et al., 2025).

Figure 4 shows the frequency of studies with human participants in AI research in science education. A total of 52 participant groups were studied in 39 articles. An analysis of Figure 4 shows that the majority of studies were conducted with K-12 students (50.98%). It can also be seen that the number of studies conducted with pre-service teachers (21.57%) and

in-service teachers (19.61%) is close. The fewest studies were carried out with college/university students (7.84%) and educators (1.96%) such as academics. It is noteworthy that only one study was identified that worked with two different categories of participants, K-12 students and in-service teachers (Chang et al., 2023). The prevalence of K-12 student samples may reflect the increasing focus on incorporating AI into compulsory educational settings, where substantial student populations provide practical contexts for implementing AI in the classroom. In contrast, limited research involving university students and educators suggests that AI integration in higher education remains less developed, possibly due to curricular constraints or slower institutional adoption.

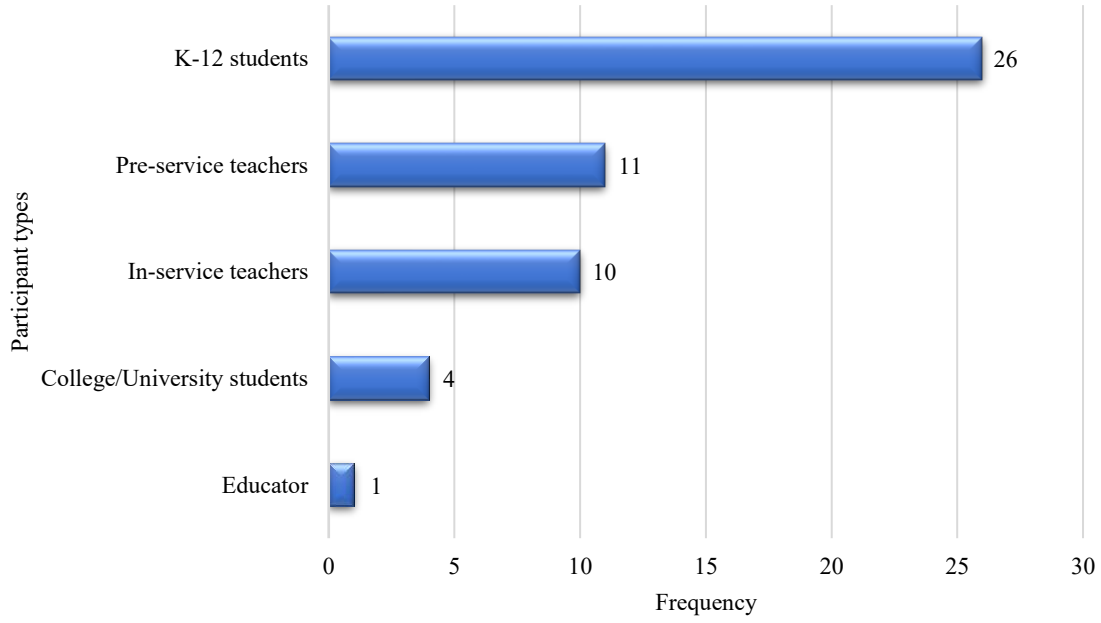



**Figure 4** Trends by Human Participant Type for AI in Science Education Research Articles

## Trends in Research Method and Data Sources

Of the 39 AI research studies in science education included in this review article, 10 articles (25.65%) used a quantitative research method, while 11 articles (28.20%) used a qualitative research method. It was found that the remaining 18 studies (46.15%) used a mixed research method.

Out of a total of 39 articles, 4 directly used AI as a participant as a data source (Cooper, 2023; Cooper & Tang, 2024; Oh & Lee, 2025; Okulu & Muslu, 2024). Figure 5 shows only the data sources in articles with human participants, such as teachers and students.

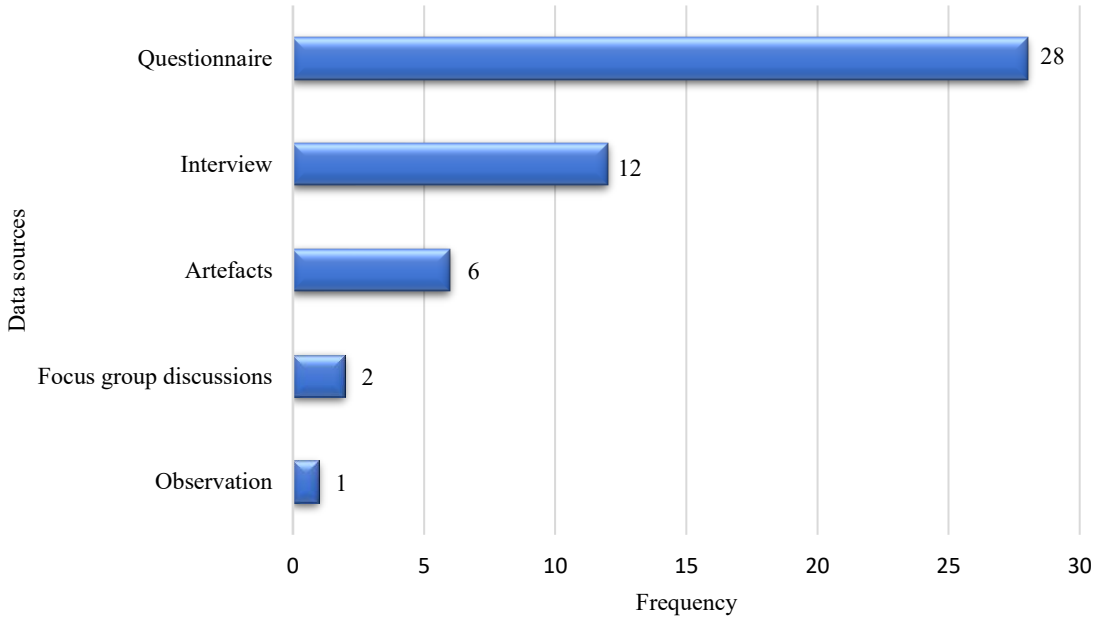



**Figure 5** Trends in Data Sources for Research Articles on AI in Science Education

As shown in Figure 5, a total of 49 data sources were used in 39 articles. It was found that questionnaires (57.14%) and interviews (24.49%) were the most commonly used data sources in studies of AI in science education. The next most used data source was artefacts (12.24%) such as drawn models, written texts, reflective notes, and teaching materials. The least used data sources were focusing group discussions (4.08%) and observations (2.04%). The extensive use of questionnaires and interviews indicates that most studies have centered on self-reported perceptions rather than classroom-based or performance-based assessments. The limited use of observations and focus groups may indicate challenges accessing classrooms with AI tools and a tendency to prioritize easily collected data over more resource-intensive methods.

**AI-Related Skills**

The AI-related skills in 39 articles in this review article were coded according to whether the AI skills were directly criticized in any part of the article or explicitly addressed in the articles. As a result of the analysis, 19 articles (48.71%) did not address AI-related skills, while 20 articles (51.29%) explicitly addressed AI-related skills.

Table 3 illustrates AI-related skills in AI articles in science education. As can be seen in Table 3, thirteen AI-related skills were identified. The skills mostly associated with AI are listed as critical thinking skills (19.23%), cognitive skills (15.38%), problem solving skills

(13.46%), communication skills (13.46%), and questioning skills (9.62%). The skills less associated with AI are listed as follows: 21st century skills (5.77%), collaboration skills (5.77%), computational thinking skills (3.85%), AI literacy skills (3.85%), writing skills (3.85%), lifelong learning skills (1.92%), self-regulated learning skills (1.92%) and time management skills (1.92%). This finding suggests that current AI research in science education emphasizes higher-order cognitive processes over AI-specific competencies. The limited attention given to computational thinking, AI literacy, and self-regulated learning suggests that many studies focus on general skill development rather than examining how AI tools support discipline-specific or technology-oriented skill sets.

**Table 3** AI-Related Skills in AI Articles in Science Education

| Skills | Total number of articles | References |
|---|---|---|
| Critical thinking skills | 10 | Alshorman, 2024; Cooper, 2023; Garofalo & Farenga, 2025; Huang et al., 2025; Jang & Choi, 2025; Lee et al., 2024; Monteiro et al., 2024; Okulu & Muslu, 2024; Wei et al., 2024; Zhai et al., 2025 |
| Cognitive skills | 8 | Alshorman, 2024; Garofalo & Farenga, 2025; Jang & Choi, 2025; Monteiro et al., 2024; Ng et al., 2024; Ramnarain at al., 2025; Wei et al., 2024; Zhai et al., 2025 |
| Problem solving skills | 7 | Alshorman, 2024; Jang & Choi, 2025; Monteiro et al., 2024; Oh & Lee, 2025; Taani & Alabidi, 2025; Wei et al., 2024; Zhai et al., 2025 |
| Communication skills | 7 | AlKanaan, 2022; Jang & Choi, 2025; Monteiro et al., 2024; Oh & Lee, 2025; Okulu & Muslu, 2024; Ramnarain at al., 2025; Wei et al., 2024 |
| Questioning skills | 5 | Huang et al., 2025; Jang & Choi, 2025; Taani & Alabidi, 2025; Wei et al., 2024; Zhang et al., 2025 |
| 21st-century skills | 3 | Alshorman, 2024; Ramnarain at al., 2025; Wei et al., 2024 |
| Collaboration skills | 3 | Jang & Choi, 2025; Oh & Lee, 2025; Wei et al., 2024 |
| Computational thinking skills | 2 | Huang & Qiao, 2024; Taani & Alabidi, 2025 |
| AI literacy skills | 2 | Ng et al., 2024; Zhai et al., 2022 |
| Writing skills | 2 | Behrens et al., 2025; Li & Ironsi, 2025 |
| Lifelong learning skills | 1 | AlKanaan, 2022 |
| Self-regulated learning skills | 1 | Ng et al., 2024 |
| Time management skills | 1 | Ng et al., 2024 |

## AI - Related Ethical Concerns

Of the 39 AI in science education articles analysed, 23 articles addressed AI-related ethical concerns (58.97%), while 16 articles did not (41.03%). Table 4 lists the ethical concern in the articles that addressed ethical concerns in AI in science education.

**Table 4** Ethical Concerns Addressed in AI Research in Science Education

| AI ethics concerns | Total number of articles | References |
|---|---|---|
| Bias | 11 | Behrens et al., 2025; Cheung et al., 2024; Cooper & Tang, 2024; Garofalo & Farenga, 2025; Jang & Choi, 2025; Li & Ironsi, 2025; Nguyen & Hayward, 2025; Oh & Lee, 2025; Ramnarain at al., 2025; Uyanik Aktulun et al., 2024; Zhai et al., 2025 |
| Privacy | 9 | AlKanaan, 2022; Alshorman, 2024; Behrens et al., 2025; Garofalo & Farenga, 2025; Li & Ironsi, 2025; Lee & Zhai, 2024; Ramnarain at al., 2025; Taani & Alabidi, 2025; Zhai et al., 2025 |
| Accuracy | 7 | Lee & Zhai, 2024; Li & Ironsi, 2025; Ng et al., 2024; Nguyen & Hayward, 2025; Oh & Lee, 2025; Ramnarain at al., 2025; Uyanik Aktulun et al., 2024 |
| Plagiarism | 5 | Behrens et al., 2025; Cooper, 2023; Garofalo & Farenga, 2025; Lee & Zhai, 2024; Monteiro et al., 2024; Ramnarain at al., 2025 |
| Content moderation | 4 | Cooper, 2023; Garofalo & Farenga, 2025; Nguyen & Hayward, 2025; Taani & Alabidi, 2025 |
| Copyright infringement | 4 | Behrens et al., 2025; Cooper, 2023; Jang & Choi, 2025; Taani & Alabidi, 2025 |
| Transparency | 3 | Cooper, 2023; Cooper & Tang, 2024; Zhai et al., 2022 |
| Accountability | 3 | Antonenko & Abramowitz, 2023; Zhai et al., 2022; Zhai et al., 2023 |
| Cheating | 3 | Behrens et al., 2025; Garofalo & Farenga, 2025; Uyanik Aktulun et al., 2024 |
| Equity | 2 | Antonenko & Abramowitz, 2023; Cooper & Tang, 2024 |
| Environmental impact of AI platforms | 2 | Cooper, 2023; Taani & Alabidi, 2025 |
| Inclusivity | 2 | Alshorman, 2024; Cooper & Tang, 2024 |
| Integrity | 2 | Behrens et al., 2025; Lee & Zhai, 2024 |
| Societal risk | 2 | Cooper & Tang, 2024; Zhai et al., 2025 |
| Reliability | 2 | Alshorman, 2024; Cheung et al., 2024 |
| Fairness | 1 | Antonenko & Abramowitz, 2023 |
| Liability | 1 | Oh & Lee, 2025 |
| Intellectual property theft | 1 | Garofalo & Farenga, 2025 |
| Autonomy risk | 1 | Ramnarain at al., 2025 |
| Functional risk | 1 | Ramnarain at al., 2025 |
| Social knowledge risks | 1 | Ramnarain at al., 2025 |
| Skill-building risks | 1 | Ramnarain at al., 2025 |

There is a total of 22 ethical concerns in 23 articles that addressed AI-related ethical concerns. The most frequently mentioned ethical concerns are bias (14.29%), privacy (11.69%), accuracy (9.09%), plagiarism (6.49%), content moderation (5.19%), copyright infringement (5.19%). Transparency, accountability and cheating were then mentioned (3.90% for each). Equity, environmental impact of AI platforms, inclusivity, integrity, societal risk, and reliability were the less mentioned ethical concerns (2.60% for each), while fairness, liability, intellectual property theft, autonomy risk, functional risk, social knowledge risk, and skill building risk (1.30% for each) were the least mentioned ethical concerns. The prominence of concerns such as bias, privacy, and accuracy suggests that discussions about AI in science education are heavily influenced by the widely recognized global risks associated with AI systems. Conversely, the limited attention given to issues such as equity, inclusivity, and societal or environmental risks suggests that the broader ethical dimensions of AI are underrepresented in the literature on science education.

## Discussion

The dual purpose of this review was to conduct a systematic review of empirical, peer-reviewed studies of artificial intelligence (AI) in science education indexed in the Web of Science (WoS) database from first mention through the end of 2024, and to address ethical concerns from a pragmatic perspective. Examining studies on AI, which is a recently trending topic, is important in terms of contributing to the science education literature and improving our understanding (Budak & Yerdelen Damar, 2025). In this study, instead of focusing on the studies of AI in education between specific years as in other reviews (e.g., Ng et al., 2023; Yusuf et al., 2024; Zhai et al., 2021), we focused on all studies of AI in science education from the date AI was first mentioned in the WoS database to the end of 2024. It is important to use the first-mention approach because it allows us to capture the earliest emergence of AI in the science education literature and to trace the field's evolution over time. Unlike existing reviews, which typically limit their scope to predetermined time ranges, this approach provides a more comprehensive and historically grounded overview.

Rather than reviewing specific areas such as the role of AI chatbots (Labadze et al., 2023), AI literacy (Casal-Otero et al., 2023; Ng et al., 2023), or the application of AI technologies in STEM education (Xu & Ouyang, 2022) or science (Heeg & Avraamidou, 2023), this study aims to be a more comprehensive review as it examines AI studies in science education from a bird's eye view regardless of the field. These findings can be interpreted through the lens of constructivist and sociocultural learning theories. These

theories suggest that learners construct knowledge through interaction with tools and peers. AI applications in classrooms provide new modalities for such interactions, thereby enhancing the learning process. This study provides empirical evidence and critical evaluations of the science education literature from a multifaceted perspective in terms of examining AI trends in science education and addressing AI ethical concerns. At the practical level, these insights can guide teachers and curriculum developers in designing classroom activities that leverage AI tools while considering students' cognitive development and ethical reasoning skills.

In this study, the analysis of empirical and peer-reviewed articles on AI in science education published in the WoS database from first mention to 2024 revealed that the first article was found in 2020 and that the number of articles tended to increase over the years. This finding is consistent with the findings that 2020 is a breakthrough year in education as reported by Casal-Otero et al. (2023) and Ng et al. (2024) in their review on AI literacy and Ng et al. (2023) in their review on AI teaching and learning. This finding is also consistent with Jia et al.'s (2024) anticipation that AI studies in science education will increase in the future and continue to attract attention (e.g., Chen and See, 2020; Hwang et al., 2020). The increasing trend in AI studies underscores the growing interest and adoption of AI in science education, indicating that educational institutions and teacher training programs need to prepare to integrate AI tools into classrooms to support student learning and engagement.

This study identified that the studies of AI in science education included in the WoS database provided data from 12 different countries. Most of the studies were from the USA and Korea. This prominence may be related to several contextual factors, including strong national education policies that support AI integration, a robust technological infrastructure, high levels of investment in digital learning, and a culture that facilitates the rapid adoption of emerging technologies. This finding is partially consistent with the results of recent review studies, which reported that most of the articles reviewed were from the USA and China (Casal-Otero et al., 2023; Ng et al., 2024). Practically, these findings suggest that policymakers and curriculum designers need to consider local contexts, resources, and readiness when implementing AI in classrooms. In addition, only two articles were found to present data from more than one country (Lee et al., 2024; Oh & Lee, 2025). This finding indicates a serious lack of comparative research on AI studies in the science education literature. These findings suggest that the science education literature needs more data on why, how, and in what contexts AI should be addressed. For this reason, we suggest that we

improve our understanding by focusing on studies of AI in science education by providing findings from different countries.

In this review study, when AI studies in science education were examined by discipline, it was found that most studies were conducted in science education, while there were very few studies in physics education, chemistry education, biology education, and STEM education. Yusuf et al. (2024) reviewed AI studies in education and found that the most common areas of research by educational discipline are general education, medical education, and language education. In parallel with the findings of this study, Yusuf et al. (2024) showed that science education and science-related fields (e.g., physics education, chemistry education, biology education, and STEM education) are underrepresented in AI studies. One probable reason for this underrepresentation is the limited integration of AI into discipline-specific teacher education programs, as well as the lack of pedagogical and technical preparation among teachers in these fields. This finding suggests a need for targeted teacher training and curriculum development in these underrepresented fields to ensure AI is effectively integrated into discipline-specific teaching practices. Additionally, conducting AI-supported activities can be challenging from a methodological standpoint in physics, chemistry, biology, and STEM contexts, which may further reduce the number of empirical studies produced. This imbalance also has implications for teaching programs because it limits the development of field-specific AI pedagogies and teachers' ability to meaningfully incorporate AI into their instruction. Therefore, there is a need to increase AI studies in science education fields, especially in underrepresented fields, and to further explore how AI is manifested in science education and teaching comparatively across different fields. Future studies should investigate the reasons for this disciplinary gap and develop discipline-specific models for integrating AI into physics, chemistry, biology, and STEM curricula to support teacher preparation and curriculum development.

One of the notable findings of the current study is that AI itself has been included as a participant in very limited studies of AI in science education (e.g., Cooper, 2023; Cooper & Tang, 2024; Oh & Lee, 2025; Okulu & Muslu, 2024). From a theoretical standpoint, including AI as a participant allows for exploration of AI as an epistemic agent, contributing to research on human-AI interaction and knowledge construction. For example, Oh and Lee (2025) only interviewed ChatGPT as a participant in their study of human episteme and ontology because it is an epistemic agent that can produce scientific knowledge. In addition, in another study, Okulu and Muslu (2024) used ChatGPT in the assessment of course design

in science education. Practically, integrating AI as a participant provides a roadmap for classroom applications and allows educators to understand both the affordances and limitations of AI in learning environments. In the studies involving humans, it was found that K-12 students were the most common participants, while college/university students and educators such as academics were the least common participants. Since it is important to study the understanding, knowledge, attitudes, and skills of college/university students toward AI, it is very important to include them as a participant group. Finally, the fact that studies with academics are very limited is an alarming report, because given the potential of academics to educate future generations, it is necessary to learn their insights, perspectives, and uses towards AI.

In the articles selected for this review, it was found that the most common research method used was mixed methods. This finding is promising in terms of trying to explore in depth with both qualitative and quantitative methods, as AI is an emerging topic in science education research. As stated by Ng et al. (2024), we suggest that future studies should be designed using quasi-experimental designs to compare the results of different groups and present the findings of the study more empirically. It was found that observation and focus group discussion were the least used data collection tools in the reviewed studies. Since observation and focus group discussion data sources have the potential to reveal participants' experiences with AI or suggestions for its integration, these data sources should be included in future research to increase our understanding of AI. Thus, incorporating diverse research methods and data sources can provide more robust evidence for AI integration strategies in classrooms.

One of the critical findings of this study is that almost half of the reviewed articles do not include AI-related skills. This highlights a gap in applying 21st-century skills frameworks in AI research, where cognitive and meta-cognitive skills are essential. AI has been found to positively contribute to higher-order thinking skills, such as problem-solving and self-regulated learning skills (Xu & Ouyang, 2022), as well as programming skills (Jia et al., 2024). In practice, this suggests that curriculum designers and teachers should intentionally include AI-related skill development to prepare students for a technologically evolving world (Jia et al., 2024).

Although AI has various benefits, the ethical issues of AI should be considered (Almasri, 2024). It is alarming that almost half of the articles included in this review did not consider AI ethical concerns such as privacy, algorithmic bias, misinformation, and issues

related to student confidentiality. This review uncovered 22 ethical concerns highlighted in AI research in science education. Practically, educators and policymakers should ensure that AI tools are implemented in ways that address privacy, fairness, and bias, while also fostering students' awareness of these ethical issues. However, while the ethical concerns of bias and privacy were frequently emphasized, the concerns of fairness, liability, intellectual property theft, autonomy risk, functional risk, social knowledge risk, and skill-building risk were least emphasized. This finding is consistent with Dignum's (2021) statement that ethical concerns about AI are mostly limited to privacy and the use of personal data. However, as Labadze et al. (2023) stated, educators need to specifically address ethical concerns of AI. When addressing the ethical concerns of AI, it is important to make efforts to digitize and transform society for the public good (Heilinger, 2022). Therefore, we need to address the ethics of AI through a pragmatic lens, from a collective perspective that benefits society for the good of all. All stakeholders - science researchers, science teachers, and curriculum makers - should strive to integrate the ethical issues of AI into education with pragmatic solutions.

## Conclusions and Suggestions

This review has important implications for science researchers, science teachers, and science curriculum makers. First, for science researchers, the study provides evidence-based benchmarks to identify knowledge gaps and to guide future research on AI in science education, particularly in areas such as pedagogical strategies, students' conceptual understanding of AI, and integration of AI into experimental and inquiry-based learning activities. Second, for science teachers, despite the growing number of AI studies, understanding of how to teach and facilitate learning with AI is still limited, highlighting the need for integrating AI tools into classroom practices, designing hands-on activities with AI, and applying strategies that actively engage students while addressing ethical considerations. Third, for curriculum developers, the findings emphasize the importance of incorporating AI content and ethical considerations into science curricula, including developing modules that cover AI applications in different scientific disciplines, embedding societal and ethical discussions, and adapting content to various grade levels to ensure meaningful learning experiences. Considering the pioneering role of AI in the digitalization of society, curriculum makers are encouraged to adopt a pragmatic approach that balances technological benefits with ethical responsibilities. Furthermore, future research should systematically investigate the practical applications and ethical challenges of AI in diverse classroom contexts, including differences across grade levels, student backgrounds, and scientific disciplines. This research

should also examine how AI influences student engagement, learning outcomes, and teachers' ethical decision-making processes, providing actionable guidance for curriculum design, teacher training, and policy development.

While this review offers significant insights, it also has some limitations. The study included only empirical, peer-reviewed publications indexed in the Web of Science (WoS) database from the date of first mention through the end of 2024, without any grade-level restriction. Another limitation is that only English-language publications were included, which may exclude relevant studies published in languages other than English. As only one database was analysed, the results should be interpreted cautiously, and generalizations should be avoided. Future studies on AI in science education should include multiple databases to allow for a more comprehensive trend analysis.

## Compliance with Ethical Standards

*Disclosure of potential conflicts of interest*

The authors declare no competing interests.

*Funding*

No funds, grants, or other support was received for conducting this study.

*CRediT author statement*

CRediT (Contributor Roles Taxonomy) was introduced with the intention of recognizing individual author contributions.

Author 1: Conceptualization, Methodology, Formal Analysis, Investigation, Writing – original draft, Writing – review and editing, visualization

Author 2: Methodology, Investigation, Writing – review and editing, supervision

*Research involving Human Participants and/or Animals*

This systematic review is based solely on previously published articles and did not involve collecting data from human participants or animals. Therefore, approval from an institutional or national research ethics committee was not necessary.

## Fen Eğitiminde Yapay Zeka (YZ) Araştırma Eğilimleri: Etik Kaygıların Eleştirel Açıdan İncelenmesi

**Özet:**

Yapay Zeka (YZ), insanlığı geliştiren ve ortak yararı ilerleten hedeflere ulaşmak için akıllı makinelerin bilim ve mühendisliğidir. Bu sistematik literatür incelemesi, Web of Science veritabanında indekslenen 39 ampirik ve hakemli bilim eğitimi çalışmasında, bilim eğitiminde yapay zekanın ilk kez bahsedildiği andan 2024 yılının sonuna kadar nasıl ele alındığını incelemek ve etik kaygıları ele almak amacıyla yapılmıştır. Fen eğitiminde YZ araştırmalarındaki eğilimleri belirlemek için içerik analizi kullanılmıştır. Bu inceleme, bilim eğitiminde YZ ile ilgili araştırmaların son yıllarda önemli ölçüde arttığını ortaya koymuştur. Araştırmaların çoğunun ABD ve Kore'de yapıldığı tespit edilmiştir. Bu çalışma, çok az sayıda çalışmanın ChatGPT gibi YZ'yi katılımcı olarak içerdiğini, çoğunun ise insanları katılımcı olarak içerdiğini göstermiştir. Öte yandan, çoğu çalışma K-12 öğrencilerine odaklanırken, en az sayıda çalışma üniversite öğrencileri ve eğitimcilere odaklanmıştır. YZ'nin gelişimi, sosyal ve etik kaygıları da beraberinde getirmektedir. İncelenen makalelerin neredeyse yarısı YZ ile ilgili etik kaygıları ele almamıştır. En sık ele alınan etik kaygılar önyargı, gizlilik ve doğruluktur. Bu çalışma, bilim eğitiminde YZ ile ilgili gelecekteki çalışmalara önemli kanıta dayalı destek sağlayabilir.

*Anahtar kelimeler*: Yapay zeka, fen eğitimi, sistematik derleme, etik.

## References

Adeyele, V. O., & Ramnarain, U. (2024). Exploring the integration of ChatGPT in inquiry-based learning: Teacher perspectives. *International Journal of Technology in Education*, *7*(2), 200-217. https://doi.org/10.46328/ijte.638

Aksit, O. (2025). Teaching science with AI: How preservice teachers perceive the role of AI tools in science education. In Kadry S. and Balasubramaniam S. (Eds.), *Artificial intelligence in education: Creating an equitable, creative, and effective learning environment*. IntechOpen. https://doi.org/10.5772/intechopen.1012592

AlKanaan, H. M. N. (2022). Awareness regarding the implication of artificial intelligence in science education among pre-service science teachers. *International Journal of Instruction*, *15*(3), 895-912. https://doi.org/10.29333/iji.2022.15348a

Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Research in Science Education*, *54*(5), 977-997. https://doi.org/10.1007/s11165-024-10176-3

Anderson, M. & Anderson, S. L. (2006). "Guest Editors' Introduction: Machine Ethics, *IEEE Intelligent Systems*, 21 (4), 10-11. https://doi.org/10.1109/MIS.2006.70

Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI magazine*, *28*(4), 15-26. https://doi.org/10.1609/aimag.v28i4.20

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

Alshorman, S. M. (2024). The readiness to use AI in teaching science: Science teachers' perspective. *Journal of Baltic Science Education*, *23*(3), 432-448. https://doi.org/10.33225/jbse/24.23.432

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Behrens, K. A., Marbach-Ad, G., & Kocher, T. D. (2025). AI in the genetics classroom: A useful tool but not a replacement for creative writing. *Journal of Science Education and Technology*, *34*, 621-635. https://doi.org/10.1007/s10956-024-10160-6

Blanchard, K., & Peale, N. V. (2011). *The power of ethical management*. Random House.

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363-376. https://doi.org/10.1162/qss_a_00018

Budak, U.S., & Yerdelen Damar, S. (2025). Research trends and emerging areas of interest from 2004 to 2024 in science education. *Journal of Turkish Science Education, 22*(4), 625-654. http://doi.org/10.36681/tused.2025.032

Casal-o, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, *10*, Article 29. https://doi.org/10.1186/s40594-023-00418-7

Chang, J., Park, J., & Park, J. (2023). Using an artificial intelligence chatbot in scientific inquiry: Focusing on a guided-inquiry activity using inquiry bot. *Asia-Pacific Science Education*, *9*(1), 44-74. https://doi.org/10.1163/23641177-bja10062

Chen, J., & See, K. C. (2020). Artificial intelligence for COVID-19: Rapid review. *Journal of Medical Internet Research, 22*(10), Article e21476. https://doi.org/10.2196/21476

Chen, P. Y., & Liu, Y. C. (2024). Impact of AI robot image recognition technology on improving students' conceptual understanding of cell division and science learning motivation. *Journal of Baltic Science Education*, *23*(2), 208-220. https://doi.org/10.33225/jbse/24.23.208

Cheung, K. K. C., Pun, J. K. H & Li, W. (2024). Students' holistic reading of socio-scientific texts on climate change in a ChatGPT scenario. *Research in Science Education*, *54*, 957-976. https://doi.org/10.1007/s11165-024-10177-2

Cheung, K. K. C., Pun, J., Kenneth-Li, W., & Mai, J. (2025). Exploring students' multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools. *Journal of Science Education and Technology*, *34*, 284-297. https://doi.org/10.1007/s10956-024-10182-0

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

Cooper, G., & Tang, K. S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. *Journal of Science Education and Technology*, *33*, 556-568. https://doi.org/10.1007/s10956-024-10104-0

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

Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education, 19*(1), 1–11. https://doi.org/10.14324/LRE.19.1.01

Erduran, S. (2023). AI is transforming how science is done. Science education must reflect this change. *Science, 382*(6677), Article eadm9788. https://www.science.org/doi/10.1126/science.adm9788

Garofalo, S. G., & Farenga, S. J. (2025). Science teacher perceptions of the state of knowledge and education at the advent of generative artificial intelligence popularity. *Science & Education*, *34*, 893-912. https://doi.org/10.1007/s11191-024-00534-y

Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications* (10th ed). Pearson.

Göktepe Körpeoğlu, S., & Göktepe Yıldız, S. (2024). Using artificial intelligence to predict students' STEM attitudes: An adaptive neural-network-based fuzzy logic model. *International Journal of Science Education*, *46*(10), 1001-1026. https://doi.org/10.1080/09500693.2023.2269291

Green, J., & Thorogood, N. (2004). *Qualitative methods for health research*. Sage.

Heeg, D. M., & Avraamidou, L. (2023). The use of Artificial intelligence in school science: a systematic literature review. *Educational Media International*, *60*(2), 125-150. https://doi.org/10.1080/09523987.2023.2264990

Heilinger, J. C. (2022). The ethics of AI ethics. A constructive critique. *Philosophy & Technology*, *35*, Article 61. https://doi.org/10.1007/s13347-022-00557-9

HLEG. (2019). High-level expert group on Artificial Intelligence. Ethics guidelines for trustworthy AI. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Huang, X., & Qiao, C. (2024). Enhancing computational thinking skills through artificial intelligence education at a STEAM high school. *Science & Education*, *33*, 383-403. https://doi.org/10.1007/s11191-022-00392-6

Huang, H. W., Teng, D. C. E., & Tiangco, J. A. N. Z. (2025). The impact of AI chatbot-supported guided discovery learning on pre-service teachers' learning performance and motivation. *Journal of Science Education and Technology*, *34*, 1286-1300. https://doi.org/10.1007/s10956-024-10179-9

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, *4*(4), 799-819. https://doi.org/10.1109/TAI.2022.3194503

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Jang, H., & Choi, H. (2025). A double-edged sword: Physics educators' perspectives on utilizing ChatGPT and its future in classrooms. *Journal of Science Education and Technology*, *34*(2), 267-283. https://doi.org/10.1007/s10956-024-10173-1

Jho, H., & Ha, M. (2024). Towards effective argumentation: Design and implementation of a generative AI-based evaluation and feedback system. *Journal of Baltic Science Education*, *23*(2), 280-291. https://doi.org/10.33225/jbse/24.23.280

Jia, F., Sun, D., & Looi, C. K. (2024). Artificial intelligence in science education (2013-2023): Research trends in ten years. *Journal of Science Education and Technology*, *33*, 94-117. https://doi.org/10.1007/s10956-023-10077-6

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), 1-27. https://doi.org/10.3390/su151612451

Karahan, E. (2023). Using video-elicitation focus group interviews to explore pre-service science teachers' views and reasoning on artificial intelligence. *International Journal of Science Education*, *45*(15), 1283-1302. https://doi.org/10.1080/09500693.2023.2200887

Karan, B., & Angadi, G. R. (2024). Potential risks of artificial intelligence integration into school education: A systematic review. *Bulletin of Science, Technology & Society*, *43*(3–4), 67–85. https://doi.org/10.1177/02704676231224705

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, *20*, Article 56. https://doi.org/10.1186/s41239-023-00426-1

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://www.jstor.org/stable/2529310

Lee, G. G., Mun, S., Shin, M. K., & Zhai, X. (2024). Collaborative learning with artificial intelligence speakers. *Science & Education*, 34, 847–875. https://doi.org/10.1007/s11191-024-00526-y

Lee, G. G., & Zhai, X. (2024). Using ChatGPT for science learning: A study on pre-service teachers' lesson planning. *IEEE Transactions on Learning Technologies*., *17*, 643-1660. https://doi.org/10.1109/TLT.2024.3401457

Lee, J., An, T., Chu, H. E., Hong, H. G., & Martin, S. N. (2023). Improving science conceptual understanding and attitudes in elementary science classes through the development and application of a rule-based AI chatbot. *Asia-Pacific Science Education*, *9*(2), 365-412. https://doi.org/10.1163/23641177-bja10070

Lee, J., Lee, G. G., & Hong, H. G. (2023). Automated assessment of student hand drawings in free-response items on the particulate nature of matter. *Journal of Science Education and Technology*, *32*, 549-566. https://doi.org/10.1007/s10956-023-10042-3

Lee, J., Suh, T., Roy, D., & Baucus, M. (2019). Emerging technology and business model innovation: The case of artificial intelligence. *Journal of Open Innovation: Technology, Market, and Complexity, 5*(3), Article 44. https://doi.org/10.3390/joitmc5030044

Li, Z., & Ironsi, C. S. (2025). The efficacy of generative artificial intelligence in developing science education preservice teachers' writing skills: An experimental approach. *Journal of Science Education and Technology*, *34*, 1260-1271. https://doi.org/10.1007/s10956-024-10148-2

Limna, P., Jakwatanatham, S., Siripipattanakul, S., Kaewpuang, P., & Sriboonruang, P. (2022). A review of artificial intelligence (AI) in education during the digital era. *Advance Knowledge for Executives*, *1*(1), 1-9. https://ssrn.com/abstract=4160798

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain intelligence: Go beyond artificial intelligence. *Mobile Networks and Applications, 23*, 368–375. https://doi.org/10.1007/s11036-017-0932-8

Martin, P. P., Kranz, D., Wulff, P., & Graulich, N. (2024). Exploring new depths: Applying machine learning for the analysis of student argumentation in chemistry. *Journal of Research in Science Teaching*, *61*(8), 1757-1792. https://doi.org/10.1002/tea.21903

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence, 171*(18), 1174–1182. https://doi.org/10.1016/j.artint.2007.10.009

Mittal, U., Sai, S., Chamola, V. & Sangwan, D. (2024). A comprehensive review on generative AI for education, *IEEE Access, 12,* 142733-142759. https://doi.org/10.1109/ACCESS.2024.3468368

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264– 269. https://10.7326/0003-4819-151-4-200908180-00135

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, *8*(5), 336-341. https://doi.org/10.1136/bmj.b2535

Monteiro, F. F., Souza, P. V. S., da Silva, M. C., Maia, J. R., da Silva, W. F., & Girardi, D. (2024). ChatGPT in Brazilian K-12 science education. *Frontiers in Education*, *9*, Article 1321547. https://doi.org/10.3389/feduc.2024.1321547

Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S., & Chu, S. K. W. (2023). A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, *28*, 8445-8501. https://doi.org/10.1007/s10639-022-11491-w

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, *55*, 1328-1353.https://doi.org/10.1111/bjet.13454

Nguyen, H., & Hayward, J. (2025). Applying generative artificial intelligence to critiquing science assessments. *Journal of Science Education and Technology*, *34*, 199-214. https://doi.org/10.1007/s10956-024-10177-x

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*, Article 42. https://doi.org/10.1186/s40561-023-00261-x

Oh, P. S., & Lee, G. G. (2025). Confronting imminent challenges in humane epistemic agency in science education: An interview with ChatGPT. *Science & Education*, *34*, 779-805. https://doi.org/10.1007/s11191-024-00515-1

Okulu, H. Z., & Muslu, N. (2024). Designing a course for pre-service science teachers using ChatGPT: What ChatGPT brings to the table. *Interactive Learning Environments*, *32*(10), 7450-7467. https://doi.org/10.1080/10494820.2024.2322462

Ramnarain, U., Ogegbo, A. A., Penn, M., Ojetunde, S., & Mdlalose, N. (2025). Pre-Service science teachers' intention to use generative artificial intelligence in inquiry-based teaching. *Journal of Science Education and Technology*, *34*, 1272-1285. https://doi.org/10.1007/s10956-024-10159-z

Resnik, D. B. (2011). What is ethics in research and why is it important. *National Institute of Environmental Health Sciences, 1*(10), 49–70. https://www.niehs.nih.gov/research/resources/bioethics/whatis

Sabourin Laflamme, A., & Bruneault, F. (2022). How to teach AI ethics? A pragmatist approach to ethical competency. *Pédagogie Collégiale, 36*(1), 61–67. https://eduq.info/xmlui/handle/11515/38675

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal, 31*(2), 47–53. https://ink.library.smu.edu.sg/sis_research/9371

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management*, *31*(2), 74-87. https://doi.org/10.4018/JDM.2020040105

Singer, A. E. (2010). Integrating ethics and strategy: A pragmatic approach. *Journal of Business Ethics*, *92*, 479-491. https://doi.org/10.1007/s10551-009-0176-z

Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and dimensions: A comparative analysis. *Scientometrics*, *126*, 5113-5142. https://doi.org/10.1007/s11192-021-03948-5

Stockman, C., & Nottingham, E. (2022). Surveillance capitalism in schools: What's the problem?. *Digital Culture & Education, 14*(1), 1–15. https://www.digitalcultureandeducation.com/s/Stockman_Nottingham_2022-mw5h.pdf

Taani, O., & Alabidi, S. (2025). ChatGPT in education: Benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology*, *56*(9), 1748-1777. https://doi.org/10.1080/0020739X.2024.2357341

Uyanik Aktulun, O., Kasapoglu, K., & Aydogdu, B. (2024). Comparing Turkish pre-service STEM and non-STEM teachers' attitudes and anxiety toward artificial intelligence. *Journal of Baltic Science Education*, *23*(5), 950-963. https://doi.org/10.33225/jbse/24.23.950

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, *15*, 398–405. https://doi.org/10.1111/nhs.12048

Wei, X., Wang, L., Lee, L. K., & Liu, R. (2024). Multiple generative AI pedagogical agents in augmented reality environments: A study on implementing the 5E model in science education. *Journal of Educational Computing Research, 63*(2), 336–371. https://doi.org/10.1177/07356331241305519

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, *9*, Article 59. https://doi.org/10.1186/s40594-022-00377-5

Yampolskiy, R. V. (2019). Predicting future AI failures from historic examples. *Foresight*, *21*(1), 138-152. https://doi.org/10.1108/FS-04-2018-0034

Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, *12*, Article e3489. https://doi.org/10.1002/rev3.3489

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, *2021*, Article 8812542. https://doi.org/10.1155/2021/8812542

Zhai, X., Haudek, K. C., & Ma, W. (2023). Assessing argumentation using machine learning and cognitive diagnostic modeling. *Research in Science Education*, *53*, 405-424. https://doi.org/10.1007/s11165-022-10062-w

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., Nyaaba, M., & Ma, W. (2025). Can generative AI and ChatGPT outperform humans on cognitive-demanding problem-solving tasks in science?. *Science & Education*, *34*, 649-670. https://doi.org/10.1007/s11191-024-00496-1

Zhang, W. X., Lin, J. J., & Hsu, Y. S. (2025). AI-assisted assessment of inquiry skills in socioscientific issue contexts. *Journal of Computer Assisted Learning*, *41*(1), Article e13102. https://doi.org/10.1111/jcal.13102

## Appendix A. The List of Reviewed Articles

| Article title | Reference |
|---|---|
| Exploring the integration of ChatGPT in inquiry-based learning: Teacher perspectives | Adeyele & Ramnarain, 2024 |
| Comparing Turkish pre-service STEM and non-STEM teachers' attitudes and anxiety toward artificial intelligence | Uyanik Aktulun et al., 2024 |
| Awareness regarding the implication of artificial intelligence in science education among pre-service science teachers | AlKanaan, 2022 |
| The readiness to use AI in teaching science: Science teachers' perspective | Alshorman, 2024 |
| In-service teachers' (mis)conceptions of artificial intelligence in K-12 science education | Antonenko & Abramowitz, 2023 |
| AI in the genetics classroom: A useful tool but not a replacement for creative writing | Behrens et al., 2025 |
| Using an artificial intelligence chatbot in scientific inquiry: Focusing on a guided-inquiry activity using inquiry bot | Chang et al., 2023 |
| Impact of AI robot image recognition technology on improving students' conceptual understanding of cell division and science learning motivation | Chen & Liu, 2024 |
| Students' holistic reading of socio-scientific texts on climate change in a ChatGPT scenario | Cheung et al., 2024 |
| Exploring students' multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools | Cheung et al., 2025 |
| Examining science education in ChatGPT: An exploratory study of generative artificial intelligence | Cooper, 2023 |
| Pixels and pedagogy: Examining science education imagery by generative artificial intelligence | Cooper & Tang, 2024 |
| Science teacher perceptions of the state of knowledge and education at the advent of generative artificial intelligence popularity | Garofalo & Farenga, 2025 |
| Enhancing computational thinking skills through Artificial intelligence education at a STEAM high school | Huang & Qiao, 2024 |
| The impact of AI chatbot-supported guided discovery learning on pre-service teachers' learning performance and motivation | Huang et al., 2025 |
| A double-edged sword: Physics educators' perspectives on utilizing ChatGPT and its future in classrooms | Jang & Choi, 2025 |
| Towards effective argumentation: Design and implementation of a generative AI-based evaluation and feedback system | Jho & Ha, 2024 |

| Article title | Reference |
|---|---|
| Using video-elicitation focus group interviews to explore pre-service science teachers' views and reasoning on artificial intelligence | Karahan, 2023 |
| Using artificial intelligence to predict students' STEM attitudes: An adaptive neural-network-based fuzzy logic model | Göktepe Körpeoğlu & Göktepe Yıldız, 2024 |
| Collaborative learning with artificial intelligence speakers | Lee et al., 2024 |
| Using ChatGPT for science learning: A study on pre-service teachers' lesson planning | Lee & Zhai, 2024 |
| Automated assessment of student hand drawings in free-response items on the particulate nature of matter | Lee et al., 2023 |
| Improving science conceptual understanding and attitudes in elementary science classes through the development and application of a rule-based AI chatbot | Lee et al., 2023 |
| The efficacy of generative artificial intelligence in developing science education preservice teachers' writing skills: An experimental approach | Li & Ironsi, 2025 |
| Exploring new depths: Applying machine learning for the analysis of student argumentation in chemistry | Martin et al., 2024 |
| ChatGPT in Brazilian K-12 science education | Monteiro et al., 2024 |
| Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study | Ng et al., 2024 |
| Applying generative artificial intelligence to critiquing science assessments | Nguyen & Hayward, 2025 |
| Adoption of artificial intelligence in science teaching: From the vantage point of the African science teachers | Nja et al., 2023 |
| Confronting imminent challenges in humane epistemic agency in science education: An interview with ChatGPT | Oh & Lee, 2025 |
| Designing a course for pre-service science teachers using ChatGPT: What ChatGPT brings to the table | Okulu & Muslu, 2024 |
| Pre-service science teachers' intention to use generative artificial intelligence in inquiry-based teaching | Ramnarain at al., 2025 |
| ChatGPT in education: Benefits and challenges of ChatGPT for mathematics and science teaching practices | Taani & Alabidi, 2025 |
| Chatbot application in a 5th grade science course | Deveci Topal et al., 2021 |
| Multiple generative AI pedagogical agents in augmented reality environments: A study on implementing the 5E model in science education | Wei et al., 2024 |

| Article title | Reference |
|---|---|
| Applying machine learning to automatically assess scientific models | Zhai et al., 2022 |
| Assessing argumentation using machine learning and cognitive diagnostic modelling | Zhai et al., 2023 |
| Can generative AI and ChatGPT outperform humans on cognitive-demanding problem-solving tasks in science? | Zhai et al., 2025 |
| AI-assisted assessment of inquiry skills in socio-scientific issue contexts | Zhang et al., 2025 |