frontiers | Frontiers in Education

TYPE Systematic Review
PUBLISHED 23 September 2025
DOI 10.3389/feduc.2025.1647573

OPEN ACCESS

EDITED BY
Konstantinos T. Kotsis,
University of Ioannina, Greece

REVIEWED BY
Leonidas Gavrilas,
University of Ioannina, Greece
Heni Mulyani,
Universitas Pendidikan Indonesia, Indonesia

*CORRESPONDENCE
Abdullah Alfarwan
✉ amalfarwan@nu.edu.sa

RECEIVED 15 June 2025
ACCEPTED 19 August 2025
PUBLISHED 23 September 2025

CITATION
Alfarwan A (2025) Generative AI use in K-12 education: a systematic review.
Front. Educ. 10:1647573.
doi: 10.3389/feduc.2025.1647573

Generative AI use in K-12 education: a systematic review

Abdullah Alfarwan*

Department of Education and Psychology, Najran University, Najran, Saudi Arabia

As generative artificial intelligence (GenAI) continues to generate interest and impact all levels of the educational system, including K-12, it is essential to understand how GenAI has been utilized at the K-12 level. This study explored the literature to understand what GenAI has been used for in K-12 and the reported impacts of its use. By investigating two prominent databases, Scopus and Web of Science, and employing snowballing techniques, this study included 30 papers after a rigorous search and applying inclusion and exclusion criteria. The findings indicate that most existing works focused on high school and preservice teachers. Researchers have explored various subject domains or areas of interest where GenAI has been used, with a significant focus on STEM-related subjects. The quasi-experimental approach emerged as the most commonly used design technique for implementing GenAI ideas in the classroom. Psychological variables were the primary measures used to gauge learning outcomes. This study further aggregates recommendations from prior works and suggests areas for future research. This paper contributes to efforts in creating frameworks and building resources needed to realize the full potential of GenAI in shaping the educational landscape, especially within the school system.

KEYWORDS

generative AI, ChatGPT, K-12 education, teachers, students

1 Introduction

Globally, discussions about generative artificial intelligence (GenAI) and its implications for the teaching and learning process have intensified, especially with ethical issues remaining central to the discourse. Another prominent topic is the lack of critical thinking among students and the reduction in problem-solving skills (Aruleba et al., 2023). While GenAI holds positive promise for education, its negative effects are also evident. Therefore, it is crucial to design policies, guidelines, and approaches to ensure responsible use of AI systems. Some popular GenAI tools include various versions of ChatGPT, Copilot, Midjourney, and Google Bard. Yu (2024) has highlighted several opportunities for ChatGPT in educational applications, including advancing educational digitization, individualized and self-directed education, pedagogical efficacy and quality, and pedagogical assessment and responsiveness. Due to its increasing relevance, GenAI has been applied in various areas such as designing learning activities, teaching, and assessment.

While substantial efforts have been made to aggregate how to teach AI and use AI to facilitate learning, there is a dearth of work within the K-12 level (Jauhiainen and Garagorry Guerra, 2024) that synthesizes empirical accounts of students' and teachers' nuanced processes of using GenAI for learning and teaching. Besides the use of GenAI for out-of-classroom duties such as lesson preparation, assessment, and administrative tasks, less has been seen in the literature about AI tools used for real-time teaching and learning, as commonly observed with other educational technologies integrated into the classroom. For instance, studies like Nyaaba et al. (2024) developed a GenAI Culturally Responsive Science Assessment (GenAI-CRSciA) framework but did not evaluate it with teachers or

students, and thus were not considered in this review. This review analyzed empirical research that shows clear use or application of any GenAI platform in teaching or learning within the K-12 education system, whether executed in formal, non-formal, or informal settings.

Research has been conducted to understand what K-12 teachers think about and do with GenAI, such as ChatGPT (Hays et al., 2024). These results merely highlight the concerns and potential impact of GenAI tools on education without specifically showing how the AI tools are incorporated into classes. Aggregating evidence of the use of GenAI tools across contexts and different domains within the K-12 context will provide examples of best practices and help researchers effectively bring GenAI to support teaching and learning practices within the school system.

Educators are seeking support and professional development opportunities to adopt GenAI for their instructional tasks (Hays et al., 2024). Investigating primary studies on GenAI use in teaching and learning practices would be helpful as a guide for teachers and curriculum drafters. Having evidence of how these tools work in practice will further aid in developing effective policies and guidance for their use. This study offers implications for the school education system by highlighting how GenAI tools have been adopted to enhance teaching and learning, which would be valuable in developing support systems and targeted policies to facilitate teachers' GenAI integration.

1.1 Generative AI in K-12 education

Generative AI has been around for more than a decade. However, the launch of ChatGPT in late 2022 significantly contributed to the popularity of GenAI. Since then, there has been a great deal written about different GenAI tools, especially ChatGPT, with various versions emerging due to continuous advancements in technology. GenAI has gained more prominence within the higher education system. Since GenAI is open to everyone with a freemium version available, and educators and researchers realizing the necessity to incorporate it into instructional practices, there has been some effort to ground GenAI in the K-12 context.

While GenAI tools such as ChatGPT have been touted as gateways to invaluable resources for educators, aiding in the differentiation of lessons, activities, reading passages, and formative assessments (Culverhouse, 2024), there is a need to document evidence of how these AI tools have been used across different contexts and domains. Although the practical implementation, testing, and assessment of generative AI in primary and secondary education remain largely unexplored (Jauhiainen and Garagorry Guerra, 2024), the available literature provides an understanding of researchers' and practitioners' efforts to incorporate GenAI into the school system. While teachers use GenAI tools for materials development, what else is there? It is useful to know how GenAI tools are being implemented beyond content generation and task refinement.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

1.2 Extant reviews

Several reviews have been conducted on empirical research regarding GenAI in education. However, they do not specifically focus on K-12 education, with their reports indicating limited work within the K-12 system. For instance, Yusuf et al. (2024) conducted a review on GenAI in education and research generally, and one of their key findings is the paucity of research and discussions on GenAI in K-12 education. However, the study did not provide information about existing work within the K-12 context. In a recent review focusing on the K-12 context, Zhang and Tur (2024) synthesized findings from 13 selected papers, encompassing a strengths, weaknesses, opportunities, and threats analysis of ChatGPT's implementation in K-12 education.

While the existing work provides an important start to building resources on how to support student learning and teacher instruction using GenAI, some limitations are involved: (1) the papers did not focus specifically on GenAI use within the K-12 context; (2) the studies are within a refined time span; (3) most of the studies restrict the search to only ChatGPT without considering other GenAI tools; and (4) the earlier efforts did not focus on how GenAI tools have been utilized in teaching and learning. Given the importance of understanding the current efforts to use GenAI in K-12 to address potential issues, this study investigates the literature to contribute to knowledge on how GenAI has been utilized in the school system in terms of approaches, study design, learning outcomes, and bottlenecks.

Considering the limitations identified in earlier studies, the purpose of this review is twofold: (1) to systematically review high-quality empirical studies on the use of GenAI in K-12 education from both research and instructional perspectives, and (2) to explore future directions for research and instruction in GenAI for K-12 education based on the reviewed papers. Thus, this paper systematically reviews the empirical evidence on GenAI use in K-12 education to answer the following two research questions:

1. What are the characteristics of research conducted on the use of GenAI in K-12 education?
2. How have studies on GenAI for K-12 education been designed and implemented?

2 Methodology

2.1 Data collection

This study follows PRISMA and Kitchenham guidelines for systematic reviews (Kitchenham and Charters, 2007). As a systematic review of primary studies, we collected data for this research from two prominent databases, including snowballing, to ensure relevant papers were included. We specifically retrieved papers from Scopus and Web of Science. These two databases are recognized as platforms that house high-quality research studies across subjects, especially STEM-related subjects that align with the interest of this study. The focus of this research—how GenAI has been utilized to learn or teach within the K-12 context—can be answered using data from both databases. To further validate the inclusion of relevant and published papers, snowballing was performed by sampling and reading through the references of the selected papers. The technique adopted for snowballing in this paper

involved carefully reading the selected papers and identifying literature cited by the authors on GenAI use in school education, especially in sections where the authors describe how their study builds on prior works or what they did differently. The specific details and number of retrieved articles can be found in the subsequent sections.

2.1.1 Inclusion criteria

Only scientific and empirical articles on the use of GenAI to teach and learn within the K-12 context were considered. We included only conference and journal papers that reported workshops or classroom interventions using any GenAI tools. This study also considers participants in vocational institutions (e.g., Florido and Hernández-Leo, 2024) because their participants share some similar characteristics with K-12 students.

2.1.2 Exclusion criteria

Articles were excluded if they focused on teaching AI literacy, experimenting with how GenAI works, or natural language processing (e.g., Noveski et al., 2024; Katuka et al., 2024). We also excluded studies that gathered the perceptions of students, teachers, or other education stakeholders (e.g., Cheah and Kim, 2024) because this study is interested in how ChatGPT has been used. Within the pre-service education context, which falls within the scope of this research as they are trained as future teachers, we excluded studies considering factors influencing their adoption of GenAI in teaching (e.g., Hu et al., 2025).

Additionally, we did not include papers that examine the priorities and challenges identified by senior educational leaders and policymakers regarding the responsible and ethical use of GenAI in school education (e.g., Bower et al., 2025). While collecting policymakers' perspectives is valuable, this study is concerned with experimenting with GenAI with teachers and students, and we believe analyzing such studies will not provide the necessary evidence. Papers that compare different versions of GenAI and their ability to generate content appropriate for educational levels (e.g., Karaca, 2024) or performance on PISA multiple-choice sample questions (Takami, 2023) were excluded. Papers that collected teachers' perspectives after a ChatGPT training program, such as the integration of ChatGPT into primary schools (Uğraş et al., 2024a), but did not report its implementation with students or effectiveness based on study evaluation, were not considered relevant for this study's goals. Conceptual papers (e.g., Blonder et al., 2024) were also excluded.

2.2 Search strategies

To arrive at the appropriate search strings to answer the questions guiding this research, we specifically utilized search terms that reflect the study aims. We searched for generative AI and considered some major GenAI tools such as ChatGPT, Google Bard, Copilot, and Midjourney. This approach differs from Zhang and Tur's (2024) review, which focused solely on ChatGPT. To achieve a more streamlined search and return relevant papers, we tested combinations of several terms, including GenAI terms and K-12 education terms or synonyms. Below are examples of our search attempts:

1. "Generative AI" OR ChatGPT OR "Google Bard" OR "Copilot" OR "Midjourney" AND "K-12" OR school OR children.
2. "Generative AI" OR ChatGPT OR "Google Bard" OR "Copilot" OR "Midjourney" AND "K-12" OR school.

The search in (1) above returned 1,058 documents in Scopus and 2,687,952 documents in the Web of Science. While it appeared more comprehensive, upon vetting the papers based on their titles, we discovered that most of them were from the surgery, medicine, or healthcare sectors (e.g., Miyake et al., 2025), likely due to the inclusion of "children" in the search string. The search in (2) returned 745 documents in Scopus and 767,713 documents in the Web of Science, but still contained mostly documents outside the scope of this review (e.g., Jones, 2025). After several trials with different search combinations, the following search string was used because it returned articles within the scope of the study:

> "Generative AI" OR GenAI OR ChatGPT OR "Google Bard" OR "Copilot" OR "Midjourney" AND "K-12 education" OR "school education."

As shown in Table 1, using the search string in both the Scopus and Web of Science databases returned 58 articles and 22,660 articles, respectively. After applying the inclusion and exclusion criteria, 30 articles were eligible for analysis, including 5 retrieved through the snowballing effort. The search was performed on February 19, 2025, and was conducted in the databases without limitation to a specific timeline to capture all relevant papers for the study's aim. Most of the papers retrieved from Web of Science were removed because they either focused on higher education or adult education or met any of the exclusion criteria. See Figure 1 for more details on the article screening process following the inclusion and exclusion criteria.

After selecting the 30 articles, data were extracted from the papers based on the study's aim and research questions and entered into an Excel worksheet for analysis. The extracted information included the year of publication, title of the articles, document type, grade level, subject/focus of the study, design, measures, GenAI tools, country representation, learning activity, and main outcomes.

3 Result

The analysis of the review presented in this section is based on the research questions guiding the study.

3.1 RQ. 1. What are the characteristics of research conducted on the use of GenAI in K-12 education?

The characteristics covered in this section include educational level covered in the selected studies, type of topics covered,

TABLE 1 Database's search and articles analyzed.

Database	Search returned	Article analyzed
Scopus	58	10
Web of Science	22,660	15
Snowballing	5	5

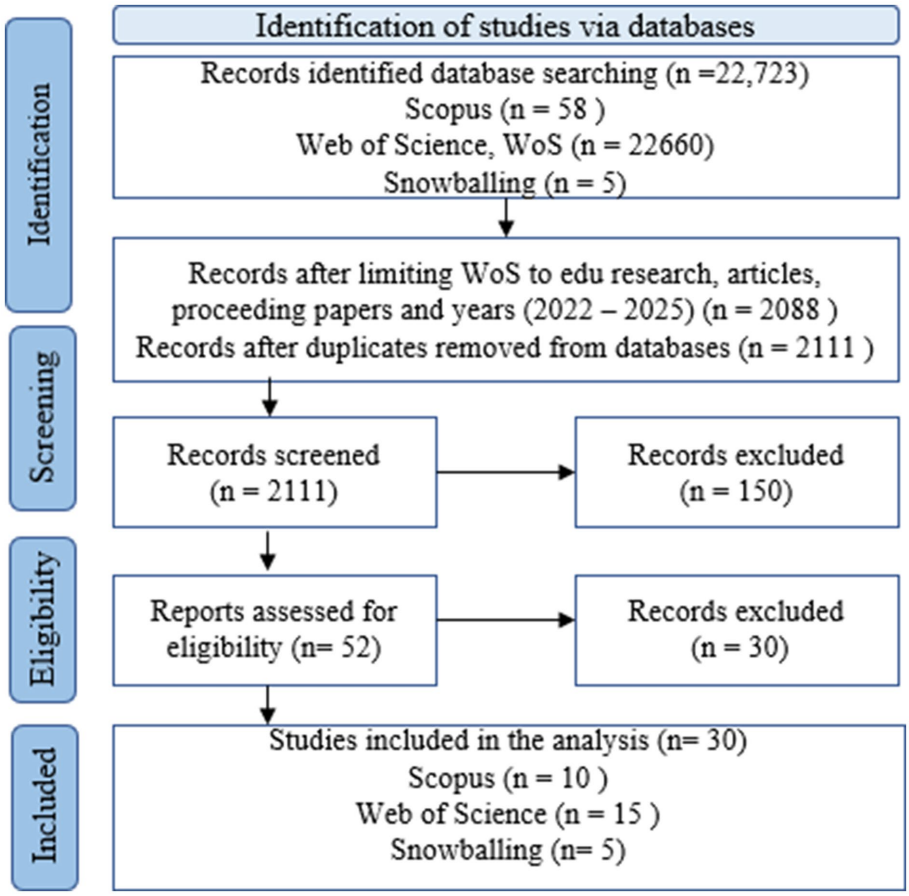

FIGURE 1
PRISMA flow diagram of the review based on database and snowballing.

country of study, GenAI used and publication outlet (journal or conference).

3.1.1 Year of publication

Figure 2 presents the years of publication. The papers we reviewed span 3 years, from 2023 to 2025. Most publications (19 out of 30) were from 2024 (Jauhiainen and Garagorry Guerra, 2024; Jang et al., 2024; Wang et al., 2024; Yuwono et al., 2024; Florido and Hernández-Leo, 2024; Han et al., 2024; Ding and Chiu, 2024; Wu et al., 2024; Tang et al., 2024; Wang et al., 2024; Ng et al., 2024; Levine et al., 2024; Gong et al., 2024; Dilling et al., 2024; Noster et al., 2024; Liao et al., 2024; Lu et al., 2024; Li and Ironsi, 2024). Six papers (Chen and Law, 2025; Song et al., 2025; Galiç et al., 2025; Lee et al., 2025a; Lee et al., 2025b; Min et al., 2025; Yang et al., 2025a) were retrieved in 2025 as of February 2025, indicating that more works are likely to be published along these lines. It is expected that papers will be limited to these years since ChatGPT (GenAI) gained prominence in late 2022. Based on the retrieved data, five papers (Jauhiainen and Guerra, 2023; Chen and Zhu, 2023; Vartiainen et al., 2023; Bitzenbauer, 2023; Alneyadi and Wardat, 2023) were published in 2023.

3.1.2 Educational level covered in the study

Figure 3 highlights the representation of studies focusing on different educational levels. The findings show that both high school students and pre-service teachers are largely covered. Middle schools have a good representation, while about two papers each focused on all students regardless of their grade bands, primary, elementary, and secondary schools. Vocational education and in-service teachers were barely targeted.

It is interesting to note that in-service teachers, who are professionals and should be provided with professional development opportunities and support to integrate innovative AI tools in their classrooms, are less represented in the reviewed papers. One of the papers that focused on elementary education involved parents, teachers, and students in their studies, making it the only paper that considered parent involvement in teaching and learning using GenAI.

3.1.3 Representation of countries studies were conducted

The findings based on Figure 4 indicate that most of the studies were conducted in Taiwan and China, with four papers each. The United States and Germany also have three papers each, followed by Korea, Australia, and Uruguay with two papers each. Overall, more of these studies were conducted in Asia (Taiwan, China, Korea, South Korea, Hong Kong). There are also representations from European countries. Only one South American country (Colombia) was part of our data, and there was no representation from Africa.

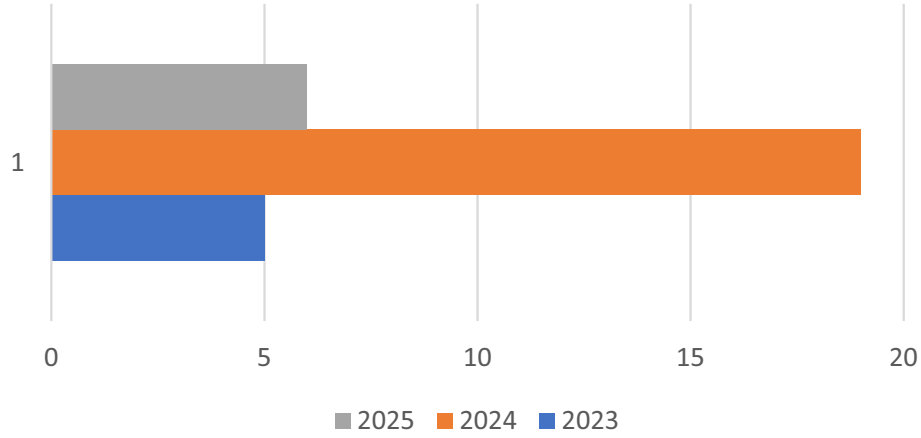

FIGURE 2
Year of publication.

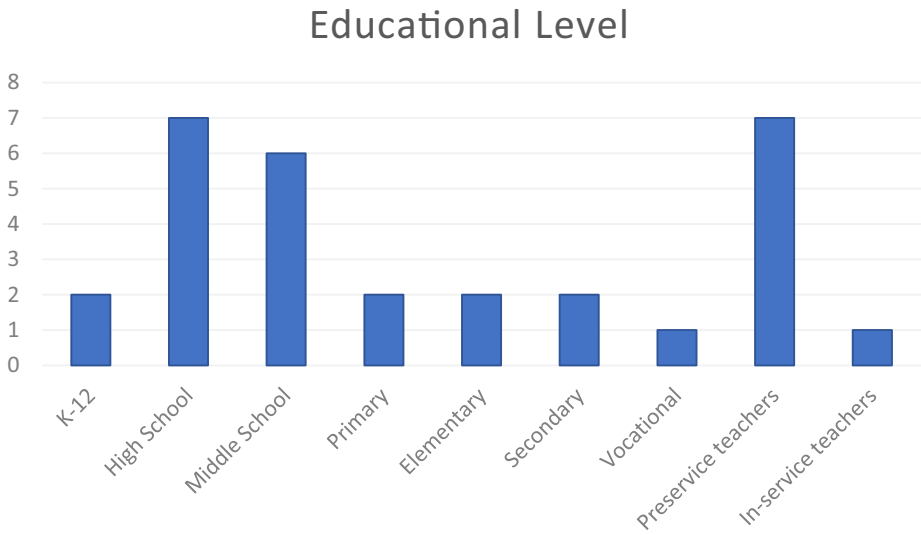

FIGURE 3
Educational level.

In the reviewed papers, we identified various subjects or areas of interest where GenAI has been used to facilitate learning, as shown in Figure 5. Mathematics was the most common subject (5 papers), followed by Physics (4 papers). Three papers evaluated the effectiveness of using ChatGPT to support high school students in programming, including writing projects. Overall, most of the papers focused on STEM-related subjects. Additionally, two papers explored how to teach AI to K-12 students. Some studies also investigated specific ideas such as knowledge building, higher-order thinking, help-seeking behaviors, and dialogic pedagogies. Non-STEM subjects, including History, Writing, and English Language, were also examined in the papers.

3.1.4 GenAI application used in the study

Based on Figure 6, various GenAI tools were used in the reviewed studies. However, ChatGPT was predominantly adopted in almost all the studies. For instance, 21 papers used only ChatGPT to understand the capabilities of GenAI in enhancing learning across various subjects or content areas. In other papers, ChatGPT was used in conjunction with tools such as Copilot, DALL-E, Bert, and Canva. Some studies utilized the ChatGPT API to power their own tools, such as integrating multimodal GenAI by leveraging OpenAI's GPT-4 and DALL·E 3.

3.1.5 Publication outlet and venues

Most of the papers that experimented with the capability of ChatGPT in practice for the K-12 population were published primarily (87%) in journals. As expected, the papers are mainly published in educational technology and AI education-related journals. As shown in the Appendix, some of these journals include Computers & Education: Artificial Intelligence, Education and Information Technologies, Contemporary Educational Technology, and IEEE Transactions on Learning Technologies.

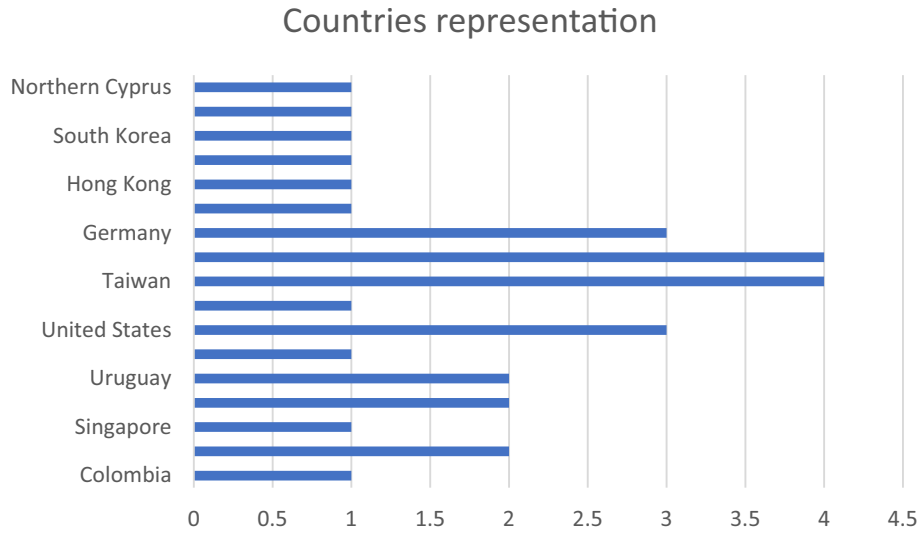

FIGURE 4
Countries representation.

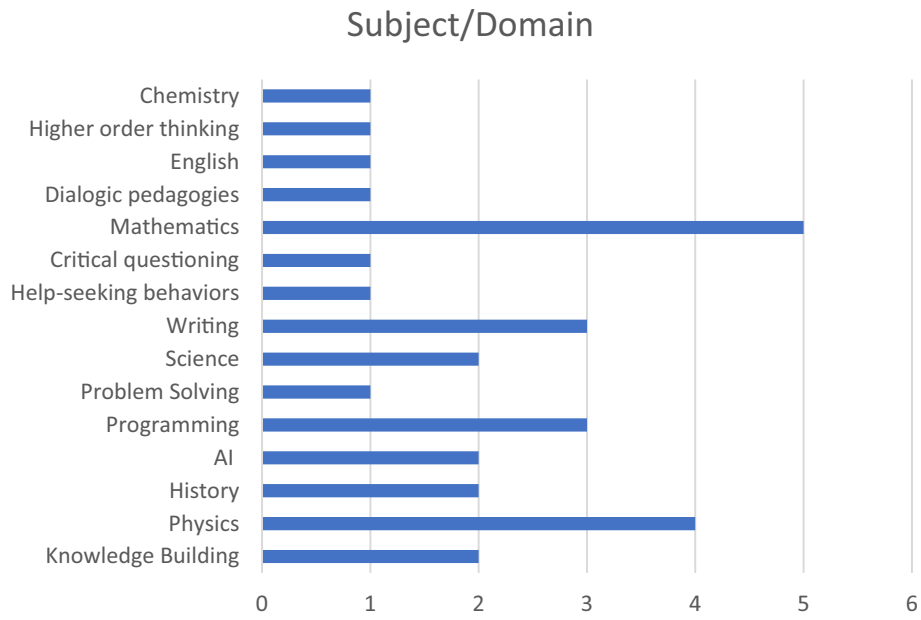

FIGURE 5
Topics focused on in the studies.

3.2 R2. How has studies on GenAI for K-12 education been designed and implemented?

This section presents the design approaches employed in the study, including the measures used by the researchers, which directly connect to the learning outcomes of those studies. Examples of learning activities performed by the researchers and/or instructors are provided to illustrate how some of the activities were carried out, along with their recommendations for future research.

3.2.1 Study design

The study design uncovered in this review pertains to the design approaches adopted by the reviewed studies. The measures used in these studies were also collected. It is important to highlight the design process of the identified studies to understand the procedures involved in learning with ChatGPT, as well as the elements measured during the teaching and learning process.

3.2.2 Design approach

Based on the reviewed papers, 9 out of 30 (30%) utilized a quasi-experimental approach to implement their ideas of GenAI

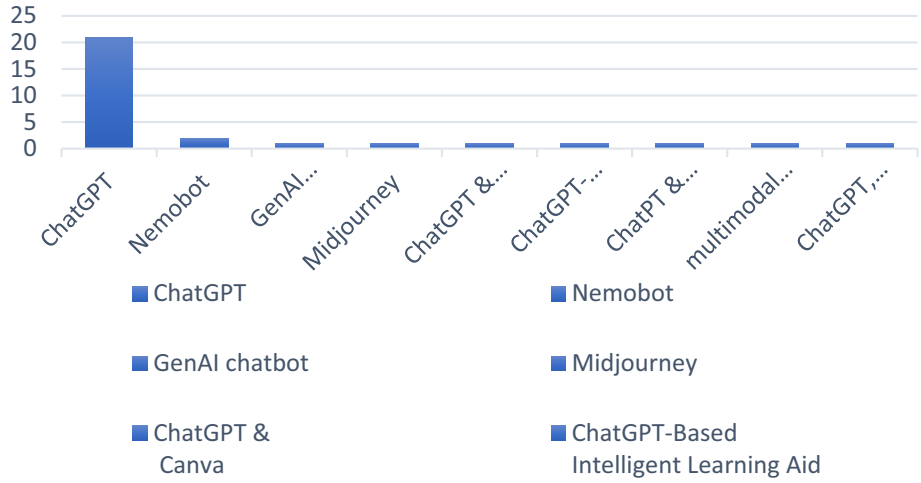

FIGURE 6
GenAI used in the study.

in the classroom. As illustrated in Figure 7, papers employing descriptive and mixed-method approaches each accounted for 20% of the analyzed data, while exploratory and case study approaches each represented 10% of the selected works. Additionally, one paper each used experimental, action research, or qualitative research approaches. All these methodologies are valuable and complement each other, providing comprehensive insights into the functionalities of GenAI for teaching and learning.

3.2.3 Measures

Researchers are interested in various measures to understand the feasibility of GenAI in classrooms for teaching and learning. These measures, primarily psychological variables, include attitudes toward AI and creativity, cognitive trust, affective trust, perceptions of usability, and continuance intention to use (Ding and Chiu, 2024); self-regulation progress using motivation, engagement, and self-efficacy as indicators (Wu et al., 2024); trust, technology acceptance, and satisfaction with ChatGPT explanations (Wang et al., 2024); self-regulated learning (Ng et al., 2024); critical thinking and problem-solving (Gong et al., 2024); scientific inquiry interaction with an AI inquiry assistant (Min et al., 2025); computational thinking and problem-solving skills (Liao et al., 2024); higher-order thinking (Lu et al., 2024); problem-solving and task completion (Feldman-Maggor et al., 2024); and flow experience, self-efficacy, and learning achievement (Yang et al., 2025b). We identified only two studies that specifically consider how GenAI supports self-regulated learning.

Aside from surveys and interviews using these psychological variables, the studies do not consider objective measures of student learning outcomes, such as standardized tests or quizzes. These measures provide insight into the outcomes resulting from exploring GenAI in teaching and learning settings. For instance, Yang et al. (2025a) found that using ChatGPT in their experimental design study reported lower levels of flow experience, self-efficacy, and learning achievement compared to those utilizing conventional methods.

3.2.4 Learning activities and experience

As previously mentioned, each of the identified papers focuses on specific ideas and demonstrates how GenAI can be used to teach those concepts. While not all the papers provide detailed walkthroughs of the processes, some of the activities presented in the literature can be helpful for teaching and learning with GenAI by following existing examples. For instance, Ding and Chiu (2024) used "Sustainable Energy" as a collaborative learning project, where participants utilized various generative AI tools to collect data and organize knowledge. Jang et al. (2024) found that structuring high school physics class content to incorporate ChatGPT for analyzing experimental data helped students analyze data and solve problems.

An example of an intervention study is provided by Vartiainen et al. (2023). The project, conducted in a middle school context, consisted of a series of three workshops, each lasting 3 hours and held over the course of 3 weeks. The educational strategy was designed to introduce students to artificial intelligence (AI) through three distinct perspectives. The third workshop focused on generative AI. It began with an overview of image recognition systems, explaining how web scrapers collect image-text pairs to train large-scale models. These datasets, originally intended for recognition tasks, also serve as the foundation for text-to-image generation. A popular generative AI tool, Midjourney, was used to demonstrate this capability. Students were introduced to the basics of prompt crafting and shown visual examples of prompts alongside the images generated by Midjourney. Following this, students were divided into small teams and provided with computers to collaboratively explore and experiment with the tool, allowing them to engage hands-on with generative AI concepts.

Generally, most studies do not detail the learning activities in which GenAI has been used. It is important to present these details to serve as resources for other researchers or practitioners to adopt in their classrooms or courses. To avoid redundancy, this paper has not detailed the GenAI activities reported in the reviewed studies. Instead, it highlights each of the primary studies, making it easier

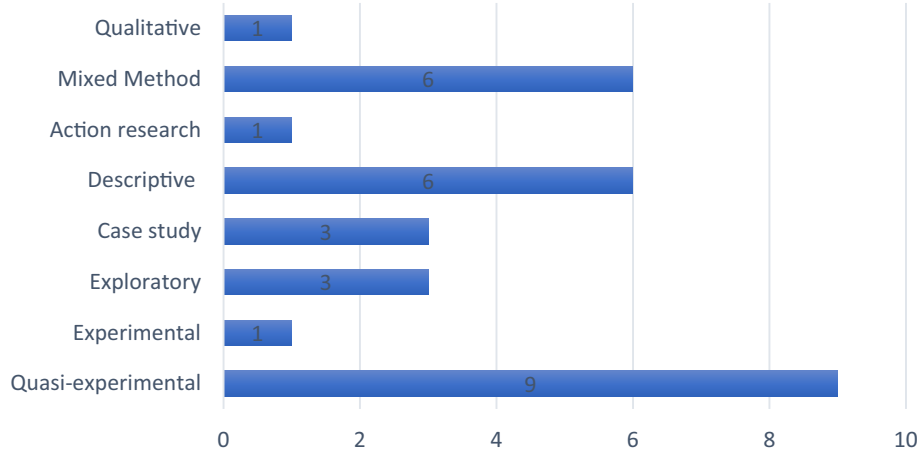

FIGURE 7
Design approaches.

for other researchers to identify which papers to consult for specific content areas or topics they are interested in using GenAI to facilitate.

3.2.5 Recommendations

Nearly all the reviewed papers (90%) emphasize the need for teachers' professional development for AI-integrated classrooms (e.g., Wang et al., 2024; Song et al., 2025). These papers highlight the importance of developing instructional approaches to build substantial understanding and experience with subject-specific AI applications through authentic learning tasks and fostering AI literacy (Song et al., 2025). In addition to teacher learning opportunities, assessing learning outcomes before and after engaging with ChatGPT (Wang et al., 2024) is recommended. Wang et al. (2024) suggest investigating the cognitive and emotional experiences of both researchers and teachers using large language models for teacher learning purposes. Chen and Zhu (2023) emphasize examining the impact of student-GenAI partnerships on learning and investigating these partnerships' effects on various learning outcomes in diverse settings.

Future research should explore not only the direct, short-term impacts of GenAI but also its indirect and long-term effects on creativity and critical thinking across different domains (Song et al., 2025). Researchers also note the benefit of conducting pre-tests to better gauge the effectiveness of GenAI intervention initiatives (Wu et al., 2024). Jauhiainen and Guerra (2023) call for longitudinal studies on the topic, including research-based evidence about generative AI's role and impacts on pupils' learning outcomes, especially when learning material has been adapted with GenAI. Based on Han et al. (2024) study results, they suggest that educational GAI platforms should support customizable teacher-in-the-loop systems to enhance the trustworthiness and content-focus of GAI systems. Experimenting with various GenAI tools may be necessary due to the occasional inaccuracies of GPT-3.5's responses (Wu et al., 2024) and the evolution of GPT and GenAI tools.

4 Discussion

GenAI has arrived, and its implications for teaching and learning, including K-12 education, cannot be overstated. K-12 education lays the foundation for future education and career success by preparing students with the knowledge, skills, and confidence needed for college and careers. It is crucial to support this demographic with emerging technologies and opportunities to ensure they are future-ready citizens. To provide timely insights into how GenAI is being used to enhance teaching and learning in K-12 settings, this paper reviews empirical and primary studies on the adoption of GenAI across the K-12 context. This article offers a window into the GenAI for K-12 world, showcasing how researchers, K-12 educators, and practitioners have used GenAI in their instructional practices, which is valuable in addressing the question of how GenAI can be effectively integrated into the school system. This section further highlights and discusses the main findings from the reviewed papers.

4.1 Characteristics of educational level

Based on the 30 selected papers, none focused on early childhood education (ECE). The papers that were identified but did not meet the inclusion criteria evaluated the integration of AI-supported ChatGPT into ECE STEM education from the teachers' perspective (Uğraş et al., 2024b; Su and Yang, 2024). Although Uğraş et al.'s work involved training 18 teachers on the use of ChatGPT, there was no evaluation of the effectiveness of the instructional exercise or how the pupils benefited from learning with the GenAI tool. Su and Yang also explored teachers' views to understand the potential benefits and challenges of using ChatGPT in early childhood education. More research is needed within the ECE context to provide insights into how GenAI can improve pupils' learning and teachers' teaching.

4.2 Parental involvement in learning with GenAI

Our results show that only one paper (Han et al., 2024) considers parents in exploring GenAI use for learning. The study included parents as part of the research demographic, which also comprised teachers and students. Parental involvement in students' learning has been considered valuable. Oyedoyin et al. (2024) observed the role of parents as a factor influencing learning and development in computing-related subjects. Oyedoyin et al.'s observation aligns with prior studies highlighting the importance of parental involvement in eradicating ICT illiteracy and increasing educational equality (Passaretta and Gil-Hernández, 2023), as well as how parental involvement in their children's learning positively affects learners' motivation, attitude, and academic achievement (Zedan, 2021). Given the value of parental involvement in learning, more research should explore the use of GenAI with parents and their children.

4.3 Teachers' use of GenAI

In this review, only one paper focused on how teachers have implemented GenAI in practice and evaluated its use. Prior works (e.g., Yang et al., 2025a) aimed to integrate GenAI into elementary literacy education by prompting ChatGPT to generate high-quality questions during a read-aloud. While Yang et al.'s study holds implications for GenAI in schools, evaluating the activities and how they enhance teaching and learning is important. The author opines that GenAI facilitates teaching and learning and can be a powerful support for teachers, but agency and decision-making must remain with the teacher. Yang et al. (2025a) asserts that unless prompted, GenAI does not know the students, backgrounds, and cultural context, making the teacher critical in implementing teaching and learning with GenAI tools. Chiu's (2024) interesting contribution also reported from the teacher's perspective the learning activities that can foster self-regulated learning (SRL) processes and self-determination theory (SDT) needs. However, how these activities support SRL and SDT in practice remains unseen. While it is a consensus that the use of any technological tool relies on teachers' expertise and initiatives, there is a need for experimenting and evidence of real-time use of GenAI tools within the school level.

4.4 Use of GenAI in different subject domains

As shown in the findings section, GenAI has been utilized in various subject domains, predominantly within STEM fields in the K-12 context. For instance, Jang and Choi (2024) explore physics educators' perceptions of the educational use of ChatGPT and future changes in society and the educational environment. Beyond the identified subject domains, there is a growing effort to utilize GenAI across disciplines. Researchers have explored the transformative potential of GenAI in geography education, focusing on its impacts on curriculum, pedagogy, and assessment through the lens of the SAMR (Substitution, Augmentation, Modification, and Redefinition) model. Evaluating the effectiveness of metrics or hypothesized approaches is needed. Feldman-Maggor et al. (2024) discuss how GenAI can be integrated into chemistry education using the TPACK framework and prompt engineering, emphasizing the types of knowledge teachers need to apply GAI effectively and the need to further develop theoretical frameworks for teachers' knowledge in the age of GAI. Clark-Fookes (2024) also discusses the possibilities of GenAI use in drama education. It is evident that GenAI has implications for all disciplines, regardless of their focus. However, detailed reports or real-time use and evaluation of its usability based on defined metrics are necessary to affirm most of the conceptual and position papers on the subject. Overall, our findings indicate that more work is required on the use of AI across multiple subject domains within the K-12 context.

4.5 Design and implementation

It is interesting to note that most of the identified studies utilized a quasi-experimental approach in designing their research. This approach allowed researchers to measure the impact of their intervention and ascertain the effectiveness of GenAI use over traditional methods or other alternatives. More research using true experimental designs is necessary. Experimental research is considered a powerful tool for establishing causal relationships between educational interventions and student outcomes, allowing researchers to test the effectiveness of new teaching methods or programs by randomly assigning students to control and experimental groups. In measuring students' knowledge and learning with GenAI, the majority of studies have designed and developed self-reported questionnaires to assess their learning and understanding using self-perceptions, which may not be accurate measures. Future work should consider objective measures such as test items. Using self-reported measures in research such as surveys or interviews can limit both the validity and generalizability of findings compared to objective measures like standardized tests. Self-reports are vulnerable to biases such as social desirability and inaccurate recall, which can distort the data and threaten construct validity, the extent to which a tool accurately measures the intended concept (Paulhus and Vazire, 2007). Additionally, variability in how individuals interpret questions can introduce measurement error, reducing internal validity (Clark and Watson, 2024). From a generalizability standpoint, self-reports may not perform consistently across different populations or cultural contexts, limiting external validity. For example, cultural norms can influence how people report emotions or behaviors, making it difficult to apply findings broadly (Paulhus and Vazire, 2007). In contrast, objective measures offer standardized, replicable data that are less prone to individual interpretation, enhancing both the reliability and validity of research outcomes (Clark and Watson, 2024).

Additionally, the activities implemented in research studies on how GenAI supports teaching and learning in the classroom should ensure scalability. Valcea et al. (2024) assert that without sensible AI use guidelines, AI is likely to have a net-negative effect on learning based on trials with ChatGPT on various cognitive tasks organized around Bloom's Taxonomy of learning. Valcea et al.'s exposition further shows the need for more trials to understand how AI tools can be effectively used to enhance learning. Researchers have reported a dearth of published work on detailed reports of GenAI use within normal school curriculum contents and its testing with pupils in the real learning context of school classes (Jauhiainen and Guerra, 2023).

Nevertheless, there are some efforts in this regard, conducting research with school pupils and populations within the K-12 context. This study reports on the synthesized work produced by researchers and practitioners attempting to enhance their teaching and students' learning using different GenAI tools.

4.6 Study limitations and future research

There are other GenAI tools that were not covered in our search terms, which may have been explored by researchers or practitioners. For instance, Microsoft Bing, Google Bard, Adobe Firefly, Education Copilot, Khanmigo, Magic AI, Midjourney, Research Rabbit, and SlidesAI. Future research should explore these tools or other commercial or open-source GenAI tools available to understand how they can be effectively integrated into the classroom. Our study shows that a few GenAI tools (such as ChatGPT and Copilot) are primarily implemented with different education stakeholders despite the numerous tools available. Future research could explore and compare different tools to provide insights into how these tools could enhance teaching and learning. With clear evidence from our review and researchers' submissions (e.g., Jauhiainen and Garagorry Guerra, 2024) of limited research experimenting with GenAI tools by teachers and students in school or out-of-school settings, there is a need for high-quality research and intervention studies along these lines. Our research only considers papers published in English, which implies that valuable findings from papers published in other prominent languages like Arabic, French, and Spanish may be missing. This review notes a lack of studies from Africa, the Middle East, and South America from the selected papers. Only English-language publications meets the inclusion criteria. As a result of these geographic and linguistic limitations, the author of this work acknowledges the potential of how these factors might bias the synthesis or exclude potentially important findings from non-English speaking regions or different educational contexts. While the papers reviewed in this study are limited to 30, they provide useful insights into how GenAI can be integrated within the K-12 system. Fewer papers have been used to perform systematic reviews in previous research. For instance, Zhang and Tur (2024) synthesize findings from 13 papers. However, future work should consider more scientific contributions from other sources besides journal and conference papers to provide a more nuanced understanding of GenAI applications for K-12 students and teachers in practice. Since this study focuses solely on how GenAI has been employed for teaching and learning, other papers focusing on GenAI applications, including perspectives of other education stakeholders (e.g., policymakers), should be considered. Collecting these details will provide holistic information about GenAI for K-12 education.

5 Conclusion

This research focuses on the actual experiences of teachers and students using GenAI within the K-12 context. Given the increasing accessibility of GenAI in schools, it is crucial to understand how to use GenAI responsibly and effectively to enhance learning. Despite significant concerns about authorship, authenticity, and ethical implications that dominate discussions on GenAI use in education, the tool is here to stay. Consequently, there have been attempts to develop guidance for using GenAI across different subject domains (e.g., Sentance et al., 2024), which involve gathering views and experiences from various education stakeholders. Beyond sharing opinions about the implications of these powerful AI tools, evidence of how GenAI has been applied in practice, within classrooms or intervention studies conducted with teachers and students, will serve as concrete guidelines for its use in the K-12 context.

As the future of K-12 education will feature personalized learning journeys, inclusivity, and empowered educators (Sharma et al., 2024), efforts must be directed toward creating frameworks and building resources to realize the full potential of GenAI in shaping the educational landscape. Current findings indicate that most research on GenAI use focuses on higher education populations and adult learners, with empirical exploration centered on perspectives, conceptions, and experiences rather than actual implementation of AI tools. Some perspective-gathering papers provide short training or workshops on GenAI use (for education leaders, in-service, or pre-service teachers) before collecting their views. Future reviews should synthesize these findings and detail the training elements. This study also observed a lack of research on students and teachers with special needs or learning disabilities, highlighting the need for research on GenAI and inclusive education in the K-12 context. Additionally, literature has not explored differences in GenAI tool use across various contexts, such as the rural–urban dichotomy, and how practices differ between students and teachers in these settings. As revealed in the study, gaps exist in the context of ECE, special needs education, and rural context in relation to GenAI use which are important areas that should be considered for future research agendas. Finally, although the papers explored how GenAI has assisted in writing and English, there is still much to learn about how GenAI supports language teaching and learning, particularly within the K-12 system.

Author contributions

AA: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Alneyadi, S., and Wardat, Y. (2023). ChatGPT: revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in emirates schools. *Contemp. Educ. Technol.* 15:ep448. doi: 10.30935/cedtech/13417

Aruleba, K., Sanusi, I. T., Obaido, G., and Ogbuokiri, B. (2023). Integrating chatgpt in a computer science course: students perceptions and suggestions. *arXiv preprint* arXiv:2402.01640. doi: 10.48550/arXiv.2402.01640

Bitzenbauer, P. (2023). ChatGPT in physics education: a pilot study on easy-to-implement activities. *Contemp. Educ. Technol.* 15:ep430. doi: 10.30935/cedtech/13176

Blonder, R., Feldman-Maggor, Y., and Rap, S. (2024). Are they ready to teach? Generative AI as a means to uncover pre-service science teachers' PCK and enhance their preparation program. *J. Sci. Educ. Technol.*, 1–10. doi: 10.1007/s10956-024-10180-2

Bower, M., Henderson, M., Slade, C., Southgate, E., Gulson, K., and Lodge, J. (2025). What generative artificial intelligence priorities and challenges do senior Australian educational policy makers identify (and why)? *Aust. Educ. Res.* 52, 2069–2094. doi: 10.1007/s13384-025-00801-z

Cheah, Y. H., and Kim, J. (2024). Stem teachers' perceptions, familiarity, and support needs for integrating generative artificial intelligence in K-12 education. *Sch. Sci. Math.*. doi: 10.1111/ssm.18334

Chen, C. H., and Law, V. (2025). The role of help-seeking from ChatGPT in digital game-based learning. *Educ. Technol. Res. Dev.* 73, 1703–1721. doi: 10.1007/s11423-025-10464-3

Chen, B., and Zhu, X. (2023). Integrating generative AI in knowledge building. *Comput. Educ.* 5:100184.

Chiu, T. K. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: a case of ChatGPT. *Educ. Technol. Res. Dev.* 72, 2401–2416. doi: 10.1007/s11423-024-10366-w

Clark, L. A., and Watson, D. (2024). "Constructing validity: new developments in creating objective measuring instruments" in Methodological issues and strategies in clinical research. ed. A. E. Kazdin. *5th* ed (American Psychological Association), 317–348.

Clark-Fookes, T. (2024). To fear or not to fear: generative artificial intelligence in drama education. *Res. Drama Educ.*, 1–7.

Culverhouse, R. (2024). "OpenAI and how it can benefit K-12 education" in Artificial intelligence applications in K-12 eds. Crompton, H., and Burke, D. (New York: Routledge), 38–55. doi: 10.4324/9781003440192

Dilling, F., Herrmann, M., Müller, J., Pielsticker, F., and Witzke, I. (2024). Initiating interaction with and about ChatGPT–an exploratory study on the angle sum in triangles: Organising Committee, 145.

Ding, Z. Y., and Chiu, M. H. (2024). Can i collaborate with you? An investigation of trust in generative AI. *Proc. Assoc. Inf. Sci. Technol.* 61, 889–891. doi: 10.1002/pra2.1130

Feldman-Maggor, Y., Blonder, R., and Alexandron, G. (2024). Perspectives of generative AI in chemistry education within the TPACK framework. *J. Sci. Educ. Technol.* 34, 1–12. doi: 10.1007/s10956-024-10147-3

Florido, H., and Hernández-Leo, D. (2024). Generative AI collaboration in the orchestration of supervised classroom problem solving. In International Conference on Collaboration Technologies and Social Computing (257–264). Cham: Springer Nature Switzerland.

Galiç, S., Urhan, S., Dost, Ş., and Lavicza, Z. (2025). Examining mathematics teachers noticing the rationality: scenario-based training with AI Chatbot. *Sci. & Educ.*, 1–32. doi: 10.1007/s11191-025-00618-3

Gong, X., Li, Z., and Qiao, A. (2024). Impact of generative AI dialogic feedback on different stages of programming problem solving. *Educ. Inf. Technol.* 30, 9689–9709. doi: 10.1007/s10639-024-13173-1

Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., et al. (2024). Teachers, parents, and students' perspectives on integrating generative AI into elementary literacy education. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, HI, USA. ACM, New York, NY, USA, 1–17. doi: 10.1145/3613904.3642438

Hays, L., Jurkowski, O., and Sims, S. K. (2024). ChatGPT in K-12 education. *TechTrends* 68, 281–294. doi: 10.1007/s11528-023-00924-z

Hu, L., Wang, H., and Xin, Y. (2025). Factors influencing Chinese pre-service teachers' adoption of generative AI in teaching: an empirical study based on UTAUT2 and PLS-SEM. *Educ. Inf. Technol.* 30, 12609–12631. doi: 10.1007/s10639-025-13353-7

Jang, H., and Choi, H. (2024). A double-edged sword: physics educators' perspectives on utilizing ChatGPT and its future in classrooms. *J. Sci. Educ. Technol.* 34, 267–283. doi: 10.1007/s10956-024-10173-1

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

Jauhiainen, J. S., and Garagorry Guerra, A. (2024). Generative AI and education: dynamic personalization of pupils' school learning material with chatGPT. *Front. Educ.* 9:1288723.

Jauhiainen, J. S., and Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: evidence from a school lesson. *Sustainability* 15:14025. doi: 10.3390/su151814025

Jones, E. P. (2025). Use of large language model (LLM) to enhance content and structure of a school of dentistry LibGuide. *J. Med. Libr. Assoc.* 113, 96–97. doi: 10.5195/jmla.2025.2084

Karaca, M. F. (2024). Is artificial intelligence able to produce content appropriate for education level? A Review on ChatGPT and Gemini. In Proceedings of the Cognitive Models and Artificial Intelligence Conference (pp. 208–213).

Katuka, G. A., Chakraburty, S., Lee, H., Dhama, S., Earle-Randell, T., Celepkolu, M., et al. (2024). Integrating natural language processing in middle school science classrooms: an experience report. In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (639–645).

Kitchenham, B., and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering version 2.3. *Engineering* 45:1051. doi: 10.1145/1134285.1134500

Lee, J., Cimová, T., Foster, E. J., France, D., Krajňáková, L., Moorman, L., et al. (2025a). Transforming geography education: the role of generative AI in curriculum, pedagogy, assessment, and fieldwork. *Int. Res. Geogr. Environ. Educ.* 34, 237–253. doi: 10.1080/10382046.2025.2459780

Lee, S., Jeon, J., and Choe, H. (2025b). Generative AI (GenAI) and pre-service teacher agency in ELT. *ELT J.* 79:ccaf005. doi: 10.1093/elt/ccaf005

Levine, S., Beck, S. W., Mah, C., Phalen, L., and Pittman, J. (2024). How do students use ChatGPT as a writing support? *J. Adolesc. Adult. Lit.* 68, 445–457. doi: 10.1002/jaal.1373

Li, Z., and Ironsi, C. S. (2024). The efficacy of generative artificial intelligence in developing science education preservice teachers' writing skills: an experimental approach. *J. Sci. Educ. Technol.*, 1–12.

Liao, J., Zhong, L., Zhe, L., Xu, H., Liu, M., and Xie, T. (2024). Scaffolding computational thinking with ChatGPT. *IEEE Trans. Learn. Technol.* 17, 1628–1642. doi: 10.1109/TLT.2024.3392896

Lu, J., Zheng, R., Gong, Z., and Xu, H. (2024). Supporting teachers' professional development with generative AI: the effects on higher order thinking and self-efficacy. *IEEE Trans. Learn. Technol.* 17, 1267–1277. doi: 10.1109/TLT.2024.3369690

Min, T., Lee, B., and Jho, H. (2025). Integrating generative artificial intelligence in the design of scientific inquiry for middle school students. *Educ. Inf. Technol.* 30, 15329–15360. doi: 10.1007/s10639-025-13410-1

Miyake, Y., Retrosi, G., and Keijzer, R. (2025). Artificial intelligence and pediatric surgery: where are we? *Pediatr. Surg. Int.* 41, 1–8.

Ng, D. T. K., Tan, C. W., and Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: a pioneering pilot study. *Br. J. Educ. Technol.* 55, 1328–1353. doi: 10.1111/bjet.13454

Noster, N., Gerber, S., and Siller, H. S. (2024). Pre-service teachers' approaches in solving mathematics tasks with chatGPT. *Digit. Exp. Math. Educ.* 10, 543–567. doi: 10.1007/s40751-024-00155-8

Noveski, G., Jeroncic, M., Velard, T., Kocuvan, P., and Gams, M. (2024). Comparison of large language models in generating machine learning curricula in high schools. *Electronics* 13:4109. doi: 10.3390/electronics13204109

Nyaaba, M., Zhai, X., and Faison, M. Z. (2024). Generative AI for culturally responsive science assessment: a conceptual framework. *Educ. Sci.* 14:1325. doi: 10.3390/educsci14121325

Oyedoyin, M., Sanusi, I. T., and Ayanwale, M. A. (2024). Young children's conceptions of computing in an African setting. *Comput. Sci. Educ.* 35, 378–413. doi: 10.1080/08993408.2024.2314397

Passaretta, G., and Gil-Hernández, C. J. (2023). "The early roots of the digital divide: socioeconomic inequality in childrens ICT literacy from primary to secondary schooling" in Research handbook on digital sociology (Edward Elgar Publishing), 307–327.

Paulhus, D. L., and Vazire, S. (2007) The self-report method. In Robins, R. W., Fraley, R. C., and Krueger, R. F., Handbook of research methods in personality psychology 224–239 New York Guilford Press.

Sentance, S., Watson, S., Addo, S. A., Shi, S., Waite, J., and Yu, B. (2024). Developing computing teacher guidance on GenAI. In Proceedings of the 2024 Conference on United Kingdom & Ireland Computing Education Research Manchester, United Kingdom. ACM, New York, NY, USA, 1. doi: 10.1145/3689535.3689538

Sharma, D. M., Ramana, K. V., Jothilakshmi, R., Verma, R., Maheswari, B. U., and Boopathi, S. (2024). "Integrating generative AI into K-12 curriculums and pedagogies in India: opportunities and challenges" in Facilitating global collaboration and knowledge sharing in higher education with generative AI, 133–161.

Song, Y., Kim, J., Xing, W., Liu, Z., Li, C., and Oh, H. (2025). Elementary school students' and teachers' perceptions toward creative mathematical writing with generative AI. *J. Res. Technol. Educ.*, 1–23. doi: 10.1080/15391523.2025.2455057

Su, J., and Yang, W. (2024). Powerful or mediocre? Kindergarten teachers' perspectives on using ChatGPT in early childhood education. *Interact. Learn. Environ.* 32, 6496–6508. doi: 10.1080/10494820.2023.2266490

Takami, K. (2023). Exploring ChatGPT performance on PISA multiple choice sample questions comparing English and Japanese expression. In 31th International Conference on Computers in Education Conference Proceedings (pp. 27–32).

Tang, K. S., Cooper, G., Rappa, N., Cooper, M., Sims, C., and Nonis, K. (2024). A dialogic approach to transform teaching, learning & assessment with generative AI in secondary education: a proof of concept. *Pedagogies* 19, 493–503. doi: 10.1080/1554480X.2024.2379774

Uğraş, H., Uğraş, M., Papadakis, S., and Kalogiannakis, M. (2024a). ChatGPT-supported education in primary schools: the potential of ChatGPT for sustainable practices. *Sustainability* 16:9855. doi: 10.3390/su16229855

Uğraş, H., Uğraş, M., Papadakis, S., and Kalogiannakis, M. (2024b). Innovative early childhood STEM education with chatGPT: teacher perspectives. *Technol. Knowl. Learn.* 30, 809–831. doi: 10.1007/s10758-024-09804-8

Valcea, S., Hamdani, M. R., and Wang, S. (2024). Exploring the impact of ChatGPT on business school education: prospects, boundaries, and paradoxes. *J. Manag. Educ. 48*, 915–947. doi: 10.1177/10525629241261313

Vartiainen, H., Tedre, M., and Jormanainen, I. (2023). Co-creating digital art with generative AI in K-9 education: socio-material insights. *Int. J. Educ. Through Art* 19, 405–423. doi: 10.1386/eta_00143_1

Wang, Y., Guo, S., Ling, L., and Tan, C. W. (2024) Nemobot: crafting strategic gaming LLM agents for K-12 AI education. In Proceedings of the Eleventh ACM Conference on Learning@ Scale (pp. 393–397).

Wang, D., Zheng, Y., and Chen, G. (2024). ChatGPT or Bert? Exploring the potential of ChatGPT to facilitate preservice teachers' learning of dialogic pedagogy. *Educ. Technol. Soc.* 27, 390–406.

Wu, T. T., Lee, H. Y., Li, P. H., Huang, C. N., and Huang, Y. M. (2024). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *J. Educ. Comput. Res. 61*, 1539–1567. doi: 10.1177/07356331231191125

Yang, T. C., Hsu, Y. C., and Wu, J. Y. (2025b). The effectiveness of ChatGPT in assisting high school students in programming learning: evidence from a quasi-experimental research. *Interact. Learn. Environ.* 33, 3726–3743. doi: 10.1080/10494820.2025.2450659

Yang, S., Trainin, G., and Appleget, C. (2025a). Teacher use of generative AI for read-aloud question prompts. *Read. Teach.* 78, 230–235. doi: 10.1002/trtr.2366

Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: new demands for teachers' roles. *Heliyon* 10:e24289. doi: 10.1016/j.heliyon.2024.e24289

Yusuf, A., Pervin, N., Román-González, M., and Noor, N. M. (2024). Generative AI in education and research: a systematic mapping review. *Rev. Educ.* 12:e3489. doi: 10.1002/rev3.3489

Yuwono, E. I., Tjondronegoro, D., Riverola, C., and Loy, J. (2024). Co-creation in action: bridging the knowledge gap in artificial intelligence among innovation champions. *Comput. Educ.* 7:100272. doi: 10.1016/j.caeai.2024.100272

Zedan, R. (2021). Parental involvement as a predictor of classroom climate, motivation for learning, and learning achievements. *J. Classr. Interact.* 56, 56–74.

Zhang, P., and Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *Eur. J. Educ.* 59:e12599. doi: 10.1111/ejed.12599