Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute
InformingScience.org

JITEResearch.org

Volume 24, 2025

THE ROLE OF CHATGPT IN EDUCATION: APPLICATIONS, CHALLENGES: INSIGHTS FROM A SYSTEMATIC REVIEW

Maria Dimeli*	Postgraduate Student, University of the Aegean, Rhodes, Greece	maridimeli@gmail.comr
Apostolos Kostas	Dept. of Primary Education, University of the Aegean, Rhodes, Greece	apkostas@aegean.gr

* Corresponding author

ABSTRACT

Aim/Purpose	The purpose of this systematic review is to identify and analyze the current findings of empirical research on the use of ChatGPT in school and higher education.
Background	As AI reshapes education, the adoption of ChatGPT has the potential to revolutionize teaching and learning in school and higher educational settings. Meanwhile, substantial ethical questions and practical challenges are raised by the implementation of such technology at these educational levels, which must be carefully considered.
Methodology	To address the research questions, a systematic literature review (SLR) was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol. As part of the SLR, articles published between January 2023 and January 2024 were sought. The search query consisted of the various Boolean operators and search terms. The search was conducted in Scopus, Web of Science, ProQuest, SSRN, ERIC and DOAJ, Science Direct, Springer Link, Taylor & Francis, and IEEEXplore. Additionally, a manual search was carried out in scientific journals focusing on the field of emerging technologies in education. Of the 1,653 articles identified, 77 were selected through the application of inclusion and exclusion criteria. After reviewing the abstracts of the selected studies, 50 articles were included in the review.

Accepting Editor Lalitha Jonnavithula | Received: October 24, 2024 | Revised: December 12, December 21, 2024 | Accepted: December 26, 2024.
Cite as: Dimeli, M., & Kostas, A. (2025). The role of ChatGPT in education: Applications, challenges: insights from a systematic review. *Journal of Information Technology Education: Research, 24,* Article 2.
https://doi.org/10.28945/5422

Contribution	This SLR presents an innovative exploration of ChatGPT's potential in both university and school education. By examining its performance, ethical implications, and impact on student outcomes, the review provides a valuable resource for educators and researchers. It not only updates existing knowledge but also provides new insights into educational practices and use cases.
Findings	The study revealed that while ChatGPT can enhance students' cognitive performance and critical thinking skills, its capacity for deep, creative, and complex problem-solving is limited. Additionally, ethical challenges such as academic integrity violations, copyright infringements, and the propagation of biased content were identified.
Recommendations for Practitioners	Practitioners are encouraged to foster a culture of prompt engineering and AI literacy among educators and students, enabling the effective integration of AI conversational agents into educational settings while addressing potential limitations and ethical concerns.
Recommendations for Researchers	Researchers should direct their efforts towards more empirical research evidence in the domain of K-12 education with a specific focus on both secondary and primary education.
Impact on Society	ChatGPT and similar tools have the potential to revolutionize education. Their effective integration can create more engaging and effective learning experiences, preparing students for the future.
Future Research	Further investigation is needed in terms of the use of advanced conversational AI models in the areas of primary education and ethical frameworks, which are underrepresented as well as in non-formal, informal, and special education settings.
Keywords	artificial intelligence, chatbot, ChatGPT, formal education, ethical issues, institutional framework, systematic literature review, PRISMA

INTRODUCTION

In recent years, our daily lives have been inundated with a multitude of Artificial Intelligence (AI) applications and artifacts, often without our conscious awareness. Programs such as internet search engines, smartphone digital assistants, automated translation tools, GPS navigation systems, and autonomous vehicles rely on AI and contribute to facilitating our lives (European Commission, 2022; European Parliament, 2023).

As an integral part of society, education is undergoing a digital transformation as it harnesses the power of AI (Bozkurt et al., 2021). Combining theoretical knowledge with practical application, the science of Artificial Intelligence in Education (AIED) focuses on developing foundational theories for integrating AI into educational systems while simultaneously creating tools aimed at enhancing the learning experience and improving learning outcomes (Holmes & Porayska-Pomsta, 2022).

While the initial use of AIED can be traced back to the late 1970s (Self, 1974), significant advancements in the field emerged in the early 2000s (Villan & dos Santos, 2023). However, factors such as improved computational infrastructure, algorithm development, advancements in CPUs and cloud computing, increased research investments, and growing demand from industries integrating AI technologies have led to a surge in research interest in AIED in the late 2010s (İpek et al., 2023). Moreover, the urgent need for advanced, intelligent, and flexible technological tools to facilitate the pressing digital and remote learning process during the COVID-19 pandemic has acted as a catalyst for rapid advancements in AIED (Churi et al., 2022; Kostas et al., 2023).

AIED leverages methods such as cloud-based big data storage, big data analytics systems, and machine learning algorithms employing deep learning techniques to provide personalized support and create more flexible and effective learning environments, opening new possibilities for learning (Yu & Lu, 2021). Towards this end, the development of tools based on these techniques, such as Natural Language Processing (NLP), which uses algorithms to enable computers to understand, process, and generate human language, and computer vision, which allows computers to 'see' and understand visual data, and sensors that capture various physical states and convert them into electronic signals, has been pursued (UNESCO, 2022).

Based on the aforementioned techniques and tools, AIED finds application in three primary categories of educational AI programs (European Commission, 2018):

- Learning support systems tailored to student needs, such as Intelligent Tutoring Systems (ITS), Learning Management Systems (LMS), virtual and augmented reality, chatbots, assistive technologies for children with disabilities, and robotic tutors (Churi et al., 2022; OECD, 2020; UNESCO, 2022; Yu & Lu, 2021).
- Teaching support systems designed to meet the needs of educators, including AI tutors, learning analytics systems, automated assessment programs, and Large Language Models (LLMs) like ChatGPT (Churi et al., 2022; Lee, 2023; UNESCO, 2021).
- Educational system support systems include methods for collecting and processing learning data and the integration of AI into school and academic libraries (Churi et al., 2022).

Despite the abundance of AI tools, the widespread implementation of AI systems, particularly in education, raises ethical and deontological concerns that necessitate the establishment of an ethical framework.

The emerging field of AI ethics, situated within applied ethics and the philosophy of technology, aims to examine the ethical dilemmas and implications of developing and deploying AI systems (Waelen, 2022). According to Tzimas (2021), AI ethics focuses on creating ethical AI systems and establishing an ethical framework for human-machine interaction in an ever-evolving technological landscape. A significant trend is the increasing use of the term "ethically aligned design," which encompasses design processes for AI systems that explicitly integrate human values (IEEE, 2018, in World Commission on the Ethics of Scientific Knowledge and Technology, 2019). Similarly, the principle of human-centric AI pervades many international recommendations and agreements on AI use, emphasizing that AI should be under human control and work for the benefit of humanity while safeguarding human rights and upholding fundamental ethical principles (European Commission, 2018; OECD, 2020).

Key principles and ethical values identified in the literature regarding the development and use of AI systems include:

(a) *Transparency*, linked to security, involves access to and understanding of data, algorithms, and decision-making processes (Tzimas, 2021), as well as awareness of personal data collection and processing and the need to ensure privacy (Holmes et al., 2022; UNESCO, 2021; Waelen, 2022).

(b) *Justice and equality* are challenged by unequal access to AI and the exacerbation of the digital divide (UNESCO, 2021; Waelen, 2022), as well as the perpetuation of biases and stereotypes reinforced by AI (Holmes et al., 2022; Tzimas, 2021) – in educational contexts, academic integrity is threatened when students use AI-generated outputs as their own (Churi et al., 2022; Holmes et al., 2022).

(c) *Promoting social and environmental well-being, sustainable development, and avoiding harm* are examples of beneficial AI, including systems for predicting natural disasters or saving energy. Conversely, the use of nuclear or autonomous weapons powered by AI can be harmful or even highly dangerous (Tzimas, 2021; Waelen, 2022).

(d) *Accountability* is a cornerstone of AI development. While AI systems can make decisions with significant human impact, their inability to assume moral responsibility means that humans must be held accountable for their actions (Waelen, 2022). Applications such as student assessment and job recruitment using AI require careful oversight, as errors can have serious consequences for individuals (Churi et al., 2022).

While efforts are underway to establish ethical frameworks for AI by international organizations, governments, and private entities, the rapid advancements in the field and the emergence of new AI systems necessitate the continuous updating of guidelines to address emerging ethical challenges and ensure the optimal utilization of emerging AI technologies. Such an emerging AI tool is conversational agents or chatbots, which mimic human conversation (UNESCO, 2022) and serve as digital assistants, providing information and personalized support to learners without time or location constraints (UNESCO, 2021; Lee, 2023).

In late 2022, ChatGPT marked a milestone in chatbot development. This large language (LLM) is based on Generative Pre-trained Transformer (GPT) technology (UNESCO, 2023). After being trained on a massive dataset (Wolfram, 2023) using machine learning techniques (Javaid et al., 2023), ChatGPT can engage in written conversations using coherent natural language that simulates human communication Skrabut, 2023). Its ability to perform Natural Language Processing (NLP) tasks, such as creating customized educational materials and exercises, providing immediate feedback on tasks like summarizing texts, proofreading, and generating original essays (Skrabut, 2023), highlights its potential as a valuable tool for both teachers and students.

The model's significant impact is mirrored by a surge in research activity, focusing on both the benefits and ethical concerns of integrating ChatGPT, primarily in higher education (Athanassopoulos et al., 2023), as well as on the model's performance across various disciplines (Elkhatat, 2023; Kortemeyer, 2023; Meo et al., 2023; Vázquez-Cano et al., 2023).

A preliminary review of the literature for this research revealed a limited number of reviews focusing on the application of ChatGPT across the entire spectrum of formal education. Of these, some examine articles published in the first few months after the model's release (AlBadarin et al., 2023; İpek et al., 2023; Lo, 2023; Lo & Hew, 2023), while others primarily analyze theoretical articles (İpek et al., 2023; Montenegro-Rueda et al., 2023). Consequently, it was deemed necessary for the Systematic Literature Review (SLR) to focus exclusively on empirical studies, incorporating the latest data on the integration of ChatGPT into educational settings.

Additionally, most of the previous Systematic Literature Reviews examine the ways of utilizing, the limitations, and the ethical implications of the model in education in general or with a focus on higher education (Lo, 2023; Vargas-Murillo et al., 2023), with very few addressing K-12 education (Zhang & Tur, 2023). Therefore, it was considered important to examine various aspects of the topic through a comparative analysis of data from the application of ChatGPT in specific subject areas, focusing on the practical challenges and ethical concerns such as academic integrity and student dependency on AI, both in higher education and in K-12 education.

Furthermore, the literature review revealed studies exploring scenarios for using the model in various scientific disciplines and courses. Additionally, studies were identified that evaluate the quality of ChatGPT's responses in different subject areas, as well as its impact on students' performance and skills. However, these dimensions of the topic have not been sufficiently analyzed in existing literature reviews. Consequently, there is a need for further research in these areas.

Therefore, this SLR enriches the literature by aiming to collect and analyze recent data from empirical studies focusing on the integration of ChatGPT into educational settings in both K12 and higher education, to highlight the ways of application, limitations, and the effects of using the model on stu-

dents' performance and skills. This work can be a valuable tool for educators, researchers, and policy-makers considering the use of ChatGPT or similar chatbots in the field of education. To achieve the above objectives, this study will attempt to answer the following research questions:

- RQ1: What is the identity and methodological design of the studies under investigation?
- RQ2: How is ChatGPT used by teachers and students? Indicative best practices and use cases in K-12 and higher education.
- RQ3: What is the performance of ChatGPT during exploratory use in various subject areas?
- RQ4: What limitations and ethical issues arise from the use of ChatGPT in formal education?
- RQ5: What is the impact of using ChatGPT on students' performance, higher-order thinking skills, and motivation?

Subsequent sections detail the SLR methodology employed, the findings obtained for each research question, a discussion comparing these results with previous studies, and concluding remarks

METHODOLOGY

To address the research questions, a systematic literature review (SLR) was conducted, following a clear, strictly predefined methodology with specific implementation stages (Page et al., 2021), elements that constitute its validity criteria (Fink, 2005; Lame, 2019). This methodology, based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol (PRISMA), involves a 27-item checklist and a revised three-stage flow diagram as illustrated in Figure 1. Strict inclusion and exclusion criteria (Haddaway et al., 2022; Page et al., 2021) were applied to identify and select relevant articles.

As part of the SLR, articles published between January 2023 and January 2024 were sought. The search query consisted of the following Boolean operators and search terms: [ChatGPT OR "conversational agents" OR "chatbot AI"] AND ["empirical research"] AND [education OR learning practices OR didactic strategies OR "classroom intervention" OR "case scenarios"]. The search was conducted in the general databases Scopus, Web of Science, ProQuest, and SSRN, in the educational databases ERIC and DOAJ, and in the publishers Science Direct, Springer Link, Taylor & Francis, and IEEEXplore. Additionally, a manual search was carried out in scientific journals focusing on the field of emerging technologies in education.

Of the 1,653 articles identified, 77 were selected through the application of inclusion and exclusion criteria, as shown in Table 1. After reviewing the abstracts of the selected studies, 50 articles were ultimately included in the review.

The data were managed with the help of the open-source reference management software Zotero (Trinoskey et al., 2009). Additionally, based on the research questions of this study, a Google Sheets spreadsheet was developed for the extraction, organization, and categorization of information provided by the selected articles. This resulted in a table, where the first column contained the article references, and the subsequent columns recorded data related to the identity of the studies (location and time of the study, publication journal, and database of identification), their methodological design (target group, educational context, and subject area of ChatGPT application, purpose and research questions, data collection and analysis tools), as well as the results and conclusions of the studies included in this work. This was followed by the narrative synthesis of the data (Petticrew & Roberts, 2006), combined with the creation of tables to highlight the connection between the results and the evidence from the SLR. The coding of data within systematic reviews is different than coding primary research, as participant data and author analysis are interpreted to provide third-order constructs, as Crompton et al. (2021) suggested. Two types of coding were used in this study, a priori coding and grounded coding, following Crompton's et al. (2021) approach. A priori coding was used for articles'

identity, educational level, education type, and methodology extraction categories, and grounded coding was used for research extraction categories, initially using in-vivo coding (Saldana, 2015). Finally, regarding the limitations of this study, no separate analysis of quantitative and qualitative studies was conducted, and no interrater reliability was calculated for the content analysis.

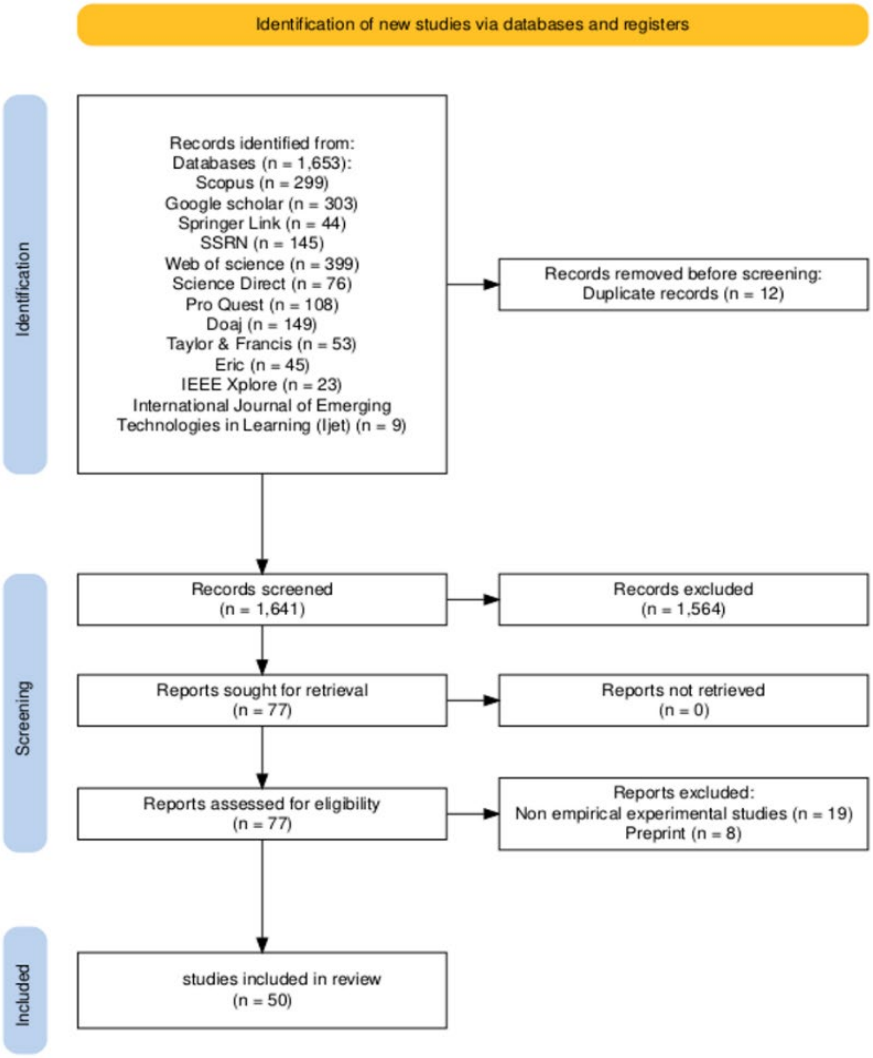

Figure 1. PRISMA flow diagram

Table 1. Inclusion and exclusion criteria for research articles

Category	Inclusion criteria	Exclusion criteria
Publication type	Research articles in scientific journals and reputable digital libraries, following peer-review	Articles not peer-reviewed, position papers, theoretical reports, literature reviews
Publication year	From January 2023 to January 2024	Before January 2023 and after January 2024
Research objective	Highlighting best practices, usage scenarios of ChatGPT, and identification of ethical and ideological limitations in the use of ChatGPT in educational contexts	Applications and uses of ChatGPT that are not related to the educational process

Category	Inclusion criteria	Exclusion criteria
Educational content	School and university education	Non-educational contexts, non-formal and informal education
Methodological design	Empirical studies of any methodological design	Theoretical articles
Language	English	Languages other than English
Accessibility	Open or institutional access to the full text of the articles	Theoretical articles; limited or paid access; no full access to the complete text

RESULTS

This section provides a comprehensive review of the data, drawing on the analysis and synthesis of 50 papers using the PRISMA methodology. The five research questions are outlined, accompanied by their corresponding findings.

RQ1: WHAT IS THE IDENTITY AND METHODOLOGICAL DESIGN OF THE STUDIES UNDER INVESTIGATION?

Studies identity

Most studies originate from the US, accounting for 12% of all articles, followed by China with 10%. Saudi Arabia contributes 6%, while Greece, Germany, Spain, Turkey, India, and Australia each contribute two articles (4%). Numerous other countries, primarily from Asia, as well as from Europe, America, and Africa, participate in the SLR, each with one article. Figure 2 presents a global map illustrating the distribution of articles by country of origin.

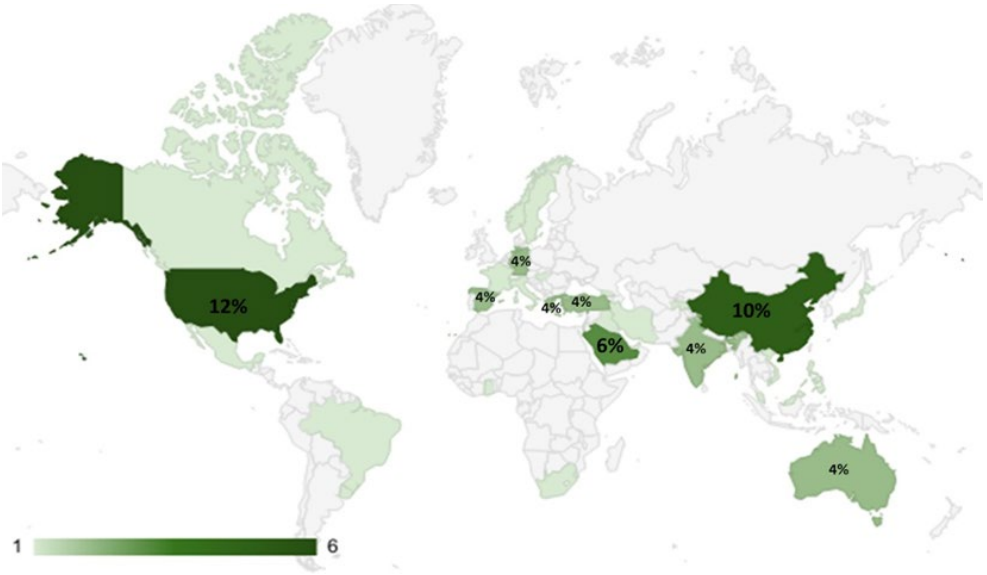

Figure 2. Geographical distribution of articles

Regarding the temporal distribution of articles, a gradual increase in publication is observed over the months, culminating in a peak during July and autumn of 2023, as illustrated in Figure 3. In terms of the journals in which the studies are published, a wide dispersion of results is evident, with most publications (6%) appearing in *Computers and Education: Artificial Intelligence and Education Sciences.*

Concerning the educational context, most studies, representing 62% of all records (31 articles), are situated within the higher education sector. Secondary and primary education follow with 14% and 10%, respectively, while a single study is found in preschool education. A total of 12% of the articles, corresponding to six publications, made general references to education or encompassed multiple educational levels, as shown in Table 2.

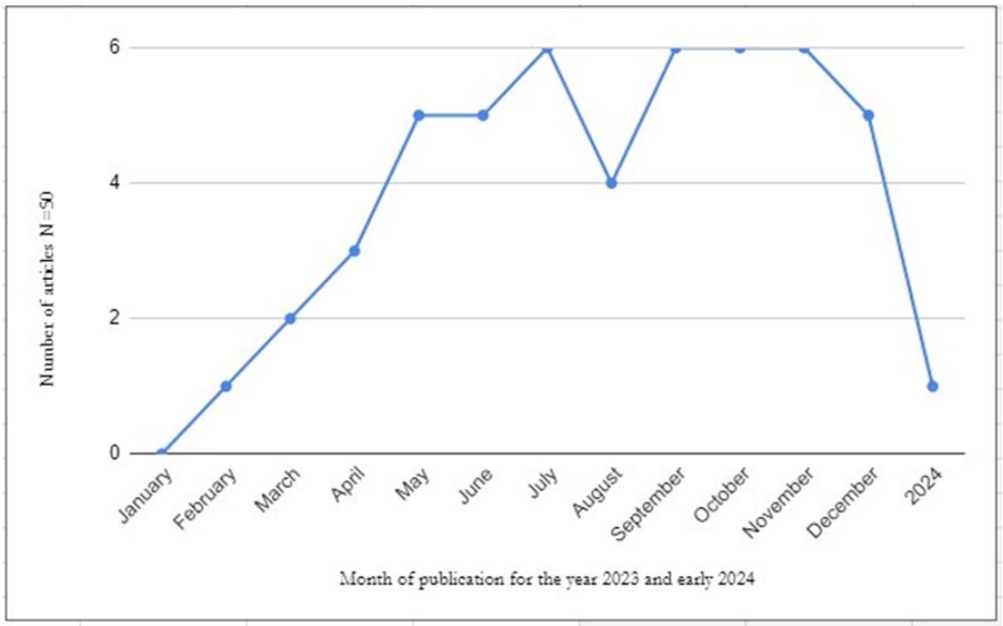

Figure 3. Timeline of article publications

Table 2. Educational level of selected articles

Educational level	References
Preschool education (1 article)	(Luo et al., 2024)
Primary education (5 articles)	(Abdelghani et al., 2024; Jauhiainen & Guerra, 2023; Jeon & Lee, 2023; Yan, 2023; Young & Shishido, 2023)
Secondary education (8 articles)	(Alneyadi & Wardat, 2023; Athanassopoulos et al., 2023; Bitzenbauer, 2023; Chen et al., 2023; Forman et al., 2023; Javier & Moorhouse, 2024; Villan & dos Santos, 2023; Waltzer et al., 2023)
Higher education (30 articles)	(Akiba & Fraboni, 2023; Albdrani & Al-Shargabi, 2023; Albert & Li, 2023; Al-Garaady & Mahyoob, 2023; Al-Obaydi et al., 2023; Aydin Yildiz, 2023; Baglivo et al., 2023; Escalante et al., 2023; Essel et al., 2024; Farazouli et al., 2023; Guo & Lee, 2023; Han et al., 2023; Ho et al., 2023; Irfan et al., 2023; Kong et al., 2023; Küchemann et al., 2023; Matzakos et al., 2023; Meron & Tekmen Araci, 2023; Michalon & Camacho-Zuñiga, 2023; Nguyen, 2023; Niu & Xue, 2023; Roy & Putatunda, 2023; Sánchez-Ruiz et al., 2023; Shaikh et al., 2023; Swargiary, 2023; Tirado-Olivares et al., 2023; Uddin et al., 2023; van den Berg & du Plessis, 2023; Xiao & Zhi, 2023; Yilmaz & Karaoglan Yilmaz, 2023)
General education or different levels of education (6 articles)	(Cooper, 2023; Daher et al., 2023; Ghafouri, 2024; Li et al., 2023; Tlili et al., 2023; Wardat et al., 2023)

Pertaining to the target group, university students constitute the target population in most articles, with 24 recorded instances, while secondary school students follow with seven recorded instances.

Subsequently, a mixed sample of university students and educators is employed in six studies, while a purely university educator sample is used in four studies. In three studies, the sample consists of both secondary school students and educators simultaneously. The fewest instances are observed among primary school students and educators.

Regarding the subject areas in which ChatGPT is implemented, the highest frequency, with a total of 15 articles, is observed in the field of language learning, with English as a foreign language being the dominant focus. This is followed by applications in the natural sciences, with Physics appearing in four articles, while Mathematics, Chemistry, Computer Science, and Education each have three records. Additionally, five articles make general references to education without specifying a particular subject area. Table 3 provides a detailed overview of the target groups and subject areas where ChatGPT was applied in the studies.

Table 3. Distribution of subject area by target group

Subject area	Target group	References
Foreign languages (15 articles)	Primary school teachers	(Jeon & Lee, 2023)
	Higher education teachers	(Nguyen, 2023)
	Secondary school students	(Athanassopoulos et al., 2023; Javier & Moorhouse, 2024; Li et al., 2023; Young & Shishido, 2023)
	Undergraduate students - Higher education	(Al-Obaydi et al., 2023; Aydin Yildiz, 2023; Escalante et al., 2023; Xiao & Zhi, 2023)
	Undergraduate and postgraduate students – Higher education	(Shaikh et al., 2023)
	Secondary school students and teachers	(Ghafouri, 2024; Waltzer et al., 2023)
	Higher education students and teachers	(Al-Garaady & Mahyoob, 2023; Roy & Putatunda, 2023)
General education (5 articles)	Higher education teachers	(Luo et al., 2024; Swargiary, 2023)
	Primary school students	(Abdelghani et al., 2024)
	Secondary school students	(Forman et al., 2023)
	Higher education students and teachers	(Tlili et al., 2023)
Physics (4 articles)	Secondary school teachers	(Cooper, 2023)
	Secondary school students	(Alneyadi & Wardat, 2023; Bitzenbauer, 2023)
	Undergraduate students - Higher education	(Küchemann et al., 2023)
Mathematics (3 articles)	Secondary school teachers	(Wardat et al., 2023)
	Undergraduate students - Higher education	(Sánchez-Ruiz et al., 2023)
	Higher education students and teachers	(Matzakos et al., 2023)
Chemistry (3 articles)	Secondary school students	(Daher et al., 2023)
	Undergraduate students - Higher education	(Guo & Lee, 2023)
	Higher education students and teachers	(Kong et al., 2023)
Computer science (3 articles)	Undergraduate students - Higher education	(Albdrani & Al-Shargabi, 2023; Niu & Xue, 2023; Yilmaz & Karaoglan Yilmaz, 2023)

Subject area	Target group	References
Elementary education (3 articles)	Higher education teachers	(Meron & Tekmen Araci, 2023)
	Undergraduate students - Higher education	(Akiba & Fraboni, 2023; van den Berg & du Plessis, 2023)
Diverse fields (2 articles)	Higher education teachers	(Farazouli et al., 2023)
	Undergraduate students - Higher education	(Swargiary, 2023)
History (2 articles)	Primary school students	(Jauhiainen & Guerra, 2023)
	Undergraduate students - Higher education	(Tirado-Olivares et al., 2023)
Scientific research (2 articles)	Primary school students and teachers	(Villan & dos Santos, 2023)
	Undergraduate students - Higher education	(Essel et al., 2024)
Medicine (2 articles)	Undergraduate students - Higher education	(Baglivo et al., 2023)
	Higher education students and teachers	(Han et al., 2023)
International relations (1 article)	Undergraduate students - Higher education	(Michalon & Camacho-Zuñiga, 2023)
Journalism (1 article)	Undergraduate students - Higher education	(Irfan et al., 2023)
World religions (1 article)	Secondary school students and teachers	(Chen et al., 2023)
Construction industry (1 article)	Undergraduate students - Higher education	(Uddin et al., 2023)
Business administration (1 article)	Undergraduate and postgraduate students – Higher education	(Albert & Li, 2023)

Methodological design

Regarding the methodological design, 21 studies (42%) employ a mixed research approach, 19 (38%) use qualitative methods, and 10 articles (20%) apply quantitative methods. Sixteen studies (32%) are exploratory, making this the most frequent research type in the SLR. Case studies with a single experimental group are used in 10 studies (20%), while experimental studies with a within-group design and between-group design are found in seven (14%) and six (12%) articles, respectively. Additionally, five studies (10%) applied field research, while smaller percentages are recorded for action research and case studies with a between-group design, as shown in Table 4.

Table 4. Classification of articles according to research type and case study

Case study	References according to research type		
	Quantitative methods	Qualitative methods	Mixed research approach
Exploratory studies (16 articles)	(Forman et al., 2023; Matzakos et al., 2023; Shaikh et al., 2023; Young & Shishido, 2023)	(Akiba & Fraboni, 2023; Cooper, 2023; Han et al., 2023; Ho et al., 2023; Jeon & Lee, 2023; Kong et al., 2023; Luo et al., 2024; Meron & Tekmen Araci, 2023; van den Berg & du Plessis, 2023; Xiao & Zhi, 2023)	(Daher et al., 2023; Nguyen, 2023)

Case study	References according to research type		
	Quantitative methods	**Qualitative methods**	**Mixed research approach**
Case studies with a single experimental group (10 articles)	(Athanassopoulos et al., 2023)	(Javier & Moorhouse, 2024; Roy & Putatunda, 2023; Tlili et al., 2023; Wardat et al., 2023)	(Albert & Li, 2023; Baglivo et al., 2023; Farazouli et al., 2023; Jauhiainen & Guerra, 2023; Sánchez-Ruiz et al., 2023)
Experimental studies with a within-group design (7 articles)	(Uddin et al., 2023)	(Al-Obaydi et al., 2023)	(Escalante et al., 2023; Guo & Lee, 2023; Irfan et al., 2023; Niu & Xue, 2023; Tirado-Olivares et al., 2023)
Experimental study with a between-groups design (6 articles)	(Aydin Yildiz, 2023; Ghafouri, 2024; Yilmaz & Karaoglan Yilmaz, 2023)		(Essel et al., 2024; Küchemann et al., 2023; Swargiary, 2023)
Field research (5 articles)	(Bitzenbauer, 2023)	(Abdelghani et al., 2024; Li et al., 2023)	(Al-Garaady & Mahyoob, 2023; Waltzer et al., 2023)
Action research (3 articles)		(Chen et al., 2023)	(Michalon & Camacho-Zuñiga, 2023; Villan & dos Santos, 2023)
Case studies with a between-group design (2 articles)			(Albdrani & Al-Shargabi, 2023; Alneyadi & Wardat, 2023)
Multiple methods qualitative research (1 article)		(Yan, 2023)	

RQ2: How is ChatGPT Used by Teachers and Students? Indicative Best Practices and Use Cases in School and University Education.

The research revealed various practices and use cases from preschool to university education. Due to the large volume of studies in the field of foreign language education within formal education, a separate analysis is dedicated to the data generated in this area, as shown in Table 5.

Applications of ChatGPT in school education

Teachers in preschool and primary education use it for creating personalized materials (Jauhiainen & Guerra, 2023; Luo et al., 2024). In primary education, it is generally used by teachers to create quizzes (Tlili et al., 2023), while in physics, it is employed for designing lesson plans and assessment rubrics (Cooper, 2023).

Preschool students use it as a conversation agent to develop interactive dialogues (Luo et al., 2024), while in primary education, it is generally used as a personal assistant for completing assignments and preparing for exams (Forman et al., 2023). In physics and mathematics, students use it to receive immediate assistance in clarifying complex concepts, finding examples, and solving problems (Alneyadi & Wardat, 2023; Wardat et al., 2023). In the study by Chen et al. (2023), students use it as a tool in a scenario for knowledge-building in a course on world religions, while in Bitzenbauer's (2023) study, the tool is applied through a structured think-pair-share activity in Physics.

Applications of ChatGPT in university education

In university education, in the field of mathematics, educators use it to find steps for problem-solving, although they perform calculations in Mathematica or via the Wolfram plugin (Matzakos et al., 2023; Sánchez-Ruiz et al., 2023). In physics, computer science, and medicine, it is used for creating assessment exercises (Han et al., 2023; Küchemann et al., 2023; Niu & Xue, 2023). In business administration, it serves as a post hoc grader, with educators applying a three-stage scenario for developing appropriate prompts (Albert & Li, 2023).

Students, on the other hand, leverage ChatGPT for comprehending mathematical and data science concepts (Albdrani & Al-Shargabi, 2023; Sánchez-Ruiz et al., 2023), solving memorization-based chemistry problems (Daher et al., 2023) primarily, finding the steps in mathematical problem-solving (Sánchez-Ruiz et al., 2023), and aiding in programming tasks such as code generation and debugging (Yilmaz & Karaoglan Yilmaz, 2023). Future teachers use it for creating lesson plans and materials (Meron & Tekmen Araci, 2023; van den Berg & du Plessis, 2023) and as an academic advisor (Akiba & Fraboni, 2023). In medicine, it is utilized for drafting medical reports (Ho et al., 2023), while in business administration, it enriches and reviews assignments (Albert & Li, 2023). In chemical engineering, Kong et al. (2023) apply a use scenario of ChatGPT in a mass transfer course, while in chemistry, Guo and Lee (2023) implement a scenario for generating an essay.

Applications of ChatGPT in foreign language education

In foreign language teaching, educators use it to create assessment tests, dialogues (Aydin Yildiz, 2023; Jeon & Lee, 2023; Young & Shishido, 2023), short stories, and personalized educational materials, as well as for the automatic evaluation of assignments (Al-Garaady & Mahyoob, 2023; Nguyen, 2023). Learners, on the other hand, use it as a conversational partner for language practice (Javier & Moorhouse, 2024; Li et al., 2023; Roy & Putatunda, 2023) and as a mentor for providing immediate feedback during essay writing (Xiao & Zhi, 2023) and for improving assignments (Athanassopoulos et al., 2023; Li et al., 2023). In English language learning, Ghafouri (2024) applies a use scenario for implementing a Relationship Building Protocol, Yan (2023) combines laboratory lessons with collaborative practices within a structured educational approach, while Roy and Putatunda (2023) follow a process for creating high-level critical knowledge using ChatGPT in an English literature course. Furthermore, Athanassopoulos et al. (2023) describe the stages of integrating ChatGPT as an assessment tool to support German writing, while Li et al. (2023) outline the process of supporting autonomous learning in academic Chinese writing. Finally, Escalante et al. (2023) implement a gradual role assignment process for GPT-4 as a grader in an academic reading and writing course for English as a new language.

Table 5. Best practices and use scenarios by subject area for educators and students in preschool, primary, university, and foreign language education

Educational level	Subject area	Best practices for educators	Best practices for students	References
Preschool education		Smart assistant (lesson planning, creation of personalized materials, ideas for classroom management)	Conversational agent that creates social stories and develops interactive dialogues	(Luo et al., 2024)
School education	General primary education	Creation of quizzes	Assistant in assignments, exam preparation, research, finding ideas for essays, and understanding concepts	(Forman et al., 2023; Tlili et al., 2023)

Educational level	Subject area	Best practices for educators	Best practices for students	References
	History (Primary)	Development of personalized educational material		(Jauhiainen & Guerra, 2023)
	Physics (Primary and Secondary) Mathematics (Secondary)	Creation of lesson plans, assessment rubrics, and knowledge evaluation tests	Immediate personalized assistance for concept comprehension, example identification, rapid task completion, and problem-solving. Productive and critical utilization of the model through a structured approach involving think-pair-share activities	(Alneyadi & Wardat, 2023; Bitzenbauer, 2023; Cooper, 2023; Wardat et al., 2023)
	Scientific research		Pedagogical mediator, co-advisor.	(Villan & dos Santos, 2023)
	World religions (Secondary)	Integration of ChatGPT into the curriculum involving the application of "prompt engineering" techniques and the verification of outputs	Assignment review aid and tool in a use case for knowledge building	(Chen et al., 2023)
University education	Mathematics	Leveraging ChatGPT for identifying problem-solving steps and then executing mathematical calculations in Mathematica or via a Wolfram add-on	Assistant for understanding and solving math problems	(Matzakos et al., 2023; Sánchez-Ruiz et al., 2023)
	Physics	Developing quality assessment tasks		(Küchemann et al., 2023)
	Chemistry, Chemical Engineering		Problem-solving assistant, mass transfer application scenario, essay development scenario	(Daher et al., 2023; Guo & Lee, 2023; Kong et al., 2023)
	Training of future teachers		Academic advisor providing course material development, syllabus design, and assessment task creation	(Akiba & Fraboni, 2023; Meron & Tekmen Araci, 2023; van den Berg & du Plessis, 2023)
	Journalism		Assistance in improving writing skills and idea generation	(Irfan et al., 2023)

Educational level	Subject area	Best practices for educators	Best practices for students	References
	Data science and Programming	Development of personalized exercises	Coding snippet generation and debugging	(Albdrani & Al-Shargabi, 2023; Niu & Xue, 2023; Yilmaz & Karaoglan Yilmaz, 2023)
	Medicine	Curriculum development and assessment assistant	Simulated public health dialogue partner scenario for creating medical reports	(Baglivo et al., 2023; Han et al., 2023; Ho et al., 2023)
	Construction industry		Hazard recognition assistant	(Uddin et al., 2023)
	Business administration	Post-assessment grader and three-phase scenario for prompt engineering	Task enrichment and assessment support	(Albert & Li, 2023)
	International relations		Prompt optimization and conversation quality metrics	(Michalon & Camacho-Zuñiga, 2023)
	Language learning German English (K12)	Creating dialogues and short stories, personalizing materials, identifying errors, creating assessment tests, and contributing to a supportive learning environment	Language practice, personalized feedback, essay improvement before submission, productive and critical use	(Athanassopoulos et al., 2023; Ghafouri, 2024; Javier & Moorhouse, 2024; Jeon & Lee, 2023; Xiao & Zhi, 2023; Young & Shishido, 2023)
	Chinese English (higher education)	Suggestions for students' essays, creating assessment exercises, automatic grading of written work, creating personalized educational materials, learning resources, and assignments	Vocabulary and grammar practice, provider of cultural information, immediate feedback for works, adapting text to formal or informal writing, experiential learning through written dialogues, partner in a scenario for the development of high-level critical thinking	(Al-Garaady & Mahyoob, 2023; Al-Obaydi et al., 2023; Aydin Yildiz, 2023; Escalante et al., 2023; Li et al., 2023; Nguyen, 2023; Roy & Putatunda, 2023; Shaikh et al., 2023; Yan, 2023)

RQ3: What is the Performance of ChatGPT During Exploratory Use in Various Subject Areas?

The analysis of the studies reveals ChatGPT's strong performance across diverse fields within both the theoretical and empirical sciences, as shown in Table 6. In Chemistry, the model excels in theoretical problem-solving (Daher et al., 2023). In Mathematics, it provides detailed steps for problem-

solving and can solve simple calculations, although limitations are observed in numerical computations (Matzakos et al., 2023; Sánchez-Ruiz et al., 2023; Wardat et al., 2023). In the field of Medicine, it outperforms medical students in answering complex questions related to vaccination (Baglivo et al., 2023).

Additionally, in History, the tool demonstrates high performance in developing argumentative historical texts according to most dimensions of historical thinking (Tirado-Olivares et al., 2023). In the field of Humanities and Social Sciences, it achieves high scores in the context of homework assignments (Farazouli et al., 2023).

Furthermore, in the research by Abdelghani et al. (2024), the GPT-3 model is highly effective in creating prompts that can help primary school children formulate divergent questions. Moreover, in the field of English as a Foreign Language, ChatGPT, as evidenced by the results of the SLR, can produce high-level literary and proverbial essays (Waltzer et al., 2023) as well as reference dialogues based on the criteria of readable writing (Young & Shishido, 2023).

Table 6. ChatGPT's performance in exploratory applications across various disciplines

Subject area	ChatGPT's performance	References
General Education	Effectiveness of GPT-3 to generate prompts that aid children in formulating divergent questions	(Abdelghani et al., 2024)
Chemistry	High performance in conceptual understanding and reasoning in chemistry problem-solving depth difficulties and errors in numerical calculations	(Daher et al., 2023)
Chemical Engineering	Guide with general instructions for designing a distillation column: challenges in providing accurate and detailed results	(Kong et al., 2023)
History	Effective performance in developing argumentative historical text based on most dimensions of historical thinking	(Tirado-Olivares et al., 2023)
Mathematics	High reliability in the theoretical solution of mathematical problems. Reliable results in simple operations and symbolic computations. Limitations in complex calculations	(Matzakos et al., 2023; Sánchez-Ruiz et al., 2023; Wardat et al., 2023)
Medicine	High performance in complex medical questions in the field of vaccination (in Italian) compared to medical students.	(Baglivo et al., 2023)
Humanities and Social Sciences	High scores on home-based assignments	(Farazouli et al., 2023)
English	Generation of reference dialogues adhering to readability criteria	(Young & Shishido, 2023)
	Generation of high-quality essays incorporating literary and proverbial elements	(Waltzer et al., 2023)

RQ4: What Limitations and Ethical Issues Arise From the Use of ChatGPT in Formal Education?

Despite the benefits and numerous applications of ChatGPT, this SLR highlights the potential challenges of using ChatGPT in education. The key concerns identified in these studies include the model's inherent limitations and the ethical implications arising from its integration.

Most researchers (22.6%, 17 articles) highlight issues with the quality of ChatGPT responses, including inaccuracies, irrelevance, and lack of depth, as well as unreliable or fabricated references. Additionally, 12.5% (9 articles) of studies note limitations in the model's ability to exhibit higher-order cognitive functions, such as creative thinking, critical reasoning, and emotional expression. Concerns regarding academic integrity and potential copyright infringement are raised in 11.1% (8 articles) and 9.7% (7 articles) of studies, respectively. The risk of students becoming overly dependent on AI, negatively impacting critical thinking skills, as well as the ethical implications arising from unequal access to the tool and the potential perpetuation of biased content, are identified in 6.9% (5 articles) of the studies, respectively. Concerns about privacy violations and the model's inaccurate calculations in STEM problems occupy 5.6% (4 articles) of the research, respectively. In 4.2% (3 articles) of cases, researchers focus on each of the risks of malicious use of the tool, inaccurate student assessments, and its inability to analyze and produce images and graphical representations. Finally, two studies, each representing 2.8% (2 articles) of the total articles, address issues of reliability in personalized exercises and ChatGPT's omissions in curriculum design. Table 7 presents a comprehensive overview of the challenges identified in the SLR regarding the utilization of ChatGPT across various educational levels.

Table 7. Limitations and ethical considerations in the use of ChatGPT in education

Limitations and ethical considerations	Educational level	References
Inaccurate, incomplete, irrelevant, outdated, superficial responses, and unreliable or fabricated citations (17 articles)	Preschool education	(Luo et al., 2024)
	Secondary education	(Chen et al., 2023; Javier & Moorhouse, 2024)
	Higher education	(Akiba & Fraboni, 2023; Albert & Li, 2023; Essel et al., 2024; Farazouli et al., 2023; Guo & Lee, 2023; Han et al., 2023; Ho et al., 2023; Michalon & Camacho-Zuñiga, 2023; Nguyen, 2023; Sánchez-Ruiz et al., 2023; Xiao & Zhi, 2023)
	General education	(Li et al., 2023; Tlili et al., 2023; Wardat et al., 2023)
Deficiencies in emotional intelligence, creative and deep thinking, critical reasoning, and problem-solving (9 articles)	Preschool education	(Luo et al., 2024)
	Secondary education	(Chen et al., 2023)
	Higher education	(Al-Garaady & Mahyoob, 2023; Guo & Lee, 2023; Han et al., 2023; Meron & Tekmen Araci, 2023; Tirado-Olivares et al., 2023)
	General education	(Daher et al., 2023; Tlili et al., 2023)
Threat to academic integrity (8 articles)	Preschool education	(Luo et al., 2024)
	Primary education	(Yan, 2023)
	Secondary education	(Chen et al., 2023; Waltzer et al., 2023)
	Higher education	(Al-Obaydi et al., 2023; Nguyen, 2023; Xiao & Zhi, 2023)
	General education	(Tlili et al., 2023)

Limitations and ethical considerations	Educational level	References
Lack of citations for the provided information (copyright infringement) (7 articles)	Preschool education	(Luo et al., 2024)
	Secondary education	(Chen et al., 2023)
	Higher education	(Albert & Li, 2023; Al-Garaady & Mahyoob, 2023; Farazouli et al., 2023; Ho et al., 2023; Roy & Putatunda, 2023)
Risk of excessive dependence - Negative impact on students' critical thinking (5 articles)	Secondary education	(Chen et al., 2023)
	Higher education	(Nguyen, 2023; Roy & Putatunda, 2023; Swargiary, 2023)
	General education	(Li et al., 2023)
Ethical issues: The digital divide and biased AI content perpetuating social inequalities (5 articles)	Preschool education	(Luo et al., 2024)
	Primary education	(Yan, 2023)
	Secondary education	(Waltzer et al., 2023)
	Higher education	(Roy & Putatunda, 2023)
	General education	(Tlili et al., 2023)
Risk of violation of users' privacy (4 articles)	Preschool education	(Luo et al., 2024)
	Higher education	(Ho et al., 2023; Roy & Putatunda, 2023)
	General education	(Tlili et al., 2023)
Errors in computations within STEM disciplines (math, geometry, chemistry, engineering) (4 articles)	Higher education	(Matzakos et al., 2023; Sánchez-Ruiz et al., 2023)
	General education	(Daher et al., 2023; Wardat et al., 2023)
Malicious use (3 articles)	Primary education	(Jeon & Lee, 2023)
	Secondary education	(Chen et al., 2023)
	Higher education	(Irfan et al., 2023)
Inaccurate student evaluations (3 articles)	Higher education	(Albdrani & Al-Shargabi, 2023; Albert & Li, 2023; Al-Garaady & Mahyoob, 2023)
Inability to analyze images and graphic representations (3 articles)	Higher education	(Han et al., 2023)
	General education	(Daher et al., 2023; Wardat et al., 2023)
Questionable reliability of personalized exercises and quizzes due to limited content variety, high predictability of answers, and superficial depth (2 articles)	Higher education	(Niu & Xue, 2023)
	General education	(Tlili et al., 2023)
Limitations in curriculum design, including deficiencies in overall content, errors, and omissions (2 articles)	Higher education	(Han et al., 2023; Meron & Tekmen Araci, 2023)

RQ5: What is the Impact of Using ChatGPT on Student's Performance, Higher-Order Skills, and Motivation?

Nearly all articles in this SLR demonstrate that ChatGPT has a positive impact on students' learning outcomes and skills.

As illustrated in Figure 4, a comprehensive analysis of 12 (twelve) studies reveals that the integration of ChatGPT into educational settings has the most pronounced effects on cognitive performance, with seven studies primarily conducted in higher education settings and five focusing on K-12 education. In addition, seven studies specifically highlight the positive influence of ChatGPT on the development of critical thinking abilities. However, only one of these studies focuses on K-12 education, while the majority are situated within higher education settings. Moreover, the tool is found to significantly enhance student motivation, engagement, and persistence, as evidenced by 6 (six) studies. Four of these articles originate from higher education, while one is found in both school and general education, respectively. Furthermore, 5 (five) studies indicate a significant positive impact of ChatGPT on the language skills of both school and university students, with one study focusing on general education. A further three studies highlight the positive impact of ChatGPT on AI literacy, two of which are focused on higher education and one on secondary education.

Based on subject areas, performance gains from using ChatGPT have been observed in K-12 education, particularly in History, where students' knowledge was enhanced (Jauhiainen & Guerra, 2023), and in Physics, where students benefited from improved conceptual understanding (Alneyadi & Wardat, 2023; Bitzenbauer, 2023), as well as across the broader spectrum of primary (Abdelghani et al., 2024) and secondary education (Forman et al., 2023). At the tertiary level, students in Mathematics developed problem-solving skills and learned new concepts (Sánchez-Ruiz et al., 2023), while those in Computer Programming built computational thinking skills (Yilmaz & Karaoglan Yilmaz, 2023). In the construction industry, students improved their ability to identify workplace hazards (Uddin et al., 2023). Additionally, ChatGPT had a positive impact on student performance in Data Science (Albdrani & Al-Shargabi, 2023), Research Methodology (Essel et al., 2024), Business Administration (Albert & Li, 2023), and English as a Foreign Language (Aydin Yildiz, 2023).

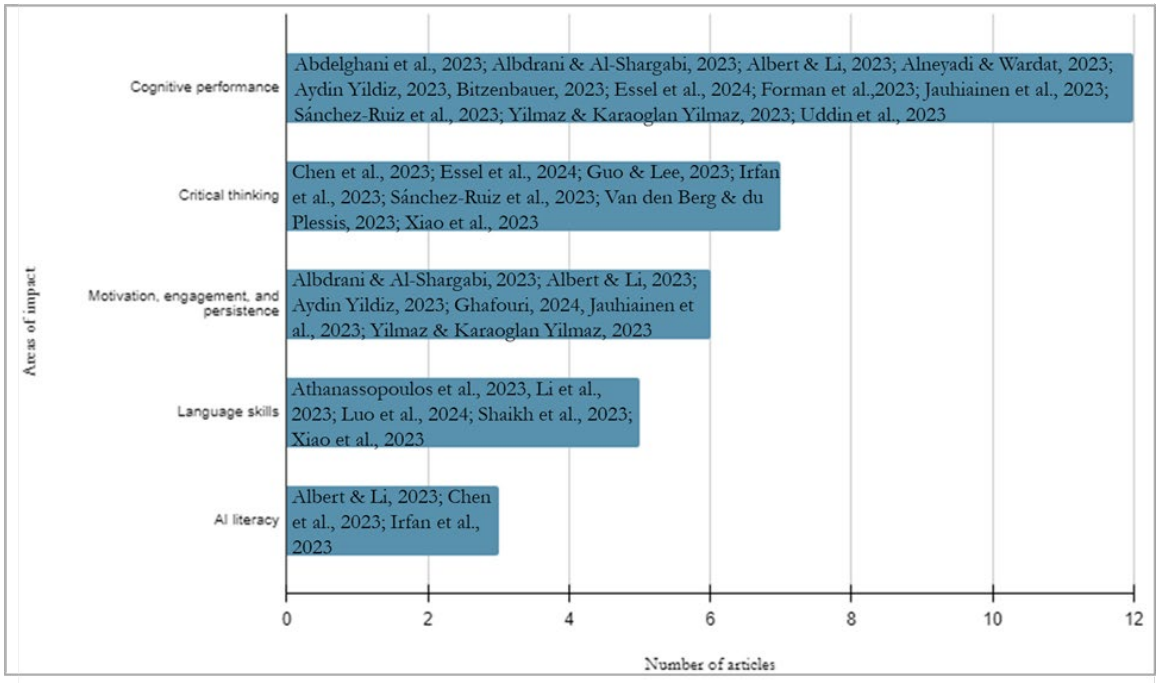

Figure 4. Areas of positive impact of ChatGPT based on SLR studies

Students' critical thinking was positively influenced by the tool in a course on world religions during knowledge construction (Chen et al., 2023), in Mathematics (Sánchez-Ruiz et al., 2023), and Chemistry (Guo & Lee, 2023) during problem-solving, in teacher education during lesson planning (van den Berg & du Plessis, 2023), as well as in Journalism (Irfan et al., 2023), Research Methodology (Essel et al., 2024), and English language learning (Xiao & Zhi, 2023).

Students demonstrate increased motivation, persistence, or enjoyment of the learning process when using ChatGPT in courses on History (Jauhiainen & Guerra, 2023), Data Science (Albdrani & Al-Shargabi, 2023), Programming (Yilmaz & Karaoglan Yilmaz, 2023), Business Administration (Albert & Li, 2023), and English (Aydin Yildiz, 2023; Ghafouri, 2024).

Improvements in students' language skills are observed in preschool education (Luo et al., 2024) and in the learning of English (Shaikh et al., 2023; Xiao & Zhi, 2023), German (Athanassopoulos et al., 2023), and Chinese (Li et al., 2023).

Additionally, ChatGPT has also facilitated the development of AI literacy in students. Its application in fields such as world religions (Chen et al., 2023), Business Administration (Albert & Li, 2023), and Journalism (Irfan et al., 2023) demonstrates its potential to foster digital competence.

In contrast to the generally positive findings regarding ChatGPT in university contexts, Swargiary (2023) found evidence of a negative correlation between the use of this tool and students' academic achievement, critical reasoning, and motivation.

DISCUSSION

In this section, a synthesis of the SLR findings is presented, followed by an in-depth comparative analysis that juxtaposes the current review's outcomes with those of previous studies. This critical evaluation enables a nuanced understanding of the research landscape and highlights the unique contributions of this study.

IDENTITY AND METHODOLOGICAL DESIGN OF THE STUDIES UNDER INVESTIGATION

While the United States accounted for a substantial portion (12%) of the studies reviewed, Asian countries exhibited a more pronounced presence, contributing 46% of the overall articles. China was particularly prominent within this group, a trend potentially attributable to demographic scale or the region's strong emphasis on technological innovation and AI in education (AIED). Most studies were published in 2023 and early 2024, coinciding with the initial deployment of ChatGPT. A gradual increase in publications was observed throughout 2023.

The primary focus of these studies was on higher education, a finding consistent with previous literature reviews (Zhang & Tur, 2023). However, this review expands on prior work by including a substantial proportion of studies (25%) dedicated to K-12 education, with one study even delving into preschool settings (Luo et al., 2024).

Regarding subject areas, foreign languages, particularly English, dominated the research. Studies exploring STEM fields, while present, were less frequent, aligning with the findings of AlBadarin et al. (2023). Moreover, previous systematic literature reviews often lacked detailed descriptions of the specific domains in which ChatGPT was applied.

In terms of research methodology, a mixed-methods approach was most prevalent, closely followed by qualitative methods. Quantitative methods were comparatively less frequently employed.

Indicative Best Practices and Use Cases of ChatGPT in School and University Education

In primary education, teachers commonly used ChatGPT to create quizzes (Tlili et al., 2023). In history, it was utilized to generate personalized learning materials (Jauhiainen & Guerra, 2023), while in physics, it was employed to develop lesson plans based on the 5E model and create assessment rubrics (Cooper, 2023). In a high school world religions course, teachers integrated the tool into the curriculum, encouraging students to achieve more accurate results through prompt engineering and to verify the model's responses (Chen et al., 2023).

Primary school students, on the other hand, generally used ChatGPT as a tool to prepare for exams, find ideas for essay writing, and understand concepts (Forman et al., 2023). In physics, students found examples and supplementary explanations (Alneyadi & Wardat, 2023) and completed assignments more quickly (Bitzenbauer, 2023). Similarly, in mathematics, students used ChatGPT for immediate feedback, explanations of mathematical concepts, and problem-solving (Wardat et al., 2023).

Specific use cases were identified in two studies. Bitzenbauer (2023) implemented a scenario in a quantum physics course involving generative and critical use of the model based on the think-pair-share teaching method. Additionally, Chen et al. (2023) integrated ChatGPT into a knowledge-building instructional model in a world religions course.

Systematic literature reviews by C. K. Lo (2023) and Vargas-Murillo et al. (2023), as well as the UNESCO guidelines (2023), confirm some of the practices identified in this review for higher education. However, the current study provides additional data from specific scientific fields.

The research revealed intriguing practices to address ChatGPT's limitations in providing accurate mathematical calculations. Specifically, the model was leveraged to generate the steps or instructions for solving problems, and subsequently, calculations were performed using software like Mathematica or Maxima. Alternatively, the Wolfram Alpha plugin was employed for more precise mathematical computations (Matzakos et al., 2023; Sánchez-Ruiz et al., 2023). In Physics, educators seemed to utilize it for creating high-quality assessment tasks (Küchemann et al., 2023). In Computer Science, it was used to generate personalized exercises in conjunction with the Rasch model (Niu & Xue, 2023). In Medicine, it was found to assist in creating lesson plans and assessment tests (Han et al., 2023). In Business Administration, it was employed as a co-grader for assignments (Albert & Li, 2023).

Students, on the other hand, used it as an aid to explain mathematical concepts and provide steps for problem-solving (Sánchez-Ruiz et al., 2023). In Chemistry, it assisted in solving problems, primarily those involving memorization (Daher et al., 2023). Future teachers employed it to create lesson plans, learning materials, and practice and assessment tasks (van den Berg & du Plessis, 2023). In journalism, it contributed to improving writing skills and generating ideas quickly (Irfan et al., 2023). In the field of programming, it was utilized to generate code snippets and assist in identifying and correcting code errors (Yilmaz & Karaoglan Yilmaz, 2023). In Medicine, students used it as a medical writing tool (Ho et al., 2023).

Regarding use cases, Kong et al. (2023) detailed a process for utilizing the model in a mass transfer course within a Chemical Engineering curriculum. In Chemistry, Guo and Lee (2023) presented steps for effectively employing the model in creating an essay. Ho et al. (2023) implemented a scenario for drafting medical reports. Finally, Michalon and Camacho-Zuñiga (2023) described a series of preparatory activities to train students in the field of international relations, enabling them to provide suitable prompts and have more effective conversations with ChatGPT.

In language learning, educators have employed ChatGPT to create dialogues and short stories, generate exercises and assessments, produce personalized learning materials and tasks, and automate written assessment and error detection (Al-Garaady & Mahyoob, 2023; Aydin Yildiz, 2023; Escalante et al., 2023; Jeon & Lee, 2023; Nguyen, 2023; Young & Shishido, 2023). Learners have utilized

ChatGPT as a conversational partner, aiding in language practice and contributing to the improvement of vocabulary and grammar (Javier & Moorhouse, 2024; Li et al., 2023; Roy & Putatunda, 2023), as well as a personal tutor for receiving personalized feedback and checking written work before final submission (Athanassopoulos et al., 2023; Li et al., 2023; Xiao & Zhi, 2023).

Use cases of the model in language learning involve the implementation of the Relationship Building Protocol by Ghafouri (2024), the structured pedagogical approach combining laboratory sessions and collaborative practices by Yan (2023), the method of assigning the role of a grader to GPT-4 by Escalante et al. (2023), the stages of integrating ChatGPT as an assessment tool to support German writing by Athanassopoulos et al. (2023), the process of creating high-level critical thinking on English literature by Roy and Putatunda (2023), and the practice of Li et al. (2023) to support autonomous learning in academic Chinese writing.

Focusing on the potential key differences in ChatGPT use across educational levels it is observed a) differences in the complexity of use: the cases of use evolve from simple interaction and personalized assistance in elementary education (e.g. conversational agent and personal assistant for assignments and exam preparation or as a tool for understanding complex concepts and solving problems) to more structured, collaborative, and advanced problem-solving roles in higher education (e.g. assistant for advanced concepts comprehension in mathematics, data science, and programming, including code generation and debugging b) differences in the level of autonomy: In higher education, ChatGPT is often used for facilitating autonomy in learning (e.g. mentor for assignment improvement and grader or an advisor in academic writing), whereas in lower levels, it is more teacher-directed (e.g. assistant in mathematics and essay writing and examples generator) c) differences in teacher's role: While teachers at all levels can benefit from using ChatGPT to create materials and assess students, the focus of their use shifts from creating basic materials (creating personalized learning materials and quizzes) to designing more complex learning experiences (developing prompts for students to use ChatGPT effectively, using ChatGPT as a grading assistant).

The above differences demonstrate that ChatGPT's role adapts to the developmental and academic demands of each educational stage, fostering both teaching efficiency and student independence.

Performance of ChatGPT in Various Subject Areas

Within the context of English as a Foreign Language (EFL) course, the tool demonstrated the capacity to produce high-quality essays (Waltzer et al., 2023) as well as reference dialogues that met high standards of readability, as measured by various metrics (Young & Shishido, 2023).

In *Chemistry*, the model exhibited strong performance in conceptual understanding and problem-solving reasoning within the domain of Materials Science, although it encountered difficulties with depth and made errors in numerical calculations (Daher et al., 2023).

In *History*, it demonstrated superior performance in generating argumentative historical text compared to pre-service elementary school teachers, based on most dimensions of historical thinking (Tirado-Olivares et al., 2023).

In *Mathematics*, it was able to solve simple operations and mathematical problems quickly and easily, providing detailed steps; however, there were instances of errors in numerical calculations (Matzakos et al., 2023; Sánchez-Ruiz et al., 2023; Wardat et al., 2023).

In *Medicine*, it appeared to outperform medical students on complex questions related to vaccination (Baglivo et al., 2023).

Finally, ChatGPT achieved high scores in the context of take-home assignments in the humanities and social sciences (Farazouli et al., 2023).

Rudolph et al. (2023) confirm the model's good performance in simple mathematical problems and its adequacy in historical knowledge. However, while Lo (2023) corroborated the model's difficulties with

mathematical calculations, they also found it to perform poorly in medical education, a finding that our study does not corroborate.

Limitations and Ethical Issues

The most frequent limitation (cited in 17 studies) across a wide range of scientific fields in the SLR was the presence of inaccurate or outdated responses, as well as unreliable or fabricated references. Numerous SLRs, including those by Labadze et al. (2023), Lo (2023), and Zhang and Tur (2023), consistently indicate a high likelihood of models generating incorrect, unreliable, or even factually inaccurate responses, often referred to as 'model hallucinations' (AlBadarin et al., 2023; Montenegro-Rueda et al., 2023; Vargas-Murillo et al., 2023).

The model's deficiencies in expressing emotions and engaging in creative, deep, and critical thinking were also frequently cited (in 9 studies), a finding less common in previous reviews. Only Zhang and Tur (2023) corroborated ChatGPT's limitations in higher-order thinking, highlighting its inability to address complex application-level problems.

Subsequently, eight studies highlighted the risk of academic integrity violations. Consistently, the vast majority of previous reviews have pointed to the risk of compromising academic integrity through the use of ChatGPT in educational settings (AlBadarin et al., 2023; İpek et al., 2023; Labadze et al., 2023; Lo, 2023; Vargas-Murillo et al., 2023; Zhang & Tur, 2023).

A novel finding emerged from seven articles, which noted the absence of references, indicating potential copyright infringement. This issue had not been a focal point in prior literature reviews.

Furthermore, findings from five studies revealed, on the one hand, the risk of excessive student reliance with a negative impact on a student's critical thinking and, on the other, the risk of social division due to unequal access to the tool and the potential for creating and reinforcing biased content. Concerns regarding the negative impact of over-reliance on ChatGPT on students' higher-order thinking skills are corroborated by Labadze et al. (2023), Vargas-Murillo et al. (2023), and Zhang and Tur (2023). Additionally, İpek et al. (2023) highlighted the risk of reinforcing biased results due to algorithmic biases inherent in the model's training data.

Additionally, the risk of privacy violations (cited in 4 studies) and the possibility of malicious use of the tool (cited in 3 studies) were demonstrated. The ethical implications of using ChatGPT in education, including concerns about personal data privacy, model misuse, and student safety, are corroborated by the findings of İpek et al. (2023), Labadze et al. (2023), and Zhang and Tur (2023).

Impact of ChatGPT on Student's Performance, Skills, and Motivation

While there are limited exceptions, most studies reviewed support the notion that ChatGPT can be a powerful tool for enhancing learning outcomes and developing critical thinking skills, a finding consistent with previous reviews (AlBadarin et al., 2023; İpek et al., 2023; Montenegro-Rueda et al., 2023).

Cognitive performance was most significantly impacted in various academic disciplines, including primary school History (Jauhiainen & Guerra, 2023), secondary school Physics (Alneyadi & Wardat, 2023; Bitzenbauer, 2023), and generally in primary and secondary education (Abdelghani et al., 2024; Forman et al., 2023). In higher education, student performance improved in Mathematics (Sánchez-Ruiz et al., 2023), Programming (Yilmaz & Karaoglan Yilmaz, 2023), Data Science (Albdrani & Al-Shargabi, 2023), Research Methodology (Essel et al., 2024), Business Administration (Albert & Li, 2023), Construction (Uddin et al., 2023), and English (Aydin Yildiz, 2023).

Furthermore, critical thinking was enhanced when students used the model in a secondary-level world religions course as well as in the university fields of Journalism (Irfan et al., 2023), Research Methodology (Essel et al., 2024), STEM (Guo & Lee, 2023; Sánchez-Ruiz et al., 2023), teacher education programs (van den Berg & du Plessis, 2023), and English language learning (Xiao & Zhi, 2023).

Furthermore, ChatGPT positively influenced student motivation and engagement, a finding corroborated by previous reviews (AlBadarin et al., 2023; Montenegro-Rueda et al., 2023). For instance, in English language learning, the tool increased motivation and sustained student persistence (Aydin Yildiz, 2023; Ghafouri, 2024). Similarly, in academic fields such as data science, programming, and business administration, ChatGPT appeared to positively impact student engagement.

Numerous studies in this review confirm the enhancement of language skills in the areas of vocabulary, grammar, discussion, and written expression among students who utilize ChatGPT while learning foreign languages (Athanassopoulos et al., 2023; Li et al., 2023; Shaikh et al., 2023; Xiao & Zhi, 2023).

Additionally, ChatGPT had a positive effect on students' digital literacy (Albert & Li, 2023; Chen et al., 2023; Irfan et al., 2023).

An exception to the positive impact of ChatGPT is the research of Swargiary (2023), which found a negative effect on student performance, critical thinking, and motivation. The literature review by Vargas-Murillo et al. (2023) confirms that excessive reliance on ChatGPT can hinder students' critical thinking skills.

Overall, emerging technologies like ChatGPT, with numerous applications for educators and learners, as revealed by the present SLR, have the potential to be a powerful tool for enhancing education. However, concerns about privacy, academic integrity, and the potential for misuse and bias linked to AI ethical principles of transparency and security, justice and equality, as well as the avoidance of harm (Holmes et al., 2022; Tzimas, 2021; Waelen, 2022) necessitate a cautious and responsible approach from educational stakeholders. Educational institutions must prioritize teaching critical thinking, digital literacy, and the ethical use of AI to ensure that these technologies are used for good.

CONCLUSION

This study employed the stages of a Systematic Literature Review and adhered to the PRISMA protocol to collect and analyze data from 50 empirical studies focused on the application of the advanced conversational AI model, ChatGPT, in K-12 and higher education.

The findings revealed a diverse range of practices and use cases for the model among educators and students, primarily in higher education, which accounted for 62% of the articles. Also, in K-12 education, where 25% of the evidence was found, the research gap was addressed to some extent, given the rarity of research in the K-12 domain.

Numerous applications of the model were identified in language learning within formal education settings. These include creating dialogues, generating exercises, automated grading of written work for educators, language practice, and personalized feedback through its use as a conversational partner and personal tutor by students.

While these findings align with the results of previous reviews, this SLR enriches the literature with additional use cases of the model in K-12 and higher education, as well as in language learning. The results of this research demonstrate that ChatGPT exhibited strong performance in various academic domains, including English literature essay writing, conceptual understanding and problem-solving in Chemistry, and Mathematics, historical argumentation, answering complex medical queries, and humanities and social sciences homework assignments. Rudolph et al. (2023) reached similar conclusions regarding ChatGPT's performance.

Significant concerns regarding ChatGPT's limitations and the ethical implications of its use in education, as identified in the literature (Books, 2023; UNESCO, 2023) and previous reviews, are corroborated by the findings of this study. The results reveal a high prevalence of inaccurate or outdated information, as well as unreliable citations. Additionally, the model demonstrated notable deficiencies in expressing emotions, engaging in creative and deep thinking, and critical reasoning. This finding, less common in previous reviews, may be attributed to the demand for creativity and critical thinking in the

theoretical cognitive fields examined and the focus of previous studies on lower educational levels where emotional interaction is more pronounced.

Consistent with previous reviews, this research demonstrated a positive impact of ChatGPT on student performance and skills, particularly in knowledge acquisition, critical thinking, motivation, engagement, language proficiency, and digital literacy across various academic disciplines at both the school and university levels.

Significant constraints arose from the application of the rigorous PRISMA protocol criteria, including limiting the time frame of publications to the first year of the model's application (January 2023-January 2024), focusing the search on data within the context of school and university education, restricting the search to English-language articles, and limiting the scope to articles accessible through open or academic access. Another limitation was the exclusive focus on ChatGPT, given the proliferation of other GPT-based CAI models (Rudolph et al., 2023) that could yield valuable data for educational research. Additionally, while studies focusing on ChatGPT in school education were included, primary education was underrepresented in the target groups of this study.

Finally, research examining the use of ChatGPT in special education was not included.

This SLR identified several limitations that underpinned the formulation of recommendations for future research in the field of Conversational Artificial Intelligence (CAI) in education. Considering the potential benefits and ethical considerations associated with using conversational AI like ChatGPT in education, as evidenced by this SLR and previous research, a strategic and cautious approach is essential to maximize the positive impact of this technology on students and educators in both school and university settings. To address the challenges and limitations identified in this study, educational institutions should establish clear ethical guidelines for AI use, emphasizing transparency, fairness, and accountability. Furthermore, comprehensive training on AI- ethics, responsible use, and critical thinking skills should be provided to both educators and students.

Additionally, equipping educators with the skills necessary to effectively integrate AI tools into their teaching practices is crucial. Governments and relevant stakeholders, like local, regional, and national educational authorities, must prioritize promoting AI literacy among members of the educational community to empower them to critically assess AI-generated content, identify biases, and use AI tools responsibly. Finally, fostering continuous collaboration among researchers, AI developers, and educators is vital to creating innovative and appropriate applications that address emerging challenges. By implementing these recommendations and regularly evaluating the impact of AI tools on student learning outcomes, educational systems can harness the potential of AI to deliver more engaging and effective learning experiences.

REFERENCES

Abdelghani, R., Wang, Y.-H., Yuan, X., Wang, T., Lucas, P., Sauzéon, H., & Oudeyer, P.-Y. (2024). GPT-3-driven pedagogical agents to train children's curious question-asking skills. *International Journal of Artificial Intelligence in Education, 34*, 483-518. https://doi.org/10.1007/s40593-023-00340-7

Akiba, D., & Fraboni, M. C. (2023). AI-Supported academic advising: Exploring ChatGPT's current state and future potential toward student empowerment. *Education Sciences, 13*(9), 885. https://doi.org/10.3390/educsci13090885

AlBadarin, Y., Tukiainen, M., Saqr, M., & Pope, N. (2023). *A systematic literature review of empirical research on ChatGPT in education* (SSRN Scholarly Paper No. 4562771). https://doi.org/10.2139/ssrn.4562771

Albdrani, R. N., & Al-Shargabi, A. A. (2023). Investigating the effectiveness of ChatGPT for providing personalized learning experience: A case study. *International Journal of Advanced Computer Science and Applications, 14*(11), 1208-1213. https://doi.org/10.14569/IJACSA.2023.01411122

Albert, D., & Li, T. (2023). *Insights from teaching with AI: How ChatGPT can enhance experiential learning and assist instructors* (SSRN Scholarly Paper No. 4516801). https://doi.org/10.2139/ssrn.4516801

Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals, 9,* 3-17. https://doi.org/10.24093/awej/call9.1

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

Al-Obaydi, L. H., Pikhart, M., & Klimova, B. (2023). ChatGPT and the general concepts of education: Can artificial intelligence-driven chatbots support the process of language learning? *International Journal of Emerging Technologies in Learning, 18*(21), 39-50. https://doi.org/10.3991/ijet.v18i21.42593

Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research, 3*(2), 818-824. https://doi.org/10.25082/amler.2023.02.009

Aydin Yildiz, T. (2023). The impact of ChatGPT on language learners' motivation. *Journal of Teacher Education and Lifelong Learning, 5*(2), 582-897. https://doi.org/10.51535/tell.1314355

Baglivo, F., Angelis, L. D., Casigliani, V., Arzilli, G., Privitera, G. P., & Rizzo, C. (2023). Exploring the possible use of AI chatbots in public health education: Feasibility study. *JMIR Medical Education, 9*, e51421. https://doi.org/10.2196/51421

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology, 15*(3), ep430. https://doi.org/10.30935/cedtech/13176

Books, T. J. (2023). *Mastering ChatGPT: Create highly effective prompts, strategies, and best practices to go from novice to expert.* TJ Books.

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

Chen, B., Zhu, X., & Díaz del Castillo, H. F. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence, 5*, 100184. https://doi.org/10.1016/j.caeai.2023.100184

Churi, P. P., Joshi, S., Elhoseny, M., & Omrane, A. (Eds.). (2022). *Artificial intelligence in higher education: A practical approach* (1st ed.). CRC Press. https://doi.org/10.1201/9781003184157

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

Crompton, H., Burke, D., Jordan, K., & Wilson, S. (2021). Learning with technology during emergencies: A systematic review of K-12 education. *British Journal of Educational Technology, 52*(4), 1554-1575. https://doi.org/10.1111/bjet.13114

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

Elkhatat, A. M. (2023). Evaluating the authenticity of ChatGPT responses: A study on text-matching capabilities. *International Journal for Educational Integrity, 19*, Article 15. https://doi.org/10.1007/s40979-023-00137-0

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*, Article 57. https://doi.org/10.1186/s41239-023-00425-2

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence, 6*, 100198. https://doi.org/10.1016/j.caeai.2023.100198

European Commission. (2018, March 9). *Statement on artificial intelligence, robotics and 'autonomous' systems.* https://data.europa.eu/doi/10.2777/531856

European Commission. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators.* https://data.europa.eu/doi/10.2766/153756

European Parliament. (2023). *What is artificial intelligence and how is it used?* https://www.europarl.europa.eu/news/en/headlines/society/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education, 49*(3), 363-375. https://doi.org/10.1080/02602938.2023.2241676

Fink, A. (2005). *Conducting research literature reviews: From the internet to paper*. Sage.

Forman, N., Udvaros, J., & Avornicului, M. S. (2023). ChatGPT: A new study tool shaping the future for high school students. *International Journal of Advanced Natural Sciences and Engineering Research, 7*(4), 95-102. https://doi.org/10.59287/ijanser.562

Ghafouri, M. (2024). ChatGPT: The catalyst for teacher-student rapport and grit development in L2 class. *System, 120*, 103209. https://doi.org/10.1016/j.system.2023.103209

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education, 100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews, 18*(2), e1230. https://doi.org/10.1002/cl2.1230

Han, Z., Battaglia, F., Udaiyar, A., Fooks, A., & Terlecky, S. R. (2023). An explorative assessment of ChatGPT as an aid in medical education: Use it with caution. *Medical Teacher, 46*(5), 657-664. https://doi.org/10.1080/0142159X.2023.2271159

Ho, W. L. J., Koussayer, B., & Sujka, J. (2023). ChatGPT: Friend or foe in medical writing? An example of how ChatGPT can be utilized in writing case reports. *Surgery in Practice and Science, 14*, 100185. https://doi.org/10.1016/j.sipas.2023.100185

Holmes, W., & Porayska-Pomsta, K. (2022). *The ethics of artificial intelligence in education: Practices, challenges, and debates*. Routledge. https://doi.org/10.4324/9780429329067

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence Education*, *32*, 504-526. https://doi.org/10.1007/s40593-021-00239-1

İpek, Z. H., Gözüm, A. İ. C., Papadakis, S., & Kallogiannakis, M. (2023). Educational applications of the ChatGPT AI System: Systematic review research. *Educational Process International Journal, 12*(3), 26-55. https://doi.org/10.22521/edupij.2023.123.2

Irfan, M., Murray, L., & Ali, S. (2023). Integration of Artificial intelligence in academia: A case study of critical teaching and learning in higher education. *Global Social Sciences Review, VIII*(I), 352-364. https://doi.org/10.31703/gssr.2023(viii-i).32

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability, 15*(18), 14025. https://doi.org/10.3390/su151814025

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3*(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115

Javier, D. R. C., & Moorhouse, B. L. (2024). Developing secondary school English language learners' productive and critical use of ChatGPT. *TESOL Journal, 15*(2), e755. https://doi.org/10.1002/tesj.755

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies, 28*, 15873-15892. https://doi.org/10.1007/s10639-023-11834-1

Kong, Z. Y., Adi, V. S. K., Segovia-Hernández, J. G., & Sunarso, J. (2023). Complementary role of large language models in educating undergraduate design of distillation column: Methodology development. *Digital Chemical Engineering, 9*, 100126. https://doi.org/10.1016/j.dche.2023.100126

Kortemeyer, G. (2023). Could an artificial-intelligence agent pass an introductory physics course? *Physical Review Physics Education Research, 19*(1), 10132. https://doi.org/10.1103/PhysRevPhysEducRes.19.010132

Kostas, A., Paraschou, V., Spanos, D., & Sofos, A. (2023). Emergency remote teaching in K-12 education during COVID-19 pandemic: A systematic review of empirical research in Greece. In T. Bratitsis (Eds.), *Research on E-Learning and ICT in Education* (pp. 235-260). Springer. https://doi.org/10.1007/978-3-031-34291-2_14

Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Can ChatGPT support prospective teachers in physics task development? *Physical Review Physics Education Research, 19*(2), 20128. https://doi.org/10.1103/PhysRevPhysEducRes.19.020128

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

Lame, G. (2019). Systematic literature reviews: An introduction. *Proceedings of the Design Society: International Conference on Engineering Design, 1*(1), 1633-1642. https://doi.org/10.1017/dsi.2019.169

Lee, S. (2023). *AI toolkit for educators.* https://www.ai4t.eu/wp-content/uploads/2023/10/AI-Toolkit-for-Educators_v2_def_for_now.pdf

Li, J., Ren, X., Jiang, X., & Chen, C.-H. (2023). Exploring the use of ChatGPT in Chinese language classrooms. (2023). *International Journal of Chinese Language Teaching, 4*(3), 36-55. https://doi.org/10.46451/ijclt.20230303

Lo, C. K. (2023). What Is the Impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Lo, C. K., & Hew, K. F. (2023). A review of integrating AI-based chatbots into flipped learning: New possibilities and challenges. *Frontiers in Education, 8*, 1175715. https://doi.org/10.3389/feduc.2023.1175715

Luo, W., He, H., Liu, J., Berson, I. R., Berson, M. J., Zhou, Y., & Li, H. (2024). Aladdin's Genie or Pandora's Box for early childhood education? Experts chat on the roles, challenges, and developments of ChatGPT. *Early Education and Development, 35*(1), 96-113. https://doi.org/10.1080/10409289.2023.2214181

Matzakos, N., Doukakis, S., & Moundridou, M. (2023). Learning mathematics with large language models: A comparative study with computer algebra systems and other tools. *International Journal of Emerging Technologies in Learning, 18*(20), Article 20. https://doi.org/10.3991/ijet.v18i20.42979

Meo, S. A., Al-Masri, A. A., Alotaibi, M., Meo, M. Z. S., & Meo, M. O. S. (2023). ChatGPT knowledge evaluation in basic and clinical medical sciences: Multiple choice question examination-based performance. *Healthcare, 11*(14), 2046. https://doi.org/10.3390/healthcare11142046

Meron, Y., & Tekmen Araci, Y. (2023). Artificial intelligence in design education: Evaluating ChatGPT as a virtual colleague for post-graduate course development. *Design Science, 9,* e30. https://doi.org/10.1017/dsj.2023.28

Michalon, B., & Camacho-Zuñiga, C. (2023). ChatGPT, a brand-new tool to strengthen timeless competencies. *Frontiers in Education, 8*, 1251163. https://doi.org/10.3389/feduc.2023.1251163

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers, 12*(8), Article 8. https://doi.org/10.3390/computers12080153

Nguyen, T. T. H. (2023). EFL Teachers' perspectives toward the Use of ChatGPT in writing classes: A case study at Van Lang University. *International Journal of Language Instruction, 2*(3), 1-47. https://doi.org/10.54855/ijli.23231

Niu, Y., & Xue, H. (2023). Exercise generation and student cognitive ability research based on ChatGPT and Rasch Model. *IEEE Access, 11*, 116695-116705. https://doi.org/10.1109/ACCESS.2023.3325741

OECD. (2020). *Trustworthy AI in education: Promises and challenges.* https://doi.org/10.1787/a6c90fa9-en

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M.

M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine, 18*(3), e1003583. https://doi.org/10.1371/journal.pmed.1003583

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell. https://doi.org/10.1002/9780470754887

Roy, D., & Putatunda, T. (2023). From textbooks to chatbots: Integrating AI in English literature classrooms of India. *Journal of E-Learning and Knowledge Society, 19*(3), 65–73. https://doi.org/10.20368/1971-8829/1135860

Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching, 6*(1), 364-389. https://doi.org/10.37074/jalt.2023.6.1.23

Saldana, J. (2015). *The coding manual for qualitative researchers*. Sage.

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences, 13*(10), 6039. https://doi.org/10.3390/app13106039

Self, J. A. (1974). Student models in computer-aided instruction. *International Journal of Man-Machine Studies*, *6*(2), 261-276. https://doi.org/10.1016/s0020-7373(74)80005-2

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education, 13*(9), 1937-1960. https://doi.org/10.3390/ejihpe13090140

Skrabut, S. (2023). *80 ways to use ChatGPT in the classroom: Using AI to enhance teaching and learning*. https://zoboko.com/book/2d2ngvje/80-ways-to-use-ChatGPT-in-the-classroom-using-ai-to-enhance-teaching-and-learning

Swargiary, K. (2023). *The impact of ChatGPT on student learning behavior: An experimental research study*. SSRN. https://doi.org/10.2139/ssrn.4611213

Tirado-Olivares, S., Navío-Inglés, M., O'Connor-Jiménez, P., & Cózar-Gutiérrez, R. (2023). From human to machine: Investigating the effectiveness of the conversational AI ChatGPT in historical thinking. *Education Sciences, 13*(8), 803. https://doi.org/10.3390/educsci13080803

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x

Trinoskey, J., Brahmi, F. A., & Gall, C. (2009). Zotero: A product review. *Journal of Electronic Resources in Medical Libraries, 6*(3), 224-229. https://doi.org/10.1080/15424060903167229

Tzimas, T. (2021). *Legal and ethical challenges of artificial intelligence from an international law perspective*. Springer. https://doi.org/10.1007/978-3-030-78585-7_4

Uddin, S. M. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging CHATGPT to aid construction hazard recognition and support safety education and training. *Sustainability, 15*(9), 7121. https://doi.org/10.3390/su15097121

UNESCO. (2021). *AI and education: Guidance for policymakers*. https://unesdoc.unesco.org/ark:/48223/pf0000376709

UNESCO. (2022). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. https://unesdoc.unesco.org/ark:/48223/pf0000380602

UNESCO. (2023). *ChatGPT and Artificial Intelligence in Higher Education: Quick start guide*. https://unesdoc.unesco.org/ark:/48223/pf0000385146

van den Berg, G., & du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences, 13*(10), 998. https://doi.org/10.3390/educsci13100998

Vargas-Murillo, A. R., de la Asuncion Pari-Bedoya, I. N. M., & de Jesús Guevara-Soto, F. (2023). Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research, 22*(7), 122-135. https://doi.org/10.26803/ijlter.22.7.7

Vázquez-Cano, E., Ramírez-Hurtado, J. M., Sáez-López, J. M., & López-Meneses, E. (2023). ChatGPT: The brightest student in the class. *Thinking Skills and Creativity, 49*, 101380. https://doi.org/10.1016/j.tsc.2023.101380

Villan, F., & dos Santos, R. P. (2023). ChatGPT as co-advisor in scientific initiation: Action research with project-based learning in elementary education. *Acta Scientiae, 25*(6), 60-117. https://doi.org/10.17648/acta.scientiae.7474

Waelen, R. (2022). Why AI ethics is a critical theory. *Philosophy and Technology*, *35*, Article 9. https://doi.org/10.1007/s13347-022-00507-5

Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the ability of teachers and students to differentiate between essays generated by ChatGPT and high school students. *Human Behavior and Emerging Technologies, 2023*, Article 1923981. https://doi.org/10.1155/2023/1923981

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), em2286. https://doi.org/10.29333/ejmste/13272

Wolfram, S. (2023). *What is ChatGPT doing … and why does it work?* https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/

World Commission on the Ethics of Scientific Knowledge and Technology (COMEST). (2019). *Preliminary study on the ethics of artificial intelligence.* https://ircai.org/wp-content/uploads/2020/07/PRELIMINARY-STUDY-ON-THE-ETHICS-OF-ARTIFICIAL-INTELLIGENCE.pdf

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3), 212. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on learners in an L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*, 13943-13967. https://doi.org/10.1007/s10639-023-11742-4

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence, 4*, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating Chatbot's dialogue for English as a foreign language learning. *International Journal of Advanced Computer Science and Applications, 14*(6), 65-72. https://doi.org/10.14569/IJACSA.2023.0140607

Yu, S., & Lu, Y. (2021). *An introduction to artificial intelligence in education.* Springer. https://doi.org/10.1007/978-981-16-2770-5

Zhang, P., & Tur, G. (2023). A systematic review of ChatGPT use in K-12 education. *European Journal of Education, 59*(2), e12599. https://doi.org/10.1111/ejed.12599

AUTHORS

Maria Dimeli holds a degree from the Department of Early Childhood Education at the National and Kapodistrian University of Athens. She has been working in a public kindergarten under the Greek Ministry of Education for 18 years. She also holds a master's degree in education sciences with a specialization in *Education with the Use of New Technologies* from the University of the Aegean. Her thesis was titled *Applications, Limitations, and Impact of ChatGPT in Education: A Systematic Literature Review*. Her research interests focus on the integration of new technologies into the educational process, particularly the utilization of AI tools for creating educational materials and designing interactive learning activities. Dimeli Maria is actively involved in training educators on the use of digital tools and their application in classroom settings.

Dr. Apostolos Kostas is an Assistant Professor in the Department of Primary Education, University of the Aegean, Greece. His research interests include e-learning, AI in Education, Distance Education, Communities of Practice and Social Networks in Education, Initial Teacher Education, and Teacher Professional Development. He has participated in numerous European and national research projects.