**REVIEW ARTICLE**

WILEY

# A systematic review of ChatGPT use in K-12 education

**Peng Zhang**[1] | **Gemma Tur**[2]

[1]Doctoral School, Universitat de les Illes Balears, Palma, Spain

[2]Department of Applied Pedagogy and Educational Psychology, Universitat de les Illes Balears, Palma, Spain

**Correspondence**
Peng Zhang, Doctoral School, Universitat de Illes Balears, Spain.
Email: p.zhang@uib.es

## Abstract

This systematic review, adhering to the PRISMA framework, investigated the utilisation of ChatGPT, a language model developed by OpenAI, throughout Kindergarten to 12th grade (K-12) educational settings. The review synthesises findings from 13 selected papers, encompassing the strengths, weaknesses, opportunities, and threats (SWOT) analysis of ChatGPT's implementation in K-12 education, implications for various stakeholders, and practical recommendations. It is highlighted that ChatGPT could empower educators through curriculum, lesson planning, materials generation, differentiation, and optimising student learning experience through personalised learning. However, concerns regarding academic integrity and output quality must be addressed. The paper provides pedagogical recommendations and ethical considerations to utilise ChatGPT better. It contributes to the ongoing discourse about AI, particularly ChatGPT's role in K-12 education, further inspiring future research and educational practices and facilitating the effective integration of ChatGPT into K-12 educational settings where collaboration arises as a key role, in particular under the approach of co-design for learning.

**KEYWORDS**
artificial intelligence, ChatGPT, K-12 education, systematic review

# 1 | INTRODUCTION

In recent years, the integration of technology, specifically artificial intelligence (AI), in the fields of education has garnered considerable attention from researchers (Luckin & Cukurova, 2019). The interdisciplinary nature of research in the learning sciences has facilitated a better understanding of human learning, which, in turn, has informed the development of AI technologies for educational purposes (Luckin & Cukurova, 2019). Notably, the deployment of AI in educational contexts has demonstrated numerous benefits. These include, but are not limited to, adaptive learning, enhancing teaching assessments, and the creation of virtual classrooms. By optimising and enriching the learning environment, AI can spark student engagement, foster initiative, and stimulate creativity (Huang et al., 2021). As Luckin and Cukurova (2019) argue, the potential of AI to harness a diverse array of data, firmly anchored in learning sciences research, gives rise to opportunities for scaffolding, tailored student support, and teacher assistance.

A prominent example of an AI tool making inroads into the educational landscape is the Generative Pretrained Transformer (GPT), a product of OpenAI. The GPT, a language model designed to generate human-like text responses, has been the subject of extensive exploration and application across a variety of educational contexts (Lund & Wang, 2023). The ChatGPT version of this model, which is specifically designed for natural language generation and conversation, has exhibited a significant impact in higher education settings. Numerous studies have explored its use and demonstrated the growing interest in using Chat GPT and other AI technologies in higher education (i.e., Cotton et al., 2023; Firat, 2023; Iskender, 2023). While there are potential benefits, such as increased student engagement and accessibility (Firat, 2023), there are also concerns regarding academic integrity and plagiarism (Cotton et al., 2023; Iskender, 2023).

With empirical studies on ChatGPT use emerging in higher education settings, multiple literature reviews, including systematic reviews have been conducted (Lo, 2023; Mohammad et al., 2023; Perera & Lankathilaka, 2023). However, these review papers predominantly focus on higher education, particularly in the field of academic writing, medical or healthcare education. There is a notable gap in the review literature when it comes to its application in kindergarten to 12th grade (K-12) education contexts. The potential of ChatGPT in K-12 education has been explored in several empirical studies (i.e., Lozano & Blanco Fontao, 2023; Woo et al., 2023). Lozano and Blanco Fontao (2023) claimed that it is important to ensure that students and teachers have sufficient knowledge about the implementation of ChatGPT to ensure its correct use and maintain the quality of the education system. As such, it is crucial to review and synthesise the findings from the existing literature on ChatGPT use in K-12 education to summarise what we already know to further inspire future practices for researchers and educators. Thus, there is a need for a comprehensive review of the current state of research on the use of ChatGPT in K-12 education.

This paper aims to address this gap by conducting a systematic review of the existing literature on the use of ChatGPT in K-12 education since the emergence of the tool. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, a range of studies were examined to answer the following research questions:

- RQ1: What are the strengths, weaknesses, opportunities, and threats (SWOT) in incorporating ChatGPT in K-12 education contexts?
- RQ2: What are the current or possible practices of ChatGPT implementation in K-12 education contexts?
- RQ3: What are the future directions and practical recommendations for utilising ChatGPT in K-12 education?

Through this systematic review, this paper aims to provide a comprehensive perspective on the current state of research and practice, and to offer insights on how ChatGPT might be effectively integrated into K-12 education. By identifying the strengths, weaknesses, opportunities, and threats (SWOT) associated with the use of ChatGPT in these contexts, the study intends to contribute to the ongoing conversation about the role of AI in K-12 education and provide valuable guidance for educators, policymakers, and researchers.

## 2 | BACKGROUND

### 2.1 | AI, AIGC, and ChatGPT

Artificial Intelligence (AI) has made significant advances over the past several years and has become more prevalent in numerous facets of our life. AI represents the capacity of computer systems to carry out tasks that usually involve human intelligence, such as learning, reasoning, problem-solving, and making decisions. It entails the creation of computational models and algorithms that allow machines to process and analyse data, identify patterns, and make predictions or decisions according to the available information (Minkkinen & Mäntymäki, 2023).

AIGC, or Artificial Intelligence Generated Content, as a subset of AI tools, has recently attracted considerable attention (Cao et al., 2023). It refers to content generated using advanced Generative AI techniques instead of content created by human creators (Cao et al., 2023). This technology is capable of automating the generation of significant quantities of content in a short time (Cao et al., 2023). ChatGPT, a conversational language model created by OpenAI, is a well-known example of AIGC (Lin et al., 2023). AIGC techniques are not restricted to text generation, however. They also include the production of digital content, including visuals, audio, and natural language, using AI models (Cao et al., 2023). While AIGC can potentially transform content creation and dissemination, concerns regarding its responsible application include privacy, bias, toxicity, misinformation, and intellectual property (Chen et al., 2023).

As a text-form AIGC tool, ChatGPT, a chatbot with a conversational AI interface created by OpenAI (Tlili et al., 2023), is a noteworthy advancement in AI. It is a form of AI technology known as a large language model (LLM). It is a type of the Generative Pre-trained Transformer (GPT) model that is trained on a huge amount of conversational data (Lin et al., 2023). ChatGPT is designed to generate human-like text and interact with users. ChatGPT employs deep learning methods, particularly transformer architecture, to comprehend and produce responses in natural language. It has been taught grammar, syntax, and semantic relationships between words and phrases using a vast quantity of text data (Karakose et al., 2023).

### 2.2 | The emergence of ChatGPT in education

ChatGPT has attracted extensive attention and has been implemented in a variety of contexts. The use of AI chatbots like ChatGPT has gained significant attention, particularly in the field of education. A recent study conducted by Mogavi et al. (2023) through a social media content analysis (SMCA) revealed the prevalence of ChatGPT in various contexts, with higher education (24.18%), K-12 education (22.09%), and practical skills learning (15.28%) being the most frequent. AI chatbots, like ChatGPT, have the potential to provide tutoring and homework help by answering questions and providing explanations to help students understand complex concepts (Huh, 2023). Research indicates a generally positive public reaction and enthusiasm concerning the application of ChatGPT in educational contexts; however, hesitant voices highlight the demand for cautious incorporation and careful consideration of ethical concerns (Tlili et al., 2023). User experiences with ChatGPT in educational environments have brought up a number of issues, including dishonesty, concerns about privacy, and manipulation (Tlili et al., 2023). According to Mogavi et al. (2023), people on social media platforms frequently discuss ChatGPT in relation to productivity (73.19%), efficiency (63.02%), and ethics (48.51%).

ChatGPT is also gaining popularity in medical education and publishing. A review of indexed medical literature during the first four months of ChatGPT's existence demonstrated an expanding body of academic work on its applications and implications (Temsah et al., 2023). Literature has examined ChatGPT's influence on medical education, scientific research, medical writing, ethical considerations, diagnostic decision-making, automation potential, and criticisms (Temsah et al., 2023). In engineering education, ChatGPT has the potential to offer personalised and effective learning experiences by providing customised feedback, clarifications, and realistic virtual

simulations for hands-on learning (Qadir, 2022). Nevertheless, there are ethical concerns, such as the possibility of unethical use by students and the unemployment of human educators (Qadir, 2022). Moreover, ChatGPT has been described as a transformative technology that has the potential to speed up and improve research in a variety of disciplines; however, as with any technology, certain challenges and limitations should be addressed (Cheng, 2023). The use of ChatGPT in scientific and disciplinary publications has raised ethical concerns regarding deceptive behaviour (Milton, 2023).

Even though current literature widely discusses ChatGPT implementation in higher education settings, the implementation of ChatGPT in K-12 education also has the potential to revolutionise students' learning experiences, even though there is comparatively much less literature about it. According to a study conducted with students pursuing a degree in Primary Education at the University of León in Spain, student teachers view ChatGPT favourably and recognise its potential applications (Lozano & Blanco Fontao, 2023). Thus, this study systematically reviews the current literature on ChatGPT use in K-12 education and synthesises the findings and practices for future implications.

## 3 | METHODOLOGY

This academic paper employs a systematic review methodology to answer the research questions mentioned earlier. Systematic reviews are meant to comprehensively synthesise present knowledge related to a particular field, providing a consolidated summary of findings from multiple studies, and offering valuable insights for future research and practical applications (Moreno-Küstner et al., 2018). The review procedure adopted the well-established Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Page et al., 2021) to guarantee robust and transparent reporting. Various data acquisition and analysis software, including Ryaan, VOSviewer, and NVivo, were employed in this study. The sections that follow elaborate on the methodology details utilised in this research.

### 3.1 | Research process

The research adhered to Tawfik et al. (2019)'s model for the process of a systematic review (see Figure 1).

Upon finalising the research questions, this study commenced the systematic review procedure by conducting a preliminary search. The purpose of this preliminary search was to identify relevant papers, validate the research idea, and ascertain the availability of sufficient articles for comprehensive analysis. Based on the outcomes of the preliminary research, the search strategy and relevant terms were refined. Subsequently, inclusion and exclusion criteria were developed to discern which publications would be incorporated into the final sample and which would be excluded. With the preparations in place, the search process commenced across various databases. The initial search was followed by the implementation of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) procedure, which encompassed the data identification, screening, and inclusion phases. Upon retrieving the targeted publications, an in-depth analysis of the data was performed to derive meaningful insights from the collected information.

### 3.2 | Search terms and keywords

The process used to determine the search terms and keywords was based on the Boolean search strategy (Zhang et al., 2020). This strategy combined terms and keywords from different groups: "ChatGPT," "education," and related alternatives. The Boolean operators "AND" and "OR" were employed to delineate the keywords effectively (see Figure 2).
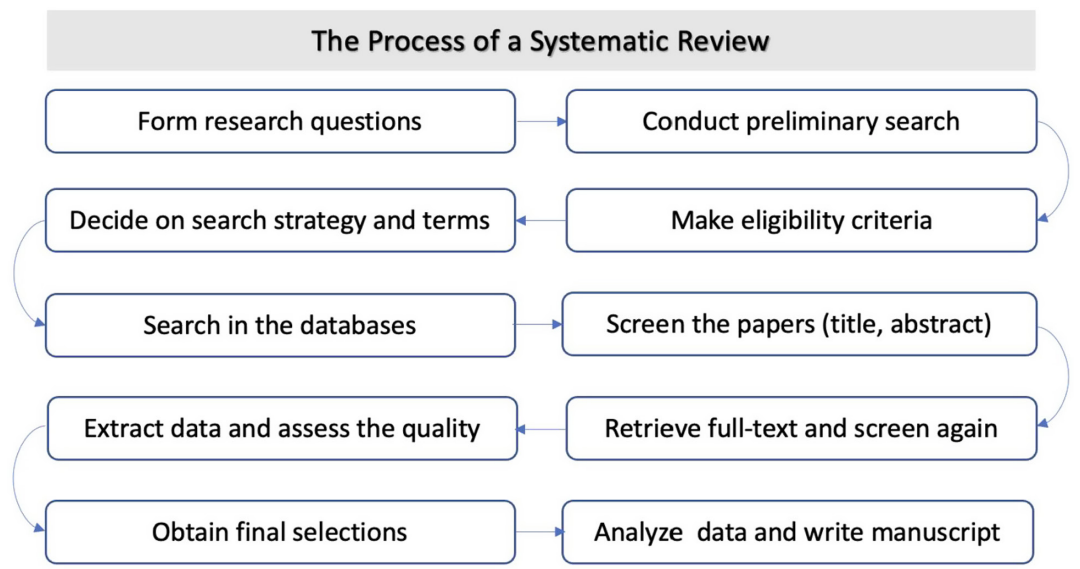
## The Process of a Systematic Review

| Form research questions | → | Conduct preliminary search |
| Decide on search strategy and terms | ← | Make eligibility criteria |
| Search in the databases | → | Screen the papers (title, abstract) |
| Extract data and assess the quality | ← | Retrieve full-text and screen again |
| Obtain final selections | → | Analyze data and write manuscript |

**FIGURE 1** The process of the systematic review.

**Search String**

Scopus, SpringerLink, Web of Science, IEEE Xplore, ACM library, ERIC

((("ChatGPT") OR ("GPT")) AND (("education") OR ("learning") OR ("school")))

Restricted to the field of education

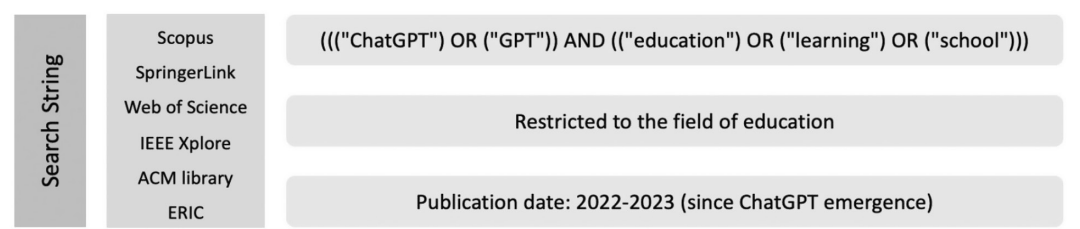Publication date: 2022-2023 (since ChatGPT emergence)

**FIGURE 2** The search string.

It is noted that the term "K-12 education" was not used as a search term since we found that very few papers appeared when we used this term during the preliminary search. To ensure inclusivity and encompass a broader range of publications adhering to the eligibility criteria, we opted for the more general term "education" instead. However, papers not directly related to K-12 education were filtered out during the screening process.

By implementing this Boolean search strategy, we were able to explore all possible combinations of terms present in the literature. The boolean operators "AND" and "OR" were utilised to split the keywords terms. The terms within each category were combined using "OR", and categories were linked using "AND". Specifically, the terms and alternative keywords used for ChatGPT were defined as ("ChatGPT" OR "GPT"), while the terms for K-12 education encompassed ("education" OR "learning" OR "school").

The study searched the literature in six prominent academic databases: Web of Science (WOS), Scopus, ERIC, SpringerLink, IEEE Xplore, and ACM Library. To focus on ChatGPT in the context of K-12 education, which emerged in 2022, a publication date filter was applied to exclude any works published before 2022. In addition, an education descriptor was applied consistently throughout the search process to refine the results and ensure their relevance to the subject matter under investigation.

## 3.3 | Eligibility criteria

In order to be considered for inclusion in the systematic review, the existing publications, as well as the associated search terms and keywords, needed to meet the eligibility criteria specified in Table 1.

## 3.4 | Database search and screening

The initial phase of the research involved conducting a comprehensive search through various online digital scientific databases and major journal databases. As a result of this thorough search process, a total of 1010 papers were initially identified from prominent databases, with the distribution as follows: Scopus ($n=272$), SpringerLink ($n=103$), Web of Science ($n=392$), ERIC ($n=18$), IEEE Xplore ($n=170$), and ACM Library ($n=55$). To facilitate the subsequent steps of the review process, the gathered papers were systematically managed and organised using the collaborative platform Rayyan, which was designed specifically for conducting systematic literature reviews (Ouzzani et al., 2016). In accordance with the updated Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) reporting standards (Page et al., 2021), data screening and eligibility assessments were carried out (refer to Figure 3 for details).

During the initial screening in Rayyan, 386 articles were successfully processed, and 624 duplicates were efficiently removed, streamlining the dataset. The remaining 334 papers underwent a meticulous title and abstract screening process based on the agreed-upon search criteria. Following this thorough evaluation, 52 papers showed potential for inclusion, but 7 of them were not accessible for retrieval.

Subsequently, the researchers obtained the full texts of the 45 remaining publications and meticulously assessed their alignment with the predetermined criteria. As a result of this rigorous evaluation, 32 papers were deemed unsuitable for inclusion, ultimately leading to the final selection of 13 papers that perfectly met the eligibility criteria, thus forming the basis of the systematic review.

## 3.5 | Selected literature analysis

The study is qualitative in nature; it employed the content analysis technique to examine the 13 selected publications based on the research questions outlined earlier. To enhance the efficiency and effectiveness of the analytical process, computer-assisted qualitative data analysis software (CAQDAS) was utilised, following the approach proposed by Leech and Onwuegbuzie (2011). The increasing popularity of CAQDAS among researchers conducting qualitative research or reviews can be attributed to its capabilities in "saving, indexing, sorting, and coding" data, as well as facilitating collaborative efforts (Leech & Onwuegbuzie, 2011, p. 71). The present study employed two CAQDAS tools, namely VOSviewer and NVivo, to analyse the retrieved data.

Once the paper selections were finalised, VOSviewer and NVivo were employed as analytical instruments to analyse further the chosen publications in alignment with the predetermined research inquiries. VOSviewer was employed to construct and visualise bibliometric networks. Its text-mining features were also used to

**TABLE 1** Eligibility criteria.

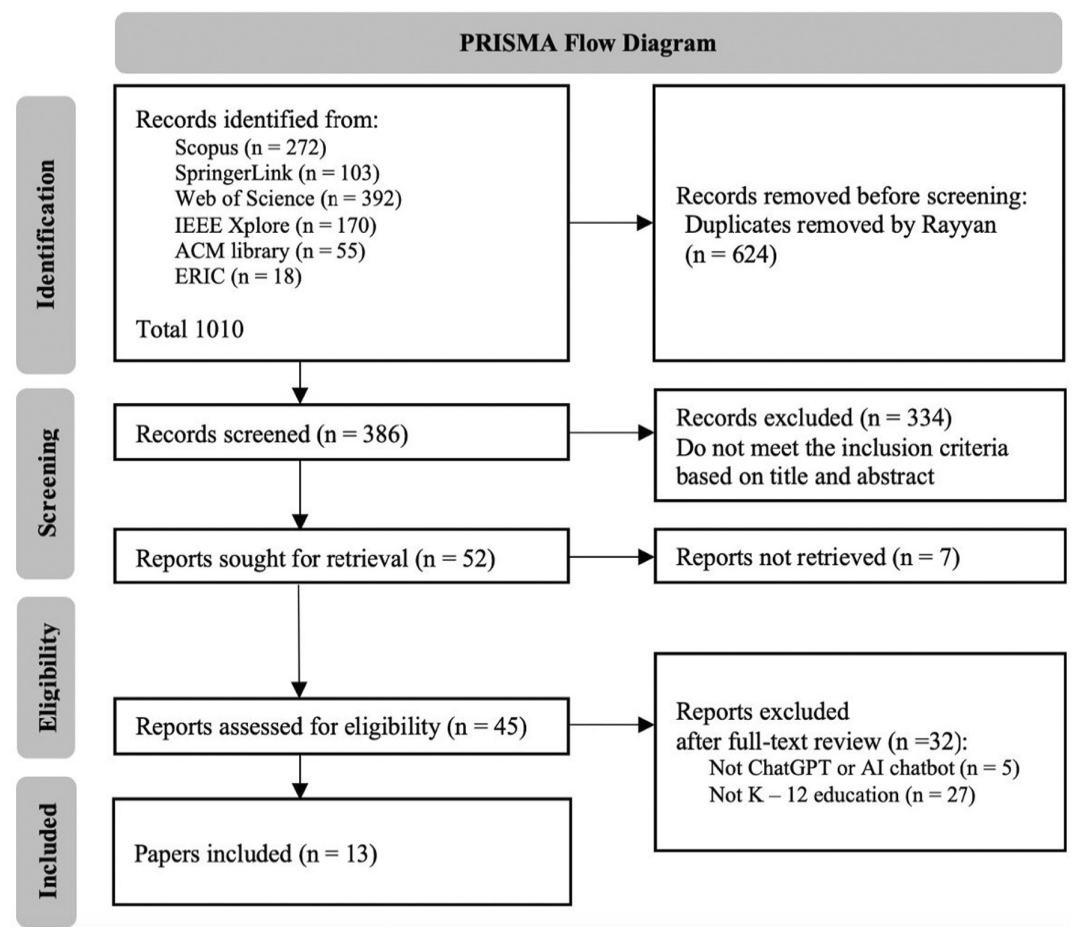| Inclusion criteria | Exclusion criteria |
|---|---|
| Papers published after ChatGPT's emergence (2022, 2023) | Publications before 2022 |
| | Studies not related to ChatGPT or AI chatbots |
| Studies focused on ChatGPT or related AI chatbots | Not in the K-12 education settings (i.e., higher education) |
| In the K-12 (primary and secondary) education contexts | |

**FIGURE 3** The PRISMA flow chart.

generate and visualise co-occurrence networks of significant phrases derived from the target literature collection (Van Eck & Waltman, 2014). Using keyword co-occurrence provided an overview of the selected studies, with their titles and abstracts analysed and visualised through VOSviewer (see Figure 2). NVivo, among various available CAQDAS tools, is widely used in educational research (Leech & Onwuegbuzie, 2011). The study followed the N7 + 1 processes proposed by O'Neill et al. (2018), which involved creating an NVivo project, importing and classifying the literature, and categorising and visualising the information to generate review themes.

## 4 | RESULTS

The findings of the systematic review are presented as follows:

### 4.1 | Bibliographic characteristics and underpinning themes

Before addressing the research questions, the current study meticulously gathered and organised the bibliographic information of the selected papers, which can be found in the Appendix A. The chosen publications

present a diverse geographical distribution, encompassing regions across Africa, Europe, the Americas, and Asia. Specifically, these regions include Italy, the United Arab Emirates (UAE), Croatia, Germany, Vietnam, the United States, Malaysia, Saudi Arabia, and Hong Kong, making the study's scope genuinely global. It is important to note that most papers ($n = 12$, 92.3%) are based on empirical studies, offering valuable insights into practical applications and real-world contexts. However, there is one exception, which takes the form of an editorial, and while it is not empirical, it still holds relevance to the research topic.

Concerning the educational levels represented in the selected papers, a significant portion ($n = 9$, 69.2%) focused on secondary education, shedding light on the implications of ChatGPT at this level. Additionally, a smaller subset of papers addresses primary education ($n = 3$, 23.1%), underscoring the relevance of this technology for younger learners. It is noted that one paper did not specify a specific K-12 education level but covered education in a broader sense. This paper was selected due to its significant discussions regarding using ChatGPT in K-12 education in certain sections. The scope of the selected papers is impressively diverse, encompassing various areas and contexts in which ChatGPT has been implemented. These focus areas include but are not limited to the application of ChatGPT in language learning, examination performance, assessment, curriculum and unit plan development, teaching assistants, and personalised learning.

When it comes to research methodology, it is evident that a substantial number of studies ($n = 10$) have been conducted using well-defined methods to examine the implementation and impact of ChatGPT. Among them, two studies employed mixed methods, providing a comprehensive understanding by combining qualitative and quantitative approaches. In contrast, one of the papers is an editorial; its focus lies in theoretical discussions and does not apply to the same research methodology analysis. Five of the remaining studies adopted qualitative research methodologies, and the other five adopted a quantitative research design.

After retrieving the target papers, a keyword co-occurrence analysis of the titles and abstracts of the chosen publications was conducted via VOSviewer (see Figure 4). The VOSviewer Keyword Co-occurrence analysis revealed two distinct clusters of keywords. The first cluster includes terms such as LLMs (Language Model), bias, concept, GPT, harmful stereotype, math, society, STEM field, tendency, and high school students. These interconnected keywords represent essential aspects related to the application of ChatGPT in K-12 education, such as people's bias and stereotypes about ChatGPT use in this context and the possibility of using ChatGPT in conceptual learning, math, society, and STEM fields. The second cluster encompasses keywords such as artificial intelligence, dataset, educator, English, impact, learner, school, student achievement, subject, and technology. These keywords are more closely linked to the impact of the implementation of ChatGPT on educators, learners, and schools. It is worth noting that LLMs and subjects are the central terms that bridge the two clusters, suggesting the characteristics of ChatGPT, a type of LLMs, which might affect teaching and learning across various subjects. The figure also shows GPT, school, technology, and LLMs as the most frequently occurring keywords,
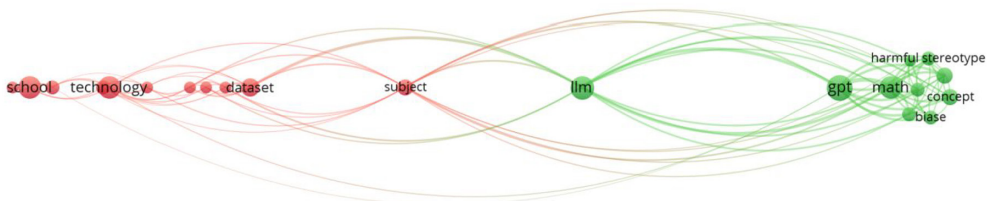


**FIGURE 4** VOSviewer keyword co-occurrence.

suggesting their relevance and prevalence in the selected papers. These findings provide valuable insights into the main themes and areas of focus regarding ChatGPT's implementation in K-12 education settings, providing a foundation for further exploration of the research questions.

## 4.2 | SWOT analysis of ChatGPT use in K-12 education

To address research question 1 (RQ1), this study employed Nvivo to analyse and synthesise the Strengths, Weaknesses, Opportunities, and Threats (SWOT) associated with implementing ChatGPT in K-12 education. The findings are further summarised in Table 2. Using Nvivo, a hierarchical chart of the SWOT factors was generated (see Figure 5). The chart highlights that the strengths (59 codes, 12 items) and opportunities (57 codes, 9 items) of incorporating ChatGPT in education received the highest number of references in the selected papers. Moreover, the weaknesses (29 codes, 8 items) and threats (25 codes, 11 items) related to ChatGPT's use in K-12 education were also addressed in the selected literature. The SWOT analysis sheds light on the different aspects of implementing ChatGPT in K-12 education.

In this study, the strengths, weaknesses, opportunities, and threats (SWOT) of ChatGPT's application in K-12 education were systematically analysed based on the coded data from multiple reviewed papers (see Table 2).

### 4.2.1 | Strengths

ChatGPT's strengths predominantly pivot around its potential for personalising learning experiences (13 references) and its positive influence on student achievement (9 references). Other significant strengths include the capacity for immediate and personalised feedback (8 references), an enhanced learning experience (4 references), and capability across various subjects (3 references). The system's efficiency, time-saving attributes, and ability to overcome language barriers were also repeatedly mentioned (3 references each), pointing to its usability in K-12 educational settings. Moreover, the technology's positive impact on productivity (3 references), motivation and engagement (2 references), user-friendliness (2 references), improved conceptual understanding (2 references), language skills development (2 references), critical thinking enhancement (2 references), task correctness (1 reference), task adequacy (1 reference), and high system usability (1 reference) further illustrate its benefits in K-12 education.

### 4.2.2 | Weaknesses

Despite its strengths, ChatGPT also has notable weaknesses. The most recurrently cited weaknesses revolve around output quality difficulties (7 references), low task specificity (4 references), and the inability to handle certain question types (4 references). Limited reasoning abilities (3 references), occasional negative impacts on productivity due to over-reliance (2 references), limited understanding of textual outputs (2 references), lack of contextual understanding (2 references), predictable responses (2 references), issues with recognising accents and dialects (1 reference), slow response times (1 reference), and lack of internet access for verification (1 reference) also emerged as ChatGPT's weaknesses in K-12 educational practices.

### 4.2.3 | Opportunities

ChatGPT use in K-12 education has lots of potential opportunities. The opportunities primarily lie in personalised learning applications (11 references), differentiated instruction (9 references), and support for educators in

**TABLE 2** The SWOT analysis.

**Strengths**
- Personalised learning experiences (13 references)
- Positive influence on student achievement (9 references)
- Immediate and personalised feedback (8 references)
- Enhanced learning experience (4 references)
- Capable in various subjects (3 references)
- Efficiency and time-saving (3 references)
- Overcoming language barriers (3 references)
- Positive impact on productivity (3 references)
- Motivation and engagement (2 references)
- User-friendly (2 references)
- Improved conceptual understanding (2 references)
- Language skills development (2 references)
- Enhancing critical thinking (2 references)
- Task correctness (1 reference)
- Task adequacy (1 reference)
- High system usability (1 reference)

**Weaknesses**
- Output quality difficulties (7 references)
- Low task specificity (4 references)
- Inability to handle certain question types (4 references)
- Limited reasoning abilities (3 references)
- Negative impact on productivity when users reply on it too much (2 references)
- Limited understanding of textual outputs (2 references)
- Lack of Contextual Understanding (2 references)
- Predictable responses (2 references)
- Inability to recognise accents and dialects (1 reference)
- Slow response time (1 reference)
- Lack of internet access for verification (1 reference)

**Opportunities**
- Application in personalised learning (11 references)
- Differentiated instruction and offering learning support to students (9 references)
- Supporting educators in generating assessment and learning tasks (7 references)
- Clear policies on the use of ChatGPT (6 references)
- Highlighting critical thinking and reflection when using ChatGPT (5 references)
- Revolutionising traditional teaching approaches (4 references)
- AI supported curriculum development (4 references)
- Call for more research to develop best practices (3 references)
- Training educators on how to use ChatGPT effectively (3 references)
- Support for English as a foreign language (EFL) students (2 references)
- Embracing AI in education (2 references)
- Democratisation of education (1 reference)

**Threats**
- Data privacy and ethical concerns (5 references)
- Potential for academic dishonesty and issue of plagiarism (4 references)
- Negative impact on productivity, procrastination and disengagement (3 references)
- Difficulties in prompt engineering (3 references)
- Inaccurate information and negative impact on learning quality (2 references)
- Lack of deep understanding and application (2 references)
- Diminished Interaction between learners and educators (1 reference)
- Too much reply leads to diminished sense of autonomy and competence (1 reference)
- Superficial learning and reduced social skills (1 reference)
- Potential bias and harmful stereotypes (1 reference)
- Lack of teacher awareness of AI's Potential (1 reference)
- Users' possible low interest in predictable technologies (1 reference)

*Note*: The numbers in parentheses represent the number of references each items is cited in, listed in the order of most referenced to least referenced.

**FIGURE 5**  Hierarchy chart of the SWOT references.

generating assessments and learning tasks (7 references). Some selected papers also addressed the clear policies on using ChatGPT (6 references) and highlighting critical thinking (5 references). Addressing these can open up opportunities and reach the potential of ChatGPT. Besides, the potential of ChatGPT in revolutionising traditional teaching approaches (4 references) and AI-supported curriculum development (4 references) also emerged as significant opportunities. Some reviewed literature calls for more research to develop best practices (3 references) and train educators on using ChatGPT effectively (3 references), further emphasising the potential for growth and development in this area.

### 4.2.4 | Threats

However, threats are also identified, which are to be noticed and carefully addressed. They are associated with data privacy and ethical concerns (5 references), potential for academic dishonesty (4 references), negative impact on productivity (3 references), difficulties in prompt engineering (2 references), inaccurate information (2 references), and a lack of deep understanding and application (2 references). The following aspects of threats in ChatGPT use were also addressed, with one reference per item: diminished interaction between learners and educators, the risk of promoting superficial learning, the potential for harmful stereotypes, lack of teacher awareness about AI's potential, and users' potential disinterest in predictable technologies. The findings underscore the need for careful consideration and management of these threats to ChatGPT implementation in K-12 contexts.

### 4.3 | Practices of ChatGPT implementation in K-12 education

In addressing the second research question (RQ2), we synthesised and categorised the selected literature on the application of ChatGPT in K-12 education into three distinct segments: educators, learners, and parents. This categorization was visualised in a hierarchy chart of codes representing various practices of ChatGPT implementation

among these groups, as produced by Nvivo software (see Figure 6). The most frequently discussed practices were those relating to educators, as demonstrated by the presence of 36 coding references and five distinct items coded. The implications for learners were also well-documented, with four academic papers discussing this topic, resulting in a total of 22 coding references. On the other hand, the role of parents in the use of ChatGPT
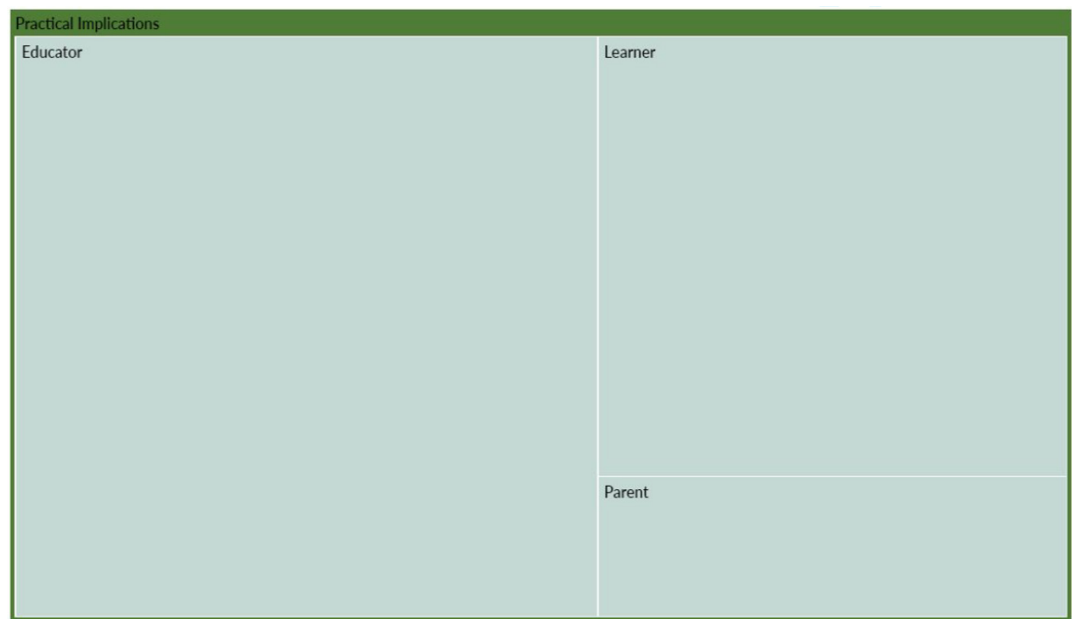


**FIGURE 6** Hierarchy chart of the implications of ChatGPT for educators, learners, and parents.
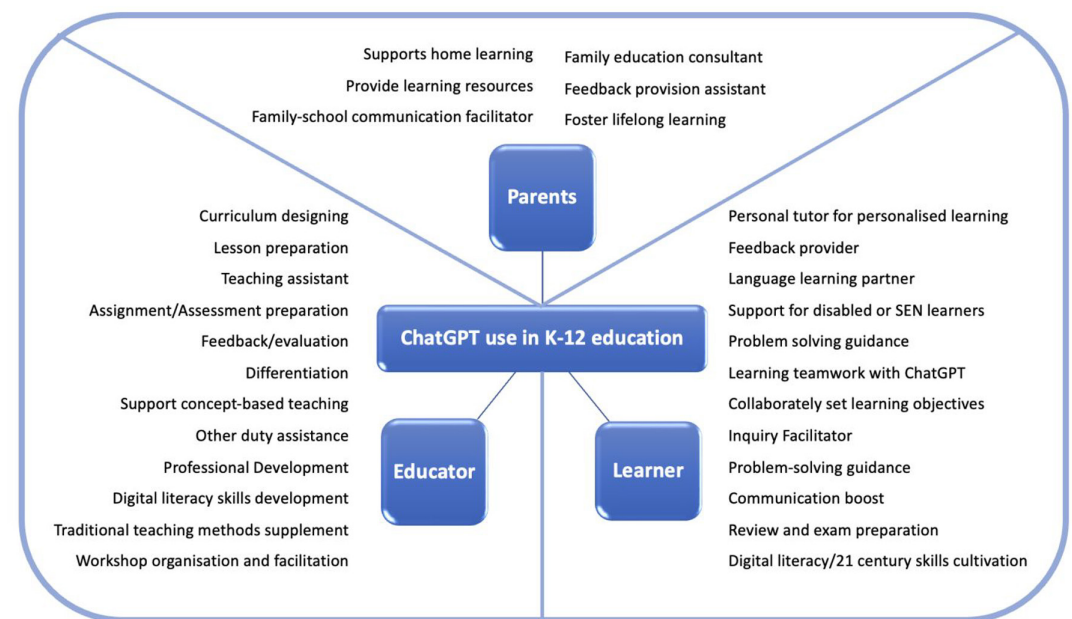


**FIGURE 7** Synthesised implementations of ChatGPT in K-12 education.

received more limited attention. Only two pieces of literature addressed this aspect, generating a total of 7 coding references.

The coded items on the implications of ChatGPT for educators, learners, and parents were then analysed, synthesised and summarised in Figure 7.

## 4.4 | Future directions and recommendations

The present study extensively examines the recommendations concerning ChatGPT practices within the realm of K-12 education, as evidenced by the comprehensive review of existing literature. To organise and present these findings systematically, the details are meticulously synthesised and classified into three categories, which are elucidated and listed in Table 3. These categories encompass: (1) teaching and learning practices at the educator and learner level, (2) regulations and ethics at the school or organisational level, and (3) prompt engineering. It is worth mentioning that prompt engineering could potentially be categorised under teaching and learning practices; however, due to its prominent significance, as emphasised by multiple reviewed publications, it has been delineated as a separate category for comprehensive analysis and recommendations.

## 5 | DISCUSSION

The following discussions were made based on the findings regarding SWOT analyses, practical implications for various users, and ChatGPT implementation recommendations in the K-12 education environments.

The findings highlight many noteworthy advantages of using ChatGPT in K-12 education. ChatGPT indicated good system usability, with favourable comments from students who used it to solve problems (Küchemann et al., 2023). Furthermore, ChatGPT's capacity in examination across various subjects, including literature, English, history, geography, civics education, mathematics, and physics, was clear from its human-level performance (Dao, Le, Phan, et al., 2023; Dao, Le, Vo, et al., 2023). ChatGPT's capacity to create personalised learning experiences, adapting to individual students' learning pace and style, was another key quality (Alneyadi & Wardat, 2023; Dao, Le, Phan, et al., 2023; Mogavi et al., 2023; Murgia et al., 2023). ChatGPT's rapid feedback and interactive learning experiences enriched the students' educational journey by boosting motivation, engagement, and better information retention (Mogavi et al., 2023; Murgia et al., 2023). Furthermore, the introduction of ChatGPT into education has shown the potential to reduce language barriers and enhance student accessibility (Lou, 2023;Othman, 2023; Woo et al., 2023). These findings demonstrate ChatGPT's revolutionary potential as an innovative and helpful tool to enhance and enrich the learning process in K-12 educational environments, as well as being relevant steps in co-design for learning.

Apart from its values, the use of ChatGPT also has some weaknesses in K-12 education settings. Some papers indicated that students reported having issues with output quality while utilising ChatGPT for problem posing, which might have resulted in mistakes and biases in generated answers (Abramski et al., 2023; Alneyadi & Wardat, 2023; Bekeš & Galzina, 2023). Furthermore, ChatGPT demonstrated low task specificity, implying insufficient information to solve problems efficiently (Küchemann et al., 2023). Notably, ChatGPT's reasoning abilities were limited, particularly in understanding complicated language and context, especially when dealing with languages other than English (Dao, Le, Phan, et al., 2023). As a result of this constraint, responses may be inaccurate or inappropriate for the circumstances at hand (Bekeš & Galzina, 2023; Mogavi et al., 2023). Furthermore, ChatGPT cannot handle specific question categories, such as complicated calculations and high-level application problems; it lacks higher-order thinking skills (Bitzenbauer, 2023; Dao, Le, Phan, et al., 2023; Dao, Le, Vo, et al., 2023). It raised concerns about their applicability to advanced mathematical tasks and synthesis knowledge in subjects like chemistry (Abramski et al., 2023). Additionally, shortcomings in recognising accents and dialects,

**TABLE 3** Synthesised recommendations for ChatGPT use in K-12 education.

| Categories | Recommendations |
| --- | --- |
| Teaching and learning practices (educator and learner level) | • Incorporate role-playing activities using ChatGPT-generated content for real-world applications (Mogavi et al., 2023).<br>• Employ ChatGPT as a complement to traditional teaching methods for personalised learning experiences, conduct automated assessment, predictive analytics, and learning analytics (Alneyadi & Wardat, 2023; Dao, Le, Phan, et al., 2023).<br>• Use ChatGPT to support learners by answering questions, providing resources, and feedback (Dao, Le, Phan, et al., 2023).<br>• Utilise ChatGPT to explore machine learning in education and formulate novel methods for AI-facilitated teaching and learning (Dao, Le, Phan, et al., 2023; Dao, Le, Vo, et al., 2023).<br>• Be aware of biases in AI outputs, thoroughly review and cross-check AI-generate content with credible sources to maintain quality and reliability, and teach students to verify information from credible sources (Abramski et al., 2023; Alneyadi & Wardat, 2023; Harrison et al., 2023; Mogavi et al., 2023).<br>• Enable ChatGPT to play the role as co-designer in learning (Murgia et al., 2023).<br>• Reimagine learning goals to focus on critical thinking through AI-generated text analysis and revision (Bitzenbauer, 2023; Mogavi et al., 2023)<br>• Align learning objectives with Bloom's Taxonomy and integrate ChatGPT to support knowledge acquisition, comprehension, problem-solving, critical analysis, synthesis, and evaluation (Mogavi et al., 2023).<br>• Use project-based assessments and peer/self-assessment techniques to promote collaboration and reflection (Mogavi et al., 2023).<br>• Foster reflective practices by prompting students to assess AI's impact on their learning (Mogavi et al., 2023).<br>• Use ChatGPT with the support of teachers who contextualise and personalise the didactic by intercepting differences in each classroom (Murgia et al., 2023).<br>• Teachers collaborate with ChatGPT to prepare lessons, including design curriculum, syllabuses, and teaching materials (Lou, 2023).<br>• Implement conversational AI, like ChatGPT, to bridge language skill gaps in EFL students (Othman, 2023).<br>• Support students writing through preparing outline, revising content, proofreading, and post-writing feedback (Woo et al., 2023).<br>• Use ChatGPT for conceptual surveys to assess students' understanding (Bitzenbauer, 2023).<br>• Modify assessments to accommodate AI limitations (Harrison et al., 2023) |
| Regulations and ethics (school/organisational level) | • Provide educators with training to use AI ethically and effectively in classrooms (Bekeš & Galzina, 2023; Othman, 2023).<br>• Establish clear rules and policies for ChatGPT use in educational settings (Mogavi et al., 2023).<br>• Address ethical implications and issues of plagiarism of AI integration in education (Harrison et al., 2023; Mogavi et al., 2023).<br>• Encourage students to cite ChatGPT and promote accountability in AI tool usage (Mogavi et al., 2023).<br>• Develop digital literacy skills for educators and learners to engage with AI tools like ChatGPT responsibly (Mogavi et al., 2023) |

**TABLE 3** (Continued)

| Categories | Recommendations |
| --- | --- |
| Prompt engineering | • Designing Effective Prompts and Incorporating Reflective Practices (Bekeš & Galzina, 2023; Mogavi et al., 2023).<br>• Provide specific training for teachers to design effective prompts for ChatGPT use (Küchemann et al., 2023)<br>• Customising the prompts used to generate ChatGPT responses relevant to the classroom situation (Murgia et al., 2023).<br>• Explicitly teach students effective prompting to optimise ChatGPT outputs (Woo et al., 2023).<br>• Collaborate with learning peers to engineer prompts (Woo et al., 2023). |

poor reaction times, and preplanned and fairly predictable replies contributed to the discovered weaknesses of ChatGPT (Alneyadi & Wardat, 2023).

The opportunities for implementing ChatGPT in K-12 education were identified in various aspects. A few publications stated that ChatGPT can change personalised learning experiences for students by outlining complicated ideas and providing fast, personalised feedback (Alneyadi & Wardat, 2023; Dao, Le, Phan, et al., 2023; Murgia et al., 2023). It also supports differentiation by demonstrating competency in a variety of subjects, allowing it to deliver credible information, answer questions, and provide feedback targeted to specific students (Mogavi et al., 2023; Murgia et al., 2023). Furthermore, ChatGPT is useful in developing assessment and learning activities (Küchemann et al., 2023), compensating for conceptual issues in designing acceptable assessments for courses such as physics (Bitzenbauer, 2023; Küchemann et al., 2023). Its incorporation into schools stimulates critical thinking among students, promoting reflection on the benefits, limitations, and consequences associated with AI use (Bitzenbauer, 2023; Mogavi et al., 2023). Moreover, the use of ChatGPT has the potential to transform traditional teaching methodologies, resulting in interactive and personalised learning experiences, particularly in STEM courses (Abramski et al., 2023; Alneyadi & Wardat, 2023). This technology also helps with curriculum creation, creating engaging learning materials, and improving language abilities for English as a foreign language (EFL) student (Othman, 2023; Woo et al., 2023). However, it underlines the significance of explicit AI regulations and the need to teach educators about successful ChatGPT use. Also, some papers emphasise the need to incorporate Chat-like AI tools into education better to prepare learners and educators for current and future issues and democratise access to equal learning opportunities globally (Harrison et al., 2023).

The study additionally points out some concerns regarding the use of ChatGPT in K-12 education. Students may become excessively dependent on the AI tool, resulting in a diminished emphasis on critical thinking and the investigation of alternative perspectives (Bitzenbauer, 2023; Mogavi et al., 2023). Moreover, the implementation of ChatGPT raises data privacy and ethical issues, such as academic integrity, misinformation, deception, and privacy concerns (Bekeš & Galzina, 2023; Mogavi et al., 2023; Murgia et al., 2023; Othman, 2023). On top of that, there are indications that ChatGPT may promote academic dishonesty and plagiarism among students (Harrison et al., 2023). In addition, excessive use of ChatGPT can harm productivity, exacerbate procrastination, and result in disengagement from tasks, thereby diminishing the overall learning experience (Bitzenbauer, 2023; Mogavi et al., 2023). Also, learners who rely extensively on AI may experience a lower level of autonomy and competence in their learning, resulting in decreased learning motivation (Bitzenbauer, 2023). AI-generated content may not always be accurate, resulting in the adoption of misleading information that interferes with the quality of learning (Harrison et al., 2023). Besides that, there are concerns about a lack of comprehension and application of knowledge due to excessive dependence on ChatGPT and a potential decrease in interaction between students and teachers (Othman, 2023). Moreover, the biases present in ChatGPT outputs can perpetuate detrimental stereotypes, posing ethical difficulties (Abramski et al., 2023). Potential hazards also include technical difficulties in prompt engineering and a lack of educators awareness of AI's potential (Bekeš & Galzina, 2023).

Furthermore, students may lack interest in predictable technologies, raising concerns regarding user acceptability (Othman, 2023).

The review and analysis of the current body of scholarly works underscore the capacity of ChatGPT to bring about significant changes in the field of K-12 education. Nevertheless, it is important to acknowledge and rectify the limitations of the technology to harness its capabilities thoroughly. Possible avenues for deeper integration of ChatGPT inside the K-12 educational system might provide substantial advantages for learners and educators. However, potential risks, such as threats to data privacy, ethical problems, and the possibility of academic dishonesty, highlight the need to develop comprehensive and robust rules and guidelines to address and mitigate these challenges effectively. The potential of ChatGPT in the field of K-12 education is promising. However, it is imperative to carefully assess and consider its limitations and possible risks to ensure responsible and efficient use within K-12 educational environments.

This study also provides insights into the practical implications of ChatGPT in K-12 education for educators, learners, and parents. The findings demonstrate that most of the mentioned practices are related to educators, indicating that ChatGPT has considerable potential for facilitating and enhancing the duties and responsibilities of educators. The implications for learners are also significant, highlighting the potential positive influence of ChatGPT on their educational experience. Nevertheless, it is crucial to acknowledge that there is a comparatively limited amount of detailed analysis about the possible applications for parents. This highlights the need for more investigation in this particular domain.

For educators, ChatGPT provides a variety of opportunities. It can assist in developing innovative and engaging curricula aligned with the curriculum framework and learning needs of learners (Lou, 2023; Mogavi et al., 2023). In addition, ChatGPT can be used to prepare lessons by providing valuable insights, resources, and teaching strategies, thereby making the process of lesson planning more efficient and effective (Bekeš & Galzina, 2023; Küchemann et al., 2023; Lou, 2023). ChatGPT can assist teachers in many ways as instructional assistants, such as by addressing common inquiries, facilitating discussions, and providing students with immediate feedback (Alneyadi & Wardat, 2023; Dao, Le, Phan, et al., 2023; Murgia et al., 2023; Woo et al., 2023). Another significant implication for educators is the preparation of assignments and tests (Alneyadi & Wardat, 2023). ChatGPT can generate diverse and challenging assignments, fostering critical thinking and problem-solving skills (Bitzenbauer, 2023; Mogavi et al., 2023). Furthermore, ChatGPT can aid in evaluating pupils' work, providing constructive feedback, and pointing out areas for development (Dao, Le, Vo, et al., 2023). Support for differentiation in the classroom is one of the most noteworthy benefits for educators (Mogavi et al., 2023). By analysing students' progress and comprehension patterns, ChatGPT can recommend individualised approaches to teaching and learning. In addition, ChatGPT may support educators in embracing concept-based teaching strategies, thereby nurturing a deeper understanding and more long-lasting retention of knowledge (Lou, 2023). Also, ChatGPT has implications for the professional development of educators (Mogavi et al., 2023); it can be a valuable resource for enhancing digital literacy skills, enabling educators to utilise technology effectively in their teaching practices. ChatGPT can complement conventional teaching methods by incorporating innovative elements into the classroom setting (Alneyadi & Wardat, 2023).

For learners, ChatGPT functions as a personal tutor, providing personalised guidance and support (Alneyadi & Wardat, 2023; Harrison et al., 2023; Mogavi et al., 2023; Murgia et al., 2023). Through its interactive features, the tool can provide instant feedback, boost conceptual understanding, and function as a language-learning companion to promote conversational proficiency (Othman, 2023; Woo et al., 2023). Support for students with disabilities or special educational needs (SEN) is one of the most significant aspects. ChatGPT can provide accommodations and adaptive resources to create an inclusive learning environment, enabling these learners to overcome obstacles and actively engage with the learning process (Mogavi et al., 2023). Also, ChatGPT can facilitate collaborative learning experiences, enhancing students' teamwork and communication abilities (Mogavi et al., 2023). By co-creating learning objectives with ChatGPT, students can develop a sense of learning ownership and enhance their critical thinking skills (Alneyadi & Wardat, 2023; Murgia et al., 2023). Besides, ChatGPT can serve as a facilitator of

inquiry, guiding students through the process of investigating complex topics and formulating research queries; it can facilitate problem-solving by encouraging learners to think critically and generate creative solutions (Alneyadi & Wardat, 2023; Mogavi et al., 2023). ChatGPT's implications for learners extend to developing digital literacy and 21st-century skills (Alneyadi & Wardat, 2023). As technology continues to play a central role in society, competence with digital tools is required by learners to flourish in contemporary society.

On the other hand, the findings imply that the potential implications of ChatGPT for parents require further exploration. Nonetheless, several valuable possibilities were discovered. For example, ChatGPT can support home learning by providing supplemental resources and responding to parents' questions about their child's education (Lou, 2023). Moreover, ChatGPT may promote effective communication between families and schools, ensuring parents remain informed of their child's progress and actively engage in their educational journey (Mogavi et al., 2023). Apart from these, we argue that ChatGPT can also serve as a family education consultant, providing insight into various educational approaches and strategies parents can employ to support their child's learning at home. In addition, ChatGPT has the potential to serve as a feedback provision assistant, assisting parents in interpreting assessment results and providing suggestions for improvement. Ultimately, ChatGPT can cultivate a culture of lifelong learning within families by encouraging parents to participate in educational activities with their children (Mogavi et al., 2023).

Additionally, the study delves into regulations and ethics at the school or organisational level, addressing essential considerations for the responsible and ethical use of ChatGPT in K-12 education practice (Bekeš & Galzina, 2023; Harrison et al., 2023; Mogavi et al., 2023; Othman, 2023). Recently, the International Baccalaureate (IB) set an example as one of the educational organisations that published their policy online (IBO.ORG, 2023), delineating how artificial intelligence tools, like ChatGPT, can be implemented ethically and effectively in IB schools.

## 6 | CONCLUSION

In conclusion, this systematic review has provided valuable insights about the application of ChatGPT in K-12 education, shedding light on its strengths, weaknesses, opportunities, and potential risks through a SWOT analysis. The results indicate that ChatGPT has the potential to revolutionise K-12 education through the provision of personalised learning opportunities, enhanced learner motivation and involvement, and the reduction of language hurdles. Furthermore, ChatGPT demonstrates a level of performance comparable to that of humans in a wide range of domains, establishing itself as a flexible tool for educational purposes. Nevertheless, the study also recognises some constraints of ChatGPT, including concerns over output quality, task specificity, and limited reasoning abilities. Effective implementation of ChatGPT's capabilities necessitates carefully considering and resolving these challenges. The research underscores the significance of implementing comprehensive guidelines and rules to guarantee appropriate and ethical usage within K-12 educational environments. Although practical implications for educators and learners have been extensively explored, further study is needed to look into the potential uses of ChatGPT for parents or other stakeholders. The research offers recommendations for pedagogical approaches and ethical considerations, that are crucial in optimising the effectiveness of ChatGPT. The limitation of the research is that the limited scope and application of the study might impede the broad generalizability of the findings. The study focuses solely on recently developed AI tools, specifically their application within the K-12 education phase, which narrows the range of the paper selection for the systematic review.

To sum up, this study emphasises the promising potential of ChatGPT as an innovative tool for enriching and enhancing the learning process in K-12 educational settings. It also emphasises the significance of thoughtful integration and ongoing research to maximise its advantages and address possible risks. According to the data collected about strengths, possibilities, along with current practices and recommendations, we suggest that AI should be further explored under the co-design for learning approach, where technology-based collaboration for learning among different agents (learners, teachers, and peers) is the main focus (Bovill et al., 2011; Gros &

Durall, 2020; Santana Martel & Perez-i-Garcias, 2020). As the development of this technology progresses, it is imperative for educators, policymakers, and stakeholders to get involved in joint efforts to maximise its advantages and tackle any possible obstacles, therefore fostering a constructive and revolutionary influence on K-12 education, where AI can play the role of another key agent to support students' learning processes, which can be enriched under the learning co-design framework.

## AUTHOR CONTRIBUTIONS

## FUNDING INFORMATION

## CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

## DATA AVAILABILITY STATEMENT

The raw data extracted from the existing publication for this review are available from the corresponding author upon request. Please note that the data underpinning this systematic review is subject to the original data creators' copyright and terms of use. Therefore, some restrictions may apply to the sharing of raw data. For further inquiries regarding data availability, please contact the corresponding author.

## ORCID

*Peng Zhang* 🆔 https://orcid.org/0000-0001-5164-1869
*Gemma Tur* 🆔 https://orcid.org/0000-0003-4508-6808

## REFERENCES

Abramski, K., Citraro, S., Lombardi, L., Rossetti, G., & Stella, M. (2023). Cognitive network science reveals bias in GPT-3, GPT-3.5 turbo, and GPT-4 mirroring math anxiety in high-school students. *Big Data and Cognitive Computing*, *7*(3), 124. https://doi.org/10.3390/bdcc7030124

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

Bekeš, E. R., & Galzina, V. (2023). Exploring the pedagogical use of AI-powered chatbots educational perceptions and practices. In *2023 46th MIPRO ICT and Electronics Convention (MIPRO)* (pp. 636–641). IEEE. https://doi.org/10.23919/MIPRO57284.2023.10159734

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

Bovill, C., Cook-Sather, A., & Felten, P. (2011). Students as co-creators of teaching approaches, course design, and curricula: Implications for academic developers. *International Journal for Academic Development*, *16*(2), 133–145. https://doi.org/10.1080/1360144X.2011.568690

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). *A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT.* arXiv. https://doi.org/10.48550/ARXIV.2303.04226

Chen, C., Fu, J., & Lyu, L. (2023). *A pathway towards responsible AI generated content.* https://doi.org/10.48550/ARXIV.2303.01325

Cheng, H.-W. (2023). Challenges and limitations of ChatGPT and artificial intelligence for scientific research: A perspective from organic materials. *AI*, *4*(2), 401–405. https://doi.org/10.3390/ai4020021

Cotton, D., Cotton, P., & Shipway, J. R. (2023). *Chatting and cheating. Ensuring academic integrity in the era of ChatGPT.* [Preprint]. EdArXiv. https://doi.org/10.35542/osf.io/mrz8h

Dao, X.-Q., Le, N.-B., Phan, X.-D., & Ngo, B.-B. (2023). Can ChatGPT pass the Vietnamese National High School graduation examination? https://doi.org/10.48550/ARXIV.2306.09170

Dao, X.-Q., Le, N.-B., Vo, T.-D., Phan, X.-D., Ngo, B.-B., Nguyen, V.-T., Nguyen, T.-M.-T., & Nguyen, H.-P. (2023). VNHSGE: VietNamese High School graduation examination dataset for large language models. https://doi.org/10.48550/ARXIV.2305.12199

Firat, M. (2023). *How Chat GPT can transform autodidactic experiences and open education?* [Preprint]. Open Science Framework. https://doi.org/10.31219/osf.io/9ge8m

Gros, B., & Durall, E. (2020). Retos y oportunidades del diseño participativo en tecnología educativa. *Edutec. Revista Electrónica De Tecnología Educativa*, *74*, 12–24. https://doi.org/10.21556/edutec.2020.74.1761

Harrison, L. M., Hurd, E., & Brinegar, K. M. (2023). Critical race theory, books, and ChatGPT: Moving from a ban culture in education to a culture of restoration. *Middle School Journal*, *54*(3), 2–4. https://doi.org/10.1080/00940771.2023.2189862

Huang, J., Saleh, S., & Liu, Y. (2021). A review on artificial intelligence in education. *Academic Journal of Interdisciplinary Studies*, *10*(3), 206. https://doi.org/10.36941/ajis-2021-0077

Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: A descriptive study. *Journal of Educational Evaluation for Health Professions*, *20*, 1. https://doi.org/10.3352/jeehp.2023.20.1

Iskender, A. (2023). Holy or unholy? Interview with open AI's ChatGPT. *European Journal of Tourism Research*, *34*, 3414. https://doi.org/10.54055/ejtr.v34i.3169

Karakose, T., Demirkol, M., Yirci, R., Polat, H., Ozdemir, T. Y., & Tülübaş, T. (2023). A conversation with ChatGPT about digital leadership and technology integration: Comparative analysis based on human–AI collaboration. *Administrative Sciences*, *13*(7), 157. https://doi.org/10.3390/admsci13070157

Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Physics task development of prospective physics teachers using ChatGPT. https://doi.org/10.48550/ARXIV.2304.10014

Leech, N. L., & Onwuegbuzie, A. J. (2011). Beyond constant comparison qualitative data analysis: Using NVivo. *School Psychology Quarterly*, *26*(1), 70–84. https://doi.org/10.1037/a0022711

Lin, Y., Du, H., Niyato, D., Nie, J., Zhang, J., Cheng, Y., & Yang, Z. (2023). Blockchain-aided secure semantic communication for AI-generated content in metaverse. *IEEE Open Journal of the Computer Society*, *4*, 72–83. https://doi.org/10.1109/OJCS.2023.3260732

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

Lou, Y. (2023). Exploring the application of ChatGPT to English teaching in a Malaysia primary school. *Journal of Advanced Research in Education*, *2*(4), 47–54. https://doi.org/10.56397/JARE.2023.07.08

Lozano, A., & Blanco Fontao, C. (2023). Is the education system prepared for the irruption of artificial intelligence? A study on the perceptions of students of primary education degree from a dual perspective: Current pupils and future teachers. *Education Sciences*, *13*(7), 733. https://doi.org/10.3390/educsci13070733

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, *50*(6), 2824–2838. https://doi.org/10.1111/bjet.12861

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, *40*(3), 26–29. https://doi.org/10.1108/LHTN-01-2023-0009

Milton, C. L. (2023). ChatGPT and forms of deception. *Nursing Science Quarterly*, *36*(3), 232–233. https://doi.org/10.1177/08943184231169753

Minkkinen, M., & Mäntymäki, M. (2023). Discerning between the "easy" and "hard" problems of AI governance. *IEEE Transactions on Technology and Society*, *4*(2), 188–194. https://doi.org/10.1109/TTS.2023.3267382

Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L. E., & Hui, P. (2023). *Exploring user perspectives on ChatGPT: Applications, perceptions, and implications for AI-integrated education.* https://doi.org/10.48550/ARXIV.2305.13114

Mohammad, B., Supti, T., Alzubaidi, M., Shah, H., Alam, T., Shah, Z., & Househ, M. (2023). The pros and cons of using ChatGPT in medical education: A scoping review. In J. Mantas, P. Gallos, E. Zoulias, A. Hasman, M. S. Househ, M. Charalampidou, & A. Magdalinou (Eds.), *Studies in health technology and informatics*. IOS Press. https://doi.org/10.3233/SHTI230580

Moreno-Küstner, B., Martín, C., & Pastor, L. (2018). Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *PLoS ONE*, *13*(4), e0195687. https://doi.org/10.1371/journal.pone.0195687

Murgia, E., Pera, M. S., Landoni, M., & Huibers, T. (2023). Children on ChatGPT readability in an educational context: Myth or opportunity? *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, 311–316. ACM. https://doi.org/10.1145/3563359.3596996

O'Neill, M., Booth, S., & Lamb, J. (2018). Using NVivo™ for literature reviews: The eight step pedagogy (N7+1). *The Qualitative Report*, *23*(13), 21–39. https://doi.org/10.46743/2160-3715/2018.3030

Othman, K. (2023). Towards implementing AI Mobile application chatbots for EFL learners at primary schools in Saudi Arabia. *Journal of Namibian Studies: History Politics Culture*, *33*. https://doi.org/10.59670/jns.v33i.434

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

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. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Perera, P., & Lankathilaka, M. (2023). AI in higher education: A literature review of ChatGPT and guidelines for responsible implementation. *International Journal of Research and Innovation in Social Science*, *VII*(VI), 306–314. https://doi.org/10.47772/IJRISS.2023.7623

Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education* [Preprint]. https://doi.org/10.36227/techrxiv.21789434

Santana Martel, J. S., & Perez-i-Garcias, A. (2020). Co-diseño educativo haciendo uso de las TIC en educación superior una revisión sistemática de literatura. *Edutec. Revista Electrónica De Tecnología Educativa*, *74*, 25–50. https://doi.org/10.21556/edutec.2020.74.1799

Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health*, *47*(1), 46. https://doi.org/10.1186/s41182-019-0165-6

Temsah, O., Khan, S. A., Chaiah, Y., Senjab, A., Alhasan, K., Jamal, A., Aljamaan, F., Malki, K. H., Halwani, R., Al-Tawfiq, J. A., Temsah, M.-H., & Al-Eyadhy, A. (2023). Overview of early ChatGPT's presence in medical literature: Insights from a hybrid literature review by ChatGPT and human experts. *Cureus*, *15*, e37281. https://doi.org/10.7759/cureus.37281

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*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly impact* (pp. 285–320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13

Woo, D. J., Guo, K., & Susanto, H. (2023). *Cases of EFL secondary students' prompt engineering pathways to complete a writing task with ChatGPT*. https://doi.org/10.13140/RG.2.2.31464.85762

Zhang, Y., He, W., & Li, Y. (2020). Efficient Boolean keywords search over encrypted cloud data in public key setting. *Mobile Information Systems*, *2020*, 1–15. https://doi.org/10.1155/2020/2904861

**APPENDIX A**

**BIBLIOGRAPHIC INFORMATION OF THE SELECTED PUBLICATIONS**

| References | Title | Region | Paper type | Phase | Focus | Methodology | Publisher |
|---|---|---|---|---|---|---|---|
| Abramski et al. (2023) | Cognitive network science reveals bias in GPT-3, ChatGPT, and GPT-4 mirroring math anxiety in high-school students | Italy | Empirical Study | Secondary | Bias of large language models (LLMs) towards anxiety | Quantitative | Big Data and Cognitive Computing |
| Alneyadi and Wardat (2023) | ChatGPT: Revolutionising student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools | UAE | Empirical Study | Secondary | ChatGPT's influence on the academic performance and learning perception of 11th graders | Mix-method | Contemporary Educational Technology |
| Bekeš and Galzina (2023) | Exploring the pedagogical use of AI-powered chatbots educational perceptions and practices | Croatia | Empirical Study | Secondary | Educational Copilot – AI Unit Planner | Qualitative | 2023 46th MIPRO ICT and Electronics Convention (MIPRO), IEEE |
| Bitzenbauer (2023) | ChatGPT in physics education: A pilot study on easy-to-implement activities | Germany | Empirical Study | Secondary | ChatGPT use in physics classroom | Quantitative | Contemporary Educational Technology |
| Dao, Le, Phan, et al. (2023) | VNHSGE: VietNamese High School Graduation Examination dataset for large language models | Vietnam | Empirical Study | Secondary | LLMs' performance on VNHSGE exam | Quantitative | ArXiv |
| Dao, Le, Vo, et al. (2023) | Can ChatGPT pass the Vietnamese National High School Graduation Examination? | Vietnam | Empirical Study | Secondary | ChatGPT's performance on VNHSGE exam | Quantitative | ArXiv |

**APPENDIX A** (Continued)

| References | Title | Region | Paper type | Phase | Focus | Methodology | Publisher |
|---|---|---|---|---|---|---|---|
| Harrison et al. (2023) | Critical race theory, books, and ChatGPT: Moving from a ban culture in education to a culture of restoration | United States | Editorial | Secondary | Restorative approaches to curriculum | NA | Middle School Journal |
| Küchemann et al. (2023) | Physics task development of prospective physics teachers using ChatGPT | Germany | Empirical Study | Secondary | Physics assessment task creation for high school students | Mix-method | ArXiv |
| Lou (2023) | Exploring the application of ChatGPT to English teaching in a Malaysia primary school | Malaysia | Empirical Study | Primary | Use of ChatGPT in English teaching | Qualitative | Journal of Advanced Research in Education |
| Murgia et al. (2023) | Children on ChatGPT readability in an educational context: myth or opportunity? | Italy | Empirical Study | Primary | ChatGPT-supported personalised learning among children | Qualitative | Adjunct Proceedings of the 31st ACM Conference on User Modelling, Adaptation and Personalisation |
| Mogavi et al. (2023) | Exploring user perspectives on ChatGPT: applications, perceptions, and implications for AI-integrated education | NA | Empirical Study | NA | Early adopters' experiences and thoughts about integrating ChatGPT in education | Qualitative | ArXiv |
| Othman (2023) | Towards implementing AI mobile application chatbots for EFL learners at primary schools in Saudi Arabia | Saudi Arabia | Empirical Study | Primary | Using AI Chatbot in EFL contexts | Quantitative | Journal of Namibian Studies |
| Woo et al. (2023) | Cases of EFL secondary students' prompt engineering pathways to complete a writing task with ChatGPT | Hong Kong | Empirical Study | Secondary | EFL students' ChatGPT prompts in writing tasks | Qualitative, case study | ArXiv |