



ChatGPT in education: Methods, potentials, and limitations

Bahar Memarian^{*}, Tenzin Doleck

Simon Fraser University, Faculty of Education, Vancouver, British Columbia, Canada

ARTICLE INFO

Keywords:

ChatGPT
Large language models
Education
Artificial intelligence
Machine learning
Data science
Pedagogy

ABSTRACT

ChatGPT has been under the scrutiny of public opinion including in education. Yet, less work has been done to analyze studies conducted on ChatGPT in educational contexts. This review paper examines where ChatGPT is employed in educational literature and areas of potential, challenges, and future work. A total of 63 publications were included in this review using the general framework of open and axial coding. We coded and summarized the methods, and reported potentials, limitations, and future work of each study. Thematic analysis of reviewed studies revealed that most extant studies in the education literature explore ChatGPT through a commentary and non-empirical lens. The potentials of ChatGPT include but are not limited to the development of personalized and complex learning, specific teaching and learning activities, assessments, asynchronous communication, feedback, accuracy in research, personas, and task delegation and cognitive offload. Several areas of challenge that ChatGPT is or will be facing in education are also shared. Examples include but are not limited to plagiarism deception, misuse or lack of learning, accountability, and privacy. There are both concerns and optimism about the use of ChatGPT in education, yet the most pressing need is to ensure student learning and academic integrity are not sacrificed. Our review provides a summary of studies conducted on ChatGPT in education literature. We further provide a comprehensive and unique discussion on future considerations for ChatGPT in education.

1. Introduction

Conditional Generative Pre-trained Transformer or ChatGPT (Open AI, 2023), a publicly accessible AI model, has witnessed tremendous uptake since its launch in November 2022. Due to its ability to synthesize large volumes of text and generate content, it has become a popular choice for teaching and learning (Baidoo-Anu & Owusu Ansah, 2023). Ever since, ChatGPT has been tested for a diverse array of tasks, with the most prominent being writing. Other and perhaps more innovative areas are writing essays and predicting historical events (McGee, 2023), talking, drawing, and editing with visual models (Wu et al., 2023), and examining the facts and myths in ChatGPT responses (Alkaiissi & McFarlane, 2023).

Most published articles apropos ChatGPT tend to qualitatively explore the benefits and warn of the potential harms of ChatGPT. Yet, we find the contexts in which ChatGPT is studied need to be reviewed to understand its true performance and characteristics. Present reviews explore the characteristics of ChatGPT models (Roumeliotis & Tselikas, 2023), the impact of ChatGPT on education (Lo, 2023), and its applications (Ray, 2023). Despite the popularity and attention, less work has been done in the literature to examine how ChatGPT is studied in

educational contexts.

A summary of literature reviews on ChatGPT and education is presented in Table 1. The few existing reviews on ChatGPT reflect both the limitations and potentials ChatGPT may offer in different domains of education.

For some, COVID-19 together with the emergence of ChatGPT has created continuous issues with assessment practices. These issues have necessitated the reuse of alternative methods of evaluation such as oral examinations (Gardner & Giordano, 2023). Yet, alternative methods such as oral examinations may be more demanding in terms of execution time and analysis.

For some others, the use of chatbots has presented a potential for re-imagining teaching and learning (Gentile et al., 2023; Wu & Yu, 2023). For example, in the meta-analysis of 24 studies, Wu and Yu (2023) found AI chatbots to play an important role in student learning outcomes, especially when implemented as a short, rather than long-term intervention. Their analysis, however, may have had many questions and over-reliance on statistical outcomes from a rather small sample size.

An area often explored is the descriptive envisioning of where ChatGPT may come to use in teaching, research, and professional activities (Emenike & Emenike, 2023; Lo, 2023). Review work in

^{*} Corresponding author.

E-mail address: bmemaria@sfu.ca (B. Memarian).

<https://doi.org/10.1016/j.chbah.2023.100022>

Received 15 August 2023; Received in revised form 9 October 2023; Accepted 13 October 2023

Available online 14 October 2023

2949-8821/Crown Copyright © 2023 Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1
Summary of literature reviews on ChatGPT and education.

Reference	Focus	Limitations
Currie (2023)	Challenges and potentials of ChatGPT in nuclear medicine	Healthcare focused
Eggmann et al. (2023)	Implications of ChatGPT in dental medicine	Healthcare focused
Emenike and Emenike (2023)	Potential uses of ChatGPT for teaching, research, and professional activities	A method of development is not provided
Eysenbach (2023)	Impact of ChatGPT on medical education	Healthcare focused
Gardner and Giordano (2023)	Oral examinations	Costly deployment
Gentile et al. (2023)	Critical dimensions related to the teacher figure	Dimensions may need to be contextualized
Miao and Ahn (2023)	Impact of ChatGPT on nursing education	Healthcare focused
Lo (2023)	Impact of ChatGPT on education	Majority of papers on preprints
Ray (2023)	Impact of ChatGPT on nursing education	Healthcare focused
Sallam (2023)	ChatGPT utility in health education	Broad and often distant topics in healthcare, the majority of papers on preprints
Thurzo et al. (2023)	Use of AI in dental education	AI feeling anatomy versus AI memorizing anatomy
Wu and Yu (2023)	Meta-analysis of the effects of AI chatbots on students' learning outcomes.	Codification of research questions and reliance on statistical outcomes for 24 studies

healthcare by some scholars (e.g., Currie, 2023; Eggmann et al., 2023; Eysenbach, 2023; Sallam, 2023) examines uses of ChatGPT in a broad array of healthcare topics, many of which may have non-generalizable or non-transferable findings. Similarly, Ray (2023) and Miao and Ahn (2023) examine the impact of AI in nursing education. Thurzo et al. (2023) examine the application of AI in dental education, yet they find AI to fall short of understanding human anatomy. It appears that AI can memorize all there is to knowledge but not feel it the same way a human dentist, doctor, or engineer may experience in their work field.

Examination of the reviewed studies shows that more work needs to be done to summarize the state of the art with ChatGPT in education, specifically for non-medical fields that may not have constrained outcomes (i.e., understanding the physiology of the human body and maintaining ideal bodily conditions as much as possible).

We conduct a literature review and aim to cover studies of ChatGPT and explore their approaches. In doing so we explore the following research questions:

1. What methods are used when studying ChatGPT in education literature?
2. What are the potentials for ChatGPT in education literature?
3. What are the limitations and challenges for ChatGPT in education literature?

This research contributes to our understanding of ChatGPT and builds its capacity for research in educational contexts. In the methods section that follows, we describe our search process and codification scheme which mainly comprises open and axial coding as seen in grounded theory research. We purposefully used grounded theory as our methodical approach to enable analysis and discussion of emergent research on ChatGPT. The results section provides a summary of findings surrounding our three research questions noted above. In the discussion, we further provide potential challenges and future potentials of ChatGPT in education based on our synthesis of the reviewed studies. The key contribution of this research is in offering an in-depth synthesis of literature and providing a discussion on future considerations for ChatGPT in education.

2. Methods

We examine published studies to create an overview of the literature. A search process is followed and coding from grounded theory is used to chart the review papers. The key components of coding in grounded theory are open coding, and axial coding, followed by selective coding. During the open coding, we label and summarize the key studies conducted, potentials, and challenges reported by the literature. In the axial coding, we attempted to compare the summarized items to find areas of similarity and grouping into more underlying themes. We did not employ selective coding, as we wanted to keep and present all themes emerging, even if their count was 1. We used Web of Science (WoS) and Scopus to search for papers. Our search included publications written in English and for an unlimited time frame. In both databases, the following string was searched:

ChatGPT AND (education OR teaching OR learning OR pedagogy)

Inclusion criteria contain articles focusing on the study or role of ChatGPT in education that were published in English. Exclusion criteria contain articles focusing on the study or role of ChatGPT in medical education and industry and any publication not in English.

We extracted and charted data surrounding methods, uses, outcomes, population, limitations, and future work of each study. We used open coding to classify the data extracted into discrete parts. Our thematic analysis includes following the common method for qualitative analysis (Creswell, 1994) and open coding the reported challenges in each reviewed study followed by connecting similar themes and conducting a count of the themes. To answer our research questions. We:

1. Categorized and examined the methods adopted by the reviewed studies as either a commentary, literature review and comments, qualitative Q and A with ChatGPT, the new model with ChatGPT, study with students/teachers, or other.
2. Thematized the themes of potentials for ChatGPT in education literature.
3. Thematized the themes of limitations and challenges for ChatGPT in education literature.
4. Synthesized and offered future considerations for each reviewed article.

We downloaded the full records of all the papers and uploaded the full records to Covidence. We used Covidence to identify duplicate studies and removed them and reviewed the title and abstract of each article and decide if relevant or not (following inclusion and exclusion criteria). We then extracted and charted the methods noted and then thematized potentials, limitations, and challenges. Further, we synthesized findings and discuss future recommendations and considerations.

3. Results

A total of 321 articles were found in WoS (71), Scopus (129), and IEEE (121). Of these 54 were duplicates and removed, leading to a total of 267 records screened. Of these, 200 studies were excluded as they predominantly explored ChatGPT in a specialized industry or medical and healthcare context. Of the 67 remaining studies sought for retrieval, 4 could not be retrieved due to a lack of open access to the publication. A total of 63 studies were thus assessed for eligibility. The screening process was conducted using Covidence software and led to a total of 63 studies included in this analysis.

3.1. What methods are used when studying ChatGPT in education literature?

Most studies from the pool assessed ChatGPT from a qualitative lens, providing comments on its potential, limitations, and threats. A summary of the methods adopted by the reviewed studies is presented in

Table 2

Summary of the methods adopted by the reviewed studies.

Type of paper	n	References
Commentary	17	(Alabool, 2023; Anders, 2023; Benuyenah, 2023; Duha, 2023; Emenike & Emenike, 2023; Harrison et al., 2023; Hwang & Chen, 2023; Kovačević, 2023; Peres et al., 2023; Qadir, 2023; Rospigliosi, 2023; Seghier, 2023; Stokel-Walker, 2022; Stokel-Walker & Van Noorden, 2023; van Dis et al., 2023; Yang, 2023; Yinping & Yongxin, 2023)
Literature review with comments	23	(Ahmad et al., 2023; Bahrini et al., 2023; Bekeš & Galzina, 2023; Costello, 2023; Cotton et al., 2023; Crawford et al., 2023; Farrokhnia et al., 2023; Garcia-Penalvo, 2023; Gašević et al., 2023; Gentile et al., 2023; Halaweh, 2023; Karaali, 2023; Kasneci et al., 2023; Kooli, 2023; Laato et al., 2023; Lin et al., 2023; Lo, 2023b; Lund et al., 2023; Miao et al., 2023; Neumann et al., 2023; Perkins, 2023; Ray, 2023; Silva et al., 2023)
New model with ChatGPT	7	(Alamleh et al., 2023; Chan et al., 2023; Cingillioglu, 2023; Dai et al., 2023; Elsayed, 2023; Ibrahim et al., 2023; Su & Yang, 2023)
Other	0	–
Qualitative Q and A with ChatGPT	11	(Ahmed & Sharo, 2023; Banić et al., 2023; Chen et al., 2023; Cooper, 2023; Elder et al., 2023; Fergus et al., 2023; Geerling et al., 2023; Humphry & Fuller, 2023; Jalil et al., 2023; O'Leary, 2023; Spasić & Janković, 2023)
Study with students and teachers	5	(Moon et al., 2023; Shoufan, 2023; Speth et al., 2023; Tili et al., 2023; Yan, 2023)

Table 2.

3.1.1. Commentary

Commentary articles are the second most common method used by the reviewed studies for examining ChatGPT in education. All unanimously explored considerations for ChatGPT in education and the need for better understanding rather than banning AI. Alabool (2023) conducted a SWOT analysis to offer insight into ChatGPT's strategic management in the education sector. Various issues that affect the different stakeholders in education are also identified. In Anders (2023), the author suggested adding instructions in specific assignments and rubrics regarding customized and supervised ChatGPT and AI integration rather than a general ban. Similarly, Duha (2023), Harrison et al. (2023), and van Dis et al. (2023) highlight the need to adapt our teaching and learning practices and roles in light of ChatGPT. In Emenike and Emenike (2023) and Peres et al. (2023) the authors noted that the use of text-based generative AI programs is inevitable and so more thought needs to be put into how to use it. In Kovačević (2023), the author offers the use of ChatGPT for the design, delivery, and evaluation of teaching units for English for Specific Purposes. Benuyenah (2023) considers AI may be the evolutionary hallmark of higher education.

Hwang and Chen (2023) bring forth an interesting lens toward the characterization of AI in education. They suggest classifying the levels at which the learner can collaborate with AI, namely:

- “Level 1 - None: The learner always waits for the teacher or others' commands or instructions.
- Level 2 - A little: The learner always asks the wrong questions.
- Level 3 - Average: The learner knows how to ask the right questions.
- Level 4 - A lot: The learner knows how to ask the right questions in logical sequences using a conversational approach.
- Level 5 - Super: The learner treats ChatGPT as a teammate and always works together as a good teammate with ChatGPT” (p. 16).

Qadir (2023) shares some of the issues and ambiguities that may emerge in teaching and learning as a result of using ChatGPT. The authors argue that while ChatGPT can present numerous benefits, those benefits may carry flaws or issues that need to be addressed before

wide-scale use. Rospigliosi (2023) and Seghier (2023) bring to attention areas of bias in using AI programs like ChatGPT. In Stokel-Walker (2022); Stokel-Walker and Van Noorden (2023) and Yan (2023), authors offer strategies to mitigate bias and better utilize AI in everyday teaching and learning. Yinping and Yongxin (2023) explore the value connotation of the digital transformation of education and offer mechanisms and approaches in which ChatGPT can be used to benefit digital education.

3.1.2. Literature review with comments

Literature review with comments is the most common method used by the reviewed studies for examining ChatGPT in education. Most of the literature review papers examine the considerations for employing ChatGPT in education.

Ahmad et al. (2023) explore the mixed decisions institutions have made regarding the implementation and use of large language models such as ChatGPT. Some show to encourage and some others ban them.

Through a review of the literature, Bahrini et al. (2023) present the applications, opportunities, and threats of ChatGPT in 10 domains. Overall, the authors find that ChatGPT may not have the same level of understanding, creativity, and empathy as humans and so may not make humans easily replaceable.

Bekeš and Galzina (2023) examine the use of AI-powered chatbots in education and share some of their advantages and disadvantages.

In Cotton et al. (2023), the authors warn universities to carefully weigh the risks and rewards of using AI tools and ensure their ethical and responsible implementation. They pose training faculty and students and using various methods to detect academic mis-integrity is needed. In Crawford et al. (2023), the authors raise the important point that while AI can help students learn, it does not substitute learning. Yet AI does in many ways provide an alternative way to learning and so its implications need to be well studied.

In Costello (2023), the author provides their opinion on the issues of AI in education.

In Gentile et al. (2023) a systematic analysis of the literature is conducted to analyze the change in the teacher's role triggered by the integration of AI into educational systems.

Silva et al. (2023) share the formative capabilities of ChatGPT and present guidelines for leveraging ChatGPT and similar generative AI models in education.

A commonality between authors is their suspicious view towards AI, yet optimism and hope that with proper supervision and policy-making, AI can come to improve teaching and learning.

Farrokhnia et al. (2023) use a SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis, a method used in organizational analysis to understand the impact of ChatGPT. They summarize internal and external factors as well as what is helpful and harmful in achieving goals. Gašević et al. (2023) bring together 11 papers that explore AI in learning and thematize potentials and limitations learners will be facing in light of AI in education. In Garcia-Penalvo (2023) the author is doubtful of the turnout of AI in education. Yet, the author notes that denying or banning it will not be helpful it may push the tsunami effect that has already begun further.

Halaweh (2023) presents an argument in favor of using ChatGPT in education and suggests the use of strategies that make the use of ChatGPT informed and ethical in education will be needed shortly. Karaali (2023) underlines that basic literacy and numeracy are as important as quantitative literacy and reasoning. Therefore, the use of AI tools such as ChatGPT should not come to deter such literacies in students. Kasneci et al. (2023) make the case that human monitoring and critical thinking are needed to guide AI tools in education. In a similar vein, Kooli (2023) advises that co-living, sustainability, and continuous adaptation to the development of AI systems will become an inevitable part of education. So AI must be viewed and collaborated with as an opportunity and ally, rather than a threat and adversary. In Laato et al. (2023), the authors examine how large language model services impact learning and teaching in higher education. Through a survey of grey

literature and the use of ChatGPT, the authors find and highlight 13 implications for students' learning in higher education.

Lin et al. (2023) review chatbot-related research from 1999 to 2022. For educational support, the authors find that AI tools improve comprehension skills, enhance teaching efficacy, and support collaborative learning.

Lo (2023) conducts a review to find its outcomes during its free three months of release. The author finds that ChatGPT's performance varies across subject domains, ranging from outstanding (e.g., economics) and satisfactory (e.g., programming) to unsatisfactory (e.g., mathematics). Lund et al. (2023) and Ray (2023) discuss the history and principles behind ChatGPT. The authors admit the major advancements made by ChatGPT and so stress the need for ethical and responsible use of such tools for teaching, learning, and publishing.

Miao et al. (2023) bring a new perspective by suggesting that it is not the student use of tools that would determine the onset of plagiarism or academic misconduct, but whether the university policies make clear use guidelines for the students. Neumann et al. (2023) review the grey literature to examine the effects of ChatGPT on higher education in the areas of software engineering and scientific writing. Perkins (2023) focuses on the great attention chatbot technology like ChatGPT has experienced in recent years and endorses a balance between AI-assisted innovation and human expertise.

3.1.3. New model with ChatGPT

Studies that examine a new model like ChatGPT are the fourth common method used by the reviewed studies for examining ChatGPT in education. Alamleh et al. (2023) assess the efficacy of algorithms in differentiating between hand-written and ChatGPT-generated text.

Chan et al. (2023) examine the use of a methodology called exercise creation methodology (ECM) to leverage recent AI technology and create ChatGPT-assisted programming exercises for beginners.

Cingillioglu (2023) attempts to distinguish between essays written by ChatGPT and humans using a proposed model that utilized techniques like Support Vector Machines or SVM, n-gram bag-of-words, and discrepancy language model. Dai et al. (2023) investigate the utility of ChatGPT to provide students with feedback for learning. The authors examine the readability of ChatGPT-generated feedback.

Elsayed (2023) explores the use of an evolutionary algorithm that helps to identify the best set of Bloom's taxonomy keywords to generate programming questions that ChatGPT has low confidence in answering. Using a selection of keywords from different cognitive levels was shown to significantly challenge the AI mode's capabilities. Ibrahim et al. (2023) examine ChatGPT's performance in two introductory and two advanced university courses. Findings show ChatGPT provides sufficient responses for introductory but provide unsophisticated responses for advanced courses.

Su and Yang (2023), on the other hand, offer a qualitative theoretical framework called "IDEE" for educative AI such as using ChatGPT and other generative AI in education. The framework includes: Identifying the desired outcomes, Determining the appropriate level of automation, Ensuring ethical considerations, and Evaluating effectiveness.

3.1.4. Qualitative Q and A

Studies that conducted qualitative questions and answers with ChatGPT are the third most common method used by the reviewed studies for examining ChatGPT in education.

Ahmed and Sharo (2023) explore the use of ChatGPT for the generation and assessment of student papers. Banić et al. (2023) share an overview of the challenges of programming education. Research findings from using pair programming with ChatGPT as an educational aiding tool are also presented. Results show that there is a positive difference in students' motivation for programming and using ChatGPT before and after conducted research, but not in the pair programming domain.

Chen et al. (2023) offer an application containing a

Question-and-Answer (QA) system, an expansion mechanism, a cache mechanism, and an error correction mechanism to serve as a mobile teaching assistant for students. The app further offers voice assistance.

In Cooper (2023), the author compares the ChatGPT responses on best practices of pedagogy with those of scholars. Elder et al. (2023) examine if ChatGPT3 can pass and complete a sophomore-level digital design laboratory. When the labs were independently graded, ChatGPT achieved a passing grade.

Fergus et al. (2023) test ChatGPT responses to knowledge and application questions in two chemistry-focused modules in year 1. The authors find ChatGPT to perform well in knowledge questions but fall short for application questions with non-text information. Geerling et al. (2023) aim to assess ChatGPT performance by having it take the Test of Understanding in College Economics (TUCE). The authors find the ChatGPT to rank in the 91st percentile, which suggests that it exceeded the mean responses of students across all institutions. Humphry and Fuller (2023) explore a chemistry report written by ChatGPT. The authors suggest that plagiarism will take a new face. Students will inevitably construct solutions with ChatGPT, so attention needs to be paid to ensure students understand the chemical principles they want an answer for, know how to ask questions, and evaluate ChatGPT's responses. Jalil et al. (2023) evaluate how well ChatGPT performs when given common questions in a software testing curriculum. The findings show that in nearly 56% of responses, the ChatGPT can provide correct or partially correct responses, and in 53% when explaining answers.

O'Leary (2023) compares several prominent AI models such as the one by Google, Facebook, and OpenAI. Spasić and Janković (2023) offer instructions for prompting ChatGPT to aid in planning lessons. The findings of their analysis with ChatGPT revealed that carefully tailored standard prompting with the additional role and seedword definitions leads to enhanced lesson planning by AI.

3.1.5. Study with students and teachers

Studies that examined students and teachers with ChatGPT are the least common method used by the reviewed studies for examining ChatGPT in education. Moon et al. (2023) compare the mentors' and ChatGPT's responses when learning a new programming language and utilize the strengths and limitations of each to inform best practices for learning programming with ChatGPT. Shoufan (2023) asks students to evaluate ChatGPT using their own words after interacting with ChatGPT to complete one learning activity. The authors then analyzed the responses to create a 27-item questionnaire and had students respond to this student-informed questionnaire three weeks later. Findings suggested that students have some reservations about using ChatGPT responses but overall are optimistic that it will be improved soon. In Speth et al. (2023), the authors evaluate the quality of exercises and best practices developed by ChatGPT and examine them through an external teaching assignment course. Tili et al. (2023) adopt a qualitative instrumental case study to examine ChatGPT in education. Work by Yan (2023), also adopted a qualitative approach to investigate students' behaviors and reflections in their exposure to ChatGPT in writing classrooms.

3.2. What are the potentials for ChatGPT in education literature?

A thematic summary and count of potentials for ChatGPT in education extracted from the reviewed studies are presented in Table 3. The most frequently noted (i.e., noted by 14 studies) potential is personalization and facilitation of complex learning, followed by (i.e., noted by 8 studies) designing specific teaching and learning activities and feedback. Below we explain the meaning of each theme:

The assessment relates to the use of AI for both delivering evaluation systems and applying evaluation on content. This could entail various contexts such as evaluation of curriculum design, delivery, student artifacts, tests, and projects.

Asynchronous communication and feedback relate to AI's ability to

Table 3
Thematic summary and count of potentials for ChatGPT in education.

Themes	n	References
Assessments	4	(Ahmad et al., 2023; Cotton et al., 2023; Perkins, 2023; Tili et al., 2023)
Asynchronous communication and feedback	4	(Cotton et al., 2023; Crawford et al., 2023; Farrokhnia et al., 2023; Qadir, 2023)
Critical thinking with/for AI	2	(Anders, 2023; Garcia-Penalvo, 2023)
Designing specific teaching and learning activities and feedback	8	(Bekeš & Galzina, 2023; Cooper, 2023; Jalil et al., 2023; Lin et al., 2023; Spasić & Janković, 2023; Speth et al., 2023; Su & Yang, 2023; Yang, 2023)
English skill improvement	4	(Bahrini et al., 2023; Kovačević, 2023; Lin et al., 2023; Perkins, 2023)
Improve digital ecosystems such as AR/VR	1	(Kasneji et al., 2023)
Improve efficiency and accuracy in research	5	(Bahrini et al., 2023; Kooli, 2023; Qadir, 2023; Seghier, 2023; van Dis et al., 2023)
Language Translation	4	(Ahmad et al., 2023; Cotton et al., 2023; Qadir, 2023; Yinping & Yongxin, 2023)
Learning to learn	3	(Bahrini et al., 2023; Karaali, 2023; Silva et al., 2023)
Make education more accessible and inclusive	4	(Alabool, 2023; Kooli, 2023; Laato et al., 2023; Lund et al., 2023)
Mental health	1	(Crawford et al., 2023)
Personalized and complex learning	14	(Ahmad et al., 2023; Ahmed & Sharo, 2023; Alabool, 2023; Banić et al., 2023; Chan et al., 2023; Elsayed, 2023; Farrokhnia et al., 2023; Humphry & Fuller, 2023; Kasneji et al., 2023; Kooli, 2023; Laato et al., 2023; Moon et al., 2023; Shoufan, 2023; Su & Yang, 2023)
Questions answering	3	(Cotton et al., 2023; Ibrahim et al., 2023; Jalil et al., 2023)
Serve different roles such as student, teacher, administrator, or co-agent	5	(Chen et al., 2023; Hwang & Chen, 2023; Lo, 2023; Perkins, 2023; Yinping & Yongxin, 2023)
Summarization	1	(Cotton et al., 2023)
Innovative potentials	2	(Alabool, 2023; Neumann et al., 2023)
Use of AI with academic integrity	2	(Alamleh et al., 2023; Cingillioglu, 2023)
Work offload	3	(Duha, 2023; Emenike & Emenike, 2023; Farrokhnia et al., 2023)

facilitate instruction in both synchronous and asynchronous means, not facing the time constraints students usually face with instructors' availability and their geographic time zones.

Critical thinking with/for AI explores the higher-level thinking potentials with AI in education. This may relate to both the development of AI's critical thinking and the scaffolding of students' critical thinking on content generated by AI.

Designing specific teaching and learning activities (e.g., science units, rubrics, and quizzes) can be facilitated with AI given appropriate pedagogical goals (e.g., constructive alignment or mapping between learning objectives and content) is achieved.

English skill improvement suggests that AI can act as an English teacher and editor. Though ChatGPT's greatest asset is in text analysis, ChatGPT may also be used for improving speech, and non-textual work in English.

Improve digital ecosystems such as AR/VR relates to ChatGPT's potential to offer services in digital mediums besides personal computers and mobile devices. In Augmented and Virtual Reality or AR/VR specifically, ChatGPT can help in creating more personalized, intuitive, and easy-to-use interfaces.

Improve efficiency and accuracy in research means that ChatGPT can serve not only as a co-author in writing, but other parts such as finding the most relevant literature, aiding in checking data analyses and calculations, and so on.

Language translation explores the potential of ChatGPT to translate speech and text in real-time, making it easier than before to connect and

socialize with individuals who may not speak English. This may be easier for everyday speech and require more work for English translation in technical and academic domains such as research collaboration. This is because the standards and naming of theories, units, and metrics may be different from one language to another. Further, each language may have different dialects in which the program needs to be careful with interpreting them and not associate a wrong word and meaning to what is said.

Learning to learn means that AI has the potential to learn how humans learn. More importantly, AI may have the capacity to study, characterize, and communicate the different types of techniques and approaches humans use during learning.

Making education more accessible and inclusive relates to ChatGPT's ability to serve as an auditor and mitigator of equity, diversity, and inclusion in instructional design, delivery, assessment, group work, and any other facet of education that is prone to bias and exclusion.

Mental health relates to ChatGPT's ability to monitor the well-being of humans and advise best practices for a healthy and balanced lifestyle. This can happen through mediums, first through informal and self-reported chats with individuals and second through monitoring physical cues from individuals such as body language through video surveillance (if given permission).

Personalized and complex learning focuses on how ChatGPT can facilitate learning that is customized and tailored to each individual's needs, knowledge base, and abilities. Complex learning which may denote ChatGPT's superior thinking abilities can be further used to augment and help human thinking.

Questions answering relates to ChatGPT's potential to act as the so-called wise man and have an answer for every question. Though the quality of answers may be questionable at times, nevertheless ChatGPT offers the potential to improve over time and like a human in training has polished and flourished knowledge over time.

Serve different roles such as student, teacher, administrator, or co-agent means that ChatGPT is not just one identity but has the potential to serve different roles and wear different thinking hats.

Summarization specifically focuses on the ChatGPT's ability to read, synthesize, and summarize large volumes of text. ChatGPT has the potential to read many volumes of text that may take years to complete and provide a high-level summary of the findings. Such summaries can play a key role in conducting more nuanced surveys of work that has been already done and schools of thought in a problem's field.

Innovative potential relates to the features and capabilities of ChatGPT that are yet to be discovered. And the potential for new releases of ChatGPT to enable further capabilities.

The use of AI with academic integrity explores the capabilities of AI to limit use types or detect any type of content creation by students and or teachers that may be plagiarized. Further, AI can teach best practices for achieving academic integrity in different contexts.

Work offload refers to the general availability and capability of ChatGPT and its potential to accomplish some of the tasks that would be traditionally done by humans. Automation is one key feature of AI programs that may be more systematic. There are however other uses with AI that may require critical thinking of AI which may not be always present.

3.3. What are the limitations and challenges for ChatGPT in education literature?

A thematic summary and count of limitations and challenges for ChatGPT in education extracted from the reviewed studies are presented in Table 4. The most frequently noted (i.e., noted by 9 studies) limitation is the potential for misuse or lack of learning with ChatGPT. The second most frequently noted (i.e., noted by 8 studies) is that technical expertise with ChatGPT is needed for appropriate use.

With educational organizations using ChatGPT in different aspects of educational programming and delivery, the notion of accountability is

Table 4

Thematic summary and count of challenges for ChatGPT in education.

Themes	n	References
Accountability	4	(Emenike & Emenike, 2023; Gašević et al., 2023; Lund et al., 2023; Tlili et al., 2023)
AI becomes the ultimate authority	2	(Cooper, 2023; Silva et al., 2023)
AI not plagiarism	2	(Anders, 2023; Duha, 2023)
All disciplines are at risk	1	(Benuyenah, 2023)
Augment learning with chatbots	1	(Geerling et al., 2023)
Availability	1	(Emenike and Emenike, 2023)
Challenges with quantitative and discipline-focused problems	2	(Humphry & Fuller, 2023; Johnkin et al., 2023)
Cost of plagiarism detection	2	(Crawford et al., 2023; Karaali, 2023)
Fake information	2	(Lo, 2023; Qadir, 2023)
Handling ambiguity	1	(Perkins, 2023)
Impact on research productivity	2	(Lund et al., 2023; Perkins, 2023)
Impersonal	2	(Ahmad et al., 2023; Lin et al., 2023)
Experiential learning projects	2	(Elder et al., 2023; Geerling et al., 2023)
Lack of higher-order skills	3	(Bahrini et al., 2023; Farrokhnia et al., 2023; Karaali, 2023)
Lack of transparency and trust	4	(Ahmed & Sharo, 2023; Bahrini et al., 2023; Jalil et al., 2023; Perkins, 2023)
Limited data quality	6	(Ahmad et al., 2023; Alabool, 2023; Elsayed, 2023; Qadir, 2023; Speth et al., 2023; Su & Yang, 2023)
Maintaining context	3	(Humphry & Fuller, 2023; Jalil et al., 2023; Perkins, 2023)
Misuse or lack of learning	9	(Duha, 2023; Elder et al., 2023; Karaali, 2023; Laato et al., 2023; Moon et al., 2023; Peres et al., 2023; Seghier, 2023; Shoufan, 2023; Yinping & Yongxin, 2023)
Multilingualism and fairness	2	(Karaali, 2023; Rospigliosi, 2023)
Plagiarism is not at word similarity level but other steps	6	(Cingillioglu, 2023; Cooper, 2023; Cotton et al., 2023; Karaali, 2023; Laato et al., 2023; Yan, 2023)
Privacy	6	(Alabool, 2023; Bahrini et al., 2023; Karaali, 2023; Lund et al., 2023; Miao et al., 2023; Stokel-Walker & Van Noorden, 2023)
Real-time multi-modal interaction lacking	1	(Rospigliosi, 2023)
Reintroduce old assessments	1	(Geerling et al., 2023)
Requires specific description	1	(Hwang & Chen, 2023)
Requires technical expertise in ChatGPT	8	(Ahmed & Sharo, 2023; Alamleh et al., 2023; Banić et al., 2023; Bekes & Galzina, 2023; Chan et al., 2023; Chen et al., 2023; Kovačević, 2023; Spasić & Janković, 2023)
Risk of bias and discrimination	6	(Ahmed & Sharo, 2023; Bahrini et al., 2023; Farrokhnia et al., 2023; Neumann et al., 2023; Stokel-Walker & Van Noorden, 2023; Yinping & Yongxin, 2023)
The reputation of the institution diminished	3	(Crawford et al., 2023; Ibrahim et al., 2023; Neumann et al., 2023)

becoming skewed. Especially when errors are made in areas where academic integrity is not violated, it becomes difficult to judge and decide who needs to take the blame or credit.

An ongoing issue with ChatGPT is whether it can ever truly have separate identities or whether it is the one source of knowledge that can take different forms and so is an ultimate epistemic authority.

AI is not plagiarism concerns about the justifications made surrounding the use of AI in design and publishing. One argument, for instance, is that AI is available to all is open access, and is also not a person, hence no plagiarism is made with the use of AI programs such as ChatGPT.

All disciplines are at risk of being changed by AI. This change may be content-wise or relate to how the manpower is needed in each discipline. With AI and automation having the ability to do classical jobs much more affordably and quickly, the academic disciplines and workforce are susceptible to becoming changed or diminished.

Augment learning with chatbots relates to the replacement of teachers and their important role in motivating and guiding students

with Chatbots. It may also relate to students' need to have chatbots and AI programs to put the puzzle pieces together and represent their learning. Without AI programs, therefore, students may have a volatile, limited, or incomplete schema of learning.

Availability relates to the access of all to ChatGPT in the future. Such programs likely become subscription-based. As there would be concerns regarding inclusivity (e.g., for low-income) and accessibility (e.g., for blind) students.

ChatGPT may currently face challenges with quantitative and discipline-focused problem-solving. Conversational or textual editing may require a general set of principles however disciplinary focused computations are quite narrow and require an understanding of a specific set of principles. The use of ChatGPT in such areas may thus be susceptible to errors and inaccuracies.

We may find that the cost of innovative technology to monitor and detect misconduct may become too high. Thus, using ChatGPT may create unforeseen costs to mitigate academic integrity and bias. This may result in hesitancy to implement ChatGPT altogether and create discussions on whether to ban or allow ChatGPT in education.

Dissemination of fake information is also another possibility with ChatGPT. Given that it largely depends on input text and data sets the quality of sources can greatly impact the truth of information developed by ChatGPT.

Handling ambiguity concerns about AI program's inability to provide satisfactory or relevant responses for ambiguous queries. AI programs may oversimplify uncertainties or not account for some of them when creating a representation of a response to a problem.

Impact on research productivity concerns researchers' way or lack of using AI programs and its impact on the landscape of research productivity score. The inconsistent use of AI by the researchers and their varied types of use can significantly impact the quality and credibility of the research.

AI is still largely non-humanistic, and humans may be inherently less motivated to communicate with a piece of technology when compared to communication with other humans. As such humans may still find AI programs impersonal and have resistance to cooperation.

Increasing experiential learning projects that artificial intelligence struggles to replicate are one way of achieving academic integrity in AI-equipped classrooms. However, AI may over time also learn to appropriately answer more open-ended questions.

Lack of higher-order cognitive skills implies that students and teachers alike may over-rely on responses from ChatGPT and lack their critical thinking skills over time. It can also signify the lack of ChatGPT at its present state to critically evaluate the content and rather model critical thinking patterns instead of generating and justifying them.

Lack of transparency and trust may result from the generation of sophisticated AI programs whose inner workings are not truly apparent and understandable by humans. With a lack of transparency in the system, humans may construct different notions of trust towards the use of ChatGPT in their daily lives and areas such as education.

Limited data quality may suggest that an AI program is as good as its input data and so limited data quality would lead to unreliable or inaccurate AI responses. A consideration is the monetization and monopoly of companies and data sources used for AI in education. Attention must be given by involved companies to not streamline biased and low-quality data sources.

Maintaining context with AI programs is another challenge. Still, responses from AI may be deemed as broad and too general. AI may be seen as having a bucket of general knowledge that it uses to provide responses when in reality it needs to create bowls of specific knowledge to better address problems under different contexts and specialized domains.

Misuse or lack of learning relates to students' use of AI in ways that do not support their true learning and create an inaccurate picture of their learning errors and state. Students may as a result create volatile learning patterns that become erased without the presence of AI

programs in education.

Multilingualism and fairness relate to a lack of application in other languages and for non-English learners, leading to discrimination due to the language individuals are fluent in. It is thus imperative to ensure ChatGPT is truly accessible to all so its intelligence accurately portrays diverse human behaviors and thinking profiles.

Plagiarism is contented to be far beyond copying ideas in the text as own. With ChatGPT, plagiarism has the potential to not necessarily be at the word similarity level, but other steps and areas are not easily detectable such as summarizing sources and conducting analyses.

Privacy may concern the use of student data, maintaining the confidentiality of data, and not allowing students' data to be compromised and leaked by hackers or other affiliate software. While attempts may be made to ensure privacy, the interaction with ChatGPT may inevitably result in compromised privacy. There could be an explanation for example on AI programs' websites on the levels of interaction and subsequent privacy threats and mitigation strategies individuals could take.

Lack of real-time multi-modal interaction concerns current ChatGPT's inability to engage through multimodal channels and have senses similar to that of a human. In other words, in many areas, AI programs do not sense feel, or experience things but rather are prescribed to do so according to text.

Reintroduce proctored, in-person assessments related to the need to bring back stringent proctoring methods in light of AI technology. While ChatGPT may become allowed in general teaching and learning, instructors may need to bring students back into the classrooms for paper-based assessments and human monitoring so that programs such as ChatGPT are not used during test-taking.

AI may still require specific goals and question descriptions. This may be because the current architecture of AI programs such as ChatGPT still largely needs specific information and prescriptions to function and produce convincing responses.

Requiring technical expertise with ChatGPT relates to having software and programming knowledge as well as being proficient in navigating the ChatGPT infrastructure, internet, and intelligent web-based tools.

The risk of bias and discrimination results from the use of AI in non-inclusive ways. There are many known and unknown ways in which bias and discrimination happen with the use of ChatGPT. It may be useful to involve ChatGPT itself in learning types of plagiarism and understanding when ethical work is violated.

The diminished reputation of academic institutions stems from a lack of monitoring protocols and policies on how to use AI in teaching, learning, and research. Institutions thus need to show accountability by reflecting on and presenting strategies for the use of ChatGPT in education.

4. Discussion

An overview of the future recommendations for the study of ChatGPT is presented in this section based on the synthesis of the reviewed studies.

4.1. Findings of review

This review paper examines areas where ChatGPT is employed in educational literature and areas of potential, challenges, and future work. Findings showed that the potentials of ChatGPT include but are not limited to the development of personalized and complex learning, specific teaching and learning activities, assessments, asynchronous communication, and feedback, accuracy in research, personas, and task delegation and cognitive offload. Findings further presented several areas of challenge that ChatGPT is or will be facing in education. Examples include but are not limited to plagiarism deception, misuse lack of learning, accountability, and privacy. Below, we provide our

recommendations based on the synthesis of the 63 reviewed studies.

4.2. Recommendations and synthesis of challenges and future considerations

An emergent concern is the potential ethical issues of ChatGPT (Ahmad et al., 2023; Ahmed & Sharo, 2023). To alleviate such issues, it is suggested to define ethical rules and make AI more transparent. A blindspot, however, could be the construction of ethics and bias itself and how both humans and AI collectively decide on what thresholds to consider for separating biased from unbiased practices. In cases where there is a dilemma, such separation of ethical and non-ethical decision-making can become particularly difficult.

When using ChatGPT, the recommendations for research, teaching, and institutional practice components may be different (Alabool, 2023). The impact each component has separately and holistically on ChatGPT's learning is unknown and difficult to untangle. The way to apply critical thinking to assess AI needs to be developed (Alamleh et al., 2023; Anders, 2023). The framework may be created by humans and AI itself and its implications need to be fully studied.

A review of applications, opportunities, and threats through a SWOT analysis presents several areas of interest and concern (Bahrini et al., 2023). Yet given the subjective nature of goal-oriented behavior, a threat at one instance may be an opportunity at another instance and vice versa when working with ChatGPT. As a result, a more transparent reflection of goal-oriented behavior is needed.

The use of ChatGPT may show promising improvements in student motivation (Banić et al., 2023). Yet such effects need to be studied in the long term and efforts need to be made to study student motivation changes based on their learning and demographic backgrounds. Rejecting the idea that the presence of AI in education in any shape and form will have no impact on higher education is perhaps naïve (Benuyenah, 2023). We may need to consider that different conceptualizations of AI will eventually become marketable and so higher education should be concerned about a range of consequences from such programs rather than from a program such as ChatGPT.

The use of ChatGPT is shown to improve question generation and banking (Chan et al., 2023; Chen et al., 2023). Yet, an emergent issue can be the computational power and storage space needed to maintain and expand such question-generation banks. Future breakthroughs may thus require further minimization of storage space and central processing needed to carry out ChatGPT's tasks.

Superior accuracy performance is seen with software in the market such as copy leaks. Yet, new models can also be investigated such as the one by Cingillioglu (2023) where the objective measures of comparison may change under different contexts.

A range of responsibility allocations from editing up to assuming co-authorship can be made with ChatGPT (Costello, 2023). The notion of language and semantics of language may serve as both an opportunity and limitation with ChatGPT (Elsayed, 2023). If we are to believe ChatGPT can lead to convergence, we may also allow for the creation of languages and semantics via AI in different domains. Ultimately, programs like ChatGPT may have the potential to serve as a medium to make a universal language and window to perception and learning.

Suggestions of AI may need to be taken with a grain of salt if the software is knowledgeable and informed enough to make both a written work and/or assess written work (Cotton et al., 2023). In the race for superior knowledge with humans or other programs, ChatGPT may learn to underscore text that was not written by itself or conversely be equipped to be more elaborative with its work.

Students are predicted to know how to use AI programs such as ChatGPT for good during learning (Crawford et al., 2023). While a skeptical view may consider all that ChatGPT can offer as cheating, a more aware view would accept that ChatGPT may lead to cheating but consider the nature of interactions accepted in the education model that is collaborative rather than one-directional with AI.

While ChatGPT can facilitate a more critical review of students' generated content, it may also lead to the students' reliance on software and result in reduced self-assessment and critical thinking among students (Dai et al., 2023; Duha, 2023; Elder et al., 2023).

ChatGPT can thus conceptualize learning with students' critical thinking in mind such as providing prompts or quizzes that would scaffold students' higher-level thinking. Given the seismic shift AI programs such as ChatGPT are making to the landscape of teaching and learning, a point of debate could become whether instructors should continue to assume some conventional duties such as assessment of student work based on clarity and format (Emenike & Emenike, 2023). The key tension here seems to be on where to draw lines between the use and misuse of such programs. Some refute the use of such programs altogether as a solution while others are more optimistic and promote the cautious use of these programs. While we continue to hear such debates and gain more data from AI programs' true performances and behaviors, we need to consider the many ways in which such seismic shifts may drift, collide, and transform areas of teaching and learning and academia, industry, and research more generally.

Developing ethical guidelines in teaching and learning is becoming more important than before in education (Farrokhnia et al., 2023). An undiscovered potential for ChatGPT is to take on different personas and assess the ethics of its generated concepts against the values of such different personas.

ChatGPT may be considered a high-risk technology tool for cheating (Fergus et al., 2023). However, this could be because we are augmenting ChatGPT on traditional types of education that had no ChatGPT in mind at the time. One area of innovation could be to let ChatGPT take the lead and offer educational interventions that also prevent or locate cheating.

The literacy of developers, users, and stakeholders of AI technology will be inevitably needed (Gašević et al., 2023). Much of one's education may thus shift from learning different contexts and conceptual domains to learning how to communicate, debug, manage, and collaborate with AI technology during teaching and learning.

There are many issues noted with the use of AI technology in education many of which are considered pressing issues (Garcia-Penalvo, 2023). An area of challenge researchers may currently face is understanding the criticality and urgency of issues noted as pressing in the literature and how to prioritize and classify such issues based on importance under different contexts.

Technology such as ChatGPT may be making room for many shades of plagiarism. As such more and more sophisticated technology is becoming needed to detect and account for unforeseen types of academic misconduct (Geerling et al., 2023). This may create a vicious cycle of needing more and more computational power and hardware to account for a range of plagiarism that may take place in research and pedagogy.

Some, however, suggest that the development of smart classrooms and new school environments that are highly integrated with AI will be the solution to rethinking the classrooms (Gentile et al., 2023). An extremely futuristic outlook in this regard may take that AI will eventually be embedded and wired into the human brain, not requiring humans to be needing to effortfully learning anymore.

In its present shape, ChatGPT may offer a more advanced understanding of proper English writing and editing, rather than the creation or synthesis of new concepts (Halaweh, 2023). The curriculum could therefore highlight the need for students to make a synthesis of their own by setting students in authentic contexts that draw upon their individualistic experiences and perceptions, something that might still be far-reaching with AI programs.

Instead of banning, some authors advise a restorative approach to education (Harrison et al., 2023). This requires gaining a critical understanding of what was appropriate and worked in the first place to pass it on to AI programs and have them cultivate such cultures.

Some of the issues of plagiarism may be present broadly and even with the absence of AI (Humphry & Fuller, 2023). An example is the use

of statistical analyses without knowing much about the theory behind it. As such, some of the plagiarism issues that exist may remain even if programs such as ChatGPT are banned.

ChatGPT can be used as a methodical approach to put to test pedagogical theories of learning (Hwang & Chen, 2023). Doing so can help us cross out outdated or incorrect theories and bring into perspective new theories of learning in light of the introduced educational technologies.

Changing text as little as moving full stops or adding typos may fool the plagiarism detection of ChatGPT-generated text (Ibrahim et al., 2023). This can have a significant impact on the honesty and integrity of work and if uncaught can diminish the reputation of institutions.

To understand how well ChatGPT performs under different contexts, common questions may be given to ChatGPT to respond to (Jalil et al., 2023). Such generality may lack the specialized focus of specific fields and as a result, may provide the false hope that ChatGPT provides adequate responses in all domains.

Quantitative literacy is deemed difficult to teach to both humans and AI programs (Johinke et al., 2023). Yet, we may also need to consider that different technologies may produce different personas of AI (Kar-aali, 2023). For example, one that is less knowledgeable but creative versus literate but rigid, and so on.

Most may interact with ChatGPT in a very narrow manner. Closed-ended questions are easier for the program to respond to. The open-ended questions are also often assessed for their language appropriacy and not necessarily criteria (Kasneci et al., 2023). Another question of debate is whether ChatGPT meets the needs of a not known and priori rubric.

The data sources used by ChatGPT are also another concern. Common datasets used are provided by large corporations, which may primarily benefit them (Kooli, 2023). A key issue with higher education can be academic integrity not only with content creation, but content privacy and sharing with third-party companies with business-focused, rather than pedagogical-focused, objectives.

The use of ChatGPT in topics such as English for Specific Purposes may enable instant feedback to students on their errors and progress (Kovačević, 2023). Since often feedback has a goal-oriented and longitudinal process it becomes important to consider and compare the possible trajectories of feedback and learning paths the student can potentially pursue.

The use of teachers and students and their artifacts often serves a role in training and aiding ChatGPT learning (Laato et al., 2023). As a result, one could argue that the performance of the ChatGPT is as good as the information fed to it. A lack of universal access and centrality to what is fed to AI programs may as a result make them deviate from their purpose which is to make assistance a uniform and universal experience.

There are generally plenty of algorithms and data sets, but fewer supervised efforts on frameworks when using ChatGPT in education. It may be time to bridge the learning theories and sciences empirically with what ChatGPT can collect and characterize from student and teacher behaviors.

The proficiency of ChatGPT is compared across domains (Lin et al., 2023), yet the researchers may have different levels of difficulty and stringency in their analyses. Not to neglect ChatGPT is a self-improving system and thus can improve in whatever field it is often prompted to work with.

Future research could explore the use of GPT/ChatGPT in conjunction with other language models or technologies to enhance their capabilities and performance (Lo, 2023). However, we may need to acknowledge that a living human-like AI might be much more different than an online computer-assisted AI.

In evaluating AI methods we may need to shift from traditional methods to ones that include human factors (Lund et al., 2023). Making AI humanistic or aware of humans' needs may require considerable input and data collection from the human side. This may not be the only concern and the inclusivity and appropriacy of data collection also become critical.

Research is beginning to account for pedagogical types of learning (e. g., active, lifelong) as part of highly technical learning types often seen with machine learning algorithms (Miao et al., 2023). These types of paradigms may help in bridging the gap and characterizing educational science theories of learning with ones that are empirically classified and measured by computers.

Besides providing insufficient or erroneous responses, ChatGPT may also provide responses that are beyond the understanding of the learner (Moon et al., 2023). Such scenarios may require the ChatGPT's backward engineering of concepts until the current level of student understanding is reached.

AI-based chatbots may have a disruptive impact on higher education (Neumann et al., 2023), impacting the notions of plagiarism, use permissions, and referencing with programs such as ChatGPT. An ongoing concern could be the many ways AI could be used in student work and the role of ChatGPT to be masked away. Examples include translating a text-generated work by ChatGPT and then reformulating it a little and translating it back again, making the identification of AI text generation almost impossible.

While ChatGPT is more under a magnifying glass, we should not neglect that there may have been assistive technologies present before the release of ChatGPT. Hence considering and comparing all forms of educational or assistive learning technologies that may have a form of intelligence will be needed (O'Leary, 2023).

The integration of AI-supported digital tools into the classroom environment is highly likely (Peres et al., 2023). There will be geopolitical considerations of such integrations and their implications must be anticipated. For example, different regions may use different versions of ChatGPT and each may cause different notions of knowledge construction.

ChatGPT may offer a cognitive offload for individuals in the education system (Perkins, 2023). No doubt demanding institutions overwork both teachers and students. One argument can be tasks that can be offloaded like the tasks we saw get replaced with machines. At some point, humans decided to use automation for some mundane tasks and did not consider such replacement as a threat to their learning. Put another way, we can consider delegating educational tasks to the AI programs that are logistical components of the traditional education system that did not offer much learning, to begin with.

ChatGPT may benefit engineering education through language editing, virtual tutoring, and problem-solving (Qadir, 2023). The concerning issue, however, is that ChatGPT and other AI programs are imperfect and make errors that are unpredictable and may have high consequences. It is therefore important to caution engineering students in training on the potential risks and effects of working mindlessly with ChatGPT.

One key purpose of ChatGPT is to interact through conversation, which involves a series of questions from users and responses from the application (Ray, 2023). The types of dialogue can be infinite, but AI program learns from only what it experiences in dialogues with humans. For AI to be inclusive of diverse thoughts, requires acceptance, tolerance, and a dialogue with different schools of thought.

Creating a culture of questions around the use of ChatGPT is needed to better assess the place and role of ChatGPT in education (Rospigliosi, 2023). While we see that many articles touch upon important concerns, less work has been done (outside of review contexts), to bring together these questions and concerns together and make systematic evaluations and refinements of ChatGPT out of them.

Using AI programs to assess its own or others' creations may in short lead to lazy work (Seghier, 2023). To provide an analogy, you are not considered to be a great teacher if you only call your teaching great. Attention must be put to not making ChatGPT a singular voice in teaching, learning, and evaluation.

Using student perceptions as a gateway to improve programs such as ChatGPT is also needed (Shoufan, 2023). However, attention must be paid to not make ChatGPT a better replacement for the student, but a

better moderator and facilitator of learning for the students.

Authors have worked with different configurations of prompting AI to get better results such as lesson preparation (Silva et al., 2023; Spasić & Janković, 2023). Given the vast possibilities of prompting AI, this task may be further automated to examine the entire solution space when working with AI programs such as ChatGPT.

In question generation, ChatGPT may show superior performance in some contexts (Speth et al., 2023). This may suggest that having more data may not be the only sufficient factor in ChatGPT's learning. More work as a result is needed to understand how and why some contexts are more readily understood and formulated by ChatGPT than others.

When should we as educators make the case to worry for students' learning is a question worth exploring (Stokel-Walker, 2022). Even without the presence of AI programs, the level of concern for student plagiarism varies from one institution to another. ChatGPT perhaps, is just bringing to light the reality check that our institutionalized certifications are at times phony and devoid of any value.

Examining the impact of ChatGPT on student motivation and engagement is necessary (Stokel-Walker & Van Noorden, 2023). This is because learning can only happen when the students put in the time and effort to spot and correct their learning errors. It would be, therefore, needed to see how interactions with chatbots and programs can elevate a sense of motivation in students.

Studying how human and AI tutors work and perform teamwork in understanding and conceptualizing learning will become necessary (Su & Yang, 2023; Tlili et al., 2023).

More work may be needed not only in different languages but also with second language learners working with ChatGPT in English (van Dis et al., 2023). If we assume that ChatGPT learns based on the quantity and frequency of interactions, we expect ChatGPT to learn the best based on the relative population of nations, for example, larger population regions to take the lead. Yet, having regionally separated interactions and experiences with ChatGPT may result in censorship, bias, and or inflated views or lead to misinformation, propaganda, and political warfare such as cold wars, among others.

How to have human students evaluate AI's work can become questionable given that students are inherently novices (Yan, 2023). Peer assessment or learning between novice learners and AI may thus need an expert human moderator.

ChatGPT has the potential to connect stakeholders in educational settings, for example aiding teachers to teach better, students to learn better, and parents to communicate (Yinping & Yongxin, 2023). An important consideration is how much the voices of different stakeholders should be accepted and factored into educational decision-making. As a result, the decision-making process of AI to make such authority allocations becomes critical.

While we come to worry about certain characteristics of ChatGPT to facilitate plagiarism, we may need to be more concerned about the divide ChatGPT may create between well and under-represented groups. A new form of gentrification may be created that divides regions and communities in unforeseen new ways. Not to neglect those who do not have the means to access technological devices such as smartphones, let alone ChatGPT may become the new isolated societies of the future.

4.3. Study limitations

We acknowledge that the literature survey comes with limitations. Namely, our review resulted in the analysis of a subset of papers specifically those that had a mention of ChatGPT in their document. Programs that are highly similar to ChatGPT but have different naming conventions were therefore omitted in this review. Furthermore, our review may lack the rigorous types of analyses that may be present in the literature review that take several years to take and examine the maturity or lack thereof of a technology over a more sensible time frame. We should not neglect that the time of analysis has a high impact on the conceptualizations of ChatGPT in education and so some of the notions

and hesitations may evolve over technological changes in programs such as ChatGPT.

5. Concluding remarks

Our discussion of future recommendations surveyed the reviewed studies and offered considerations and blind spots in the studies of ChatGPT. The potentials of ChatGPT include but are not limited to the development of personalized and complex learning, specific teaching and learning activities, assessments, asynchronous communication, and feedback, accuracy in research, personas, and task delegation and cognitive offload. Several areas of challenge that ChatGPT is or will be facing in education are also shared. Examples include but are not limited to plagiarism deception, misuse lack of learning, accountability, and privacy. We recommend future work take a more empirical rather than opinionative stance and study ChatGPT in education.

Data availability statements

Data sharing does not apply to this article as no datasets were generated or analyzed during the current study.

Declaration of competing interest

None.

Acknowledgements

This work was supported by the Canada Research Chair Program and Canada Foundation for Innovation.

References

- Ahmad, N., Murugesan, S., & Kshetri, N. (2023). Generative artificial intelligence and the education sector. *Computer*, 56(6), 72–76.
- Ahmed, Y. A., & Sharo, A. (2023). On the education effect of CHATGPT: Is AI CHATGPT to dominate education career profession?. In *2023 international conference on intelligent computing, communication, networking and services (ICCNIS)*, (pp. 79–84). <https://doi.org/10.1109/ICCNIS58795.2023.10192993>
- Alabool, H. M. (2023). ChatGPT in education: SWOT analysis approach. In *2023 international conference on information technology (ICIT)* (pp. 184–189). <https://doi.org/10.1109/ICIT58056.2023.10225801>
- Alamleh, H., AlQahtani, A. A. S., & ElSaid, A. (2023). Distinguishing human-written and ChatGPT-generated text using machine learning. In *2023 systems and information engineering design symposium (SIEDS)* (pp. 154–158). <https://doi.org/10.1109/SIEDS58326.2023.10137767>
- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2).
- Anders, B. A. (2023). Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking? *Patterns*, 4(3), 1–2. <https://doi.org/10.1016/j.patter.2023.100694>
- Bahrini, A., Khamoshifar, M., Abbasimehr, H., Riggs, R. J., Esmaeili, M., Majdabadkohne, R. M., & Pasehvar, M. (2023). ChatGPT: Applications, opportunities, and threats. In *2023 systems and information engineering design symposium (SIEDS)* (pp. 274–279). <https://doi.org/10.1109/SIEDS58326.2023.10137850>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). *Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning*. Available at: SSRN 4337484 (2023).
- Banić, B., Konecki, M., & Konecki, M. (2023). Pair programming education aided by ChatGPT. In *2023 46th MIPRO ICT and electronics convention (MIPRO)* (pp. 911–915). <https://doi.org/10.23919/MIPRO57284.2023.10159727>
- Bekes, 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). <https://doi.org/10.23919/MIPRO57284.2023.10159734>
- Benuyenah, V. (2023). Commentary: ChatGPT use in higher education assessment: Prospects and epistemic threats. *Journal of Research Innovations Teaching and Learning*, 16(1), 134–135. <https://doi.org/10.1108/JRIT-03-2023-097>
- Chan, W. K., Yu, Y. T., Keung, J. W., & Lee, V. C. S. (2023). Toward AI-assisted exercise creation for first course in programming through adversarial examples of AI models. In *2023 IEEE 35th international conference on software engineering education and training (CSE&T)* (pp. 132–136). <https://doi.org/10.1109/CSE&T58097.2023.00028>
- Chen, Y., Deng, H., Chen, C.-H., & Chung, C.-L. (2023). Efficient artificial intelligence-teaching assistant based on ChatGPT. In *2023 international conference on smart systems for applications in electrical sciences (ICSSES)* (pp. 1–5). <https://doi.org/10.1109/ICSSES58299.2023.10200077>
- Cingillioglu, I. (2023). Detecting AI-generated essays: The ChatGPT challenge. *International Journal of Information and Learning Technology*. <https://doi.org/10.1108/IJILT-03-2023-0043>
- 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>
- Costello, E. (2023). ChatGPT and the educational AI chatter: Full of bullshit or trying to tell us something? *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-023-00398-5>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education & Teaching International*. <https://doi.org/10.1080/14703297.2023.2190148>
- Crawford, J., Cawling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Creswell, J. (1994). *Research design: Qualitative and quantitative approaches*. Sage.
- Currie, G. M. (2023). Academic integrity and artificial intelligence: Is ChatGPT hype, hero or heresy? *Seminars in Nuclear Medicine*. <https://doi.org/10.1053/j.semnuclmed.2023.04.008>
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. In *2023 IEEE international conference on advanced learning technologies (ICALT)* (pp. 323–325). <https://doi.org/10.1109/ICALT58122.2023.00100>
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- Duha, M. S. U. (2023). ChatGPT in education: An opportunity or a challenge for the future? *TechTrends*. <https://doi.org/10.1007/s11528-023-00844-y>
- Eggmann, F., Weiger, R., Zitzmann, N. U., & Blatz, M. B. (2023). Implications of large language models such as ChatGPT for dental medicine. *Journal of Esthetic and Restorative Dentistry*. <https://doi.org/10.1111/jerd.13046>
- Elder, C., Pozek, G., Horine, S., Tripaldelli, A., & Butka, B. (2023). Can artificial intelligence pass a sophomore level digital design laboratory? *SoutheastCon, 2023*, 861–868. <https://doi.org/10.1109/SoutheastCon51012.2023.10115116>
- Elsayed, S. (2023). Towards mitigating ChatGPT's negative impact on education: Optimizing question design through bloom's taxonomy. In *2023 IEEE region 10 symposium (TENSYP)* (pp. 1–6). <https://doi.org/10.1109/TENSYP5890.2023.10223662>
- Emenike, M. E., & Emenike, B. U. (2023). Was this title generated by ChatGPT? Considerations for artificial intelligence text-generation software programs for chemists and chemistry educators. *Journal of Chemical Education*, 100(4), 1413–1418. <https://doi.org/10.1021/acs.jchemed.3c00063>
- Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: A conversation with ChatGPT and a call for papers. *JMIR Medical Education*, 9. <https://doi.org/10.2196/46885>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education & Teaching International*. <https://doi.org/10.1080/14703297.2023.2195846>
- Fergus, S., Botha, M., & Ostovar, M. (2023). Evaluating academic answers generated using ChatGPT. *Journal of Chemical Education*, 100(4), 1672–1675. <https://doi.org/10.1021/acs.jchemed.3c00087>
- Garcia-Penalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in Knowledge Society*, 24. <https://doi.org/10.14201/eks.31279>
- Gardner, D. E., & Giordano, A. N. (2023). The challenges and value of undergraduate oral exams in the physical chemistry classroom: A useful tool in the assessment toolbox. *Journal of Chemical Education*, 100(5), 1705–1709. <https://doi.org/10.1021/acs.jchemed.3c00011>
- Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. *Computer Education*. <https://doi.org/10.1016/j.caeai.2023.100130>
- Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in College economics: Now what? *American Economist*. <https://doi.org/10.1177/05694345231169654>
- Gentile, M., Città, G., Perna, S., & Allegra, M. (2023). Do we still need teachers? Navigating the paradigm shift of the teacher's role in the AI era. *Frontiers of Education*, 8. <https://doi.org/10.3389/educ.2023.1161777>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2). <https://doi.org/10.30935/cedtech/13036>
- 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>
- Humphry, T., & Fuller, A. L. (2023). Potential ChatGPT use in undergraduate chemistry laboratories. *Journal of Chemical Education*, 100(4), 1434–1436. <https://doi.org/10.1021/acs.jchemed.3c00006>
- Hwang, G.-J., & Chen, N.-S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society*, 26(2). [https://doi.org/10.30191/ETS.202304_26\(2\).0014](https://doi.org/10.30191/ETS.202304_26(2).0014)
- Ibrahim, H., Asim, R., Zaffar, F., Rahwan, T., & Zaki, Y. (2023). Rethinking homework in the age of artificial intelligence. *IEEE Intelligent Systems*, 38(2), 24–27. <https://doi.org/10.1109/MIS.2023.3255599>

- Jalil, S., Rafi, S., LaToza, T. D., Moran, K., & Lam, W. (2023). ChatGPT and software testing education: Promises & perils. In *2023 IEEE international conference on software testing, verification and validation workshops (ICSTW)* (pp. 4130–4137). <https://doi.org/10.1109/ICSTW58534.2023.00078>
- Johinke, R., Cummings, R., & Di Lauro, F. (2023). Reclaiming the technology of higher education for teaching digital writing in a post—pandemic world. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.01>
- Karaali, G. (2023). Artificial intelligence, basic skills, and quantitative literacy. *Numeracy*, 16(1). <https://doi.org/10.5038/1936-4660.16.1.1438>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7). <https://doi.org/10.3390/su15075614>
- Kovačević, D. (2023). Use of ChatGPT in ESP teaching process. In *2023 22nd international symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1–5). <https://doi.org/10.1109/INFOTEH57020.2023.10094133>
- Laato, S., Morschheuser, B., Hamari, J., & Björne, J. (2023). AI-assisted learning with ChatGPT and large language models: Implications for higher education. In *2023 IEEE international conference on advanced learning technologies (ICALT)* (pp. 226–230). <https://doi.org/10.1109/ICALT58122.2023.00072>
- Lin, C.-C., Huang, A. Y. Q., & Yang, S. J. H. (2023). A review of AI-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, 15(5). <https://doi.org/10.3390/su15054012>
- 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>
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581. <https://doi.org/10.1002/asi.24750>
- McGee, R. W. (2023). How would history be different if Karl Marx had never been born?. In *A ChatGPT essay*.
- Miao, H., & Ahn, H. (2023). Impact of ChatGPT on interdisciplinary nursing education and research. *Asian Pacific Island Nursing Journal*, 7. <https://doi.org/10.2196/48136>
- Miao, Q., Zheng, W., Lv, Y., Huang, M., Ding, W., & Wang, F.-Y. (2023). DAO to HANOI via DeSci: AI paradigm shifts from AlphaGo to ChatGPT. *IEEE CAA*, 10(4), 877–897. <https://doi.org/10.1109/JAS.2023.123561>. *Journal of Automatica Sinica*
- Moon, J., Yang, R., Cha, S., & Kim, S. B. (2023). chatGPT vs mentor : Programming language learning assistance system for beginners. In *2023 IEEE 8th international conference on software engineering and computer systems (ICSECS)* (pp. 106–110). <https://doi.org/10.1109/ICSECS58457.2023.10256295>
- Neumann, M., Rauschenberger, M., & Schön, E.-M. (2023). “We need to talk about ChatGPT”: The future of AI and higher education. In *2023 IEEE/ACM 5th international workshop on software engineering education for the next generation (SEENG)* (pp. 29–32). <https://doi.org/10.1109/SEENG59157.2023.00010>
- O’Leary, D. E. (2023). An analysis of three chatbots: BlenderBot, ChatGPT and LaMDA. *Intelligent Systems in Accounting, Finance and Management*, 30(1), 41–54. <https://doi.org/10.1002/isaf.1531>
- Open, A. I. (2023). *ChatGPT; personal communication*.
- Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*. <https://doi.org/10.1016/j.ijresmar.2023.03.001>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Qadir, J. (2023). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE global engineering education conference (EDUCON)* (pp. 1–9). <https://doi.org/10.1109/EDUCON54358.2023.10125121>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154.
- Rospigliosi, P. A. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments*, 31(1), 1–3. <https://doi.org/10.1080/10494820.2023.2180191>
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and open-AI models: A preliminary review. *Future Internet*, 15(6), 192.
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare (Switzerland)*, 11(6). <https://doi.org/10.3390/healthcare11060887>
- Seghier, M. L. (2023). ChatGPT: Not all languages are equal. *Nature*, 615(7951), 216. <https://doi.org/10.1038/d41586-023-00680-3>
- Shoufan, A. (2023). Exploring students’ perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2023.3268224>. PG-1-1.
- Silva, D. De, Mills, N., El-Ayoubi, M., Manic, M., & Alahakoon, D. (2023). ChatGPT and generative AI guidelines for addressing academic integrity and augmenting pre-existing chatbots. In *2023 IEEE international conference on industrial technology (ICIT)* (pp. 1–6). <https://doi.org/10.1109/ICIT58465.2023.10143123>
- Spasić, A. J., & Janković, D. S. (2023). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In *2023 58th international scientific conference on information, communication and energy systems and technologies (ICEST)* (pp. 47–50). <https://doi.org/10.1109/ICEST58410.2023.10187269>
- Speth, S., Meißner, N., & Becker, S. (2023). Investigating the use of AI-generated exercises for beginner and intermediate programming courses: A ChatGPT case study. In *2023 IEEE 35th international conference on software engineering education and training (CSEET&T)* (pp. 142–146). <https://doi.org/10.1109/CSEET58097.2023.00030>
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays — should academics worry? *Nature*. <https://doi.org/10.1038/d41586-022-04397-7>
- Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. *Nature*, 614(7947), 214–216. <https://doi.org/10.1038/d41586-023-00340-6>
- Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*. <https://doi.org/10.1177/20965311231168423>
- Thurzo, A., Strunga, M., Urban, R., Surovková, J., & Afrashtehfar, K. I. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, 13(2). <https://doi.org/10.3390/educsci13020150>
- 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). <https://doi.org/10.1186/s40561-023-00237-x>
- Wu, C., Yin, S., Qi, W., Wang, X., Tang, Z., & Duan, N. (2023). *Visual chatgpt: Talking, drawing and editing with visual foundation models*. *ArXiv Preprint ArXiv*.
- Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13334>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-11742-4>
- Yang, H. (2023). How I use ChatGPT responsibly in my teaching. *Nature*. <https://doi.org/10.1038/d41586-023-01026-9>
- Yinping, Z., & Yongxin, Z. (2023). Research on ChatGPT’s strategy to promote the digital transformation of education. In *2023 26th ACIS international winter conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD-Winter)* (pp. 28–31). <https://doi.org/10.1109/SNPD-Winter57765.2023.10223978>