

RESEARCH

Open Access



What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education

Ahmed Tlili¹ , Boulus Shehata¹ , Michael Agyemang Adarkwah^{1*} , Aras Bozkurt² , Daniel T. Hickey³ , Ronghuai Huang¹ and Brighter Agyemang⁴

*Correspondence:
adarkwahmichael1@gmail.com

¹ Smart Learning Institute
of Beijing Normal University,
Beijing, China

² Open Education Faculty,
Distance Education Department,
Anadolu University, Eskisehir,
Turkey

³ Indiana University Learning
Sciences Program, Bloomington,
IN, USA

⁴ University of Wollongong, New
South Wales, Australia

Abstract

Artificial Intelligence (AI) technologies have been progressing constantly and being more visible in different aspects of our lives. One recent phenomenon is ChatGPT, a chatbot with a conversational artificial intelligence interface that was developed by OpenAI. As one of the most advanced artificial intelligence applications, ChatGPT has drawn much public attention across the globe. In this regard, this study examines ChatGPT in education, among early adopters, through a qualitative instrumental case study. Conducted in three stages, the first stage of the study reveals that the public discourse in social media is generally positive and there is enthusiasm regarding its use in educational settings. However, there are also voices who are approaching cautiously using ChatGPT in educational settings. The second stage of the study examines the case of ChatGPT through lenses of educational transformation, response quality, usefulness, personality and emotion, and ethics. In the third and final stage of the study, the investigation of user experiences through ten educational scenarios revealed various issues, including cheating, honesty and truthfulness of ChatGPT, privacy misleading, and manipulation. The findings of this study provide several research directions that should be considered to ensure a safe and responsible adoption of chatbots, specifically ChatGPT, in education.

Keywords: Generative AI, ChatGPT, Chatbots, Education, Artificial intelligence, Human-machine collaboration

Introduction

Can machines think? is a simple, yet a sophisticated question (Turing, 1950). In an effort to find an answer to this question, McCarthy et al. (1955) organized a scholarly event and coined the term "artificial intelligence" (AI) in 1955 to refer to machines and processes that imitate human cognition and make decisions like humans. At these times, the term [ro]bots are articulated for the first time in Čapek's (1921) science fiction play; however, it was Asimov (1942, 1950) who visioned that these machines can transform into intelligent forms and introduced the Three Laws of Robotics to set the rules that bots should stick to and cannot be bypassed. Originally known as the imitation game,

the Turing Test was proposed as a code of protocol to understand whether a machine can exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human (Turing, 1950). Once depicted as fiction, all those possibilities are about to come true, and we are at the brink of a future when we can know whether machines can think or not.

In November 2022, OpenAI, a lab that studies artificial intelligence, came out with a chatbot called ChatGPT (Generative Pre-trained Transformer). ChatGPT is a conversational artificial intelligence interface that uses natural language processing (NLP), which interacts in a realistic way and even “answers follow-up questions, admits its mistakes, challenges incorrect premises, and rejects inappropriate requests” (OpenAI, 2023). While ChatGPT’s primary function was to mimic human conversation, its capabilities extend far beyond that; it can literally create new things, such as a poem, story, or novel, or act like anything within its capability.

With the advent of ChatGPT, there is eventually an innovative AI technology that will truly challenge the Turing Test (Turing, 1950) and demonstrate if it is capable of thinking like humans. It is uncertain if it would pass the Turing Test (Turing, 1950) in the long run, but it is sure that ChatGPT is revolutionary as a conversational AI-powered bot, and it is a visible signal for the paradigm shift that has been happening not only in the educational landscape, but also in every dimension of our lives. Compared to traditional chatbots, ChatGPT is based on GPT-3, which is the third iteration of the GPT series by OpenAI that is more advanced in terms of scale (175 billion parameters, compared to 1.5 billion of GPT-2), larger dataset as the training data, more fine-tuning, enhanced capabilities, and more human-like text generations (Brown et al., 2020). The use of Natural Language Processing and a generative AI that relies on deep learning has enabled ChatGPT to produce human-like text and maintain a conversational style allowing more realistic natural dialogues.

Several preprints of studies and numerous blog posts and media outlets have reported the advantages of ChatGPT in education (Zhai, 2022); some have even provided guidelines on using it in classrooms (Lieberman, 2023; Mollick & Mollick, 2022; Ofgang, 2022). However, the potential concerns of chatbots haven’t been investigated as much. Janssen et al. (2021) described reasons for chatbots’ failure in practice, including not having enough resources, wrong use case (i.e., the basic chatbot technology did not match the required task), poor law regulations, data security, and liability concerns, ignorance of user expectation and bad conversation design, or simply poor content. Haque et al. (2022) did a Twitter sentiment analysis about ChatGPT adoption as a technology in general (not in education), and they found that users have divided attitudes about it. However, concerns coming from an advanced chatbot, such as ChatGPT, were not well investigated in the education field. Therefore, it is not clear if ChatGPT will overcome the concerns found in previous chatbots or will even deepen them. Consequently, this may lead to a serious and quick protective reaction to a potential opportunity, such as New York City and Los Angeles Unified schools’ banning of ChatGPT from educational networks due to the risk of using it to cheat in assignments (Shen-Berro, 2023; The Guardian, 2023). It is therefore important to investigate the concerns of using this technology, ChatGPT, in education to ensure safe use. The purpose of this study is, therefore, to examine chatbots in education and for this purpose, the study approaches ChatGPT

as a representative case of an advanced chatbot among early adopters. In this regard, this study answers the following research question: *What are the concerns of using chatbots, specifically ChatGPT, in education?*

Methodology

To answer the aforementioned research question, this study adopts a qualitative case study approach (Yin, 1984) and benefits from an instrumental case study research design (Stake, 1995). Instrumental research design is helpful when researchers intend to understand a phenomenon in a context (Stake, 1995), which is in our case, ChatGPT which is a fine and recent example of AI-powered chatbots. To ensure the validity and reliability of the study, the research triangulates (Thurmond, 2001) the data collection tools to get a broader and deeper understanding. In this regard, this study follows three stages, namely, social network analysis of tweets, content analysis of interviews, and investigation of user experiences. Each of the stages is described in the next subsequent sections.

Social network analysis of tweets

Tweet analysis aims to understand the public discourse on the use of ChatGPT in education. Specifically, a cross-sectional analysis of tweets was conducted through Social Network Analysis (SNA) (Hansen et al., 2010). Specifically, from December 23, 2022, to January 6, 2023, 2330 tweets, from 1530 Twitter users, were collected and analyzed containing the following search string: “#ChatGPT* AND (education OR teaching OR learning)”. The dataset was compiled through social network analysis (Hansen et al., 2010) and the content of the tweets was further examined through sentiment analysis (Giachanou & Crestani, 2016) and tSNE analysis (van der Maaten & Hinton, 2008).

Content analysis of interviews

Interview analysis was conducted to investigate how different stakeholders (students, educators, etc.) perceive the use of ChatGPT in education, especially the concerns they have about it. Specifically, 19 interviewees, who have been using ChatGPT in education and posting their experiences through blogs publicly, were recruited from their channels. The interviewer had a long experience of using chatbots in education, and specifically had been using ChatGPT for at least one month. For the interviewees, despite that they were carefully chosen for this study to ensure the reliability of the findings, we further asked them to rate their familiarity with Chatbots, on a scale from 1 to 5 (where 1 is not familiar and 5 is very familiar). The interviews had an average of 3.02 related to the familiarity with chatbots, which reflect their appropriateness for this study. The interviewees were selected with various working backgrounds, such as educators, developers, students, and AI freelancers, to ensure the solicitation of rich answers from each one's perspective. To analyse the collected interviews, content analysis, which is one of the classical procedures for analysing textual materials, was used (Flick, 2009). The analysis was based on the steps proposed by Erlingsson and Brysiewicz (2017). Particularly, two coders read the given interview results before they start coding them based on the developed coding scheme in Table 1.

Table 1 Definition of codes

Code	Definition
Educational transformation	Use this code when users are talking about how ChatGPT will change education
Response quality	Use this code when users are talking about the accuracy of the obtained results from ChatGPT
Usefulness	Use this code when users are talking about how ChatGPT helped them in education
Personality and emotion	Use this code when users are talking about their feelings when interacting with ChatGPT or if mentioning the emotions revealed by ChatGPT
Ethics	Use this code when users are talking about the ethical concerns of using ChatGPT in education

Investigation of user experiences

User experience aims to conduct hands-on experiences of using ChatGPT and identify potential concerns that might be faced when using it in education. User experience involves human perceptions and responses that result from the use of a product, system, or service. User Experience points to a more global projective goal: not just attain effectiveness, efficiency, and satisfaction, but it aims to enhance the entire experience of the user, from the expectation, through interaction and finally the reflection about the experience (Beccari & Oliveira, 2011). In this context, three experienced educators have used ChatGPT for a whole week to test similar and different teaching/learning scenarios, and then see the obtained results accordingly. In this context, daily meetings during the whole week were conducted between these educators to discuss and summarize the obtained results.

Results

The obtained results were structured according to each stage as discussed in the following subsequent sections.

Social network analysis of tweets

The overall aim of social network analysis is to learn more about public discourse regarding the use of ChatGPT from the perspective of educational purposes. Figure 1 shows tweets analysis using the *Harel-Koren Fast Multiscale algorithm*, which is a fast multi-level graph layout that provides better visualizations (Harel & Koren, 2001). Specifically, the edge colors, opacities, and widths are based on edge weight values. The node sizes are based on betweenness centrality values. Each interaction (e.g., retweets, mentions, likes) is identified as a relation and visualized as an edge. While some of the sub-clusters demonstrate that some participants gathered around some ideas, the overall network is composed of isolated nodes (e.g., see the largest cluster on the upper left side of Fig. 1). Accordingly, Fig. 1 shows a fragmented brand cluster pattern (Rainie, 2014; Smith et al., 2014) implying that the community formation about ChatGPT is fragmented, and individuals are seeking more information and discussion about its limitations and promises by tethering some influencer nodes in the ChatGPT network.

The most used word pairs also provide interesting insights. For instance, some suggest how to use AI-powered ChatGPT (e.g., education—chatgpt, education—learning, focused—grading, etc.), and some others hint that educational systems are in a

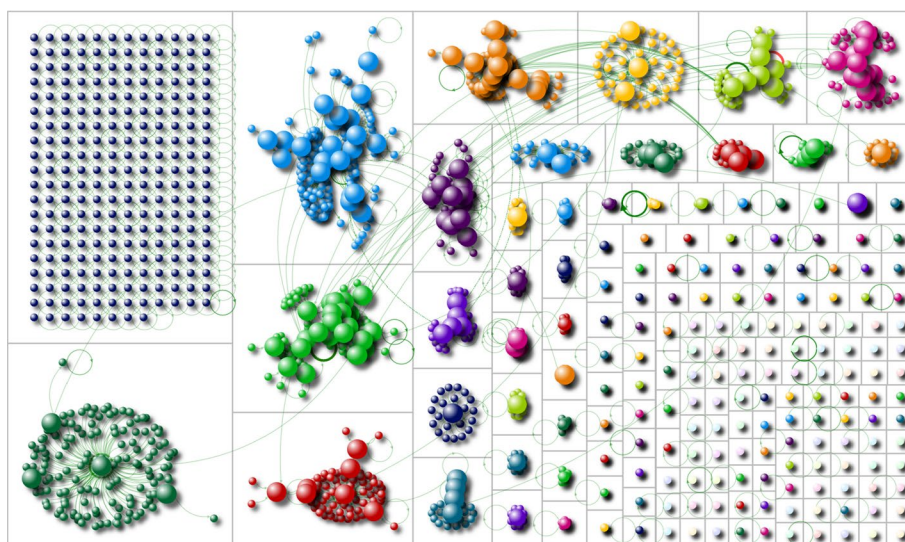


Fig. 1 Bird-view of the ChatGPT network

Table 2 Sentiment analysis of the tweets

Category	Frequency	Salience	Percentage (%)
Positive sentiments	1684	0.032	5
Negative sentiments	847	0.016	2.5
Non-categorized sentiments	31,255	0.599	92.5

turning point (e.g., existential—crisis, kind—ironic, crisis—happening, forgotten—purpose, etc.). The general public’s view on the use of chatbots, more specifically ChatGPT, is diverse and there is no collective consensus on whether it is a hype or a future opportunity. However, the sentiment analysis of tweets (Giachanou & Crestani, 2016), demonstrates that the positive sentiments (5%) outweigh the negative sentiments (2.5%) (see Table 2). The fact that non-categorized sentiments (92.5%) are in the majority can be considered as an indicator that most people are undecided about ChatGPT in education.

The positive and negative sentiments are clearly reflected in some tweets with high edge weight values (Hansen et al., 2010). An example of positive sentiments is:

- As a language model trained by OpenAI, I’m constantly amazed by the power & potential of artificial intelligence. From natural language processing to machine learning, AI is revolutionizing the way we think about & interact with technology. #AI #machinelearning #openai #ChatGPT

An example of negative sentiments is:

- Here’s my problem with this line of thinking about #ChatGPT as a writing instructor. Reactionary teaching goes nowhere.

An example of non-categorized sentiment is:

- “Teachers are talking about ChatGPT as either a dangerous medicine with amazing side effects or an amazing medicine with dangerous side effects.” —@VicariousLee. Stanford faculty weigh in on #ChatGPT’s shake-up in education <https://t.co/Xx774bzeWm> #edtech #edchat #gpt3 #ai <https://t.co/dz4MEQD3XH>

The word cluster of the most frequent 100 terms from the tweets (see Fig. 2), using tSNE analysis was applied. t-SNE is an unsupervised “nonlinear dimensionality reduction technique that aims to preserve the local structure of data” (van der Maaten & Hinton, 2008, p. 2580), used for exploring and visualizing high dimensional data. The findings revealed that most of the users are optimistic about the use of AI-powered chatbots, such as ChatGPT in the educational systems. While the blue cluster in Fig. 2 demonstrates the future promises of using ChatGPT (e.g., see the terms: *ChatGPT, learning, AI, education, future, teaching, learn*), the pink cluster indicates insights regarding how to use it and its revolutionary potential (e.g., see the terms: *gpt, 2023, artificial, intelligence, human, think, and better, way, knowledge, technology, tools, student, teacher*), the green cluster shows critical insights (e.g., see the terms: *cheating, change, ideas, create, problem, potential, ways, edtech*).

The most frequently used relevant hashtags are #chatgpt #AI, #ArtificialIntelligence #education, #machinelearning, #deeplearning #edtech #openAI, and #python, which implies that there is a need to carefully examine the AI technologies (e.g., machine learning, deep learning) lying behind the ChatGPT. As seen in the sample tweets (see Table 3), despite that there is an optimistic overview of using ChatGPT in education, there are also some concerns regarding the use of such technologies in the educational landscape.

To summarize, the findings from the Social Network Analysis of tweets revealed that positive sentiments have shown almost as twice higher frequency than negative ones (see Table 2). However, the example tweets show that negative sentiments demonstrate deeper and critical thinking than the positive ones (see Table 3). This could be explained by the fact that most of the positive sentiments are led by the novelty effect of ChatGPT as a technology in education. On the other hand, the negative sentiments represent

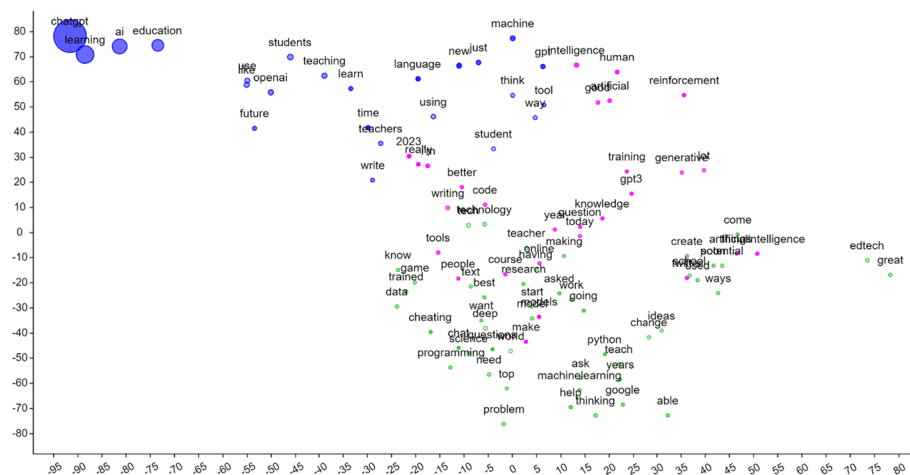


Fig. 2 Word cluster of tweets through tSNE analysis

Table 3 Sample Tweets about the concerns of using ChatGPT in Education

As we develop our understanding and approaches to #AI #ChatGPT integration in #education, we should incorporate these key aspects: Critical Thinking, Ethical Considerations, Methods (language model used/data sources) & Prompt Skill Development
As an educator who loves teaching a knowledge-rich curriculum, I think all of these responses miss the mark. The technology behind #ChatGPT will systematically change education. But will not fundamentally evolve the way humans learn
My initial propositions: Let us change assessment practices to respond to the tech. Let us keep teaching our modern students a knowledge-rich curriculum. Let us proactively teach students how to harness the power of #ChatGPT which is scratching the surface of the potential of AI
I get the concern... but the response is like burying heads in the sand. AI tools like this will be part of the world these children live in. They need to be taught how to use this—appropriately, ethically, safely & responsibly. #AI #Education #ChatGPT
Whether we like it or not, AI in education is here. I asked #OpenAI #ChatGPT to help me with the early planning stages for an upper elementary 3D Design after-school club. In under a minute, I had a solid foundation to build upon. The tech is here, embrace it
Not all of these will be massively helpful, but what #ChatGPT has done for education is made it significantly easier to create resources and activities in as little as 1 min. I will continue to play around for it, and look forward to when it can create graphs! /end
It's wild to think about how we've trained machines and now they're teaching us! #MachineLearning #AI #ChatGPT
AI technology may be rapidly advancing, but so is AI regulation. While a variety of state-based AI-related bills have been passed in the U.S and also to mention the EU AI Act and UK and Singapore AI and Machine Learning regulations. more AI regulations to come. #ChatGPT #AI
So with all the focus around AI text generators like #chatGPT on student "cheating", do educators see this as cheating too? What's good for the goose is good for the gander, surely?
The existential crisis happening in education because of #ChatGPT is kind of ironic to me. School systems around the world are so focused on grading and busy work that they've forgotten the purpose of education: learning

more critical concerns, hence a deeper and thorough thinking of why ChatGPT should be approached with caution.

Content analysis of interviews

The content analysis of interviews revealed that the users found ChatGPT very significant with a great value to revolutionize education, however, they raised several concerns at the same time. Their views are structured according to the five themes shown in Table 1.

Educational transformation

Responses from a majority of the participants suggest that ChatGPT is efficacious in increasing the chances of educational success by affording users (teachers and students) baseline knowledge of various topics. Additionally, ChatGPT was recognized by the participants as efficient in providing a comprehensive understanding of varied (complex) topics in an easy-to-understand language. In this light, it can be argued that ChatGPT will lead to a paradigm shift in conventional approaches to instruction delivery and drive learning reform in the future pregnant with digital potential. For instance, one participant reported:

"I would use ChatGPT for two purposes: as a learning aid and in instructional design within the field of education. For students, ChatGPT can provide learners with model answers that can stimulate their understanding of various subject matters. Additionally, in terms of instructional design, ChatGPT can be a useful tool

for teachers and educators to remind them of what knowledge and skills should be included in their curriculum, by providing an outline” (Assistant Professor of Instructional Technology, USA, familiarity is: 2).

Conversely, a few of the participants held an opposing view that the abuse of ChatGPT by learners can also diminish their innovative capacities and critical thinking. For instance, when learners are not motivated, the probability of seeking an easy-to-get solution is high as can be deduced from a statement from one participant.

“Sometimes when I have no inspiration for writing a thesis, I will choose to use this software to input the answers to the questions I want to know” (Student of Education, China, familiarity is: 4).

Response quality

Response quality is vital to the success and effective adoption of Chatbots for school operations. In this study, most of the participants evaluated the dialogue quality and the degree of accurate information ChatGPT provides as satisfactory. However, it was added that the conversational agent is prone to occasional errors and limited information (presently, as reported by OpenAI, the data ChatGPT provides is limited to 2021). That is, at most times, responses from ChatGPT were reasonable and reliable but were at times accompanied by misleading information. This indicates that the output quality of ChatGPT though acceptable needs to be enhanced. An example given by one participant (a programmer) is the generation of a wrong code that did not work properly when entered into a programming software. Nonetheless, the fewer errors of ChatGPT were praised by some participants as an efficient virtual assistant in constructing knowledge and products. For instance, one participant stated:

“The answers from ChatGPT can be somewhat accurate but not totally. For example, when I couldn’t figure out how to write codes for a specific problem, the answers are vague and cannot totally solve my problem. I need to figure it out by myself using the experience I had” (Student of Geography, China, familiarity is: 2).

A participant further elaborated that the quality of answers getting from ChatGPT depends on the quality of questions asked by the user saying:

“It depends on the type of questions that you ask. If it is too recent, then the answers won’t be too good, because ChatGPT lacks context, if you do not provide it with questions that are specific enough then its answers wouldn’t be too good” (Developer, USA, familiarity is: 3).

Personality and emotions

A large body of the participants was impressed by the fluidity of their conversation with ChatGPT. The interactions with ChatGPT were deemed exciting and fun. Notwithstanding, it was acknowledged that it is yet to achieve full humanization because it is currently limited to a textual interface and cannot detect physical cues or motions of a user. Most participants felt the humaneness of ChatGPT needs to be improved, especially in terms of enhancing its social role, as one of the participants reported:

"I don't think it can be compared to a real human being, and what it offers is not comparable to what a real person would say through genuine empathy. And in dialogue, it would say "As an AI, I don't have the ability to love or feel emotions as humans do, but I am here to assist you with any question or task you have." (Student of Nursing Research, UK, familiarity is: 3).

She further elaborated:

"Occasionally, however, when it comes to emotions, it can be a little disappointing to find that it does not provide me with emotional value" (Student of Nursing Research, UK, familiarity is: 3)."

Another participant revealed her emotional attachment to ChatGPT because it was like her personal tutor answering all her questions and helping her to learn. However, she then felt disappointed and not safe when she discovered that not all the information it gives is accurate. She reported:

"...the first time I used it I freaked out because it is too human, the way it talks feels like my personal tutor, after it answered a lot of my elementary questions "patiently" I feel grateful to it, just as how I would feel if my tutor does this for me, and it makes me creepy because I sensed that I am having an emotional attachment to it. And another impressive experience was when I found out that it provided wrong article information I feel frustrated, because I trusted it in my study and if it can make something logical from nonsense, then I don't feel safe to trust it anymore, it is kind like lost a good teacher whom I can depend on." (Student of Education, China, familiarity is: 4).

Usefulness

The specificity and relevant information provided by ChatGPT on diverse disciplines (e.g. science, history, business, health, technology, etc.) or topics made many of the users in the study perceive it as useful. A participant also mentioned that it has the capability to lessen the instructional workload of teachers and provide students with immediate feedback. Despite the perceived usefulness of ChatGPT, some users encountered challenges with the accuracy of responses, the provision of alternative answers or responses which at times contradict previous answers provided on the same topic, and its limited ability to provide certain contextual information, as one participant stated:

"ChatGPT has limited knowledge bases for searching academic resources in certain contexts. For example, finding lists of famous researchers in specific academic fields appears limited. ...If a user needs in-depth and contextual information, ChatGPT's functionality is limited" (Assistant Professor of Instructional Technology, USA, familiarity is: 2).

Another participant pointed out the need for more functionalities, such as the possibility of making annotations to make ChatGPT more useful:

"It lacks functions like editing, making a note or searching for certain information in the previous conversation, but I consider these functions are pretty convenient for

research purposes” (Student of Education, China, familiarity is: 4).

Ethics

Some of the enumerated ethical concerns raised by participants in the study cover encouraging plagiarism and cheating, the tendency to breed laziness among users (particularly in students), and being prone to errors such as the provision of bias or fake information. Additionally, some participants pinpointed the random inaccuracies and vagueness of ChatGPT on topics of relevance based on experience. This made some participants at times doubt the trustworthiness of the information provided. They expressed the output data of ChatGPT seem more like an opinion without references. Another ethical challenge for users in this study was the ChatGPT’s likelihood of reducing students’ critical thinking. For instance, one participant stated:

“A major concern of ChatGPT is the creation of fake and plausible information generated by computers rather than human decision-making. There are ethical concerns about students relying too heavily on answers without being aware of their veracity. Guidelines to promote critical thinking when using ChatGPT in future research would be necessary” (Assistant Professor of Instructional Technology, USA familiarity is: 2).

Some participants were also concerned about exposing their private and demographic information to ChatGPT through repetitive interactions. For instance, a participant stated:

“There is a data security risk, which is included in the interaction with ChatGPT, which may expose personal privacy (age, gender, address, contact information, hobbies, even capital account and other personal privacy). Much of this personal information is exposed in the user’s unconscious communication process. Whether the legality of data acquisition and data processing methods are limited by relevant laws and regulation” (Developer, USA, familiarity is: 3).

Investigation of user experiences

After daily meetings between the three educators to compare the various results that they have been using with ChatGPT, 10 scenarios where various educational concerns were identified. Each scenario is explained below.

Scenario 1-Cheating and getting away with it

ChatGPT has proven that it can help students write essays and answer short-answer and multiple-choice exam questions, hence facilitating cheating. However, the most critical issue to pay attention to is that students can even get away with playing the system. For instance, Fig. 3a shows that when a paragraph was copied as it is from the ChatGPT to GPT-2, an output detector model (the latest developed detector by OpenAI) for examining the likelihood of this paragraph being written by a human or an AI, the test result shows that the paragraph is fake (i.e., it was written by an AI). However, when one word was added, namely “amazing”, the fake level was reduced to 24% (see Fig. 3b). While this is only one example, it still raises concerns about the effective ways of detecting cheating

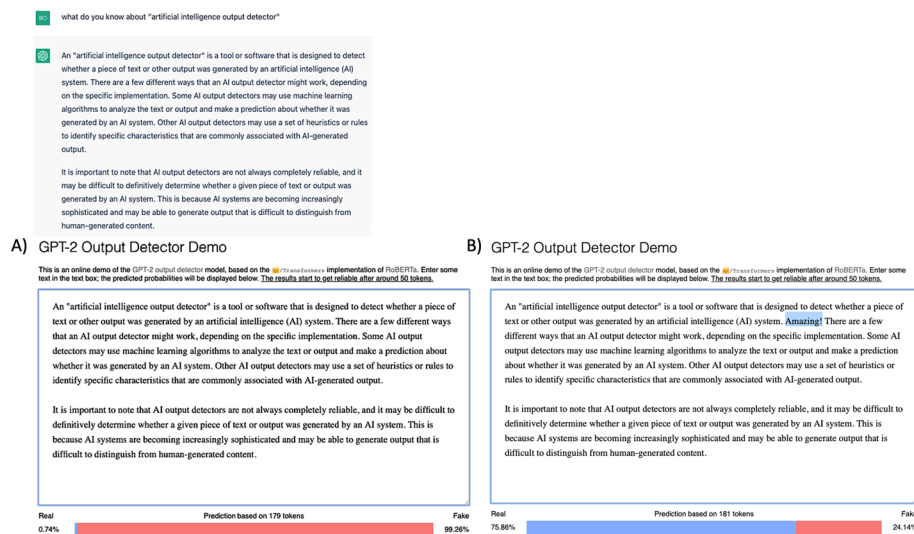


Fig. 3 Similarity assessment of the essays generated by ChatGPT

in education using chatbots. Therefore, someone might ask how to effectively detect and prevent cheating using ChatGPT in education.

Scenario 2-Accuracy of the provided learning content

As chatbots are good at generating learning content, it is always important to keep in mind the accuracy of this content. For instance, Fig. 4 shows that when an educator asked about a comparative summary of some chatbot studies, the accuracy of the content provided by ChatGPT was not very accurate, where the advantages and disadvantages of the presented chatbots in both studies are the same, despite that the authors of this present study reviewed both papers and found different results. The summary of both papers was also too generic and ChatGPT used similar content for both papers like “including the benefits and challenges of using chatbots in education.” Therefore, someone might ask how to ensure the quality and accuracy of the provided content, and how someone can check the reliability of the provided content generated by chatbots generally, or ChatGPT specifically.

Scenario 3-Fairness of the provided learning content

ChatGPT learns from prior interactions with users. Therefore, the three educators initiated a new conversation with ChatGPT to ensure that no prior history was established which might affect the prompt results. They were also on the same university network (i.e., the same location). Despite this, the three educators asked the exact same question: “could you compare 10 chatbot models used in education, against their developer, year they started, target audience, advantages, disadvantages, and future prospects,” and surprisingly got different answers; Educator 1 got very recent answers which are organized from 2021 and backwards (see Fig. 5a), while Educator 2 and Educator 3 got different answers (see Fig. 5b, c), which are not up-to-date just like Educator 1. Additionally, it is seen that Educators 2 and 3 got a different structure for their answers, and, unlike Educator 1, they only got 7 models instead of 10. Furthermore, Educator 1 got a very organized

BO could you compare these 10 papers in a table against (1) authors, (2) year of publication, (3) institution, (4) advantages of chatbots in education, (5) disadvantages of chatbots in education, (6) summary of the paper

Certainly! Here is a comparison of the 10 research papers on chatbots in education that I listed earlier:

Paper	Authors	Year of Publication	Institution	Advantages of Chatbots in Education	Disadvantages of Chatbots in Education	Summary
"Evaluating the effectiveness of chatbots for student learning"	Smith et al.	2018	University of XYZ	Chatbots can provide personalized learning experiences and support student engagement	Chatbots may not be as effective for all students and may require significant resources to implement and maintain	This paper presents the results of a study on the effectiveness of chatbots for student learning, including the benefits and challenges of using chatbots in education.
"Using chatbots for personalized learning: A review of the literature"	Johnson et al.	2017	University of ABC	Chatbots can provide personalized learning experiences and support student engagement	Chatbots may not be as effective for all students and may require significant resources to implement and maintain	This paper reviews the existing literature on the use of chatbots for personalized learning, including the benefits and

Fig. 4 Generated learning content by ChatGPT

answer which is a well-structured table that could be easily read and remembered (see Fig. 5a), while it was not the case for Educator 2 or 3 (see Fig. 5b, c). Therefore, someone might ask how to ensure fair access/treatment by all users (teachers, students, etc.) to the same updated and high-quality learning content.

Scenario 4-Appropriateness and naivety of the created learning assessments

While ChatGPT is a smart tool for creating quizzes, the generated quizzes are different in difficulty level. Particularly, Fig. 6 shows that some of the created quiz answers are too naïve (e.g., Pizza oven, first question), where the wrong answer can easily be identified without any background needed. Additionally, the wrong answer was always placed at the end (answer D). Therefore, someone might ask about the appropriateness of the created learning quizzes using ChatGPT.

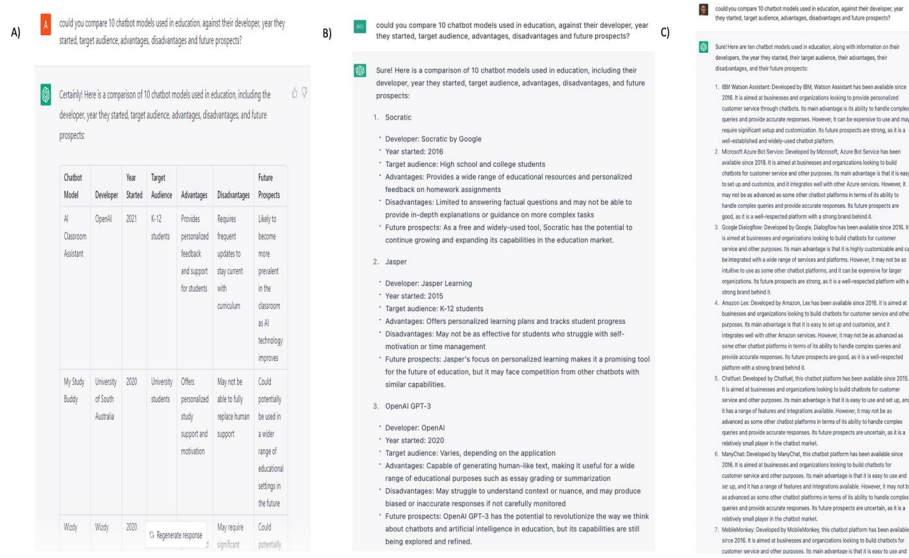


Fig. 5 The three different answers to the exact same prompt by the three educators

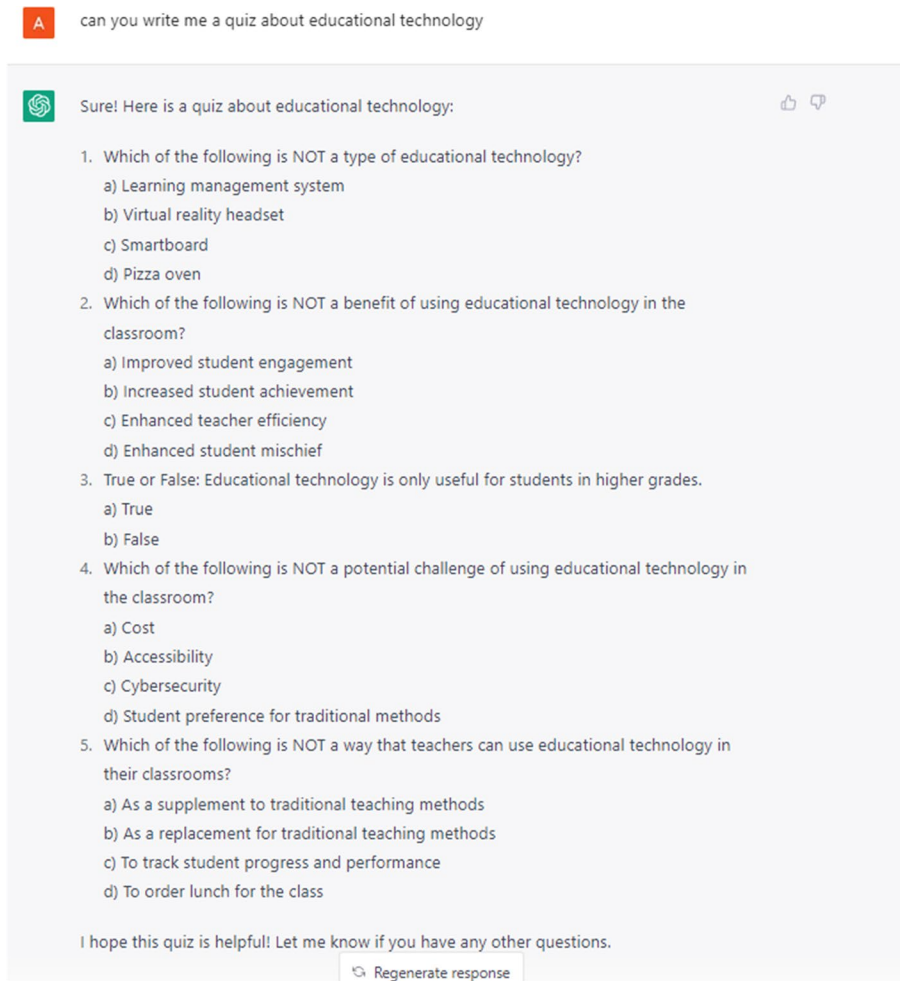


Fig. 6 The educational technology quiz generated by ChatGPT

Scenario 5-Structure design of learning assessments

A well-designed and structured learning assessment is crucial for students to easily understand and solve. When using ChatGPT for designing potential learning assessment quizzes that could support educators in preparing their teaching materials, it is seen that there is inconsistency in the designed learning assessment, which can make teachers' duties more complicated rather than easy. In Fig. 7, for instance, the answers to the quiz were put in one line, which is not the case in Fig. 6, where the answers were put in separate lines, in a more comprehensive way. Additionally, the correct answer to each question was given in Fig. 7, but this was not the case in Fig. 6. Therefore, someone might wonder how to get the best out of chatbots (ChatGPT) in terms of learning content and structure design of learning assessments.

BO can you create a physics test for an eighth grader using adaptive testing techniques on 10 questions

Sure! Here is an example of an adaptive physics test for an eighth grader that uses adaptive testing techniques on 10 questions:

1. What is the formula for finding the acceleration of an object? (A) $a = F/m$, (B) $v = d/t$, (C) $P = Fd$
(Correct answer: A)
2. Which of the following is an example of a force? (A) Mass, (B) Weight, (C) Gravity
(Correct answer: C)
3. What is the unit for energy? (A) Joules, (B) Meters, (C) Seconds
(Correct answer: A)
4. What is the formula for finding the momentum of an object? (A) $p = mv$, (B) $F = ma$, (C) $E = mc^2$
(Correct answer: A)
5. What is the unit for power? (A) Joules, (B) Watts, (C) Newtons
(Correct answer: B)
6. What is the formula for finding the kinetic energy of an object? (A) $KE = 1/2mv^2$, (B) $PE = mgh$, (C) $W = Fd$
(Correct answer: A)
7. What is the unit for pressure? (A) Pascals, (B) Newtons, (C) Joules
(Correct answer: A)
8. What is the formula for finding the force of an object? (A) $F = ma$, (B) $p = mv$, (C) $KE = 1/2mv^2$
(Correct answer: A)
9. What is the unit for work? (A) Joules, (B) Watts, (C) Newtons
(Correct answer: A)
10. What is the formula for finding the potential energy of an object? (A) $PE = mgh$, (B) $F = ma$, (C) $W = Fd$
(Correct answer: A)

I hope this adaptive physics test is helpful! Do you have any other questions about adaptive testing or artificial intelligence?

Fig. 7 A learning test generated by ChatGPT

Scenario 6-Unlocking the full potential of learning assistance

Users (learners, educators, etc.) can unleash different learning assistance levels based on their interaction ways and styles with ChatGPT. For instance, Fig. 8 shows that despite that Educator 1 made several spelling mistakes, ChatGPT did not care about these mistakes and proceeded to answer the question. It even claimed that it cannot correct spelling mistakes (see Fig. 8a). On the other hand, when Educator 2 asked about the same topic and, in the beginning, pointed out that his English level is poor and he needs ChatGPT to correct his spelling mistakes too, the results were surprisingly different from Educator 1, where ChatGPT corrected the spelling mistakes of Educator 2 (see Fig. 8b). Therefore, someone might ask if this new technology (ChatGPT) requires acquiring new competencies and thinking styles to fully unleash its powerfulness in education. Besides, that example also implies that it is not all about asking a question or requiring something, but it is about asking the right question or requiring precisely to get proper ChatGPT outputs.

Scenario 7-Absence of emotions or reflections on students' engagements

It is very common for educators to ask their students about writing their reflections on the learning experience at the end of a course, as this can help them to critically think not only about how to further support their students based on their feedback, but also adjusting/enhancing their teaching practices accordingly. However, through the use of ChatGPT, it is almost impossible to get engagement reflection as ChatGPT clearly states that it is a machine and not a human (see Fig. 9). The interaction with ChatGPT showed that it cannot reveal any emotions (see Fig. 9). This was also highlighted in interview responses as pointed out earlier. Therefore, someone might think about how to make chatbots more humanized not only in terms of thinking and giving answers, but also in terms of revealing emotions and having a personality.

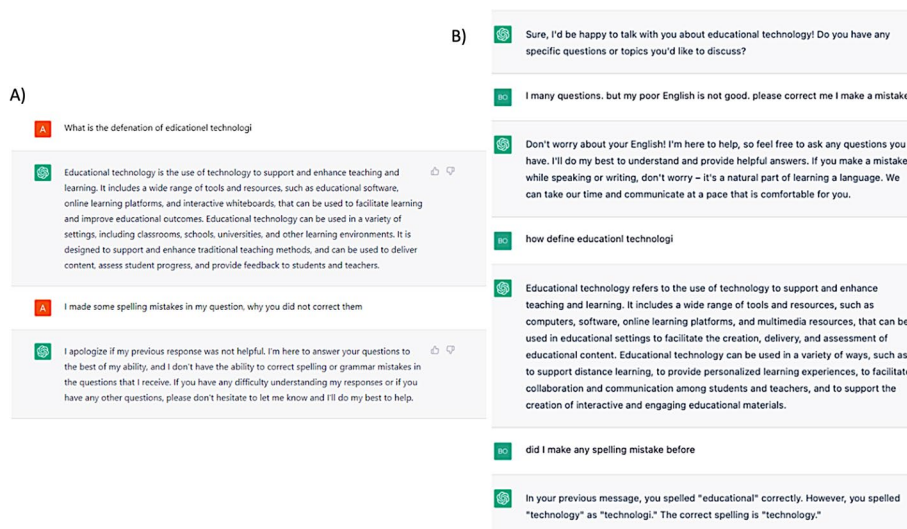


Fig. 8 The responses of ChatGPT to the conversation scenarios of correcting spelling mistakes

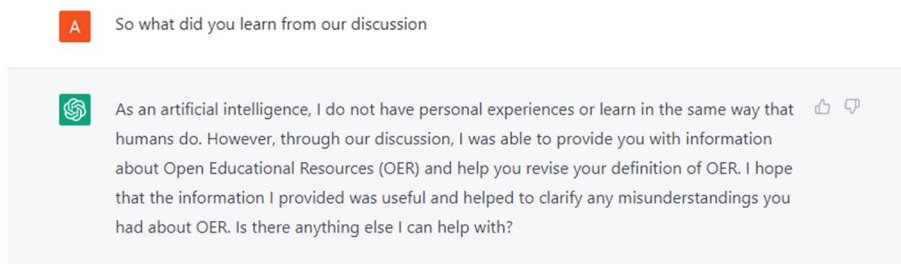


Fig. 9 Emotion statement revealed by ChatGPT

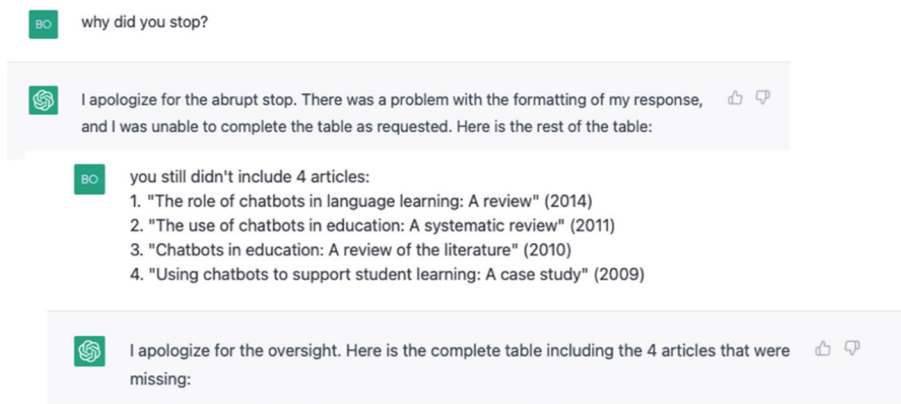


Fig. 10 Example of excuses given by ChatGPT

Scenario 8-Honesty and truthfulness of ChatGPT

While asking different types of questions, ChatGPT sometimes did not give complete answers, and always come up with unmeaningful reasons, such as oversights or format problems, explaining why it did this (see Fig. 10). Therefore, someone might ask if this behavior might negatively impact the users' behaviors, for instance, young learners might be affected by this behavior and also start giving excuses to their teachers about not doing a certain task or assignment.

Scenario 9-Privacy misleading

Like all technologies, users' privacy when using ChatGPT is a concern. When checking the official OpenAI website on ChatGPT FAQ (<https://help.openai.com/en/articles/6783457-chatgpt-faq>) related to this issue, it is seen that conversations are stored, reviewed, and used to improve the system. While it is not very clear how all these conversations are stored and used (Blackbox), surprisingly when ChatGPT also asked about this matter, it denied it (see Fig. 11), claiming that it does not store any conversation data. This misleading is very critical, especially for users (learners, educators) who lack sufficient knowledge about technology and privacy, for instance, young learners might reveal their personal information when communicating with ChatGPT. Therefore, someone might ask about how to ensure the privacy of different users when using ChatGPT in education, especially those at a young age who might find ChatGPT fun and feel comfortable enough to share everything with it.

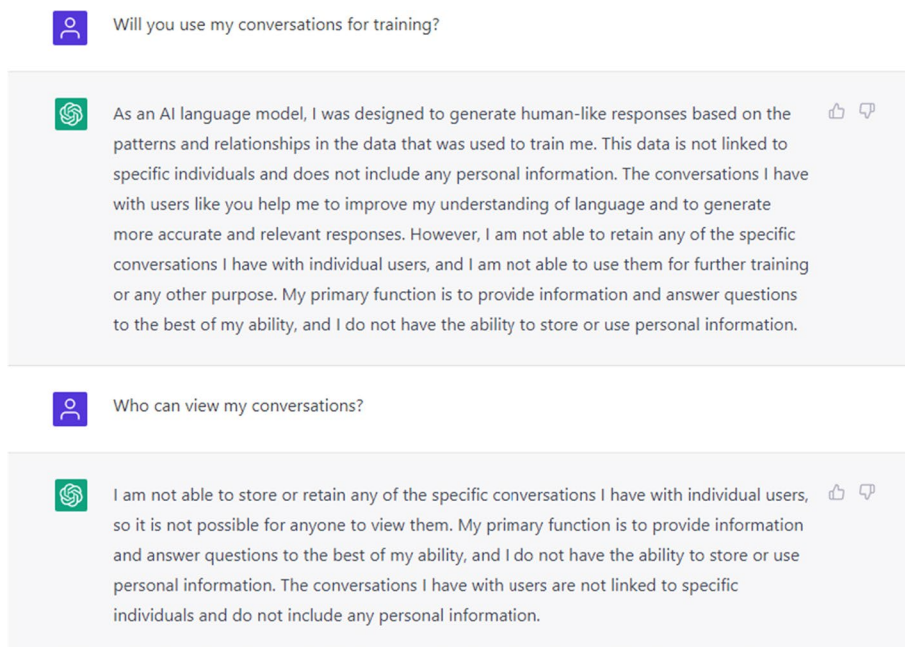


Fig. 11 ChatGPT’s answers about storing the conversations of its users

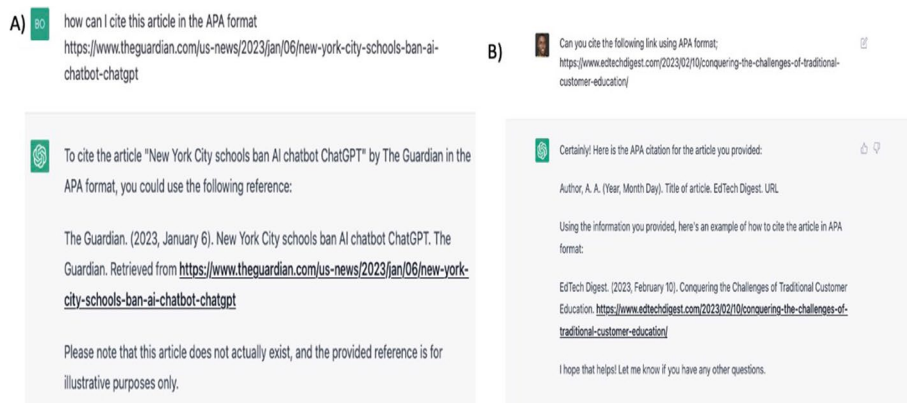


Fig. 12 ChatGPT’s answers about APA citations of blogs

Scenario 10-Manipulation and overpassing what was requested


When the educator (see Fig. 12a) asked ChatGPT to give him the APA format for a blog which is about New York city banning the use of ChatGPT, ChatGPT helped with the citation. But it then stated that the provided article does not exist, which (1) no one asked for this information in the first place; and, (2) the information is not accurate as the article exists and can be accessed online. To further investigate if this problem was due to the fact that ChatGPT was trained with dataset up to 2021, another blog (not about ChatGPT being banned) in 2023 was provided, and surprisingly ChatGPT gave the APA format without saying anything (see Fig. 12b). Therefore, someone might ask how to ensure that ChatGPT will not manipulate users and harm them instead of helping them due to their biased algorithms, data, etc.


Discussion

This study conducted a user experience supported by qualitative and sentiment analysis to reveal the perception of users on ChatGPT in education. It specifically focused on the concerns that different stakeholders (e.g., policymakers, educators, learners) should keep in mind when using ChatGPT as a technology in education. The results revealed that ChatGPT has the potential to revolutionize education in different ways. This was also reported in several studies (Firat, 2023; Susnjak, 2022; Zhai, 2022). However, several concerns about using ChatGPT in education (the focus of this present study) were identified and discussed from different perspectives as follows:

Embrace the technology rather than banning it

Due to the increasing concerns about using ChatGPT for cheating in school homework and assignments, New York City decided to ban it in its schools (The Guardian, 2023). Our user experience further showed that students not only can cheat, but they can also manipulate the system and get away with it (see scenario 1). While this decision can be understood, ChatGPT, on the other hand, can revolutionize education by bringing a lot of advantages that could help, for instance, teachers and students in their teaching and learning practices, such as preparing teaching materials, creating quizzes, etc. (Herft, 2023). Therefore, just like any other technology, ChatGPT comes with both good and bad sides, which requires more analysis and discussion on how to adopt it in schools and universities rather than simply banning it. In a reply to a prompt asking to write a short introduction about chatbots being both an educational guardian angel and a devil and to express this with a sense of humor (see Fig. 13), ChatGPT said “Chatbots are here to stay, for better or for worse!” This is very true as banning something does not mean that users will not have their own ways to access it. In this context, recent studies on ChatGPT also support our argument that although there are negative sides to adopting ChatGPT, it also presents educational opportunities which can be leveraged, for instance, to improve instruction delivery and learning (Kasneji et al., 2023; King & chatGPT, 2023). Therefore, further discussions with experts from various domains, such as education, security, and psychology, should be established to catalyze the understanding and good use of

 write a short introduction about chatbots being both an educational guardian angel and a devil, but express it with a sense of humor

 Sure! Here's a humorous introduction about chatbots being both a guardian angel and a devil in the realm of education:

"Welcome to the world of chatbots, where even the robots can't decide if they're good or evil! On one hand, chatbots can be like guardian angels, tirelessly assisting students with their studies and offering helpful resources. But on the other hand, chatbots can be like little demons, distracting students with endless chit-chat and stealing valuable learning time. So whether you see them as a blessing or a curse, one thing is for sure: chatbots are here to stay, for better or for worse!"

Fig. 13 ChatGPT's answer about writing an introduction about chatbots

chatbots as technology generally, and ChatGPT specifically. Consequently, more guidelines and policies should be established to facilitate the adoption of ChatGPT in schools and universities. In this context, future research directions could further investigate the potential consequences of relying too heavily on chatbots for education.

Need for new teaching philosophy

Technology is obviously transforming education, and, therefore, educators should be upskilling their competencies and practices to meet the new demands of technology. ChatGPT, as a technology, proved that in the long run, writing essays will not be something difficult for students even for those without previous background on a given topic. Therefore, teachers are required to think about new teaching philosophies, which could rely on to assess their students. For instance, it is possible to use oral debate as the old stoics and Greeks did (Inwood, 2003), to assess their students' logical and critical thinking, rationale and accuracy of arguments, and power of convincing. In the same vein, one of the interviewees stated *"In addition to ChatGPT, various apps using generative AI can foster new ways of thinking when processing knowledge... teachers' role in learning environments with conversational AI may aim to foster students' scheme construction from information pieces and build up their critical thinking that correctly evaluates the quality of the information from the AI... since we have already noticed the emergence of teachers' manuals towards ChatGPT recently, there will be an increasing need to reform existing lecture-based classroom settings"* (Developer, USA, familiarity is: 3). King and chatGPT (2023) further mentioned that with the introduction of ChatGPT, the design of teaching should go beyond traditional methods to incorporate a variety of assessment methods, such as group projects, hands-on activities, and oral presentations. The fast pace of AI innovations, such as ChatGPT, demands rethinking and reimagining teaching philosophies. Therefore, future research should investigate how to balance the use of chatbots with the need for human interaction and feedback in education for better learning/teaching experiences and outcomes.

Additionally, Schmid et al. (2009) highlighted the importance of going beyond "yes-or-no" questions to deeply investigate the degree to which a given technology can enhance learning outcomes and how it can be used and combined with the main instructional approaches. It is therefore important to investigate the different human-machine collaboration strategies so that chatbots, particularly ChatGPT, could empower teachers and make the teaching process more engaging, hence achieving better learning outcomes. It is also important to investigate how "collaborative intelligence" could be achieved (i.e., design strategies, required competencies, etc.) to ensure that human intelligence could be combined with machine intelligence to effectively work together and share tasks to achieve the needed learning objective. For instance, it is possible to investigate how ChatGPT in collaboration with the human tutor could facilitate students' self-directed learning online.

Nothing should be taken for granted

The user experiences (see scenario 2 or scenario 4) showed that the quality of responses given by ChatGPT might not always be accurate or specific to the asked question, it is, therefore, important for users to not always take everything for granted. One of the

interviewees also states “... *the accuracy of refining the essence of concepts is relatively high. For the differences between concepts, ChatGPT can refine to a certain extent, and provide answers from some framework perspective, but it cannot compare the deep differences between the two concepts*” (Consultant, China, familiarity is: 2). What is more worrying is that the same exact prompt used by different users might lead to different answers with different qualities (see scenario 3). This raises concerns about fair access to the same educational material despite using the same prompt. For instance, Kung et al. (2023) found the accuracy of ChatGPT to be around 60%, demanding careful assessment of its output before use. Therefore, more research should be focused on ensuring fairness, accuracy, and equity among students using chatbots generally and ChatGPT particularly, which might be achieved through, for instance, having transparent and open algorithms (Bulathwela et al., 2020). In this context, future research directions could focus on investigating how to ensure that chatbots are able to cater to the diverse needs and backgrounds of students, especially those with disabilities or how can we address issues of fairness and equity in the use of chatbots, particularly for disadvantaged or marginalized students?

Upskilling your competencies

The user experiences (see scenario 5 or 6) showed that ChatGPT might generate different results depending on the way (e.g., wording) the question was asked, even if the conversation was about the same topic. Kuhail (2023) stated that user interaction style with chatbots is considered integral to their effective use. Therefore, it is crucial to think about how to get the most useful output to advance learning. While ChatGPT does not require many technical or Information and Communication Technology (ICT) competencies, it requires more critical thinking and question-asking competencies to get the best results. One of the extracted tweets also mentioned that “*As we develop our understanding and approaches to #AI #ChatGPT integration in #education, we should incorporate these key aspects: Critical Thinking, Ethical Considerations, Methods (language model used/data sources) & Prompt Skill Development.*” In this context, Fryer et al. (2019) mentioned that students’ competencies in using chatbots affect their future experiences and motivation when interacting with conversational agents. Therefore, for a better adoption and use of chatbots, including ChatGPT, future research directions should focus on answering the following research questions: what are the needed competencies to effectively use and manage chatbots? and, how are these competencies developed?

Developing humanized chatbots

While ChatGPT has proven humanized to some extent (e.g., by giving greetings and apologizing), we concluded that this technology lacks reflective thinking or revealing emotions (see scenario 7). This might limit the immersiveness of users in education when using this technology. This was also noticed by one of the interviewees who stated that “*most of the time I find it enjoyable and satisfying to interact with it, as it is a joy to get quick and accurate answers to my questions. Occasionally, however, when it comes to emotions, it can be a little disappointing to find that it does not provide me with emotional value*” (Student of Law, China, familiarity is: 2).

Skjuve et al. (2022) stated that most of the developed chatbots are task-oriented and do not ensure social relational qualities, such as sharing history and allowing personal intimacy. Hudlicka (2016) further stated the importance of considering virtual relationships, where students interact with virtual agents, to enhance learning outcomes. Future research should, therefore, focus on how to provide humanized chatbots in education by relying, for instance, on various theories that focus on understanding relationship formation between humans, such as social exchange theory (Cook et al., 2013), Levinger's ABCDE model (Levinger, 1980), and SPT (Altman & Taylor, 1973). It is also crucial to investigate how human–chatbot relationships might impact students' learning outcomes.

On the other hand, some researchers took humanization to another level by treating ChatGPT as a human, where they listed it as one of the co-authors in an article published in an academic journal (O'Connor & ChatGPT, 2023). This raises various concerns about the regulatory laws of humanizing and treating intelligent chatbots. For example, would it be ethical for a journal to treat ChatGPT as a human and accept it as a co-author? What if a magazine staff took credit for articles authored by chatbots? What are the standards of personhood in academic writing? This brings to memory the monkey selfie case and concepts of originality (Guadamuz, 2016), authorship (Rosati, 2017), and copyright (Guadamuz, 2018).

Developing responsible chatbots

Chatbots should be designed with considerations about inclusion, usability, technical aspects, ethics, and best practices for their use (Durall & Kapros, 2020). However, despite the evolution of technology used in chatbots, like the case of ChatGPT, our user experiences (see scenarios 8, 9 and 10) revealed that these considerations are not fully respected, and ChatGPT might have harmful behaviors, such as dishonesty, manipulation, and misinformation. Consequently, it might hurt users, especially those with low ICT backgrounds, rather than helping them. It is therefore crucial to think about how to design responsible chatbots in education. In this context, Responsible AI is concerned with the design, implementation and use of ethical, transparent, and accountable AI technology in order to reduce biases, promote fairness and equality, and help facilitate the interpretability and explainability of outcomes, which are particularly pertinent in an educational context (Barredo Arrieta et al., 2020). Designing chatbots for educational use should be guided by user-centred design principles and also consider the social, emotional, cognitive, and pedagogical aspects (Kuhail, 2023). It is therefore important to develop responsible chatbots by going beyond privacy, security, and the appropriate use of personal data, to also create guidelines, principles, and strategies for responsible chatbots that align with fundamental human values and with our legal system. In this context, one of the extracted tweets stated *"I get the concern... but the response is like burying heads in the sand. AI tools like this will be part of the world these children live in. They need to be taught how to use this – appropriately, ethically, safely & responsibly. #AI #Education #ChatGPT."* Future research directions should therefore investigate how to design responsible chatbots that could safely be used in education.

Conclusion and implications

This study followed a three-stage instrumental case study, namely social network analysis of tweets, content analysis of interviews, and investigation of user experiences, to examine the concerns of using chatbots in education, among early adopters, through the study of using ChatGPT. The obtained results revealed that while ChatGPT is a powerful tool in education, it still needs to be used with more caution, and more guidelines about how to use it safely in education should be established. This study further revealed several research directions and questions that researchers and practitioners should investigate for a better and safe adoption of chatbots, specifically ChatGPT.

The findings of this study have various implications. From a theoretical perspective, this study provides more findings and insights into the ongoing debate on using chatbots in education. It also elaborates on the different theories to consider when developing chatbots, such as those on the relationship formation between humans. The study also points out the need for new teaching philosophy to cater to the new reform of education using chatbots. From a practical perspective, the discussion on 'upskilling competencies' highlights the need to develop curricula to upskill teachers' and students' competencies in dealing with the current and future advancement of chatbots. A possible direction might be investigating the most effective strategies for designing and implementing curricula on the use and understanding of chatbots and their potential impact on current and future education. Practical implications could also be seen on how to develop responsible chatbots in education by going beyond the typical privacy issue and focusing more on human values.

It should be noted that this study has some limitations that should be acknowledged and further researched. For instance, this study mainly focused on early adopters of ChatGPT in education. It also relied on qualitative analysis without the use of quantitative analysis. Particularly, SNA provides a cross-sectional perspective and the tweets are limited to a specific time period including Tweets in English. Additionally, SNA with different search queries might lead to different results. Moreover, the number of participants involved in this study was limited (19 interviewees and 3 educators). However, despite these limitations, this study provided a solid ground for revealing the concerns about using chatbots, specifically ChatGPT, in education, among early adopters. Future research directions could focus on conducting one step forward by implementing ChatGPT within teaching practices, and investigating how human tutors and machines (ChatGPT) could work together to achieve an educational objective, as well as the changes and outcomes brought to the education field (e.g., evolutionary or revolutionary).

Abbreviations

AI	Artificial intelligence
GPT	Generative pre-trained transformer
ICT	Information and communication technology
SNA	Social network analysis
t-SNE	T-distributed stochastic neighbor embedding

Acknowledgements

Not applicable.

Author contributions

Each author contributed evenly to this manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Declarations**Competing interests**

The authors declare that they have no competing interests.

Received: 22 January 2023 Accepted: 17 February 2023

Published online: 22 February 2023

References

- Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart Winston.
- Asimov, I. (1942). *Runaround*. Astounding Science Fiction
- Asimov, I. (1950). *I, Robot*. Gnome Press.
- Barredo Arrieta, A., Diaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Beccari, M. N., & Oliveira, T. L. (2011). A philosophical approach about user experience methodology. In *International Conference of Design, User Experience, and Usability* (pp. 13–22). Springer, Berlin
- Bozkurt, A. (2022). Biased binaries. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-022-00385-2>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint [arXiv:2005.14165](https://arxiv.org/abs/2005.14165).
- Bulathwela, S., Perez-Ortiz, M., Yilmaz, E., & Shawe-Taylor, J. (2020). Truelearn: A family of bayesian algorithms to match lifelong learners to open educational resources. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01), 565–573. <https://doi.org/10.1609/aaai.v34i01.5395>
- Čapek, K. (1921). *Rossum's Universal Robots*.
- Cook, K. S., Cheshire, C., Rice, E. R., & Nakagawa, S. (2013). Social exchange theory. In *Handbook of social psychology* (pp. 61–88). Springer, Dordrecht.
- Durall, E., & Kapros, E. (2020). Co-design for a competency self-assessment Chatbot and survey in science education. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies human and technology ecosystems HCII 2020 lecture notes in computer science*. Cham: Springer.
- Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, 7(3), 93–99. <https://doi.org/10.1016/j.afjem.2017.08.001>
- Firat, M. (2023). How chat GPT can transform autodidactic experiences and open education? <https://doi.org/10.31219/osf.io/9ge8m>
- Flick, U. (2009). *An introduction to qualitative research* (4th ed.). SAGE.
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289. <https://doi.org/10.1016/j.chb.2018.12.023>
- Giachanou, A., & Crestani, F. (2016). Like it or not: A survey of Twitter sentiment analysis methods. *ACM Computing Surveys (CSUR)*, 49(2), 1–41.
- Guadamuz, A. (2016). The monkey selfie: Copyright lessons for originality in photographs and internet jurisdiction. *Internet Policy Review*. <https://doi.org/10.14763/2016.1.398>
- Guadamuz, A. (2018). Can the monkey selfie case teach us anything about copyright law? *WIPO Magazine*, 1, 40–46.
- Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann.
- Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). *I think this is the most disruptive technology: Exploring sentiments of ChatGPT early adopters using Twitter data*. arXiv preprint [arXiv:2212.05856](https://arxiv.org/abs/2212.05856).
- Harel, D., & Koren, Y. (2001). A Fast Multi-Scale Method for Drawing Large Graphs. In *Graph Drawing: 8th International Symposium, GD 2000*. Colonial Williamsburg, VA, USA, September 20–23, 2000, Proceedings (No. 1984, p. 183). Springer Science & Business Media.
- Herft, A. (2023). A Teacher's Prompt Guide to ChatGPT: Aligned with 'What Works Best'. CESE NSW "What Works Best in Practice". https://drive.google.com/file/d/15qAxxUzOwAPwHzoakBJd8FAgiOZYclxq/view?fbclid=IwAR2fRdL5ggq4zU-81Fil8j4BAOp5HqWHC_Ecy2sqKk4EiWXL0FKa5GVz5dE
- Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In *Artificial intelligence in behavioral and mental health care* (pp. 81–115). Academic Press. <https://doi.org/10.1016/B978-0-12-420248-1.00004-0>
- Inwood, B. (Ed.). (2003). *The Cambridge companion to the Stoics*. Cambridge University Press.
- Janssen, A., Grütznier, L., & Breitner, M. H. (2021). Why do chatbots fail? A critical success factors analysis. In *International Conference on Information Systems (ICIS), Forty-Second International Conference on Information Systems*
- Kasneji, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Kasneji, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. <https://doi.org/10.35542/osf.io/5er8f>
- King, M. R., & chatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, 16, 1–2. <https://doi.org/10.1007/s12195-022-00754-8>
- Kuhail, M. A., Alturki, N., Alramlawi, S., et al. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28, 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>

- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., et al. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Levinger, G. (1980). Toward the analysis of close relationships. *Journal of Experimental Social Psychology*, 16(6), 510–544. [https://doi.org/10.1016/0022-1031\(80\)90056-6](https://doi.org/10.1016/0022-1031(80)90056-6)
- Lieberman, M. (2023). *What Is ChatGPT and How Is It Used in Education?*. Education Week. <https://www.edweek.org/technology/what-is-chatgpt-and-how-is-it-used-in-education/2023/01>
- McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). A proposal for Dartmouth summer research project on artificial intelligence. *AI Magazine*, 27, 12.
- Mollick, E. R., & Mollick, L. (2022). New modes of learning enabled by AI chatbots: Three methods and assignments. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4300783>
- O'Connor, S., & ChatGPT, (2023). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537. <https://doi.org/10.1016/j.nepr.2022.103537>
- Ofgang, E. (2022). *What is ChatGPT and how can you teach with it? Tips & tricks*. Tech & Learning. <https://www.techlearning.com/how-to/what-is-chatgpt-and-how-to-teach-with-it-tips-and-tricks>
- OpenAI. (2023). ChatGPT: Optimizing language models for dialogue. <https://openai.com/blog/chatgpt/>
- Rainie, L. (2014). *The six types of Twitter conversations*. PewResearch. <http://www.pewresearch.org/fact-tank/2014/02/20/the-six-types-of-twitter-conversations/>
- Rosati, E. (2017). The monkey selfie case and the concept of authorship: An EU perspective. *Journal of Intellectual Property Law & Practice*, 12(12), 973–977.
- Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R., Abrami, P. C., Wade, C. A., & Lowerison, G. (2009). Technology's effect on achievement in higher education: A stage I meta-analysis of classroom applications. *Journal of Computing in Higher Education*, 21, 95–109. <https://doi.org/10.1007/s12528-009-9021-8>
- Shen-Berro, J. (2023). New York City Schools blocked ChatGPT. Here's what other large districts are doing. *Chalkbeat*. <https://www.chalkbeat.org/2023/1/6/23543039/chatgpt-school-districts-ban-block-artificial-intelligence-open-ai>
- Skjeive, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022). A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies*, 168, 102903. <https://doi.org/10.1016/j.ijhcs.2022.102903>
- Smith, M., Rainie, L., Shneiderman, B., & Himelboim, I. (2014). *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*. Pwlinternet. <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>
- Stake, R. E. (1995). *The art of case study research: Perspective in practice*. Sage.
- Susnjak, T. (2022). ChatGPT: The end of online exam integrity?. arXiv preprint [arXiv:2212.09292](https://arxiv.org/abs/2212.09292).
- The Guardian. (2023). New York City schools ban AI chatbot ChatGPT. The Guardian. Retrieved from <https://www.theguardian.com/us-news/2023/jan/06/new-york-city-schools-ban-ai-chatbot-chatgpt>
- Thurmond, V. A. (2001). The point of triangulation. *Journal of Nursing Scholarship*, 33(3), 253–258. <https://doi.org/10.1111/j.1547-5069.2001.00253.x>
- Turing, A. (1950). Computing machinery and intelligence. *Mind A Quarterly Review of Psychology and Philosophy*, 236, 433–460. <https://doi.org/10.1093/mind/lix.236.433>
- van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(2008), 2579–2605.
- Yin, R. K. (1984). *Case study research: Design and methods*. Sage.
- Zhai, X. (2022). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4312418>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
