

Using ChatGPT in Education: Human Reflection on ChatGPT's Self-Reflection

Eugène Loos ^{1,*} , Johanna Gröpler ²  and Marie-Louise Sophie Goudeau ³ 

¹ School of Governance, Utrecht University, 3511 ZC Utrecht, The Netherlands

² TH Wildau College, University of Applied Sciences Wildau, 15745 Wildau, Germany; groepler@th-wildau.de

³ Utrecht University Library, 3508 TC Utrecht, The Netherlands; m.s.goudeau@uu.nl

* Correspondence: e.f.loos@uu.nl

Abstract: ChatGPT is a fascinating AI text generator tool. It is a language model developed by OpenAI, a research and deployment company with the mission, according to OpenAI's website: "to ensure that artificial general intelligence benefits all of humanity". ChatGPT is able to generate human-like texts. But how does it work? What about the quality of the texts it provides? And is it capable of being self-reflective? Information sources must be efficient, effective and reliable in education, in order to enhance students' learning process. For this reason, we started a dialogue with ChatGPT-3 while using, among others, a SWOT analysis it generated about its own functioning in an educational setting. This enabled us, as human authors, to analyze the extent to which this AI system is able to practice self-reflection. Finally, the paper sketches implications for education and future research.

Keywords: ChatGPT; AI; text generating; education; SWOT analysis; self-reflection; hallucinations



Citation: Loos, E.; Gröpler, J.; Goudeau, M.-L.S. Using ChatGPT in Education: Human Reflection on ChatGPT's Self-Reflection. *Societies* **2023**, *13*, 196. <https://doi.org/10.3390/soc13080196>

Academic Editor: Gregor Wolbring

Received: 29 June 2023

Revised: 7 August 2023

Accepted: 10 August 2023

Published: 21 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. What Is ChatGPT?

ChatGPT is a language model developed by OpenAI, capable of generating human-like text based on the input provided to it [1]. It uses advanced machine learning algorithms, such as deep learning, to analyze patterns and relationships in large amounts of text data and generate new text that is consistent with the input [2]. It is often used for a variety of natural language processing tasks, such as chatbots [3], language translation [4], and content generation [5].

1.2. How Does ChatGPT Work?

ChatGPT works by utilizing a deep neural network, specifically the Transformer architecture [6], to generate text based on input it has received. The model has been trained on a massive dataset of text [7], which enables it to predict the next word in a sequence of words given a prompt. This process continues until the model generates a full response or reaches a maximum length. The output is generated based on probabilities calculated by the model [8], which are influenced by patterns in the training data and the context of the input prompt [9].

2. ChatGPT in an Educational Setting

2.1. To What Extent Can ChatGPT Be Used for Education?

ChatGPT and other language models like it have the potential to be used in education in a variety of ways. Here are some examples:

1. Question Answering: ChatGPT can be used to answer questions in real-time, providing students with instant feedback and helping them to learn more effectively.

2. Tutoring: ChatGPT can be integrated into educational software to provide students with personalized, one-on-one tutoring sessions.
3. Content Generation: ChatGPT can be used to generate educational content, such as summaries, explanations, and study materials, based on input it has received.
4. Language Learning: ChatGPT can be used as an AI language tutor, helping students to practice speaking, writing, and comprehension in a foreign language.

However, it's important to note that while language models like ChatGPT have the potential to enhance education, they should not be used as a replacement for human teachers. AI systems like ChatGPT should be seen as tools to support and augment the educational process, not as a substitute for it.

2.2. SWOT Analysis for ChatGPT in an Educational Setting

Here is a SWOT analysis of using ChatGPT in an educational setting:

Strengths:

1. Instant Feedback: ChatGPT can provide students with instant feedback on their questions, helping them to learn more effectively.
2. Personalized Tutoring: ChatGPT can be used to provide students with personalized, one-on-one tutoring sessions, which can help them to learn at their own pace.
3. Improved Access to Education: ChatGPT can be used to provide educational resources and support to students in areas where access to human teachers is limited.
4. Increased Efficiency: ChatGPT can be used to automate repetitive tasks, such as grading, which can free up time for teachers to focus on other tasks.

Weaknesses:

1. Lack of Empathy: ChatGPT is an AI system and does not have emotions, which can limit its ability to understand and respond to complex situations in a human-like manner.
2. Limited Contextual Understanding: ChatGPT is limited by the context it receives, and can sometimes generate responses that are inappropriate or incorrect.
3. Bias in Training Data: The training data used to train language models like ChatGPT can contain biases that can be reflected in its responses.

Opportunities:

1. Personalized Learning: ChatGPT can be used to provide personalized learning experiences to students, tailoring the educational process to their individual needs and abilities.
2. Improved Student Engagement: ChatGPT can be used to create interactive, engaging educational experiences for students, which can help to increase their motivation and engagement.

Threats:

1. Competition from Other Technologies: There is a growing number of educational technologies being developed, which could limit the adoption of ChatGPT in the educational setting.
2. Technical Limitations: There are technical limitations associated with language models like ChatGPT, such as the need for large amounts of computing power and the inability to understand and respond to complex situations.
3. Ethical and Regulatory Concerns: The use of language models like ChatGPT raises ethical and regulatory concerns, such as data privacy, data security, and bias in AI systems.

By taking into account the strengths, weaknesses, opportunities, and threats of using ChatGPT in an educational setting, organizations and educators can make informed decisions about how to use this technology effectively and responsibly.

3. How Can Misuse of ChatGPT in Education Be Avoided?

Misuse of ChatGPT in education can be avoided by following best practices and ethical considerations. Here are some steps that can help.

1. Ensure transparency: Be transparent about the use of ChatGPT in education, including how it is being used and the limitations of the technology.
2. Monitor and evaluate results: Regularly monitor and evaluate the results of using ChatGPT in education to ensure that it is having a positive impact and is being used effectively.
3. Avoid using ChatGPT as a replacement for human teachers: ChatGPT should be seen as a tool to support and augment the educational process, not as a replacement for human teachers.
4. Address bias and ethical considerations: Take steps to address bias and ethical considerations in the use of ChatGPT in education, such as evaluating the training data for biases and implementing measures to mitigate them.
5. Ensure privacy and security: Take steps to ensure the privacy and security of student data, including following data protection and privacy regulations.
6. Regularly update the model: Regularly update the model to ensure that it continues to reflect the latest research and best practices in the field of conversational AI.

By following these steps, the potential benefits of using ChatGPT in education can be realized while avoiding potential misuses and negative consequences.

4. Method

While reading Sections 1–3, you might have wondered whether ChatGPT had generated the text. The answer is, yes—it had indeed done so. ChatGPT-3 generated the text for Sections 1–3, in response to the questions we posed on 5 February 2023 shown in each section’s heading.

ChatGPT used references in the texts it generated for Sections 1.1 and 1.2 (we adapted them to this journal’s reference style). But in Section 1.2, the reference to its own link OpenAI (2021) does not work; moreover, no references were provided for Sections 2.1, 2.2 and 3. When we asked ChatGPT to add references to the text generated in Sections 2.1 and 2.2, it replied: “I’m sorry, but I am not able to provide references for this text as it is a general statement and not based on any specific research or study”. When we nonetheless pressed ChatGPT to try again to find references, it answered: “I’m sorry, I am an AI language model and do not have the ability to find specific references as I do not have access to the internet. It is the responsibility of the author to provide references for the information presented in the text”. Section 3 also lacked references, and when we asked for clarification, ChatGPT let us know once again: “I’m sorry, I am an AI language model and do not have the ability to provide references for the information presented in the text. It is the responsibility of the author to provide references for their claims. However, it is always good practice to support statements and recommendations with relevant and credible sources to enhance the credibility and trustworthiness of the information”. Please note that ChatGPT generated refs. [1–9]. Refs. [1,2,4,6,9] are real. Refs. [3,5]: Conference proceedings exist, but the paper do not. Ref. [7]: Link does not work: “The page you were looking for doesn’t exist. You may have mistyped the address or the page may have moved.” Ref. [8]: Paper exists, but one author (Raffel) is not correct.

We have added our own references to Sections 5 and 6, based on our experiences/knowledge and examples from our educational work practices.

ChatGPT is a fascinating AI text generator tool. It is a language model developed by OpenAI, a research and deployment company with the mission of ensuring “that artificial general intelligence benefits all of humanity” <https://openai.com/about> (accessed on 29 June 2023). It is able to generate human-like texts. But what is the quality of these texts? And is ChatGPT capable of self-reflection? In education, it is of the utmost importance that information sources be efficient, effective and reliable to enhance students’ learning

process. In this paper, we, the authors—an associate professor in the field of accessible reliable digital information, an information/education specialist at Utrecht University in the Netherlands, and an academic writing trainer/expert digital infrastructure at University of Applied Sciences Wildau in Germany, will answer the following research question: To what extent is ChatGPT-3 capable of being self-reflective. We therefore started a dialogue with this AI system.

In Section 5, we focus on ChatGPT’s self-reflection capabilities by analyzing our dialogues with this AI system. In Section 6, we will present our conclusions and implications for education and future research, including a discussion on whether or not to acknowledge ChatGPT as a co-author of this paper.

5. Human Reflection on ChatGPT’s Self-Reflection

5.1. Human Reflection on Sections 1, 2.1, 2.2 and 3

5.1.1. Section 1.1 What Is ChatGPT?

ChatGPT’s answer is much in the style of a lexicon or Wikipedia entry, providing brief, general information only. It appears to be correct, but offers no further information about the development of language models and their place in the world of AI. Furthermore, terms and concepts such as “advanced machine learning algorithms” and “deep learning” are not specified, nor is it explained how these are related to “natural language processes”. We acknowledge that the quality and depth of the answers are partly due to the quality and clarity of our own prompts.

5.1.2. Section 1.2 How does ChatGPT Work?

The fact that ChatGPT always “predict[s] the next word in a sequence of words given a prompt” and that “the output is generated based on probabilities calculated by the model [8], which are influenced by patterns in the training data and the context of the input prompt”, or that pre-formulated phrases may also be part of ChatGPT’s text generating, should give us all pause for thought. See https://www.reddit.com/r/ChatGPT/comments/zujg8g/why_is_chatgpt_so_politically_correct/ (accessed on 29 June 2023) for more information on this issue. Again, wariness is in order when looking at ChatGPT’s sentences in Section 3 (see also Section 5.2.1), such as “I’m sorry, I am an AI language model and do not have the ability to provide references for the information presented in the text. It is the responsibility of the author to provide references for their claims. However, it is always good practice to support statements and recommendations with relevant and credible sources to enhance the credibility and trustworthiness of the information”. See also Section 6 where we discuss ChatGPT’s answer about co-authorship. It looks as if ChatGPT uses pre-formulated phrases, although we do not know this for sure; OpenAI is very reticent about the training data set on which ChatGPT is based. For this reason, Ref. [10] calls for a value-sensitive design ([11], p. 587) that “(…) contributes one new professional practice—called data statements—which we argue will bring about improvements in engineering and scientific outcomes while also enabling more ethically responsive NLP technology. A data statement is a characterization of a dataset that provides context to allow developers and users to better understand how experimental results might generalize, how software might be appropriately deployed, and what biases might be reflected in systems built on the software”. Ref. [12] argues: “The recent emergence and adoption of Machine Learning technology, and specifically of Large Language Models, has drawn attention to the need for systematic and transparent management of language data. This work proposes an approach to global language data governance that attempts to organize data management amongst stakeholders, values, and rights”. For more information, see also [13], who reviewed “the risks of relying on proprietary software and survey the first crop of open-source projects of comparable architecture and functionality”. ChatGPT was one of the projects they reviewed. Additionally, according to [14], there are data labelers working for ChatGPT: “In a statement, an OpenAI spokesperson confirmed that Sama employees in Kenya contributed to a tool it was building to detect toxic content, which

was eventually built into ChatGPT. The statement also said that this work contributed to efforts to remove toxic data from the training datasets of tools like ChatGPT. ‘Our mission is to ensure artificial general intelligence benefits all of humanity, and we work hard to build safe and useful AI systems that limit bias and harmful content,’ the spokesperson said. ‘Classifying and filtering harmful [text and images] is a necessary step in minimizing the amount of violent and sexual content included in training data and creating tools that can detect harmful content.’” Apart from the fact that these data labelers are paid less than USD 2 an hour, it is not clear which criteria are used to make ChatGPT less toxic.

5.1.3. Section 2.1 To What Extent Can ChatGPT Be Used for Education?

On the one hand, ChatGPT explains complex concepts in simpler words, which may be easier to understand and thus could support the learning process. On the other hand, the answers ChatGPT gives remain very unspecific. The examples provided by ChatGPT of how it could be used for education read like key points from a longer text; a summarization, in short. Let us, as human experts, now examine the examples generated by ChatGPT.

Example 1 (Question Answering): What kind of question is ChatGPT referring to? Is it referring to the preparation for an exam or the exam itself? There is potential for use in the learning process in cases where students generate questions for self-assessment, and it could provide some explanation of certain concepts students have failed to grasp. In this way, it could be a useful support in their learning process. The question is, however, what kind of training data are used by ChatGPT to provide this support (see the importance of adopting an ethically responsive NLP technology, in which data statements are crucial, as well as the adoption of global language data governance practices, as discussed above). Moreover, using ChatGPT to clarify concepts for students means that the students must also possess the skills to write prompts: only by asking the right questions will valid answers be obtained. In addition, as students are novices in their field of study, it is difficult for them to determine whether ChatGPT is hallucinating or not [15–20]; see also <https://bernardmarr.com/chatgpt-what-are-hallucinations-and-why-are-they-a-problem-for-ai-systems/> (accessed on 29 June 2023); <https://www.cspinet.org/blog/chatgpt-amazing-beware-its-hallucinations> (accessed on 29 June 2023); <https://zapier.com/blog/ai-hallucinations/> (accessed on 29 June 2023). A possibility that ChatGPT does not mention here (as it does in the SWOT analysis, points 1 and 4 under Strengths) is that teachers could also use ChatGPT for feedback on exams, or even grading, based on specific criteria, although the reliability of such feedback or grading would be debatable (see also [21–29]).

Example 2: (Tutoring): Once again, ChatGPT’s answer offers information but without further explanation. What would be the nature of this tutoring? How exactly could a control mechanism be integrated? And while language models like ChatGPT have the potential to enhance education, as ChatGPT says itself: “they should not be used as a replacement for human teachers”. The fact that a conversation with ChatGPT seems almost human, and is often polite and friendly, carries the risk of the user being unaware of being in contact with an AI system, of which the functionality and exact training data set used is unknown (see also [10,12,13]). As ChatGPT has no emotions (Weakness, point 1) and only limited contextual “understanding” (Weakness, point 2), it will not be able to interpret a student’s emotional struggles regarding the learning process in the way human teachers can (see Scenario 7 Absence of emotions or reflections on students’ engagements [28]). Moreover, ChatGPT takes the context given by the student’s prompt within a chat (up to at least 4096 tokens) into account when providing answers. So if a student feeds ChatGPT information that is not correct (<https://www.whyofai.com/blog/your-chatgpt-questions-answered> (accessed on 29 June 2023); <https://www.techradar.com/how-to/the-5-biggest-mistakes-people-are-making-with-chatgpt-and-how-to-avoid-them> (accessed on 29 June 2023)), this could lead to wrong answers and poor guidance. Biases in the data training set could also negatively affect the tutoring of students with learning disabilities or minority students, as ChatGPT’s results may not be inclusive or culturally sensitive.

Example 3 (Content Generation): ChatGPT provides examples for generating educational content (summaries, explanations, and study materials). Two other examples could include: (1) text generation for students' assignments, such as literature reviews. In that case, it is essential that the content be checked by a human to ensure the educational material is not contaminated by biases or fake information; (2) the suggestion put forward by [27] that the potential application of GPT as a teacher-facing tool "be classified into four key facets of teaching processes: planning, instruction, assessment, and interaction with other stakeholders such as families and administrators."

Example 4 (Language Learning): In terms of language learning, ChatGPT seems not to practice critical self-reflection. Since it is not (yet) able to "understand" nuances in speaking and take into consideration cultural differences, the question is whether it can really support speaking, writing and comprehension in a foreign language [30]. When reviewing the use of ChatGPT in language education, it still appeared to struggle to provide correct translations in languages that were scarce in the training set [26,31]. However, these difficulties may by now have been resolved in ChatGPT-4 [31]. The popular language learning app DuoLingo has announced that it will incorporate ChatGPT into its software (<https://blog.duolingo.com/duolingo-max/> (accessed on 29 June 2023)), confirming ChatGPT's potential to contribute at least to some extent to learning languages.

ChatGPT's four examples in Section 2.1 illustrate how this could be used for educational purposes, but it fails to discuss the potential pitfalls of such use—although it did warn that "it's important to note that while language models like ChatGPT have the potential to enhance education, they should not be used as a replacement for human teachers. AI systems like ChatGPT should be seen as tools to support and augment the educational process, not as a substitute for it". This would seem to indicate some measure of self-reflection (see also Section 2.2 on weaknesses and threats, generated by ChatGPT when asked to conduct a SWOT analysis of its own functioning in an educational setting). The disadvantages of ChatGPT's fully replacing human teachers are evident in the examples it provides. For instance, effective learning through question answering poses the risk of ChatGPT providing biased, hallucinated [15–20], incorrect or outdated information, and failing to provide reliable sources to evaluate the generated content (e.g., [26]). Incorrect answers may be generated due to issues with the prompts, as discussed above. To ensure efficient and effective learning takes place, learners need to be aware of the limitations that are inherent to language models like ChatGPT.

Finally, it is important to underline that the issues with all these examples highlight the need for an educational framework focusing on the development of a variety of academic skills. This includes, but is not limited to:

- Digital literacy. This notion is defined by the American Library Association as "the ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills" (<https://literacy.ala.org/digital-literacy/> (accessed on 29 June 2023)). Digital skills are a prerequisite for AI literacy and information literacy [32]. Also called metaliteracy [33], these focus on gaining an understanding of and acquiring the skills to collect, evaluate, process and share information [33–35]. Such skills could help to evaluate the information produced by ChatGPT. Moreover, such skills help students perceive the value of their ability as a student to collect, evaluate, process and share information, and assume the responsibilities that come with sharing this information with others [34].
- Critical thinking skills. As ChatGPT automates the writing process, students need to understand how their ability for critical thinking sets them apart from ChatGPT and why that matters. Ref. [36] argues: "As a teacher of critical thinking, I earnestly hope that you will never take my word on anything. It's my job to teach you how to think for yourself. Today, I want to help you distinguish (for yourself) between mechanical writing and sophisticated writing. Mechanical writing can be done by people or computers... but computers will be better at it. Sophisticated writing, on the

other hand, requires critical thinking skills that language-generation models do not possess. Mechanical writing is about communicating existing information. At its best, mechanical writing is clear, conventional, and correct. Sophisticated writing is about generating new insights through the writing process. At its best, sophisticated writing is thoughtful, self-aware, and creative". In other words, human teachers are needed for the development of digital literacy and critical thinking skills.

5.1.4. Section 2.2 SWOT Analysis for ChatGPT in an Educational Setting

ChatGPT showed itself capable of generating a text that follows the format of a SWOT analysis (see [37,38] for more information about its origins). Let us now have a look at the extent to which it is able to be self-reflective in a critical way in its analysis of its own strengths, weaknesses, opportunities and threats.

Strengths: What stands out immediately is that the output provided on its strengths closely resembles that of the first three examples in Section 2.1 (1. Question Answering, 2. Tutoring and 3. Content Generation) on how ChatGPT could be used for education. Is it possible that the generated text in both cases is the most probable output, because ChatGPT "predict[s] the next word in a sequence of words given a prompt"?

It is unclear what ChatGPT means by feedback on students' questions—could this perhaps mean the answers to their questions? Nor is it clear what ChatGPT means by one-to-one tutoring sessions.

Weaknesses: The points that ChatGPT mentions here clearly undermine all the strengths and opportunities listed. Ironically, the tool also points this out itself. ChatGPT does not seem to "be aware" of this, and in no way relates the four components of its SWOT analysis to one another. While ChatGPT appears self-reflective in stating that it "is limited by the context it receives and can sometimes generate responses that are inappropriate or incorrect" (Weakness, point 2), no limitations are specified and it is not clear what "sometimes" means. Hallucinations occur, amongst others, if there is bias in the training set or if the training set is limited either regarding the topic or the language in which the question is asked. They also result from prompts that are abstract or contain incorrect information [10,15–20]. As well, the following weakness should be kept in mind, a point made by [23] who argues that: "Additionally, when students rely on AI to complete their work, they may not fully understand the material and may lack a sense of ownership over their learning. This can reduce the effectiveness of education and limit their ability to apply their knowledge in the future".

We agree with the critical self-reflection by ChatGPT that it lacks empathy: "ChatGPT is an AI system and does not have emotions, which can limit its ability to understand and respond to complex situations in a human-like manner" (weakness, point 1).

As ChatGPT itself states, it has limited contextual "understanding" (weakness, point 2). Nevertheless, the inappropriate and incorrect responses it sometimes generates can appear very plausible and convincing. These hallucinations are a weakness in ChatGPT's general functioning, but in an educational setting it becomes particularly difficult for students to differentiate between when the language model is hallucinating and when it is providing a correct response. As students are novices in their field of study, differentiating between what is correct and what is incorrect becomes even harder. The awareness of this phenomenon will become even more important as ChatGPT is integrated into other software applications and search machines, camouflaging the fact that AI is involved. Fact-checking ChatGPT's responses is difficult as ChatGPT is yet not always capable of providing recent, reliable sources for its claims [15,26]. The capabilities of Elicit, a GPT-3-based tool that can make suggestions for academic articles, suggest that these problems may be solved in the future. However, it is essential for both teachers and students to understand that AI has its limitations and why this is the case. Moreover, they must be aware of the need to check the reliability of the generated texts to ensure the tool is used in a safe way in an educational setting.

Opportunities: The points ChatGPT presents here are not evidence-based. Nonetheless, despite the challenges, the use of ChatGPT and other AI tools in education seems inevitable. Students and teachers are already using these tools, and this is probably only the very beginning of a profound change in the educational field (see [39] for the labor market impact potential). It is crucial to train students and teachers in using ChatGPT in a responsible way [26,32,39,40]. The introduction of ChatGPT in education makes it more than ever necessary to focus on learning academic skills. These include skills such as writing, digital literacy (as discussed above) and critical thinking to generate new ideas (also discussed above). As well, attention must be paid to integrity, which should become an explicit part of the curriculum (see [23] and [41] about ethical implications).

Threats: It is interesting that the points raised by ChatGPT in this section describe threats to its own OpenAI developers (Threat 1), as it questions the efficient, effective and reliable use of the language model in education (Threats 2 and 3). Also, ChatGPT fails to mention the following points as social threats.

Currently, other AI tools are being released (e.g., Bard, Perplexity AI and Neuroflash; see also <https://aibusiness.com/nlp/7-language-models-you-need-to-know> (accessed on 29 June 2023)), all of which are pilots. In the race for AI dominance, this “ship-then-fix” policy seems to be the prevailing strategy. This may result in AI tools being released with problems that have not (yet) been solved [39]. Hopefully, educators are and will remain alert to the risks associated with AI, such as copyright infringement, manipulation and discrimination, in order to ensure a safe learning environment.

Privacy is another issue—see for example the concerns in Italy and the actions taken by the Italian data protection authority in this regard: <https://www.bbc.com/news/technology-65139406> (accessed on 29 June 2023); <https://www.bbc.com/news/technology-65431914> (accessed on 29 June 2023). The authors of Ref. [28] make an important observation about ChatGPT’s misleading privacy practices. “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> (accessed on 29 June 2023)) related to this issue, it is seen that conversations are stored and used (Blackbox) surprisingly, when ChatGPT also asked about this matter, it denied it (. . .), 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.”

Ethical and regulatory concerns are an issue when it comes to creating a safe learning environment. Unfortunately, the concerns threatening the use of ChatGPT in education are not limited to “data privacy, data security, and bias in AI systems”, as ChatGPT implies (Threat 3). First, users must understand that it is essential to fact-check each answer generated by ChatGPT [39], as these answers are all dependent on the context given by the user’s prompts. The large majority of ChatGPT users therefore run the risk of having incorrect information end up in educational material, such as digital textbooks. Secondly, the possibility of plagiarism (also called Aigiarism, see: <https://medium.com/@cristian.nedelcu/chatzero-is-the-best-tool-to-spot-aigiarism-737846323985> (accessed on 29 June 2023) and <https://www.theguardian.com/technology/2022/dec/31/ai-assisted-plagiarism-chatgpt-bot-says-it-has-an-answer-for-that> (accessed on 29 June 2023)) is one of the biggest concerns regarding the adaption of ChatGPT in education. This has evoked responses varying from forbidding the use of ChatGPT or adapting the assessment of students’ learning, to creative ways to embedding it in educational practices [22–24,26–28,32,42–45].

ChatGPT also does not mention the important risk of social inequality arising in the future due to groups of students not having been exposed to the technology at their educational institution, which would “provide some students with an unfair advantage

over others, particularly if not all students have access to the technology or if it is used in an unequal manner" [23]. Once ChatGPT and its probably more powerful successors are no longer able to be accessed for free, only those with enough money will be able to use these technologies. Who will have access to the best working version? Imagine a scenario where ChatGPT is not freely available to users studying at an institution that provides no access to ChatGPT, or one where students simply lack the money to gain access to the tool. How would the bias in the data be affected if only specific users are able to feed the language system with their prompts? These are important questions to address, as AI will probably continue to be integrated into educational settings.

Finally, one very important issue: What if one day ChatGPT should become the only information source we could use? We would no longer be in the position to use different sources (at this very moment, still a key capability in media literacy, see [33,35,46]) to check the reliability of the texts we read [39]—we would no longer write them ourselves, as the AI system could do it (better) for us [47].

5.1.5. Section 3 How Can Misuse of ChatGPT in Education Be Avoided?

As described in the sections above, here again the suggestions on how to avoid misconduct of ChatGPT seem more like a summarization. It is interesting that ChatGPT partially "understood" the question more from the perspective of the OpenAI developers—see points 4, 5 and 6. Points 1 and 3 refer to those who might use the tool in education, such as human teachers, while point 2 refers to both AI developers and human teachers. Here, as well, it would have been useful if ChatGPT could have gone into more detail.

When addressing the issue of avoiding the misuse of ChatGPT, the tool itself appears to lay the full responsibility on the user formulating the question (see steps 1–5 in Section 3; step 6, "regularly update the model", is not clear: how can a user update the model?). From this perspective, neither OpenAI nor the government are responsible for preventing the misuse of ChatGPT; its own responsibility is not part of the self-reflection. This position is in contrast with the debate in the British parliament on the development of their own BritGPT to guarantee national security (<https://www.theguardian.com/business/2023/feb/22/uk-needs-its-own-britgpt-or-will-face-an-uncertain-future-mps-hear> (accessed on 29 June 2023)), which is a measure that fits best with steps 5 and 6 ("ensure privacy and security" and "regularly update the model"). As a society we have to make sure that we are prepared to take the right measures regarding the issues touched on in steps 5 and 6, as we may end up with problems that we cannot fix [39].

The first four steps ("1. Ensure transparency, 2. Monitor and evaluate results, 3. Avoid using ChatGPT as a replacement for human teachers, 4. Address bias and ethical considerations") generated by ChatGPT seem to be meant to prevent the misuse of the tool by educators, and they underline the importance of digital literacy and critical thinking skills. This alone may not be sufficient to prevent the misuse of the tool in the educational setting. Awareness and transparency are certainly necessary to begin with, but understanding how a language system works, as well as being digitally literate and capable of thinking critically, are all vital for a responsible use of ChatGPT [40]. In the first place, as a user, you must always be aware that ChatGPT answers are, to a large degree, based on the quality of your prompt. Secondly, you must keep in mind that ChatGPT may hallucinate. Third, you must remember that you have to fact-check—and understand why that is important. Fourth, in order to do this, you must know how to fact-check. As a user, you should have an overview of the entire information landscape, which will help you to understand why you should not use and share ChatGPT's responses without first fact-checking. Teaching these digital skills, including critical thinking, is important, especially as low performers tend to overestimate their skills [46].


5.2. ChatGPT: Do You Have Direct Access to the Internet?

5.2.1. Introduction


According to OpenAI, ChatGPT was trained on a dataset that had as cutoff the year 2021: “ChatGPT is not connected to the internet, and it can occasionally produce incorrect answers. It has limited knowledge of world and events after 2021 and may also occasionally produce harmful instructions or biased content. We would recommend checking whether responses from the model are accurate or not. If you find an answer is incorrect, please provide that feedback by using the ‘Thumbs Down’ button”. <https://help.openai.com/en/articles/6783457-what-is-chatgpt> (accessed on 29 June 2023).

As the sources of the data training set used by ChatGPT to generate text are unknown, we checked with ChatGPT to verify whether it really had no direct internet access.

E ChatGPT, you let us know “I’m sorry, I am an AI language model and do not have the ability to find specific references as I do not have access to the internet. It is the responsibility of the author to provide references for the information presented in the text”. Please explain how it is possible that you give us references without access to the internet.


 I apologize for the confusion in my previous statement. As a language model, I have been trained on a diverse range of texts and information, which includes information that is widely considered to be common knowledge, such as the names of famous people and basic historical events. However, I don’t have access to real-time information or the internet to verify information and provide accurate and up-to-date references. The information I provide is based on my training and may not always accurate or reflect the most current understanding of a particular subject

E You’re too modest, ChatGPT. Some references really are up-to-date. Where did you find the references if not on the internet?

 As a language model, I was trained on a massive dataset of text which includes information from various sources such as books, articles, websites, and more. This training data was created and compiled by OpenAI and is stored in my memory as a large database of text. When I provide information or answer questions, I use this database of text to generate responses based on patterns and relationships found in the data.


So, while I don’t have direct access to the internet to verify information in real-time, I have been trained on a large corpus of text that includes information that is widely considered to be true and accurate. However, it is important to note that this information may be outdated or contain inaccuracies, and it is always a good idea to verify information with multiple sources.


E ChatGPT, you explained: “While I don’t have direct access to the internet to verify information in real-time, I have been trained on a large corpus of text that includes information that is widely considered to be true and accurate”. But I still have three questions: (1) Where did you find this large corpus of text, (2) who decides what is widely considered to be true and accurate, and (3) how do you define “true” and “accurate”?


 (1) The large corpus of text used to train me was provided by OpenAI and includes a diverse range of sources such as books, articles, and websites.


(2) The information included in my training data is based on what is commonly accepted by a significant portion of the scientific and scholarly community. However, it’s important to note that opinions and knowledge can change over time, and some information may be considered outdated or incorrect.


(3) The definition of “true” and “accurate” in the context of my training data is based on correspondence with objective reality and reliability in representing information. However, it’s worth noting that there may be differing opinions and interpretations of what is considered true and accurate, and information can change as new evidence becomes available.


 ChatGPT, you explained that “While I don’t have direct access to the internet to verify information in real-time, I have been trained on a large corpus of text that includes information that is widely considered to be true and accurate”. Does no access to the internet to verify information in real time mean that you had access in the past?


 No, I do not have any direct access to the internet, either in the past or in real-time. I was trained solely on the large corpus of text provided by OpenAI, and I do not have the ability to access the internet to search for or verify information.


 ChatGPT, you told us: “No, I do not have any direct access to the internet, either in the past or in real-time. I was trained solely on the large corpus of text provided by OpenAI, and I do not have the ability to access the internet to search for or verify information”, but I don’t understand this. If you don’t have any direct access to the internet, either in the past or in real-time, how did you use OpenAI to get the large corpus of text?

 The large corpus of text used to train me was collected and curated by OpenAI, and stored in their databases. This training data was then used to train my model, allowing me to generate responses based on the patterns and relationships found in that text. However, I don’t have real-time access to the internet or any external sources of information outside of the training data provided by OpenAI.


 ChatGPT, you told us that “The large corpus of text used to train me was collected and curated by OpenAI, and stored in their databases”. Where did OpenAI find the large corpus of text?

 The large corpus of text used to train me was sourced from various publicly available texts on the internet, such as websites, books, and other written works. OpenAI collected and curated this text to create the training data used to train my model. This training data was preprocessed and cleaned to ensure that it was suitable for use in machine learning, and then used to train my model so that I could generate responses based on the patterns and relationships found in that text.


 ChatGPT, you told us that “The large corpus of text used to train me was sourced from various publicly available texts on the internet, such as websites, books, and other written works”. This sounds more logical to us than your earlier statement: “No, I do not have any direct access to the internet, either in the past or in real-time. I was trained solely on the large corpus of text provided by OpenAI, and I do not have the ability to access the internet to search for or verify information”. I don’t understand: These two statements are in contradiction with each other. Do you have access to the internet: yes or no?!

 No, I do not have direct access to the internet. My responses are generated based solely on the patterns and relationships found in the training data provided by OpenAI, which was sourced from various publicly available texts on the internet. While the training data was obtained from the internet, I do not have the ability to search or access the internet in real-time to verify information or generate responses based on up-to-date information. I apologize for any confusion caused by my previous answers.


E ChatGPT, you told us: “No, I do not have any direct access to the internet, either in the past or in real-time. I was trained solely on the large corpus of text provided by OpenAI, and I do not have the ability to access the internet to search for or verify information”, but I don’t understand this. If you don’t have any direct access to the internet, either in the past or in real-time, how did you use OpenAI to get the large corpus of text?

 The large corpus of text used to train me was collected and curated by OpenAI, and stored in their databases. This training data was then used to train my model, allowing me to generate responses based on the patterns and relationships found in that text. However, I don’t have real-time access to the internet or any external sources of information outside of the training data provided by OpenAI.

E ChatGPT, you told us that “The large corpus of text used to train me was collected and curated by OpenAI, and stored in their databases”. Where did OpenAI find the large corpus of text?

 The large corpus of text used to train me was sourced from various publicly available texts on the internet, such as websites, books, and other written works. OpenAI collected and curated this text to create the training data used to train my model. This training data was preprocessed and cleaned to ensure that it was suitable for use in machine learning, and then used to train my model so that I could generate responses based on the patterns and relationships found in that text.

E Thanks for your clarification ChatGPT and for explaining: “No, I do not have direct access to the internet. My responses are generated based solely on the patterns and relationships found in the training data provided by OpenAI, which was sourced from various publicly available texts on the internet. While the training data was obtained from the internet, I do not have the ability to search or access the internet in real-time to verify information or generate responses based on up-to-date information. I apologize for any confusion caused by my previous answers”. Apologizing is not necessary, you did your best and in the end explained us that you have access to the internet, but that this access is not direct but from “various publicly available texts on the internet”.

 I apologize for any confusion. To clarify, I do not have direct access to the internet and cannot search or access information in real-time. My responses are generated based solely on the patterns and relationships found in the training data provided by OpenAI, which was sourced from various publicly available texts on the internet, but I do not have the ability to access or search the internet to verify or update information. Thank you for your understanding.

Please note that we asked ChatGPT the question about direct internet access on 8 February 2023. On 12 May 2023, OpenAI announced a beta version with direct internet access—see <https://help.openai.com/en/articles/6825453-chatgpt-release-notes> (accessed on 29 June 2023).

5.2.2. Human Reflection on ChatGPT’s Artificial Answers Characteristics

Our dialogues with ChatGPT show that this language model is able to carry on a conversation with humans. See also ([48], p.1), where a sociologist conducted a conversation with ChatGPT about its “use in knowledge production following the shape of a dialogue, generated in real time, between the author and ChatGPT. The conversation reflects on, and is a reflexive contribution to, the study of artificial intelligence from a sociology of science perspective”, discussing insights from Science and Technology Studies (STS) (Ashmore (1989) [49] and Woolgar (1988) [50]). He concludes: “What I learnt presently from ChatGPT,

however, is that it has an affective aesthetic *style* embedded in its whole conversational approach. This poses significant dangers for academic research and education. It is much harder to discern the potential scientific, social and ethical consequences of its style of talk than it is its content" ([48], p. 9). From our dialogue with ChatGPT, which, in contrast to [48], focused less on reflexivity and more on ChatGPT's *self*-reflection, we can conclude that it generates human-like speech, referring to itself as "I", "me" and "my". It is noteworthy to observe how we as users also turned into very polite beings, often prefacing our questions with a "please". Would we have done that with an actual person every single time we had a further request? We also lost patience once, when we asked: Do you have access to the internet: yes or no?!

ChatGPT seems to be a bit submissive as it often apologizes for not providing the requested output, for example when it says: I apologize for any confusion caused by my previous answers. For more information about ChatGPT's artificial answers characteristics, we refer to [48,51,52].

6. Conclusions and Implications for Education and Future Research

The answers given by ChatGPT seemed correct, but they were also superficial, more like general statements. The tool often failed to provide specific examples. Perhaps the answers could have contained more substantial information or arguments had we used more specific prompts and added follow-up questions. The analysis of our dialogues with ChatGPT has shown that it could be used as a tool for brainstorming and generating ideas for further research on some topics, but that it cannot replace a human teacher. Furthermore, it could be interesting to compare, using the same prompts, the output from ChatGPT with that of other chat-assistant based tools using similar language models, such as Bard, Perplexity AI and Neuroflash. See [13,53–55] for more information about a comparison between chatbots, a topic that is beyond the scope of our paper.

A really weak point are Its references, as ChatGPT acknowledges: "I'm sorry, I am an AI language model and do not have the ability to find specific references as I do not have access to the internet. It is the responsibility of the author to provide references for the information presented in the text". ChatGPT does not perform that well at providing reliable sources [15,26,32], and it remains unclear on which dataset it bases its statements. Note, too, that ChatGPT-3 only provides a limited number of examples, while more could have been given. Examples that might have been added include: assisting students with learning disabilities and being a writing coach for students [29], enabling teachers to create rubrics [25,27] (see also rubrics https://drive.google.com/file/d/15qAxnUzOwAPwHzoaKBJd8FAgiOZYclxq/view?usp=embed_facebook (accessed on 29 June 2023)), or helping teachers to grade students' work, such as essays [21,22,26]. For more information about how close ChatGPT's performance comes to that of human experts, we refer to [56,57].

What about ChatGPT's artificial answers' characteristics? On the one hand, ChatGPT talks like a human, apologizing (I'm sorry"), often using "I", "me" and "my", while on the other using the non-human-like "it" and referring to itself as "an AI language model". For more information about the characteristics of ChatGPT's artificial answers, see [48,51], and for an evaluation of the extent to which linguists can distinguish between ChatGPT/AI and human writing, see [52].

Regarding ChatGPT's self-reflection, we agree with this AI language model that it should not be used to replace human teachers. The disadvantages of ChatGPT in fully replacing human teachers are evident in the examples given by ChatGPT itself. The development of digital literacy and critical thinking require human teachers. Only after students have developed these skills and understand how to set the value of the content generated by ChatGPT against the content generated through their own written processing of information can ChatGPT improve their learning. Similarly, teachers need to understand how to use ChatGPT wisely to benefit from the automation of some parts of their teaching.

ChatGPT showed itself capable of generating a text following the format of a SWOT analysis [37,38]. Let us now have a look, bearing in mind a number of critical points on the way ChatGPT generated this text:


1. It is important to note that the selection of strengths, weaknesses, opportunities and threats seems random. From an academic point of view, specific arguments are necessary to explain why these aspects have been listed instead of others. The SWOT analysis could also have been more specific—although had more specific prompts been given and follow-up questions added, the answers might have contained more substantial information or arguments;
2. In the generated SWOT analyses, ChatGPT underestimates its own weaknesses and the possible threats it faces. An important weakness is the fact that ChatGPT produces hallucinations [15–20], as students cannot easily differentiate between when the system is hallucinating and when it is providing a correct response. As users, they need to have an overview of the entire information landscape, which will help them to understand why they should not use and share ChatGPT's responses without first fact-checking. Teaching digital skills and critical thinking in this way is important, especially as low performers tend to overestimate their skills [46]. As students are novices in their field of study, differentiating between correct and incorrect becomes even harder. The awareness of this phenomenon will be even more important as ChatGPT is integrated into other software applications, making it less obvious that AI is involved. Fact-checking ChatGPT's responses is difficult, as the tool is not yet always able to provide recent, reliable sources for its claims [15,26,38];
3. It should also be noted that ChatGPT's own SWOT analysis demonstrates its limited vision for education—it focuses mainly on feedback on questions given by students and not the broader learning process, including (institutional) learning goals, the learning process, assessment and outcome. ChatGPT also seems not to be “aware” that it requires a human teacher to interpret a student's emotional struggles in relation to the learning process [28];
4. It is not only important to be aware of biases in the training data. ChatGPT also seems to make politically correct pre-formulated phrases (see https://www.reddit.com/r/ChatGPT/comments/zujg8g/why_is_chatgpt_so_politically_correct/) (accessed on 29 June 2023)). This prevents ChatGPT from being offensive, but it also indicates that the software owner has an influence on biases in the data. ChatGPT was developed by OpenAI, which used to be a non-profit organization. However, this changed in 2019, coinciding with a USD 1 billion investment from Microsoft. The effects of this on biases in the ChatGPT dataset is unclear. For this reason, it is important that ChatGPT's creator OpenAI provide transparency about its data set, and use a value-sensitive design [11] in which data statements “will bring about improvements in engineering and scientific outcomes while also enabling more ethically responsive NLP technology” ([11], p. 587). See also [12], who call for the “systematic and transparent management of language data (...) an approach to global language data governance that attempts to organize data management amongst stakeholders, values, and rights”. Also interesting is a comment from [14], who quotes a spokesman of a company (Sama) as stating that there are data labelers working for ChatGPT to “limit bias and harmful content”. Apart from the fact that these data labelers are paid less than USD 2 an hour, the criteria used to make ChatGPT less toxic are not known.
5. We referred to the SWOT analysis of ChatGPT by [22] who focused on implications for educational practice and research. They are quite positive about ChatGPT's “skills” for “generating plausible responses, self-improving capability, providing personalized responses, providing real-time responses” (pp. 3–4) (strengths) and “increasing the accessibility of information, facilitating personalized learning, facilitating complex learning, decreasing” teaching workload (pp. 4–6) (opportunities). We assume a more critical stance towards these points, as their efficiency and effectiveness are not yet evidence-based. We agree with [22] that the “lack of deep understanding, difficulty


in evaluating the quality of responses, the risk of biases and discrimination, lack of higher-order thinking skills” (pp. 6–7) (weaknesses) and “lack of understanding of the context, threatening academic integrity, perpetuating discrimination in education, democratization of plagiarism, declining in higher-order cognitive skills” (pp. 7–9) (threats) should be taken into consideration if the use of ChatGPT is being considered for educational and research practices.


It was also striking to note that when addressing the issue of avoiding the misuse of ChatGPT in education, ChatGPT itself appeared to lay the full responsibility on the user, which in itself is not a strongly self-reflective statement.


This paper aimed to examine whether ChatGPT-3 is capable of being self-reflective. To that end, we started a dialogue with this AI system. Among other things, we used a SWOT analysis that generated about its own functioning in an educational setting. We make no pretense of reviewing the results of papers about the educational use of ChatGPT. For an overview of the implications for education, such as the potential of using ChatGPT for training teachers, integration with other educational software and strategies for fact-checking AI-generated information, we refer to [21–29,58,59].

Finally, we as authors were curious about ChatGPT’s opinion on co-authorship, as is the case in some papers that added ChatGPT as a co-author (see also [32,60–64] for a discussion about this issue). So, we asked it some questions and got the following answers:

 ChatGPT, You do not have the capability to receive or respond to emails, so we address the following point here. As you answered our questions we asked you for our article ‘Using ChatGPT in education: Human reflection on GPT’s self-reflection’, we would like to know if you can be co-author for our paper?

 I’m sorry, but as an AI language model developed by OpenAI, I do not have the ability to be an author of a paper or any other creative work. I am simply a tool that can assist you in generating text based on the input you provide. While I can provide information and answer questions to help inform your writing, I cannot be credited as an author.

 Thank you ChatGPT for this clear response. In the academic world there is support for your standpoint. See: Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, 613, 620–621.

 Yes, it’s common practice in the academic community to only include human authors who have made substantial contributions to a research paper or other written work. Including AI systems like me as authors would not align with traditional understandings of authorship and the role of creative and intellectual contribution in producing a work. It’s important to note that while AI tools like me can assist in generating text and provide information to inform research, it’s still the responsibility of human authors to ensure the accuracy and validity of the information, and to interpret and analyze the results in a meaningful way.

Author Contributions: All authors contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable for studies not involving humans or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained in the article.

Acknowledgments: We would like to thank Laura Koenders, educational consultant/trainer at Utrecht University, for her feedback on this paper and Mark Dingemanse, associate professor at the

Centre for Language Studies, Department of Language and Communication, Radboud University, for his useful references [10,12] related to data statements and data governance.

Conflicts of Interest: The authors declare no conflict of interest.

References

- OpenAI. Available online: <https://openai.com/n.d> (accessed on 5 February 2023).
- Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, UK, 2016; Volume 1.
- Li, J.; Gao, J.; He, X.; Deng, L. A Deep Reinforcement Learning Framework for the Generation of Conversational Responses. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 6275–6283.
- Wu, Y.; Schuster, M.; Chen, Z.; Le, Q.V.; Norouzi, M.; Macherey, W.; Reyes, O. Google’s neural machine translation system: Bridging the gap between human and machine translation. *arXiv* **2016**, arXiv:1609.08144.
- Fan, W.; Wei, F.; Liu, Y.; Tian, Q. Hierarchical reinforcement learning for content generation. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 31 October–4 November 2018; pp. 3657–3667.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
- OpenAI. OpenAI GPT-3 Model. 2021. Available online: <https://openai.com/models/gpt-3/> (accessed on 5 February 2023).
- Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Raffel, C. Language Models are Few-Shot Learners. *arXiv* **2020**, arXiv:2005.14165.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language models are unsupervised multitask learners. *OpenAI* **2019**, *8*, 9.
- Bender, E.M.; Friedman, B. Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Trans. Assoc. Comput. Linguist.* **2018**, *6*, 587–604. [CrossRef]
- Friedman, B.; Nathan, L.P.; Yoo, D. Multi-lifespan information system design in support of transitional justice: Evolving situated design principles for the long (er) term. *Interact. Comput.* **2017**, *29*, 80–96. [CrossRef]
- Jernite, Y.; Nguyen, H.; Biderman, S.; Rogers, A.; Masoud, M.; Danchev, V.; Mitchell, M. Data governance in the age of large-scale data-driven language technology. In *2022 ACM Conference on Fairness, Accountability, and Transparency*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 2206–2222.
- Liesenfeld, A.; Lopez, A.; Dingemanse, M. Opening up ChatGPT: Tracking openness, transparency, and accountability in instruction-following text generators. CUI ’23, Eindhoven, July 19–21. *arXiv* **2023**, arXiv:2307.05532.
- Perrigo, B. OpenAI Used Kenyan Workers on Less than \$2 Per Hour: Exclusive. Time, 18 January 2023. Available online: <https://time.com/6247678/openai-chatgpt-kenya-workers/> (accessed on 29 June 2023).
- Alkaiissi, H.; McFarlane, S.I. Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus* **2023**, *15*, e35179. Available online: <https://www.cureus.com/articles/138667-artificial-hallucinations-in-chatgpt-implications-in-scientific-writing> (accessed on 29 June 2023). [CrossRef]
- Azamfirei, R.; Kudchadkar, S.R.; Fackler, J. Large language models and the perils of their hallucinations. *Crit. Care* **2023**, *27*, 1–2. [CrossRef]
- Bender, E.M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S. On the Dangers of Stochastic Parrots. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. *Crit. Care* **2021**, *27*, 610–623. [CrossRef]
- Beutel, G.; Geerits, E.; Kielstein, J.T. Artificial hallucination: GPT on LSD. *Crit. Care* **2023**, *27*, 148. [CrossRef] [PubMed]
- Marcus, G. How Come GPT Can Seem so Brilliant One Minute and so Breathtakingly Dumb the Next? *The Road to AI We Can Trust*. 2022. Available online: <https://garymarcus.substack.com/p/how-come-gpt-can-seem-so-brilliant> (accessed on 29 June 2023).
- Peng, B.; Galley, M.; He, P.; Cheng, H.; Xie, Y.; Hu, Y.; Gao, J. Check your facts and try again: Improving large language models with external knowledge and automated feedback. *arXiv* **2023**, arXiv:2302.12813.
- Aluthman, E.S. The effect of using automated essay evaluation on ESL undergraduate students’ writing skill. *Int. J. Engl. Linguist.* **2016**, *6*, 54–67. [CrossRef]
- Farrokhnia, M.; Banihashem, S.K.; Noroozi, O.; Wals, A. A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innov. Educ. Teach. Int.* **2023**, *8*, 1–15. [CrossRef]
- Kooli, C. Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability* **2023**, *15*, 5614. [CrossRef]
- Rasul, T.; Nair, S.; Kalendra, D.; Robin, M.; de Oliveira Santini, F.; Ladeira, W.J.; Heathcote, L. The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *J. Appl. Learn. Teach.* **2023**, *6*, 1. Available online: <https://journals.sfu.ca/jalt/index.php/jalt/article/view/787> (accessed on 29 June 2023).
- Trust, T.; Whalen, J.; Mouza, C. Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemp. Issues Technol. Teach. Educ.* **2023**, *23*, 1–23.
- Rudolph, J.; Tan, S.; Tan, S. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* **2023**, *6*, 37074. Available online: <https://journals.sfu.ca/jalt/index.php/jalt/article/view/689> (accessed on 29 June 2023).

27. Tajik, E.; Tajik, F. A Comprehensive Examination of the Potential Application of Chat GPT in Higher Education Institutions. 2023. Available online: https://www.techrxiv.org/articles/preprint/A_comprehensive_Examination_of_the_potential_application_of_Chat_GPT_in_Higher_Education_Institutions/22589497/1 (accessed on 29 June 2023).
28. Tlili, A.; Shehata, B.; Adarkwah, M.A.; Bozkurt, A.; Hickey, D.T.; Huang, R.; Agyemang, B. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learn.* **2023**, *10*, 15. [CrossRef]
29. Zhai, X. Chatgpt for next generation science learning. *XRDS Crossroads ACM Mag. Stud.* **2023**, *29*, 42–46. [CrossRef]
30. Moqbel, M.S.S.; Al-Kadi, A.M.T. Foreign Language Learning Assessment in the Age of ChatGPT: A Theoretical Account. *J. Engl. Stud. Arab. Felix* **2023**, *2*, 71–84. [CrossRef]
31. Jiao, W.X.; Wang, W.X.; Huang, J.T.; Wang, X.; Tu, Z.P. Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv* **2023**, arXiv:2301.08745.
32. King, M.R.; ChatGPT. A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cell. Mol. Bioeng.* **2023**, *16*, 1–2. [CrossRef] [PubMed]
33. Mackey, T.P.; Jacobson, T.E. Reframing information literacy as a metaliteracy. *Coll. Res. Libr.* **2011**, *72*, 162–178. [CrossRef]
34. Bruce, C. Informed Learning. Association of College and Research Libraries/American Library Association, Chicago, 2008. Available online: <http://ebookcentral.proquest.com/lib/uunl/detail.action?docID=5888833> (accessed on 29 June 2023).
35. Bent, M.; Stubbings, R. The SCONUL Seven Pillars of Information Literacy: Core Model For Higher Education. SCONUL, 2011. Available online: <https://www.sconul.ac.uk/sites/default/files/documents/coremodel.pdf> (accessed on 29 June 2023).
36. Bishop, L. A Computer Wrote this Paper: What Chatgpt Means for Education, Research, and Writing. *Res. Writ.* **2023**. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4338981 (accessed on 29 June 2023). [CrossRef]
37. Puyt, R.; Lie, F.B.; De Graaf, F.J.; Wilderom, C.P. Origins of SWOT analysis. In *Academy of Management*; Academy of Management: Briarcliff Manor, NY, USA, 2020; p. 17416.
38. King, T.; Freyn, S.; Morrison, J. SWOT analysis problems and solutions: Practitioners' feedback into the ongoing academic debate. *J. Intell. Stud. Bus.* **2023**, *13*, 30–42. [CrossRef]
39. Eloundou, T.; Manning, S.; Mishkin, P.; Rock, D. Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv* **2023**, arXiv:2303.10130.
40. Cox, C.; Tzoc, E. ChatGPT: Implications for Academic Libraries. *Coll. Res. Libr. News* **2023**, *84*, 99. Available online: <https://crln.acrl.org/index.php/crlnews/article/view/25821> (accessed on 29 June 2023). [CrossRef]
41. Khlaif, Z.N. Ethical Concerns about Using AI-Generated Text in Scientific Research. 2023. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4387984 (accessed on 29 June 2023).
42. Cotton, D.R.; Cotton, P.A.; Shipway, J.R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innov. Educ. Teach. Int.* **2023**, *8*, 1–12. [CrossRef]
43. Kikerpill, K.; Siibak, A. App-Hazard Disruption: An Empirical Investigation of Media Discourses on ChatGPT in Educational Contexts. (In Press). Available online: https://advance.sagepub.com/articles/preprint/App-hazard_innovation_An_empirical_investigation_of_media_discourses_on_ChatGPT_in_educational_contexts/22300885 (accessed on 29 June 2023).
44. Khalil, M.; Er, E. Will ChatGPT get you caught? Rethinking of plagiarism detection, 2023. *arXiv* **2023**, arXiv:2302.04335.
45. Li, L.; Ma, Z.; Fan, L.; Lee, S.; Yu, H.; Hemphill, L. ChatGPT in education: A discourse analysis of worries and concerns on social media. *arXiv* **2023**, arXiv:2305.02201.
46. Mahmood, K. Do people overestimate their information literacy skills? A systematic review of empirical evidence on the Dunning-Kruger effect. *Commun. Inf. Lit.* **2016**, *10*, 3. Available online: <https://pdxscholar.library.pdx.edu/comminfolit/vol10/iss2/3> (accessed on 29 June 2023). [CrossRef]
47. Honegger, B.D. Warum Soll Ich Lernen, Was Die Maschine (Besser) Kann? Available online: <http://blog.doebe.li/Blog/> (accessed on 12 March 2023).
48. Balmer, A. Sociological Conversation with ChatGPT about AI Ethics, Affect and Reflexivity. *Sociology* **2023**, *9*, 00380385231169676. [CrossRef]
49. Ashmore, M. *The Reflexive Thesis: Wrihting Sociology of Scientific Knowledge*; University of Chicago Press: Chicago, IL, USA, 1989.
50. Woolgar, S. (Ed.) *Knowledge and Reflexivity: New Frontiers in the Sociology of Knowledge*; Sage: London, UK, 1988.
51. Champagne, M. Chatting with an AI, Chatting with a Human, What's the Difference? Conference Paper. Conference: Philosophers' Jam, Vancouver, Canada, 2023. Available online: https://www.researchgate.net/publication/366958150_Chatting_with_an_AI_Chatting_with_a_Human_What's_the_Difference (accessed on 29 June 2023).
52. Casal, E.; Kessler, M. Can linguist ChatGPT/AI and human writing? A study of research ethics and academic publishing. *Res. Methods Appl. Linguist.* **2023**, *2*, 100068. [CrossRef]
53. Borji, A.; Mohammadian, M. Battle of the Wordsmiths: Comparing ChatGPT, GPT-4, Claude, and Bard., June 12, 2023. Preprint SSRN Electron. J. 2023. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4476855 (accessed on 28 June 2023).
54. Rudolph, J.; Tan, S.; Tan, S. War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *J. Appl. Learn. Teach.* **2023**, *6*, 37074.
55. Ram, B.; Verma, P.V.P. Artificial intelligence AI-based Chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World J. Adv. Eng. Technol. Sci.* **2023**, *8*, 258–261.

56. Guo, B.; Zhang, X.; Wang, Z.; Jiang, M.; Nie, J.; Ding, Y.; Wu, Y. How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection. *arXiv* **2023**, arXiv:2301.07597.
57. Zhang, P. Taking Advice from ChatGPT. *arXiv* **2023**, arXiv:2305.11888.
58. Fraiwan, M.; Hasawneh, N. A Review of ChatGPT Applications in Education, Marketing, Software Engineering, and Healthcare: Benefits, Drawbacks, and Research Directions. *arXiv* **2023**, arXiv:2305.00237.
59. Kas Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Kasneci, G. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **2023**, *103*, 102274. [[CrossRef](#)]
60. Ali, M.J.; Djalilian, A. Chatbots and ChatGPT-Ethical Considerations in Scientific Publications. *Semin. Ophthalmol. Readersh. Aware. Ser.* **2023**, *38*, 403–404. [[CrossRef](#)] [[PubMed](#)]
61. Zhavoronkov, A. Rapamycin in the context of Pascal’s Wager: Generative pre-trained transformer perspective. *Oncoscience* **2022**, *9*, 82–84. [[PubMed](#)]
62. Editorials, N. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature* **2023**, *10*, 612–613. Available online: <https://www.nature.com/articles/d41586-023-00191-1> (accessed on 29 June 2023).
63. Stokel-Walker, C. ChatGPT listed as author on research papers: Many scientists disapprove. *Nature* **2023**, *613*, 620–621. [[CrossRef](#)] [[PubMed](#)]
64. Polonsky, M.J.; Rotman, J.D. Should Artificial Intelligent Agents be Your Co-author? Arguments in Favour, Informed by ChatGPT. *Australas. Mark. J.* **2023**, *31*, 91–96. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.