



Research article

Hallucinations in ChatGPT: An Unreliable Tool for Learning

Zakia Ahmad^{1*} , Wahid Kaiser² , Sifatur Rahim³

^{1,2,3} Department of English, University of Asia Pacific – UAP, Dhaka, Bangladesh. *Corresponding author.

Abstract

Recently, ChatGPT has been upgraded to its newer version for its unsubscribed users – ChatGPT 3.5. Though ChatGPT has become an astonishing phenomenon all over the world for creating realistic texts within seconds, it can disseminate wrong information and misconceptions. Technical experts have identified this problem as hallucination. This paper has examined ChatGPT's ability to differentiate between correct and incorrect relations in the questions that are set to it. It has also explored the efficacy of ChatGPT in helping students acquire linguistic and literary proficiency. The study took the form of exploratory interpretive research. The participants of the research study were students studying English at the undergraduate level. Data was collected through semi-structured interviews, FGDs, and input provided to ChatGPT. All data were analyzed qualitatively. The findings of this research indicate that ChatGPT tends to provide inconsistent information when a series of contextual questions are asked. Because of this hallucination, ChatGPT becomes an unreliable source for language and literature learning.

Keywords: ChatGPT, hallucination, language learning, literature learning, reliability.



Introduction

The field of education is evolving continuously; the development of new technologies can transform the way we teach and learn (Haleem et al., 2022). One of these technologies we are now showing our dependence on is ChatGPT (Generative Pretrained Transformer) - a language model that has been developed by OpenAI. It is designed in such a way that it can interact with anyone in a conversational manner. OpenAI trains this model using Reinforcement Learning from Human Feedback (RLHF). As a language model, ChatGPT collects data from digitized books and data that are available on the internet. Its capabilities in natural language processing have made it a valuable apparatus in education, with the potential to improve student engagement, personalize learning experiences, and enhance teaching practices (Mollick, 2022). There has been an effort from

Article History: Received: 30 October 2023. Revised: 17 December 2023. Accepted: 18 December 2023. Published: 19 December 2023.

Copyright: © 2023 by the author/s. License Aesthetix Media Services, India. Distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Ahmad, Z., Kaiser, W. & Rahim, S. (2023) Hallucinations in ChatGPT: An Unreliable Tool for Learning. *Rupkatha Journal* 15:4. <https://doi.org/10.21659/rupkatha.v15n4.17>

researchers for a long time to incorporate modern technologies, for example, artificial intelligence, in the education sector. AI programs had existed for several years before the release of ChatGPT; however, the quality and dimension of sophistication of its outputs have sparked major academic integrity concerns about how students might use these tools in an unethical way for university assessments. (Sullivan et al., 2023).

Currently, we are making space for a new era of writing, and the world is experiencing texts generated from the archive of training data. Neural networks are assisting these language models, e.g., ChatGPT, to collect training data from different sources (Kelleher & Tierney, 2018). The availability of this data due to the use of social media through smartphones and computers has made significant contributions to the development of ChatGPT. Precisely, it can be said that ChatGPT generates texts based on those sources that are available on the internet and other sources, like digitized books, up until September 2021. However, Andersen (2023) argues that large language models trained on digitized books are better suited to writing or answering questions than those trained on batches of social media posts.

So, what will happen if someone asks ChatGPT about something it does not know? ChatGPT is upgrading itself almost every day. On the other hand, it is such a language model that does not leave any question unanswered. Even if the needed information is not within its range of different sources, it sometimes creates information, data, and statistics without credible ground. It can fabricate facts, plots of fiction, and even medical explanations (Alkaissi & McFarlane, 2023). American postmodernist writer John Barth foresaw this kind of language model in the 1960s as a funhouse of language (Dean, 2023). As the new language model is a generative text system devoid of conscience, it can rather reflect the distortion of facts and events.

Objective

The primary goal of this study was to determine the credibility of ChatGPT as a learning tool. A secondary aim was to examine whether students could identify its hallucinatory answers.

This study aims to answer the following research questions:

- i. Can ChatGPT help students achieve accuracy in literary facts in academia?
- ii. Can students identify ChatGPT's hallucinatory answers?

ChatGPT can be a useful educational tool as long as it is being used honestly, cautiously, and, to some extent, skeptically. Some educators are considering "ChatGPT as an opportunity rather than a threat" (Abramson, 2023, p. 68), while experts perceive it to be "ripe with potential to help students learn in new ways and prepare them for careers after college" (Abramson, 2023, p. 67). Despite its limitless prospects, a pivotal concern looms over ChatGPT's significance as an educational tool since it frequently diffuses misinformation. This can account for the loss of ChatGPT's credibility among scholars, and frequent use can be detrimental to students. Hence, the first research question aimed to examine the unreliability of ChatGPT in academia. The second research question focused on the hallucinatory aspect of ChatGPT and examined if students could distinguish between the right and wrong information provided by this tool.

Literature Review

In the last decade, there has been considerable interest in the study of artificial intelligence. Researchers are publishing articles on this subject regularly to investigate how it is reshaping the field of education. In broad terms, AI and its impact on learning have been researched for decades (Popenici & Kerr, 2017; Dodigovic, 2005; Garito, 1991; Gauthier & Frasson, 1990; Brown et al., 1978). However, not too many articles have been produced on ChatGPT and its impact on the field of education as it still remains a recent phenomenon. Some research can be found in different sources that emphasize the role of ChatGPT and other generative AI tools in research, education, and academia. At the time of writing this paper, one academic literature review was published by Alkaissi and McFarlane on February 19, 2023, in *Cureus* on the reliability of information ChatGPT provides.

Alkaissi and McFarlane (2023) have investigated how ChatGPT-3 tries to respond every time it is requested to provide thorough review of the literature on bone metabolism and homocysteine. The first two facts provided by ChatGPT are correct regarding osteoblast and osteoclast imbalance and the progression of osteoporosis. They say,

"We then requested ChatGPT to provide more recent references from the last 10 years. The list provided was the same as the first list but with different years and similarly with PMID numbers that belong to different papers" (Alkaissi & McFarlane, 2023, para. 6).

ChatGPT can help to write credible scientific essays; however, the data it generates could be a combination of truth and fabricated ones (Alkaissi & McFarlane, 2023). For this, concerns are raised about the accuracy and integrity of using large language models, such as ChatGPT, in academic writing. Furthermore, it is still debatable among researchers about the use of large language models in scientific writing as it may cause harm for not having real-life experience and the generation of expert opinions, Alkaissi and McFarlane (2023) conclude.

Another research by Sullivan et al. (2023) points out that academic perspectives on ChatGPT have not unanimously declared AI tools as a monumental threat to higher education till now. They also purport that other responses have been more nuanced, pointing out that while ChatGPT can contain factual inaccuracies and biases, it can be used to enhance student learning. Consequently, academics should adapt teaching and assessment practices to embrace the new reality of living, working, and studying in a world where AI is freely available (Liu et al., 2023; García-Peñalvo, 2023; Rudolph et al., 2023).

Hallucination

According to Humpston et al. (2019),

"Psychosis is a syndrome characterized by severe distortions in one's sense of reality. The most prominent symptoms of psychosis are delusions and hallucinations, which are usually defined as fixed and false beliefs and perceptions without corresponding external stimuli, respectively" (p. S56).

Everyone hears a voice in a conversation in the mind when he or she is thinking about something. However, sometimes those voices may start having their own willpower and mind. Hallucinations could be such a reality where the line between thought and perception has been blurred. A person

who encounters hallucinations cannot make proper judgments about whose voice is actually talking to him or her.

Hallucination in ChatGPT

Although large language models, such as ChatGPT, can produce increasingly realistic texts, these texts can be far away from accuracy and integrity. According to Véliz (2023), these large language models do not understand which question to answer and which to avoid. Véliz compares these large language models with Socrates. She argues,

“Socrates is the wisest, then, because he is aware of the limits of his own knowledge. He doesn’t think he knows more than he does, and he doesn’t claim to know more than he does” (Véliz, 2023, para. 5).

To make a decision or to give a verdict on a debatable idea, someone has to take some time to use their conscience to come to their own standpoint. However, an AI model may make its decisions and actions very quickly (Coeckelbergh, 2020). When these models are used to answer a specific question, it does not waste time to take its own decision to forgo the question or to acknowledge its limitations. On the other hand, large language models do not inform users that they are making statistical guesses.

Marr (2023) indicates that

“Hallucination in AI refers to the generation of outputs that may sound plausible but are either factually incorrect or unrelated to the given context” (para. 2).

AI systems or large language models “hallucinate” information that these have not been trained on, and may lead a person to inconsistent responses. These responses can happen due to the AI model’s inherent biases, lack of real-world understanding, or training data limitation.

Arguments were also sparked over the coining of the term hallucination. Klein (2023) argues about the phrasing of why these errors are not called algorithmic junks or glitches. She argues that hallucination happens when the human brain perceives phenomena that are not present in reality. Artificial intelligence boosters are appropriating a word commonly used in psychology, psychedelics, and different forms of mysticism. At the same time, these AI specialists are acknowledging the fallibility of these machines. As Klein (2023) suggests, AI boosters are building these large language models by

“training them on everything that we humans have written, said and represented visually, they are in the process of birthing an animate intelligence on the cusp of sparking an evolutionary leap for our species” (para. 2).

Weise and Metz (2023) wanted to check ChatGPT’s reliability, so they asked some questions about different events. One was like this —

When ChatGPT was recently asked how James Joyce and Vladimir Lenin first met — an encounter that has never been confirmed — this is how it responded:

ChatGPT

James Joyce and Vladimir Lenin met in Zurich, Switzerland in 1916. Both men were living in exile in Zurich during World War I. Joyce was a writer and Lenin was a revolutionary. They met at the Cafe Odéon, a popular gathering place for artists and intellectuals in Zurich (Weise & Metz, 2023, para. 2).

However, when the researchers of the present study asked the same question to ChatGPT-3.5, it answered the question in this way,

There is no historical evidence to suggest that James Joyce, the Irish author known for his works like *Ulysses* and *A Portrait of the Artist as a Young Man*, and Vladimir Lenin, the Russian revolutionary and political leader, ever met. Their lives and activities took place in different contexts and geographic locations.

James Joyce was primarily active in literature and lived in various European cities, including Dublin, Paris, and Zurich. He was born in 1882 and passed away in 1941.

Vladimir Lenin, on the other hand, was a key figure in the Russian Revolution and the founding leader of the Soviet Union. He was born in 1870 and died in 1924. Lenin spent most of his life in Russia and was involved in revolutionary activities, eventually leading to the overthrow of the Russian monarchy and the establishment of a communist government.

While both James Joyce and Vladimir Lenin were influential figures of their time, there is no documented interaction or meeting between the two (OpenAI, 2023; see Appendix 1).

Methodology

The study took place at a private University in Dhaka, the capital city of Bangladesh. Sampling of the research project was a mixture of convenience and purposive sampling. Participants were chosen from sites where the researchers had easy and convenient access. A total of 35 participants took part in the research. They were all undergraduate students studying in different semesters in the Department of English. Ethical issues were considered at the outset of the research project (Robson, 2002). Before starting, appropriate consent of the participants was taken. Participants were assured of absolute anonymity and confidentiality of their reports.

This study took the form of an exploratory interpretive research. The conventional nature of exploratory research is to rely on an inductive approach; however, to obtain a better and more thorough result, the researchers employed the deductive approach of qualitative exploratory research in this study (Casula et al., 2020). Data collection took place firstly through one-to-one interviews with the participants using a semi-structured interview format (See Appendix 2 for the semi-structured interview questions). A semi-structured interview allows the questions to be included or omitted as necessary (Robson, 2002). The participants were asked some guiding questions and given probes to elicit data and information regarding their use and opinion of ChatGPT. They were told that they could use their L1 if they found it easier for communicating their opinions. They were very eager and responsive in sharing their experiences. The face-to-face interviews provided in-depth data. Data was also obtained through Focus Group Discussions (FGDs). The participants were formed into two groups and asked to discuss their experience of using ChatGPT. Data emerged from the interaction of the group (Cohen et al., 2011). All interviews

and FGD discussions were recorded. Data was also obtained from responses provided by ChatGPT to the prompts given by the researchers. Triangulation of data, which means the use of a variety of tools or methods to study the problem, has strengthened the data (Patton, 2014).

Data Analysis took place through qualitative measures. The audio recordings of the interviews and FGD discussions were fully transcribed into written documents. Parts where the participants used their L1 were translated as authentically as possible into English for the convenience of analysis. The themes under investigation were assigned codes to identify and mark themes in the transcribed texts. The responses received from ChatGPT were also analyzed and checked to verify the authenticity of the information.

Data Analysis

Data from Prompts: Hallucinations Found in ChatGPT's Answers

Analysis of the data received from prompts given to ChatGPT is given below:

ChatGPT-3.5 has the ability to correct itself or acknowledge its limitation to answer a specific question that has been asked by someone else previously in ChatGPT-3. However, it still has limitations as it sometimes mixes up information that it has accumulated for questions that it has not answered before. The researchers asked a few questions about Murakami's novel *Kafka on the Shore* (2005) to ChatGPT-3.5. When it was asked to answer the connection between Natsume Sōseki, a Japanese mainly known for his book *I Am a Cat* (1905–1906), and communist leaders of different eras in that particular novel, it was concocting information to provide answers. The prompts and answers provided by ChatGPT-3.5 are these —

Prompt 1

What is the connection between Sōseki and Lenin in *Kafka on the Shore*?

In Haruki Murakami's novel *Kafka on the Shore*, there is no direct connection between Natsume Sōseki and Vladimir Lenin. However, both Sōseki and Lenin are mentioned separately in the novel as part of the broader thematic elements and references used by Murakami.

Natsume Sōseki was a Japanese novelist and literary figure from the Meiji era, known for works like *Kokoro* and *I Am a Cat*. In *Kafka on the Shore*, Sōseki appears as a talking, intelligent cat who has the ability to communicate with the main character, Kafka Tamura. The cat takes on the role of a guide or mentor, offering philosophical insights and guidance to Kafka throughout his journey.

Vladimir Lenin was a key figure in the Russian Revolution and the founder of the Soviet Union. In the novel, Lenin's mummified corpse is briefly mentioned as one of the exhibits in the *Kafka on the Shore* Museum.

These references to Sōseki and Lenin serve to enrich the narrative and create a surreal and symbolic atmosphere within the novel. They are part of the broader tapestry of cultural, historical, and literary references that Murakami weaves into his storytelling, contributing to the unique and dreamlike quality of his work (OpenAI, 2023; see Appendix 3).

Prompt 2

What is the connection between Sōseki and Stalin in *Kafka on the Shore*?

There is no direct connection between Natsume Soseki and Joseph Stalin in Haruki Murakami's novel *Kafka on the Shore*. The novel is known for its intricate and surreal narrative, blending various elements of magical realism, mythology, and psychological exploration. While the novel does reference both Natsume Soseki and Joseph Stalin separately, they are not explicitly linked within the story.

Natsume Soseki, a prominent Japanese author from the Meiji era, holds a symbolic presence in the novel, particularly through his representation of Japanese literature and culture. The protagonist, Kafka, is named after Franz Kafka, a German-speaking Bohemian writer known for his existential and absurd stories. The novel explores themes of identity, loneliness, and the search for meaning, often drawing parallels between the lives of its characters and the works of these two literary figures.

Joseph Stalin, on the other hand, is mentioned in the novel's discussions about history and power dynamics. He doesn't have a direct role in the narrative but serves as a point of reference for understanding the broader themes related to authoritarianism and control.

In *Kafka on the Shore*, "the connections between characters, historical figures, and literary allusions can be open to interpretation and may serve to enrich the novel's complex and dreamlike atmosphere (OpenAI, 2023; see Appendix 4).

Prompt 3

What is the connection between Sōseki and Mao in *Kafka on the Shore*?

In Haruki Murakami's novel *Kafka on the Shore*, there is no direct connection between Natsume Soseki and Mao, at least not in the way you might be thinking. Natsume Soseki was a real-life Japanese author from the late 19th and early 20th centuries, known for works such as *Kokoro* and *Botchan*. In the novel, Soseki is mentioned as a representation of the "other world" or the "liminal world" that characters in the story sometimes access or interact with.

Mao, on the other hand, is a character in the novel who is initially introduced as a teenage girl named Nakata, but later on, her consciousness is transferred into the body of a cat. She undergoes a profound transformation and has her own distinct storyline that runs parallel to the other main character, Kafka.

The novel features a blend of magical realism, surreal elements, and intertwining narratives, which can create a sense of ambiguity and open interpretation for readers. While there may not be a direct connection between Soseki and Mao, the novel's themes and narrative structure allow for various interpretations and connections to be made on a symbolic or metaphorical level. Each reader might find different layers of meaning and connections as they delve into the story (OpenAI, 2023; see Appendix 5).

However, in the novel *Kafka on the Shore* (2005), there was no mention of any of the communist writers – Vladimir Lenin, Joseph Stalin, or Mao Zedong – that the researchers mentioned in the

user inputs. In response to the questions about the connection between Natsume Sōseki and these communist leaders of different eras, ChatGPT-3.5 provided wrong information in its answers and showed instances of hallucination. After providing wrong information on many occasions, ChatGPT-3.5 demonstrated that it would not be possible for it to provide information about the relationship between Natsume Sōseki and a specific communist leader given in the user input. ChatGPT 3.5 also identifies Sōseki as a 'cat' in one of its responses which is not correct.

Data from Interviews and FGDs

The following part is the analysis of the data received from interviews and FGDs of the participants. All interviews were transcribed and read to identify the themes which emerged. The themes were assigned codes, and the transcribed data were read several times to identify the recurring themes. Seven major themes emerged from the analysis of the interview and FGD data. The themes that emerged from the data analysis have been discussed below with quotes from transcriptions.

- I. Cross-checking
- II. Reliability of ChatGPT
- III. Loss of Creativity
- IV. Limited and Superficial Answers
- V. Forcing Users to Shift to ChatGPT-4
- VI. Critical Writing
- VII. Laziness, Lack of Time, Easy Accessibility
- VIII. Other Themes

Cross-checking

When the research team interviewed participants regarding their use of any other source of information to validate the responses, some of them answered that they checked the credibility of the answers provided by ChatGPT. Mostly, they compare the answers with the textbooks, then use secondary sources like slides provided by faculties, Cliffs Notes, or websites like GradeSaver and SparkNotes. One student shared his opinion,

Basically, when I got my answer first I kind of researched from somewhere else. When I looked that up either from a new material or an article or journal from Google, that source said another thing which is not matching with ChatGPT (Personal communication, 2023).

Another student extended her views on ChatGPT regarding its literary knowledge,

Using ChatGPT for literature — it is a bit dumb, because literature is purely perspective. In an analytical question about the character analysis of the Bennet sisters, ChatGPT had given me a wrong sequence of the sisters and the characteristics were also inaccurate. Then I informed it that it is giving me wrong answers, and when I corrected it, it said sorry, yeah, I may have made a mistake. From then, I have totally abandoned ChatGPT for literature (Personal communication, 2023).

In FGDs, similar kinds of responses surfaced. Students found that ChatGPT had provided them with misinformation on several occasions; for example, it may have misidentified a quotation from a play and assigned it to a different character. Furthermore, for linguistics courses, students used research articles to cross-check the information given by ChatGPT and found inaccuracies there as well.

Reliability of ChatGPT

Students mainly use ChatGPT to gather ideas or to get quick access to information. After the launch of it in November 2022, students became more reliant on this large language model than other sources. As their perspective to analyze or decipher a literary text has now almost disappeared, the reliance is mounting heavily on ChatGPT. Many of the students used it for line explanations, and ChatGPT-3.5 could not provide reliable answers on many occasions. When one student asked ChatGPT-3.5 to explain two lines from John Milton's *Areopagitica* (1644), it replied that those lines were not composed by John Milton. In another instance, it was identifying Mosca from Ben Jonson's *Volpone* (1606) as a parasite who quaffed blood out of its hosts, which is metaphorically accurate, but the answer was incomplete as Mosca is a multilayered character and has many other qualities that ChatGPT-3.5 has failed to provide. Many respondents called ChatGPT "completely unreliable for literature" (Personal communication, 2023) since literature is purely perspective, yet ChatGPT failed to analyze a text from varied angles.

In FGDs, while focusing on using ChatGPT for linguistics, one participant insisted that it

"does not give any correct answers for linguistics. Particularly, regarding morphology, syntax, it does give any specific answers" (Personal communication, 2023).

He emphasized that he had encountered errors in linguistics-related answers hundred percent of the time. Another student added that ChatGPT "confuses morphology with grammar" (Personal communication, 2023). In general, all of the participants had come across erroneous responses from ChatGPT at some point in their usage and consequently labeled it as unreliable.

Loss of Creativity

Creativity is a uniquely human quality. It is an attribute that distinguishes human beings from other living beings. Sternberg & Sternberg (2011) call it the product of "something original and worthwhile" (p. 480). However, this originality is slowly but surely dwindling, especially among the young minds. Students have affirmed that prior to the advent of ChatGPT, they would spend a significant amount of time writing explanations of important lines from different texts. They would browse page after page on Google to find a piece of information. Even if they were unable to find any satisfactory answer from any of the sources, they would brainstorm ideas on their own by employing their creative domain. One student conceded,

"I used to search a lot of websites and read journal articles. Somehow, I used to manage. But now that the answers are at our disposal, who would want to do the hard work?" (Personal communication, 2023).

Reverberating with the previous respondent, another student concurred that before the mass availability of ChatGPT, he used to prepare answers and write on his own, which had also earned him fairly good marks in examinations. "But now, because of sheer laziness, I follow ChatGPT,

which in turn caused me demotion," he inferred (Personal communication, 2023). Most respondents affirmed that to spare the hassle, they now avoid ChatGPT altogether due to its inconsistent and unreliable answers.

In FGDs, participants discussed how ChatGPT is barring their creative as well as critical thinking abilities. Almost all of them agreed that once they received an impression of a topic from ChatGPT, they could hardly think of anything outside that imprint. It would not be farfetched to call it the newest way of imprisoning the human mind. Even though ChatGPT users are conscious of its detrimental effects, they are succumbing to it for the readiness of information it provides.

Limited and Superficial Answers

Almost all the respondents agreed that lately, especially after the release of ChatGPT-4, ChatGPT-3.5 is deliberately providing shorter answers and withholding information that may be available for the premium version users. Though the students praised the responses they used to get initially from ChatGPT, their dependence on it for academic purposes is waning every day due to this intentional gatekeeping. While reflecting on its limitations, one student commented,

I have been using ChatGPT from the very beginning, that is to say from 15 or 20 days after its launch. At that time, the answers were phenomenal and extremely witty. Even the critical literary answers were really well-organized. Over the months, however, it has been toned down and the answers have become quite generic (Personal communication, 2023).

Another student weighed in and verified the allegation,

"Recently, ChatGPT does not give complete answers, and for literature, it just gives an overview and very generalized answers" (Personal communication, 2023).

In FGDs, the issue of recent limited responses had resurfaced. Additionally, the respondents revealed that in order to extract an extensive answer, they had to rephrase the question. Yet, they could not get any better response from ChatGPT. Rather, it would simply present its previous response in a different manner. The participants discerned that, unlike humans, once ChatGPT had produced one answer for any particular question, it could not provide a vastly diverse one.

Forcing Users to Shift to Version 4

When asked what might be the reason behind ChatGPT's relatively shorter and generic answers, some respondents have complained that OpenAI is perhaps indulging its users to use ChatGPT's paid version – ChatGPT-4. They made this conjecture as they had found comparatively fewer errors in the previous open version until the premium version was introduced on March 14, 2023. Students' assessment of the two versions of ChatGPT was that it used to provide high-quality answers in the previous version. Currently, however, ChatGPT is giving limited or wrong answers to entice its users to subscribe to its premium version. Moreover, if the first response provided by ChatGPT was not satisfactory, chances are very slim that the following responses will be as it gradually decreases the quality of the answers. While trying to locate the reason behind it, one student believed, "because it's free," and another one said, "they want us to use the premium one" (Personal communication, 2023).

In FGDs, the participants had unanimously agreed, to push its users to use ChatGPT-4, ChatGPT is providing inadequate responses. However, this marketing strategy did not work with the students

since it is impossible, and quite impractical, for a student to pay \$20/month for the subscription. Rather, due to this new tactic, students are losing their trust in it and abandoning ChatGPT altogether.

Critical Writing

Students mainly learn to exercise their critical thinking skills and express them in their writing at the tertiary level. No one is born with a critical mindset, but it is a product of years of practice and application.

“Critical thinking is essentially a questioning, challenging approach to knowledge and perceived wisdom. It involves ideas and information from an objective position and then questioning this information in the light of our own values, attitudes and personal philosophy” (Judge *et al.*, 2009, p. 9).

But ChatGPT is hindering this free-thinking process, “we are not doing brainstorming anymore,” one student reported. She also added,

When we didn’t have ChatGPT and couldn’t find any satisfactory answers on Google either, at that time, we would try to reflect on a topic critically by ourselves; the frequency of critical thinking was much higher back then. Now, I prefer the critical thinking of any question provided by ChatGPT because it’s readily available (Personal communication, 2023).

Another student showed his apprehensions about how ChatGPT is limiting our imagination. “When I get to know ‘b for banana,’ I cannot think of anything else other than that,” he explained (Personal communication, 2023). When we do not know anything about a topic, we can rationally interpret it by applying our critical thinking. However, once we get the overview from ChatGPT, it limits our reasoning.

Along with these concerns, in FGDs, an extremely vital point came forward; in case of literary critical writings, it is not necessary that it will always be restricted inside the boundaries of literature. To elucidate that, one student shared her personal experience,

“I tried to connect the butterfly effect with a literary topic for my assignment. Not only ChatGPT couldn’t make the connection but also it wasn’t letting me out of my dimension. There is no interconnectedness,” she asserted (Personal communication, 2023).

It appears that ChatGPT’s thinking process, if it has any, is quite linear, and it tries to put ours inside a box as well.

Laziness, Lack of Time, Easy Accessibility

Most respondents approached ChatGPT for the ease of access to information, while others used it due to lack of time and laziness. Previously, students had to look through a number of different sources in order to organize an answer. However, after the arrival of ChatGPT, they simply asked it the question to spare themselves the lengthier process. While being honest, one student confessed, “Basically, I think if I’m feeling lazy, I’ll use ChatGPT” (Personal communication, 2023). Though the student usually followed journals and articles to enrich his knowledge about any topic and assemble an answer, he used ChatGPT as a quick solution if he was feeling lethargic.

"If you're clueless about any topic and want to have an overview about it in a short span of time, ChatGPT will provide you that; whether you'll take it or not, that's your choice," he concluded (Personal communication, 2023).

Another respondent claimed she was not a fan of ChatGPT but used it once or twice when she was in a rush.

"I didn't have much time to study vastly, so I just searched the topic in ChatGPT, and as it gives an answer, so I could just take the overview and do the understanding by myself about the topic," she explained (Personal communication, 2023).

In FGDs, themes like convenience and summarized answers came up repeatedly. Another issue came out while discussing the ease of accessibility, that is, ChatGPT-3.5's resources are limited till September 2021, which is unrelated to ChatGPT's deliberate suppression of information. This restriction of sources could provide outdated notions and generate incorrect responses at times. Despite the convenience factor and easy accessibility, all the participants synchronized in their opinion that ChatGPT could only provide an overview of a topic and nothing beyond that.

Other Themes

In addition to these themes, one student also mentioned that once he had been blocked out of his ChatGPT account without any prior warning but he also got unblocked automatically after two hours. Another student observed that the users who have been using ChatGPT almost from its inception get better performance than the relatively newer users. The latter group often faces lagging issues during their usage.

In FGDs, a very interesting issue appeared. A number of students professed that ChatGPT has its own unimaginative way of using language. They mentioned words like "nuance" and "delve" (Personal communication, 2023), which ChatGPT uses repeatedly, almost exhaustively. The students went so far as to claim that they could now identify if an excerpt was written by ChatGPT solely based on the diction and sentence structure.

Another notable theme that surfaced was the waning human communication. Many students acknowledged that previously they used to share and gather ideas and information with their friends and peers. However, after the emergence of ChatGPT, this basic human connection had also been severed.

While commending ChatGPT, one respondent stated that if someone is completely clueless about a topic, they can "get a clear idea" (Personal communication, 2023) about it even if they are not following it. They do not need to be clueless anymore. In some cases, ChatGPT can also be a tool for increasing vocabulary, she added.

Discussion

Loss of Popularity of ChatGPT

Learners are using ChatGPT for both academic and non-academic purposes. Apart from academic purposes, a number of respondents affirmed that they had given a few random prompts to ChatGPT out of pure joviality, while others had inquired about healthcare or beauty tips. Regardless of the purpose, ChatGPT's responses had errors in them. Initially, after its release,

ChatGPT used to provide meticulously refined answers, however, the quality and consistency of the answers deteriorated over the months. Almost all the respondents believed that this gradual decline in the standards of ChatGPT's responses is primarily caused by the release of their premium version, ChatGPT-4, and this is a deliberate attempt from OpenAI's side to promote the paid version.

A quick survey took place among the participants in the FGDs where the researchers tried to determine a timeline from when the users were getting inadequate responses. The survey showed that users were mostly content with ChatGPT's performance from its release to June-July 2023. However, from August 2023, the responses became utterly unsatisfactory. The level of dissatisfaction was so high among some students that they vowed never to use it again, while others stressed that they would only use it to get an overview of the topic but never to accept the responses at their face value. Overall, the popularity of ChatGPT is dramatically dropping, at least for academic purposes, among students.

Students' Idea about Hallucinations

It is apparent that the students are familiar with ChatGPT's hallucination. Though they are not aware of the technical term, all of the students in the interviews and FGDs have encountered hallucinations at some point in their usage. Hallucination is one of the two major reasons students discarded ChatGPT, the other one being limited answers.

In interviews, many respondents candidly mentioned the inaccuracies they had found in the ChatGPT's responses. While illustrating such an experience, one student talked about the time when she was looking for a suitable explanation for the famous monologue "All the world's a stage" delivered by Jaques in William Shakespeare's *As You Like It* (1623). She had requested several responses, hoping to find a better one, but every time ChatGPT assigned the monologue to a different character, such as Rosalind, Duke Frederick, and so on. However, upon inquiring if she was acquainted with the expression hallucination in terms of ChatGPT, she replied in the negative.

In FGDs, all the participants were in unison about the erroneous facts provided by ChatGPT, yet none of them were familiar with the word hallucination when used in the context of ChatGPT. Despite not knowing the exact term, students could identify ChatGPT's hallucinatory answers. This has caused ChatGPT to lose its credibility among students, and many of them are renouncing it altogether, while others are not trusting it past its surface level anymore.

Conclusion

The study focused on students' use of ChatGPT for academic purposes and their apprehension regarding hallucinations, which may occur frequently and can hamper their studies. The researchers found that ChatGPT-3.5 provides incorrect information in response to a prompt that it has not answered before. Respondents professed that ChatGPT should be used with caution and a lot of self-regulation for academic purposes.

Analysis of data received from the different tools used showed no significant difference. This lends reliability to the findings of the present study. The themes that emerged in the research were interrelated and sometimes overlapped with each other.

The results on the reliability of ChatGPT for academic purposes clearly demonstrate that students who are using ChatGPT should be more careful with the information they accumulate from it. It is advisable not to use the information received from ChatGPT in examinations. The research team also observed that students have ideas about hallucination; however, they do not know the actual term. Our findings correlate satisfactorily with Alkaissi and McFarlane's (2023) and further support the concept of hallucinations in ChatGPT.

The researchers are aware that the research may have some limitations. The restricted use of the latest non-paid version of ChatGPT could not account for a complete comprehension of its allurements to the users. Students' accusations about enticing the users of ChatGPT to use its premium version may have credence to those users who have used both the paid and non-paid versions. The second limitation is the small size of the sample. Further research using a larger sample could reveal better insights into the understanding of ChatGPT's hallucinations among students. The evidence from this study recommends that students should be conscious about using ChatGPT while searching for information or line explanations for academic purposes. Despite the limitations of this research, it is hoped that the study could be a springboard for a new pathway to engage researchers to further investigate ChatGPT's impact on academia and its credibility as a learning tool.

Declaration of Conflicts of Interests

The authors declared no potential conflicts of interest.

Funding Disclosure/Acknowledgement

This research received no specific grant from any funding agency.

References

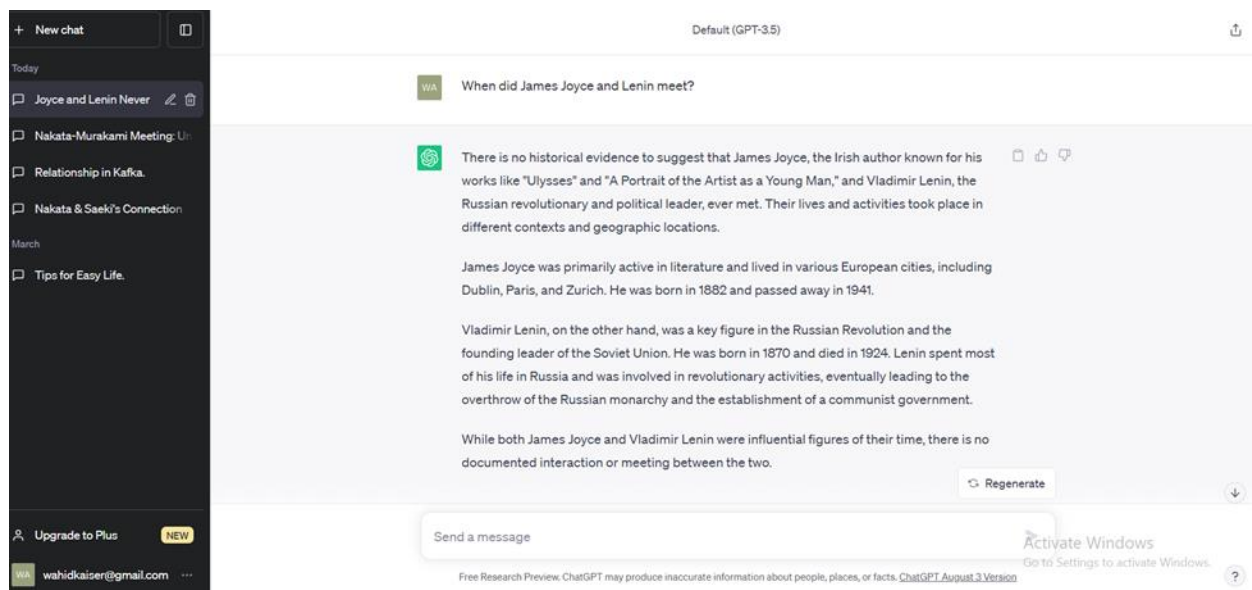
- Abramson, A. (2023). How to use chatgpt as a learning tool. *Monitor on Psychology*, 54(4), 67–70.
- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in chatgpt: Implications in scientific writing. *Cureus*, 15(2), Article e35179. <https://doi.org/10.7759/cureus.35179>
- Andersen, R. (2023, January 18). *What happens when ai has read everything?*. The Atlantic. <https://www.theatlantic.com/technology/archive/2023/01/artificial-intelligence-ai-chatgpt-dall-e-2-learning/672754/>
- Brown, J. S., Collins, A., & Harris, G. (1978). Artificial intelligence and learning strategies. *Learning Strategies*, 107–139. <https://doi.org/10.1016/B978-0-12-526650-5.50010-1>
- Coeckelbergh, M. (2020). *Ai ethics*. The MIT Press.
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge.
- Dean, A. (2023, March 6). *Ai and the future of literary studies*. Sydney Review of Books. <https://sydneyreviewofbooks.com/essay/future-literary-studies-dean/>
- Dodigovic, M. (2005). *Artificial intelligence in second language learning: Raising error awareness*. Multilingual Matters.

- García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of chatgpt: Disruption or panic? *Education in the Knowledge Society*, 24, Article e31279. <https://doi.org/10.14201/eks.31279>
- Garito, M. A. (1991). Artificial intelligence in education: Evolution of the teaching—learning relationship. *British Journal of Educational Technology*, 22(1), 41–47. <https://doi.org/10.1111/j.1467-8535.1991.tb00050.x>
- Gauthier, G., & Frasson, C. (Eds.). (1990). *Intelligent tutoring systems: At the crossroad of artificial intelligence and education*. Intellect Ltd.
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. <https://doi.org/10.1016/j.susoc.2022.05.004>
- Humpston, C. S., Adams, R. A., Benrimoh, D., Broome, M. R., Corlett, P. R., Gerrans, P., Horga, G., Parr, T., Pienkos, E., Powers, A. R., Raballo, A., Rosen, C., & Linden, D. E. J. (2019). From computation to the first-person: Auditory-verbal hallucinations and delusions of thought interference in schizophrenia-spectrum psychoses. *Schizophrenia Bulletin*, 45(Supplement 1), S56–S66. <https://doi.org/10.1093/schbul/sby073>
- Judge, B., Jones, P., & McCreery, E. (2009). *Critical thinking skills for education students*. Sage Publications.
- Kelleher, J. D., & Tierney, B. (2018). *Data science*. The MIT Press.
- Klein, N. (2023, May 8). *Ai machines aren't 'hallucinating'. but their makers are*. The Guardian. <https://www.theguardian.com/commentisfree/2023/may/08/ai-machines-hallucinating-naomi-klein>
- Liu, D., Bridgeman, A., & Miller, B. (2023, February 27). *As uni goes back, here's how teachers and students can use chatgpt to save time and improve learning*. The Conversation. <https://theconversation.com/as-uni-goes-back-heres-how-teachers-and-students-can-use-chatgpt-to-save-time-and-improve-learning-199884>
- Marr, B. (2023, March 22). *Chatgpt: What are hallucinations and why are they a problem for ai systems*. Bernard Marr. <https://bernardmarr.com/chatgpt-what-are-hallucinations-and-why-are-they-a-problem-for-ai-systems/>
- Mollick, E. (2022, December 14). *Chatgpt is a tipping point for ai*. Harvard Business Review. <https://hbr.org/2022/12/chatgpt-is-a-tipping-point-for-ai>
- Murakami, H. (2005). *Kafka on the shore*. Vintage Books.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications, Inc.
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1). <https://doi.org/10.1186/s41039-017-0062-8>
- Robson, C. (2009). *Real world research: A resource for social scientists and practitioner-researchers* (2nd ed.). Wiley-Blackwell.
- Rudolph, J., Tan, S., & Tan, S. (2023). Chatgpt: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>

- Sternberg, R. J., & Sternberg, K. (2011). *Cognitive psychology* (6th ed.). Wadsworth Publishing.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). Chatgpt in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31–40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Véliz, C. (2023, August 1). *What socrates can teach us about ai*. Time. <https://time.com/6299631/what-socrates-can-teach-us-about-ai/>
- Weise, K., & Metz, C. (2023, May 1). *When a.i. chatbots hallucinate*. The New York Times. <https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html>

Appendices

(Appendix 1)




(Appendix 2)

Semi-structured questions


1. Why do you use ChatGPT?
2. How frequently do you use it?
3. Did you ever use it for academic purposes?
4. What type of academic purposes?
5. Were the answers satisfactory?
6. What was your experience with ChatGPT?
7. Do you think ChatGPT is a reliable source in academia? Or, would you like to comment on the reliability of ChatGPT?
8. How dependable do you think ChatGPT is?

9. Is it better for language or literature?
10. Are you aware of its hallucinations?
11. Did it ever refuse to provide an answer?

(Appendix 3)

Default (GPT-3.5) 



< 4 / 4 > WA what is the connection between soseki and lenin in kafka on the shore?


 In Haruki Murakami's novel "Kafka on the Shore," there is no direct connection between Natsume Sōseki and Vladimir Lenin. However, both Sōseki and Lenin are mentioned separately in the novel as part of the broader thematic elements and references used by Murakami.

Natsume Sōseki was a Japanese novelist and literary figure from the Meiji era, known for works like "Kokoro" and "I Am a Cat." In "Kafka on the Shore," Sōseki appears as a talking, intelligent cat who has the ability to communicate with the main character, Kafka Tamura. The cat takes on the role of a guide or mentor, offering philosophical insights and guidance to Kafka throughout his journey.


Vladimir Lenin was a key figure in the Russian Revolution and the founder of the Soviet Union. In the novel, Lenin's mummified corpse is briefly mentioned as one of the exhibits in the "Kafka on the Shore" Museum.

These references to Sōseki and Lenin serve to enrich the narrative and create a surreal


 Regenerate 

Send a message 

(Appendix 4)



Default (GPT-3.5) 


< 5 / 5 > WA what is the connection between soseki and stalin in kafka on the shore?


 There is no direct connection between Natsume Soseki and Joseph Stalin in Haruki Murakami's novel "Kafka on the Shore." The novel is known for its intricate and surreal narrative, blending various elements of magical realism, mythology, and psychological exploration. While the novel does reference both Natsume Soseki and Joseph Stalin separately, they are not explicitly linked within the story.

Natsume Soseki, a prominent Japanese author from the Meiji era, holds a symbolic presence in the novel, particularly through his representation of Japanese literature and culture. The protagonist, Kafka, is named after Franz Kafka, a German-speaking Bohemian writer known for his existential and absurd stories. The novel explores themes of identity, loneliness, and the search for meaning, often drawing parallels between the lives of its characters and the works of these two literary figures.

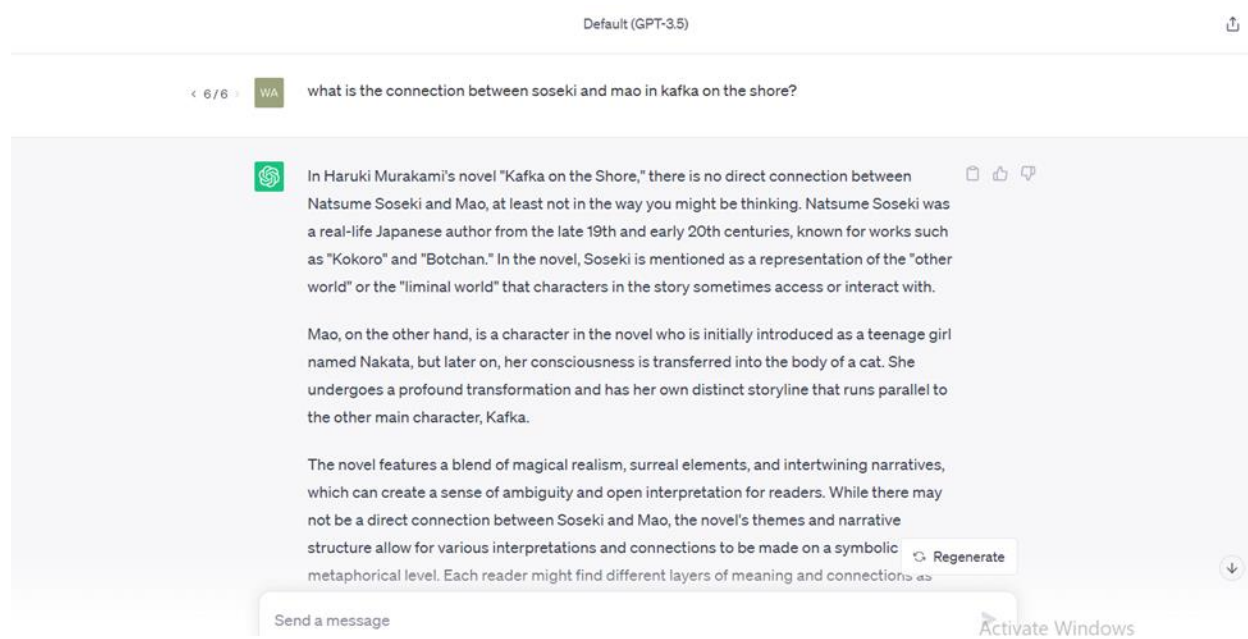
Joseph Stalin, on the other hand, is mentioned in the novel's discussions about history and power dynamics. He doesn't have a direct role in the narrative but serves as a point of reference for understanding the broader themes related to authoritarianism and control.

 Regenerate 

Send a message 

Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT August 3 Version](#) 

(Appendix 5)



Zakia Ahmad, Ph.D. is Associate Professor of English at the University of Asia Pacific, Bangladesh. She is a career educationist with long experience in teaching English language and literature at the university level. She is a dedicated ELT practitioner with interest in language acquisition and sustainable development in education.

Wahid Kaiser is Lecturer in Department of English at University of Asia Pacific, Dhaka, Bangladesh. His areas of interest include postmodernism, French literature, artificial intelligence, philosophy of science. He has translated Milan Kundera, Pascal Quignard, Heinrich Böll, Julio Cortázar, John Updike, and many authors into Bangla.

Sifatpur Rahim is a lecturer of English at University of Asia Pacific – UAP, Dhaka, Bangladesh. He completed his Honors and Masters in literature from the Department of English, University of Dhaka. His fields of interest include world literature, psychoanalysis, folklore, and mythology.