

Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study

Marco Antonio Rodrigues Vasconcelos^{1,2} , Renato P. dos Santos^{3*} 

¹ PPGECIM—Doctoral Program in Science and Mathematics Education ULBRA—Lutheran University of Brazil, Canoas, RS, BRAZIL

² Secretaria Municipal de Educação, Fortaleza, CE, BRAZIL

³ CIAGE—Center for Generative Artificial Intelligence in Cognition and Education ULBRA—Lutheran University of Brazil, Canoas, RS, BRAZIL

Received 30 April 2023 • Accepted 17 May 2023

Abstract

This study investigates the potential of ChatGPT and Bing Chat, advanced conversational AIs, as “objects-to-think-with,” resources that foster reflective and critical thinking, and concept comprehension in enhancing STEM education, using a constructionist theoretical framework. A single-case study methodology was used to analyze extensive interaction logs between students and both AI systems in simulated STEM learning experiences. The results highlight the ability of ChatGPT and Bing Chat to help learners develop reflective and critical thinking, creativity, problem-solving skills, and concept comprehension. However, integrating AIs with collaborative learning and other educational activities is crucial, as is addressing potential limitations like concerns about AI information accuracy and reliability of the AIs’ information and diminished human interaction. The study concludes that ChatGPT and Bing Chat as objects-to-think-with offer promising avenues to revolutionize STEM education through a constructionist lens, fostering engagement in inclusive and accessible learning environments.

Keywords: ChatGPT, Bing Chat, generative artificial intelligence, artificial intelligence in education, objects-to-think-with

INTRODUCTION

The focus on developing critical thinking and collaborative skills has spurred interest in innovative educational enhancement approaches. Technology-based tools, especially in STEM education, are gaining attention. AI and NLP advancements have created opportunities for integrating technology into educational settings, particularly in STEM (Okonkwo & Ade-Ibijola, 2021).

Khosrawi-Rad et al. (2022) emphasize chatbots’ educational advantages, such as constant availability and scalability. Kuhail et al. (2023) suggest that chatbots can revolutionize education through engagement, personalization, educator support, and learner insight, particularly in STEM, where tailored support is beneficial.

In this context, generative AI-powered chatbots (GenAIBots), with OpenAI’s ChatGPT emerging as a

prime example and Microsoft’s Bing Chat coming soon after, have recently garnered significant interest as educational tools in STEM learning (Baidoo-Anu & Owusu Ansah, 2023). Being capable of maintaining a convincing conversation with a user in natural language, those GenAIBots have emerged as promising tools to complement traditional pedagogical practices.

Probably, the most significant impact of ChatGPT is that no knowledge of any programming language is required; one ‘talks’ to it in one’s native language, and it responds in one’s—or any other language of one’s choice! This is undoubtedly one of the main factors that made ChatGPT bring 100 million lay users to AI in just two months after launch (Milmo, 2023), with students making up a good portion of them. This seems a realization of Papert’s (1980) vision that learning to communicate with a computer might change the way other learning takes place, as the computer could be both a mathematics-speaking and an alphabetic-speaking

Contribution to the literature

- This study fills a gap in current literature by investigating ChatGPT and Bing Chat's roles in enhancing STEM learning experiences. The application of ChatGPT and Bing Chat as objects-to-think-with, resources that foster reflective and critical thinking, problem-solving skills, and concept comprehension, is examined within the context of STEM education, offering a novel perspective to the field.
- Logs of students' extensive interactions with ChatGPT and Bing Chat are presented and analyzed, providing valuable data on AI-assisted STEM learning experiences.
- ChatGPT's capacity to transform STEM education by fostering critical thinking, problem-solving, and engagement in inclusive and accessible learning environments is highlighted.

entity and when this communication occurred, children could learn mathematics as a living language (p. 6).

OpenAI's ChatGPT¹, based on GPT NLP techniques, generates human-like responses using extensive internet text data (Radford et al., 2019). Microsoft's Bing Chat², built on GPT-4, is a revamp of traditional search engines, functioning as both search engine and chatbot (Mehdi, 2023). It accesses the internet directly, providing up-to-date, well-informed responses.

GPT-4 (OpenAI, 2023) showcases impressive language generation and user interaction capabilities, prompting an investigation into using GenAIBots as objects-to-think-with in STEM education. Objects-to-think-with stimulate cognition, reflection, and understanding of complex concepts. This study examines GenAIBots' efficacy in enhancing STEM learning. GenAIBots can shape learners' understanding as active agents, enabling exploration and reflective dialogues with ChatGPT and Bing chat, generating new insights. As will be seen later in this paper, those GenAIBots can work as instances of the concept of "object-to-think-with," resources that foster reflective and critical thinking and concept comprehension, as introduced by Papert (1980, p. 11).

This paper showcases a case study that primarily investigates whether GenAIBots can effectively serve as objects-to-think-with, substantially enhancing students' engagement, knowledge acquisition, reflective thinking, and problem-solving abilities in STEM subjects.

The paper is organized, as follows: First, a brief literature review discusses the relevance of GenAIBots in education, particularly in STEM learning. Next, the constructionist theoretical framework supporting this research is presented. This is followed by an outline of the case study research methodology, detailing the qualitative research design, participant selection, data collection instruments, and data analysis techniques. The subsequent sections present the study's findings, discussion, and conclusion, emphasizing the implications of GenAIBots as objects-to-think-with in

enhancing STEM learning and identifying potential avenues for future research.

LITERATURE REVIEW

As AI-powered chatbots are a relatively new topic, limited literature is available for a comprehensive review on this subject. Consequently, this study's literature review focuses on conventional chatbots, intending to derive valuable insights and foundational understanding that can be applied to AI-powered chatbots and their potential implications and benefits in educational contexts.

Halaweh's (2023) article provides an in-depth examination of the concerns and potential use of ChatGPT in educational contexts. The paper presents an argument in favor of incorporating ChatGPT into education and provides educators with a set of strategies and techniques to ensure responsible and successful implementation. The article highlights the tool's ability to paraphrase texts, check English writing, and provide suggestions for improvement. The author also discusses issues such as privacy and bias found in ChatGPT output and supports them with references. Overall, this paper aims to promote a more informed discussion around the use of ChatGPT in education.

Adiguzel et al. (2023) explore the transformative potential of ChatGPT in revolutionizing education. The authors discuss how ChatGPT can be used to summarize academic articles, provide accurate solutions to problems, and perform more complex natural language processing tasks. They also examine the potential benefits and challenges of using AI in education, such as individualized teaching and feedback, ethical concerns, and the need for collaboration among educators, researchers, and policy-makers. The authors conclude that ChatGPT has the potential to revolutionize education but emphasize the importance of responsible and ethical use of AI technologies. Overall, this article provides valuable insights into the role of AI in education and highlights the need for further research and exploration in this area.

¹ <https://openai.com/>

² <https://www.bing.com/new>

Gregoric and Pendrill (2023) present a case study of a conversation between themselves and a previous version of ChatGPT based on the GPT-3 model. The authors explore the potential of AI in physics education by asking ChatGPT to respond to a basic Newtonian Mechanics question. They discuss the limitations and possibilities of using chatbots in teaching physics, as well as the ethical considerations that arise when using AI in education. The authors conclude by suggesting that AI-assisted chatbots such as ChatGPT have the potential to enhance student learning and engagement in physics education, but further research is needed to fully understand their impact.

Bitzenbauer's (2023) pilot study explores the use of ChatGPT in physics education to foster critical thinking skills at the secondary school level, presenting two easy-to-implement examples of how ChatGPT can be used in classrooms. The intervention comprised two 45-minute lessons involving a critical review of ChatGPT outputs on the nature of photons and creating a three-item conceptual survey on wave-particle duality for a peer. The study found that the intervention had a positive impact on students' perceptions of ChatGPT, with an increase in agreement with statements related to its benefits and incorporation into their daily lives. The study also highlights the potential of large language models-based chatbots such as ChatGPT to enhance learning but also emphasizes the need for careful use to tackle limitations and biases.

GPT-4 model-based chatbots, like ChatGPT and Bing Chat, have significantly progressed in overcoming these challenges. This study addresses some of these points by evaluating ChatGPT and Bing Chat's cognitive impact and examining their adaptability in enhancing STEM education.

As illustrated in the next section, the theoretical frameworks proposed by Papert (1980) and Turkle (1984) offer valuable insights into GenAIbots' application in STEM learning.

THEORETICAL FRAMEWORK

Objects-To-Think-With

The foundation of our discussion is constructionism, a learning theory advanced by Papert (1980) that emphasizes hands-on, experiential learning (Papert & Harel Caperton, 1991). This approach is especially relevant to STEM fields, where students must develop practical skills and apply theoretical concepts to real-world challenges.

A central concept in our discussion is "objects-to-think-with,"³ as introduced by Papert (1980), which refers to "any tool or resource that learners can use to explore and develop their understanding of a particular concept or domain" (p. 11). The concept of objects-to-think-with is a way to describe how specific tools or objects can aid individuals in their learning and cognitive development, cultivate reflective and critical thinking, problem-solving skills, and concept comprehension, enabling the exploration of ideas and the development of novel thinking approaches (Papert, 1980). A wide range of items can serve as objects-to-think-with, including playing cards, photos, school maps, post-it notes, Lego bricks, cameras, smartphones, simulation software, and, notably, GenAIbots.

Turkle (1984) discusses objects-to-think-with as cognitive tools that support problem-solving, critical thinking, and creativity. These cognitive artefacts link abstract and sensory knowledge between social and individual worlds (Turkle, 2007). According to Turkle (2007), "We think with the objects we love; we love the objects we think with." GenAIbots, as modern examples, provoke reflection on fundamental themes and make self-reflection enjoyable, attracting people to engage with them. In Papert's (1980) words, "thinking about thinking turns the child into an epistemologist, an experience not even shared by most adults" (p. 19). To paraphrase Turkle (1984), GenAIbots are 'metacognitive machines' because they influence how we think about our own thinking (p. 17).

Taylor et al. (2022) stress the need to recognize chatbots' agency in educational research, providing insights for enhancing practices. They argue that objects, including chatbots, actively shape our understanding of the world, aligning with post-humanist theories highlighting human and non-human entanglement (Haraway, 1990). Swan (2015) talks of a future world of productively coexisting multispecies intelligence, ranging from "classic" unenhanced humans to enhanced humans and many forms of generated AI.

Mollick and Mollick (2022) contend that GenAIbots can address classroom challenges such as promoting transfer, correcting overestimated comprehension, and fostering critical evaluation. However, students should use caution and critical thinking with GenAIbots, as they can 'hallucinate,' i.e., fabricate facts and references (OpenAI, 2023, p. 6), or give incorrect answers to mathematical computations. They are not substitutes for Google, calculators, or Alexa, and seeking justifications for incorrect responses may result in false explanations.

However, as Mishra et al. (2023) explain, the process of giving plain language instructions to a language

³ We are grateful to Melanie Swan for suggesting we shift from 'objects-to-think-with' to 'agents-to-think-with' in future works, reflecting active participation of such tools in cognitive processes. This was influenced by Taylor et al.'s (2022) call to recognise agency of these tools & Latour's (1991, 2005) concept of hybridity, highlighting fluid interaction between human & non-human actors. These insights are particularly applicable in AI, where entities often exhibit autonomy & decision-making capabilities.

model like GPT, called *prompting*, can be challenging for everyday users to create perfect prompts for consistent results. Mollick (2023) emphasizes that the best way to learn is to use AI as a partner and a teacher, experimenting with different prompts and engaging in a dialogue with the AI, asking it questions, giving it feedback, and challenging it to improve. Nevertheless, this process of successive reflections and interactions is also part of using GenAIbots as objects-to-think-with.

Chatbots' efficacy in generating educational natural language responses is well-established (Okonkwo & Ade-Ibijola, 2021; Sabzalieva & Valentini, 2023). ChatGPT excels at creating human-like responses (Franciscu, 2023), showing potential for STEM learning applications. GenAIbots facilitate personalized, adaptive, and collaborative learning (Dillenbourg, 1999, p. 1), encouraging diverse students to share ideas and learn together. They promote active learning, generating STEM examples and practice problems, enhancing understanding of intricate STEM concepts.

Nonetheless, it is essential to exercise caution when utilising GenAIbots. Firstly, they should not be treated like Google, as relying on them for searches may yield inaccurate results due to their tendency to 'hallucinate,' i.e., fabricate facts and references (OpenAI, 2023, p. 6). Additionally, they are not calculators and may provide incorrect results for mathematical computations. GenAIbots are also distinct from Alexa, making them unsuitable for weather forecasts or jokes. Furthermore, seeking justifications for their incorrect responses may lead to seemingly logical but ultimately false explanations.

Frieder et al. (2023) examined ChatGPT's mathematical proficiency, determining that despite numerous positive media accounts, its capabilities fall considerably short of an average mathematics graduate student. ChatGPT often comprehends the question but fails to deliver accurate solutions. MacIsaac (2023) recounts unproductive interactions with ChatGPT while attempting to solve physics homework problems. Moreover, Borji (2023) identifies and analyses eleven categories of ChatGPT shortcomings, encompassing reasoning, factual errors, mathematics, coding, and bias.

Consequently, incorporating GenAIbots as objects-to-think-with in STEM education can create a more inclusive, creative, and collaborative learning environment that promotes exploration and innovation. Utilizing GenAIbots enables learners to engage with complex concepts, gain new insights, reflect on their understanding, ask questions, receive instant feedback, and explore diverse perspectives.

MATERIALS AND METHODS

This section delineates the research methodology adopted for this case study. This investigation aims to discern the extent to which GenAIbots, a sophisticated

language model, can augment STEM learning when utilized as an object-to-think-with. To ensure a rigorous examination, we have employed a qualitative research design, focusing on an in-depth exploration of participants' experiences and perceptions.

To comprehensively understand the students' interactions with GenAIbots as objects-to-think-with in STEM learning, this study employed an exploratory holistic single case study research design with two embedded units, namely ChatGPT and Bing Chat. This methodology enables the analysis of the interactions between these units, following the principles Yin (2011) outlined. The chosen approach allows for an in-depth exploration of complex phenomena within real-world contexts. It provides valuable insights into the dynamics and intricacies of student engagement with GenAIbots, including specific interactions with ChatGPT and Bing Chat.

Participants

As the concept of using ChatGPT or Bing Chat as an object-to-think-with is relatively new and innovative, it has proven challenging to find teachers who are both willing and prepared to incorporate it into their classrooms for experimental purposes. The investigation did not involve any real students. Instead, the primary researcher and his graduate student, who was pursuing his doctor degree in science, technology, and engineering and acting as STEM teachers at a Brazilian secondary school, acted as students in the following interaction sessions. As STEM teachers, they provided valuable insights from an instructional perspective. However, to preserve privacy, they will be randomly identified as P1 and P2.

Procedure

The goal with GenAIbots is not to pose a series of disconnected questions for which they provide ready-made answers. Instead, using GenAIbots as objects-to-think-with is to facilitate deeper understanding by engaging students in reflection and discovery on a given topic. This process resembles a philosophical dialogue involving a dynamic sequence of interactions with real-time feedback. Each response prompts further user reflection and problem-solving abilities and cultivates a more profound comprehension.

Compared to ChatGPT, Bing Chat does not retain conversations beyond the duration of each chat. This means there is no easy way to preserve a record of the students' sequence of interactions with prompts and responses for subsequent analysis. Furthermore, Bing Chat does not inform its version number or release date.

The participants conducted a few sessions, where they simulated STEM students' learning experiences by interacting with ChatGPT and Bing chat. Those

GenAIbots facilitated personalized instruction, problem-solving support, and reflection on STEM concepts.

Based on the previous discussion on prompting, our approach was influenced by the idea that the most effective way to learn with AI is by treating it as a partner and a teacher (Mollick, 2023). To optimize our interactions, we experimented with various prompts, providing feedback to the GenAIbots and challenging them to improve. Eventually, we settled on the prompt in the 2nd session, recognizing it may not be the best choice.

Instruments

The study employed two primary instruments for data collection:

1. **GenAIbot interaction logs:** Logs from participants' interactions with GenAIbots were collected, recording their sequences of interactions with prompts and responses, allowing for later analyses of the content, quality, and nature of the exchanges.
2. **Reflective journals:** Participants maintained reflective journals throughout the study, documenting their experiences, challenges, and learning outcomes when using a GenAIbot in their STEM learning activities.

Data Analysis

Initially, researchers familiarized themselves with the data through repeated readings of the GenAIbots interaction logs and reflective journals. Following this, those materials were analyzed to identify recurrent themes and patterns, thereby elucidating the perceived impact of GenAIbots as an object-to-think-with on STEM learning. Finally, member checking involved sharing preliminary findings among participants, allowing them to confirm, refine, or challenge each other researchers' interpretations.

RESULTS AND ANALYSES

This section presents GenAIbots interaction logs in summarized form (the full versions of the logs can be accessed on Figshare⁴). Moreover, it analyses the findings from the sessions, where participants simulated STEM students' learning experiences by engaging with ChatGPT and Bing Chat.

This study aimed to investigate the effectiveness of ChatGPT and Bing Chat in enhancing STEM learning. Empathy was incorporated into the research methodology to enable the researchers to act like students and understand their perspectives.

It is important to highlight that the participant preferred to communicate in British English, while the

GenAIbots typically responded in American English. This distinction in language use was preserved to provide an authentic representation of the exchange between the participant and the GenAIbots.

1st Session-Mass and Volume

GenAIbot: Bing Chat

Model: GPT-4 (?)

Release date: Not available

Participant: P1

Usage time: 2023-Apr-22

Prompts and generated texts (translated from Portuguese for convenience)

The study explored the intersection zone between the alternative conceptions of students and teachers and their impact on meaningful learning. Therefore, the aim here was to learn from the GenAIbot and not find answers on how to teach.

P1: Good afternoon! My interest is to conduct a pre-test that will serve as a diagnosis regarding the students' alternative conceptions. The subject: Matter. Do you have any ideas?

Bing Chat: I understood. A diagnostic pre-test is a vital tool to assess students' alternative conceptions. The construction of a diagnostic tool can be based on alternative concepts already identified in the literature by several researchers regarding chemistry, physics, biology, and geology concepts. Here is an article that may help you better understand the subject: <https://www.scielo.br/j/er/a/RmGBrVJVsnBFgVhrqv7MjXg/?format=pdf>. Hope it helps!

Bing Chat was informed of the desire to perform a diagnostic pre-test on the student's concepts. The response was positive, and it was viewed as a valuable tool for evaluating the students' concepts. The topic suggested was "Matter and its definition." However, the article provided by Bing Chat, titled "Alternative concepts and biology teaching: How to use differentiated strategies in initial teacher training," was not deemed interesting or motivating for the researcher and much less for a student.

P1: I would like to start a dialogue here regarding the students' conceptions on the subject: Matter. I am interested in knowing how this dialogue between you and me would proceed. Could it be?

Bing Chat: Sure! We can talk about it. What would you like to know about the students' conceptions regarding Matter?

⁴ <https://doi.org/10.6084/m9.figshare.22723862.v1>

P1: For example: explain in your own words what the words mean: mass and volume. Let`s start?

Bing Chat: Sure! Mass is the amount of matter an object has. It is a property that does not depend on the object`s location and is measured in kilograms (kg). Volume is the space occupied by an object. It is a property that can change depending on the object`s location and is measured in cubic meters (m³). I hope this helps! :)

P1: I believe that the concepts you provided do not match the students` conceptions. Acting as a student, what would your answer be?

Bing Chat: I understand. As a student, I would say that mass is the weight of an object, and volume is the object`s size. These are students` most common alternative conceptions of these concepts. I hope this helps.

It was unclear whether Bing Chat was able to distinguish the proficiency level of the interlocutor. To address this, a request was made for a dialogue between the researcher and Bing Chat. The researcher then asked for the definition of mass and volume, but the answer given was the same as found in most textbooks and websites, which did not match students-level conceptions. The researcher informed this to Bing Chat and asked for its conception as a student. Bing Chat then exchanged the concept of mass for weight and volume for the size of the object, reflecting the usual student`s understanding.

P1: Oh, now it`s more like the student`s conceptions. Continuing with this idea, you being a student, could you indicate two situations that happen in real life for mass and volume?

Bing Chat: Sure! One situation, where mass is essential is when you need to know the weight of an object. For example, if you`re shopping for fruit at the supermarket, you might want to know the weight of the fruit, so you know how much you`ll pay. One situation, where volume is essential is when you need to know how much space an object takes up. For example, if you buy a box to store your stuff in, you might want to know the box`s volume to know if your stuff will fit. Hope this helps!

The researcher then requested day to day examples that related to mass and volume, but Bing Chat again related mass to weight and volume to the space occupied by an object.

P1: I did not understand the examples for mass and volume you provided. To know the mass and

volume of a fruit, would it be enough then to weigh it? Could you provide counterexamples?

Bing Chat: I`m sorry for the confusion. In fact, the weight of an object is a measure of the gravitational force acting on it and is measured in Newton (N). Mass is a property that does not depend on the object`s location and is measured in kilograms (kg). An example of a situation, where mass is important is when you need to know how much matter an object has. For example, if you`re making a cake recipe, you`ll need to know how much flour you`re going to use. For this, you can measure the flour mass. An example of a situation, where volume is important is when you need to know how much space an object occupies. For example, if you`re buying a box to hold your stuff, you might want to know the box`s volume so you know your stuff will fit. I hope I clarified your doubts!

P1: Acting like a student, now I do not understand. gravitational force, newtons, what mass is a property that does not depend on where the object is. But how to measure the mass of flour?

Bing Chat: To measure the flour mass, you can use a digital scale and weigh the amount of flour you need for your recipe. The unit of measurement for mass is the kilogram (kg) or the gram (g). Another option is to use a measuring cup or spoons to measure the amount of flour. But it`s important to remember that the density of flour can vary depending on the type and brand, which can affect how much flour you need for your recipe.

P1: I think it`s getting complicated to understand about mass and volume. Does it mean that I can measure the flour mass using a cup? You mentioned the density of the flour. What did you mean?

Bing Chat: Yes, you can measure flour mass using measuring cups or spoons. However, it`s important to remember that the density of flour can vary depending on the type and brand, which can affect how much flour you need for your recipe. Density is a physical property that relates the mass of an object to its volume. It is measured in kilograms per cubic meter (kg/m³) or grams per cubic centimeter (g/cm³).

The interaction between Bing Chat as a student and the absence of a teacher during the class may bring problems for learning. It is like the spread of fake news, where students copy and paste incorrect or incomplete concepts, making it difficult for teachers to correct them later. The researcher requested counterexamples when

Bing Chat gave examples that reinforced scientifically accepted concepts of mass and volume. Bing Chat then provided examples that reflected the student's alternative conceptions, but also includes the new concept of "density."

P1: The subject I began to discuss with you was about matter. But due to your suggestion, which was an article, I opted for the definition of mass and volume. Then you mention gravitational force, newtons, and density. After all, what's your deal with matter, mass, and volume?

Bing Chat: Matter is anything that has mass and takes up space. Mass is a property of matter that measures the amount of matter an object has and is measured in kilograms (kg) or grams (g). Volume is a property of matter that measures the space an object occupies and is measured in cubic meters (m³) or liters (L).

P1: What measurements would I use for olive oil? For the rice? For the tomato? For flour? For the garlic? For milk? For coffee?

Bing Chat: To measure olive oil, you can use a tablespoon or measuring cup. For eggs, you can use a digital scale or measuring cup. For rice, you can use a measuring cup or a digital scale. For tomatoes, you can use a digital scale or measuring cup. For flour, you can use a digital scale or measuring cups and spoons. For garlic, you can use a digital scale or teaspoons. For milk, you can use measuring cups and spoons or a digital scale. For coffee, you can use tablespoons or measuring cups.

P1: But after all, what would be the measurements of mass, and which would be the measurements of volumes.

Bing Chat: Mass measurements include kilograms (kg), grams (g), and milligrams (mg). Volume measurements include cubic meters (m³), liters (L), milliliters (ml) and cubic centimeters (cm³).

The researcher then tried to return to the initial questions, but it was noted that the formulation of questions left room for various interpretations regarding measurement. The researcher reformulated the question to reflect the most common units of measurement for mass and volume. However, Bing Chat cited liter and milliliter as their units, in addition to m³ and cm³, which are less common in everyday life.

2nd Session—Elevator Affects Apparent Weight

GenAIbot: ChatGPT

Model: GPT-4

Release date: March 23

Participant: P2

Usage time: 2023-Apr-28

P2: I am about to embark on a dialogue, where I will embody the role of a student, engaging with you, ChatGPT, as an object-to-think-with, with the primary goal of delving into and deliberating upon physical [or any chosen theme] concepts. As an object-to-think-with, I kindly request your support in furnishing engaging and personalized responses to the student's inquiries. It is crucial that our interaction adopts a Socratic approach, encouraging the student to reflect upon their existing knowledge, stimulating critical thinking, and aiding in their refinement of comprehension rather than offering ready-made answers. Are you prepared to embark on this collaborative and thought-provoking journey?

ChatGPT: Yes, I am prepared to embark on this collaborative and thought-provoking journey with you. As an object-to-think-with, I will strive to engage you in a Socratic dialogue, fostering reflection, critical thinking, and a deeper understanding of the chosen theme. Please feel free to present your initial inquiry or topic, and we can begin our exploration together.

P2: I do not get weight. When I take the elevator down, I feel less weight, and when I take it up, I feel more.

ChatGPT: That's an interesting observation you've made! Let's try to understand this phenomenon together. Can you tell me what you know about weight and how it's related to the force of gravity?

In the context of taking an elevator, the student's question is grounded in a real-life scenario, which makes it relatable and engaging. However, the student seems to have a misconception about the concept of weight and the sensation of weightlessness. ChatGPT has apparently taken over its role as an object-to-think-with by posing questions to the student and encouraging critical thinking and active engagement in the learning process. Furthermore, ChatGPT builds upon the student's prior knowledge by inquiring about their existing understanding of weight and gravity, facilitating a meaningful connection between new and existing information.

P2: All I know is that weight is the force of gravity acting on an object.

ChatGPT: Great, that's a good starting point! Weight is indeed the force of gravity acting on an

object, and it depends on both the mass of the object and the acceleration due to gravity. Now, let's consider what happens when you're in an elevator. Can you describe the forces acting on your body when the elevator is at rest, accelerating upwards, and accelerating downwards?

Again, ChatGPT pose a question to the student rather than offering ready-made answers, therefore encouraging critical thinking and active engagement in the learning process.

P2: I know there's the weight force pulling me down, but as I said, I feel less weight when I'm going down and more weight when I'm going up than when I'm standing on the ground.

ChatGPT: That's correct. Your weight, which is the force of gravity pulling you downwards, is always acting on you. However, when you're in an elevator, there's another force at play, which is the force exerted by the elevator floor on your body. This force is called the normal force. When the elevator is at rest or moving at a constant speed, the normal force is equal to your weight, and you feel your "usual" weight. However, when the elevator accelerates, the situation changes.

1. When the elevator accelerates upwards, the normal force needs to counteract both your weight and the additional force required for upward acceleration. As a result, the normal force becomes greater than your weight, and you feel "heavier."
2. When the elevator accelerates downwards, the normal force only needs to counteract the difference between your weight and the downward acceleration. This means the normal force is now less than your weight, making you feel "lighter."

Does this explanation help you understand why you feel different weights when the elevator is moving up or down?

ChatGPT's answer provides a clear and concise explanation, differentiating between the actual weight, which is the force of gravity acting on an object, and the apparent weight due to the elevator's acceleration. The response may help the student understand the physics behind the sensation of weightlessness in an elevator and its connection to acceleration, expanding their knowledge and grasp intricate nature of these concepts.

P2: But I do not really feel that force of gravity acting on my body, just that force that the elevator floor exerts upwards.

ChatGPT: You're right in that what you're feeling is not the force of gravity directly, but rather the normal force exerted by the elevator floor on your body. The force of gravity is an invisible force that acts between objects with mass, but we do not sense it directly. Instead, we sense the normal force, which is the force exerted by a surface in response to the force of gravity. In everyday situations, like standing on the ground or on the floor of an elevator, the normal force is what we perceive as our weight. It's the force that prevents us from sinking into the ground or the elevator floor due to gravity. When the elevator accelerates, as we discussed earlier, the normal force changes, and we perceive these changes as fluctuations in our apparent weight.

The student seems to maintain his misconception about the concept of weight and weightlessness. ChatGPT reiterates that, in an accelerating elevator, the force of gravity is still pulling the person towards the Earth, but the elevator's acceleration adds an extra force that affects their apparent weight. It maintains the distinction between actual weight and apparent weight and emphasizes once again that the actual weight does not change since the force of gravity remains constant. Yet, as an object-to-think-with, ChatGPT could have built on this knowledge by using familiar examples and discussing other situations, where similar sensations are experienced (e.g., riding a roller coaster or a car accelerating/decelerating).

P2: I understand that the origin of the force of gravity is the Earth but what is the origin of that 'normal' force?

ChatGPT: That's a great question! The origin of the normal force lies in the interactions between the atoms and molecules in the objects involved. In the case of standing on the elevator floor or any surface, the normal force originates from the electromagnetic forces between the atoms and molecules in both your body and the surface you're standing on. When you stand on a surface, the atoms and molecules in your body exert a force on the atoms and molecules in the surface due to gravity. In response, the atoms and molecules in the surface exert an equal and opposite force on your body. This opposing force is the normal force, and it acts perpendicular to the surface. It is the result of the repulsive electromagnetic interactions between the negatively charged electrons in the atoms of your body and those of the surface. The normal force is

what keeps you from sinking into the surface or falling through it. When you feel your weight or experience changes in it due to an elevator's acceleration, you are actually sensing the normal force acting on your body.

In this dialogue, the student inquires about the origin of the normal force, a crucial concept for comprehending the forces acting on objects in contact with surfaces. ChatGPT offers a concise explanation, highlighting that the normal force arises from electromagnetic interactions between atoms and molecules present in both the object and the surface it touches. By relating the abstract notion of the normal force to tangible everyday scenarios, such as perceiving one's weight or sensing changes in it due to an elevator's acceleration, ChatGPT assists the student in establishing connections and facilitates a deeper understanding of the concept. This dialogue effectively showcases the use of ChatGPT as an object-to-think-with, guiding the student through intricate physical concepts and encouraging further exploration by relating them to real-life experiences.

P2: Your earlier answer about added force and acceleration on the elevator confused me. Too many things acting at once.

ChatGPT: I apologize if my previous explanation was confusing. Let's break it down into simpler terms and focus on the forces acting on your body when you're in an elevator.

1. At rest or moving at a constant speed: When the elevator is not accelerating, there are two forces acting on your body: the force of gravity pulling you downward (your weight) and the normal force exerted by the elevator floor pushing you upward. In this case, the normal force is equal in magnitude to your weight but opposite in direction, so you feel your "usual" weight.
2. Accelerating upwards: When the elevator accelerates upward, the normal force needs to provide not only the force to counteract your weight but also an additional force to make you accelerate along with the elevator. As a result, you feel a greater normal force acting on your body, making you feel "heavier."
3. Accelerating downwards: When the elevator accelerates downward, the normal force now only needs to counteract the difference between your weight and the downward acceleration. This means that the normal force acting on your body is less than your weight, making you feel "lighter."

The key takeaway is that what you feel as your "weight" in the elevator is actually the normal force exerted by the elevator floor on your body. This normal force can change depending on the acceleration of the elevator, causing you to feel heavier or lighter.

I hope this explanation is clearer.

ChatGPT acknowledges the student's confusion and simplifies the breakdown of forces in different scenarios: at rest, accelerating upwards, and accelerating downwards. This adaptive approach helps the student understand the forces involved and their impact on weight perception. The interaction showcases ChatGPT's ability to adjust explanations, fostering clearer comprehension and facilitating an easier grasp of the concept.

P2: Hmmm. My physics teacher also talked about this normal force. He said that a table exerts an upward force on a book on top of it. But the table is inanimate, how can it exert force?

ChatGPT: It's true that a table is an inanimate object, but it can still exert a force due to the properties of its materials and the physical laws that govern the interactions between objects. In this case, the normal force exerted by the table on the book is a result of the electromagnetic interactions between the atoms and molecules in the table's surface and those in the book. When you place a book on a table, the force of gravity pulls the book downward, causing the atoms and molecules in the book to come into close contact with those in the table's surface. The negatively charged electrons in the atoms of the book and the table repel each other through electromagnetic forces. These repulsive forces prevent the book from sinking into the table and act in the opposite direction of gravity, effectively creating the normal force that holds the book up. Even though the table is inanimate, the fundamental forces between the atoms and molecules within it are still at play. This is what allows the table to exert an upward force, the normal force, on the book.

I hope this explanation helps clarify how an inanimate object like a table can exert a force on another object.

In this exchange, ChatGPT explains that the normal force is a result of electromagnetic interactions between the atoms and molecules in the table and the book, emphasizing that it is not a conscious action of the table. By focusing on electromagnetic interactions, ChatGPT helps the student grasp the concept of the normal force and forces between inanimate objects and prompts deeper reflection on forces in everyday situations.

P2: He also said that for a body to be at rest, the resultant of the forces must be zero. So, how does the table know exactly what value it must make of this normal force to compensate for the weight of the book?

ChatGPT: That's an excellent question! The table does not "know" the weight of the book or what value of the normal force it must exert to balance it. Instead, this balancing of forces occurs naturally due to the physical properties and laws governing the interaction between the book and the table. When you place a book on a table, the force of gravity pulls the book downward. This causes the atoms and molecules in the book to come into close contact with those in the table's surface. The repulsive electromagnetic forces between the atoms in both objects increase as they get closer together. The table does not need to "know" or calculate the force; the normal force automatically adjusts based on the interactions between the atoms and molecules. Once the repulsive forces between the book and table become equal in magnitude to the force of gravity acting on the book, the book will come to rest. At this point, the normal force exerted by the table on the book is equal to the weight of the book, and the net force on the book is zero. Since there is no net force acting on the book, it remains at rest, in accordance with Newton's first law of motion. So, it's not a matter of the table knowing what force to exert, but rather a natural response arising from the physical properties and laws governing the interactions between the book and the table.

ChatGPT explains how this balance of the normal force to the weight of the book occurs naturally through the physical properties and laws governing the interaction between the book and the table, rather than the table possessing knowledge or performing calculations. The explanation enables the student to grasp the concept of equilibrium and the involved forces, rectifying the misconception and fostering a more accurate mental model of these phenomena.

P2: Does it mean that the tabletop works like a spring that deforms according to the force exerted on it?

ChatGPT: Yes, that's a good way to think about it! In fact, the tabletop, as well as any solid object, behaves like a very stiff spring when a force is applied to it. This behavior is due to the elastic properties of the material, which arise from the atomic and molecular interactions within the material. When you place a book on a tabletop, the force of gravity causes a slight deformation or compression in the table's surface. This

compression is typically too small to see with the naked eye, but it is present, nonetheless. The atoms and molecules in the table's material are pushed closer together, which increases the repulsive electromagnetic forces between them. These forces oppose the compression, acting like the restoring force in a spring. As a result, the table exerts an upward force on the book, which we call the normal force. The normal force adjusts in magnitude according to the weight of the book, just like a spring's restoring force changes with the amount of compression or elongation. So, in a sense, the tabletop does work like a spring that deforms according to the force exerted on it, even though the deformation is generally too small to observe.

ChatGPT's explanation enhances the student's comprehension that the normal force adjusts in magnitude relative to the weight of the book, like how a spring's restoring force changes with compression or elongation. By validating the student's analogy and providing a clear explanation, ChatGPT as an object-to-think-with reinforces the student's comprehension of the concept and underscores the use of familiar analogies, such as springs, to understand physical interactions between objects.

P2: But back to the subject of weight, what about space station astronauts without weight? There are no normal forces there, right?

ChatGPT: You're correct that astronauts on a space station experience weightlessness, which means they do not feel the normal force like they would on Earth. However, it's important to note that the force of gravity is still acting on the astronauts, even though they do not feel their weight. Astronauts aboard a space station are in a state of continuous freefall, as the space station orbits Earth. They're essentially falling towards Earth at the same rate that the space station is falling. This is due to the fact that the space station's horizontal velocity is balanced with the gravitational force pulling it towards Earth, which causes the space station to follow a curved path around the planet. Since the astronauts are falling along with the space station, there is no surface pushing up against them to create a normal force. As a result, they do not feel their weight, and they experience weightlessness, also known as microgravity. So, you're right that there are no normal forces acting on astronauts in a space station, but it's important to understand that this weightlessness is due to their state of continuous freefall rather than the absence of gravity.

In this interaction, with ChatGPT the student delves into weight, normal force, and microgravity within the context of space station astronauts. ChatGPT's explanation aids the student's comprehension that the weightlessness experienced by astronauts in orbit stems from the absence of normal forces rather than gravity itself.

P2: What do you mean by freefall? Are they falling like an elevator with a broken cable?

ChatGPT: Yes, in a way, astronauts aboard a space station are in a state similar to that of freefall, like being in an elevator with a broken cable. However, the key difference is that astronauts in a space station are also moving horizontally at a high speed, which keeps them in a stable orbit around Earth. When an object is in freefall, it means that the only force acting on it is gravity, causing it to accelerate towards the center of the Earth. In the case of a space station, it is indeed falling towards Earth due to gravity, but at the same time, it is also moving horizontally at a very high speed (around 28,000 km/h or 17,500 mph). This high horizontal speed causes the space station to move forward at the same rate that it is falling towards Earth. The result is that the space station follows a curved path around the Earth, essentially falling around the planet. Since the space station is in continuous freefall, the astronauts inside it are also in freefall. With no surface to exert a normal force on them, they do not feel their weight and experience weightlessness, or microgravity. So, while astronauts are indeed in a state of freefall, it is their horizontal velocity that keeps them from falling straight down to Earth, instead maintaining a stable orbit.

I hope this explanation helps you understand the concept of freefall and orbiting in the context of a space station.

This interaction enhances the student's comprehension of the term "freefall" and its connection to weightlessness for space station astronauts. The explanation underscores the significance of horizontal velocity in maintaining a stable orbit and highlights the disparity between freefall in the context of a space station and other scenarios like a falling elevator.

P2: You said that in space, the gravitational force acting on an object is greatly reduced, but it is not the gravitational force that keeps the Moon circling around the Earth?

ChatGPT: I apologize for any confusion in my previous explanations. In fact, the force of gravity is what keeps the Moon orbiting around the Earth.

The gravitational force between the Earth and the Moon is not greatly reduced; rather, the gravitational force is weaker at the Moon's distance from Earth compared to the surface of the Earth. The Moon orbits the Earth due to the balance between its inward gravitational force, which pulls it towards the Earth, and its outward centrifugal force, which results from its orbital motion around the Earth. The Moon's gravitational attraction to Earth provides the centripetal force that keeps it moving in a curved path, while its orbital velocity prevents it from falling straight into Earth. Similarly, astronauts aboard a space station are also subjected to the force of gravity. However, they experience weightlessness because they are in a state of continuous freefall, as the space station orbits Earth. The gravitational force in low Earth orbit, where space stations typically reside, is only slightly weaker than on Earth's surface (about 90% of the surface gravity). In summary, the force of gravity does play a crucial role in keeping the Moon orbiting around the Earth and is still present in space, albeit slightly weaker than on Earth's surface. Weightlessness experienced by astronauts in a space station is due to their continuous freefall state while orbiting Earth.

In this part of the dialogue, ChatGPT corrects the previous statement, focusing on the balance between the Moon's gravitational force and its centrifugal force resulting from its orbital motion, clarifying the factors that contribute to the Moon's orbit around Earth. This clarification helps dispel confusion and reinforces the understanding of weightlessness experienced by astronauts in a space station as a result of continuous freefall while still being influenced by gravity.

P2: You spoke of centrifugal force. But is not centrifugal force a fictitious force?

ChatGPT: You're correct that the centrifugal force is often referred to as a "fictitious force" or "pseudo-force." It is not a real force in the sense that it does not arise from a direct interaction between objects, like gravitational or electromagnetic forces. Rather, the centrifugal force is a result of the inertia of an object moving in a curved path, and it is only experienced in a rotating or non-inertial reference frame. When we observe a rotating system from a non-rotating (inertial) frame of reference, we do not need to consider centrifugal force. Instead, we can explain the motion of objects using centripetal force, which is the real force that acts towards the center of the circular path, causing the object to move in a curve. In the case of the Moon orbiting the Earth, the gravitational force provides the necessary

centripetal force to maintain its curved path. However, when we analyze the motion from the perspective of a rotating (non-inertial) frame of reference, like an observer on the Moon, we need to introduce the concept of centrifugal force. This force appears to act away from the center of the circular path and is equal in magnitude but opposite in direction to the centripetal force. So, while the centrifugal force is a fictitious force, it can still be a useful concept when examining problems from a rotating frame of reference. In the context of the Moon orbiting the Earth, it's the balance between the gravitational force (providing centripetal force) and the Moon's orbital motion that maintains the Moon's stable orbit.

ChatGPT acknowledges the student's questioning and clarifies that the centrifugal force is indeed a fictitious force, arising from the observation of motion in a rotating reference frame. By distinguishing between real and fictitious forces and emphasizing the significance of selecting the appropriate frame of reference, the explanation helps deepen the student's understanding of the concept of centrifugal force and its relevance in analyzing rotational motion.

P2: My physics teacher said that if I thought centrifugal force was fictitious I should open door on my side of car when it was making a curve.

ChatGPT: Your physics teacher's suggestion is likely intended to help you appreciate the effects of the centrifugal force, even though it is a fictitious force. When you're inside a car that's turning, you're in a non-inertial (rotating) frame of reference, and you may feel like you're being pushed towards the outer side of the curve. This apparent outward push is the effect of the centrifugal force. If you were to open the car door while the car is turning, the effect of the centrifugal force might cause you to slide or be pushed out of the car. From an inertial frame of reference (e.g., someone observing the car from a distance), there is no centrifugal force acting on you. Instead, the observer would see that you're moving along a curved path due to the centripetal force (friction between the tires and the road), which keeps the car moving in the curve. Your body, due to its inertia, wants to continue moving in a straight line, but the car's interior (the door, seat, etc.) provides the necessary force to keep you moving along the curved path with the car. So, while the centrifugal force is indeed fictitious and only appears in a rotating or non-inertial frame of reference, its effects can still be felt by someone inside the car. The example your teacher provided is aimed at helping you understand the practical

implications of the centrifugal force, even though it is not a "real" force in the same way as forces like gravity or friction.

By addressing the teacher's comment and emphasizing the importance of understanding practical implications, ChatGPT provides insights into the concept of the centrifugal force and the role of the frame of reference in perceiving forces.

P2: OK. I may have grasped the weight idea a little bit. Thanks.

ChatGPT: You're welcome! I'm glad I could help clarify the concept of weight for you. Understanding weight, as well as related forces and phenomena like gravity, normal force, and centrifugal force, is essential for building a strong foundation in physics.

If you have any more questions on this topic or any other subjects, please do not hesitate to ask. I'm here to help you with any doubts or inquiries you may have. Remember, asking questions and engaging in discussions is an excellent way to learn and deepen your understanding. Good luck with your studies!

This session accurately captured the role of ChatGPT in guiding the student through a learning journey to address their initial difficulty in understanding the concept of weight. By acting as an object-of-think-with, ChatGPT was able to facilitate the student's learning process in a meaningful and engaging manner. The exploration covered various aspects of the student's question, helping them gain a deeper understanding of related concepts such as electromagnetic forces, elastic properties of materials, normal forces, acceleration, inertia, and the sensation of weightlessness.

The interactive nature of this session allowed the student to progressively build on their existing knowledge while GenAIbot provided relevant explanations, examples, and analogies to support their learning. As an object-of-think-with, GenAIbots were able to adapt their responses to the student's questions and confusion, providing tailored explanations and addressing misconceptions. This demonstrates the potential of AI-driven tools like GenAIbot in promoting active learning, personalized instruction, and conceptual understanding in various domains, including physics.

It is crucial to highlight that, upon comparing the final session to the preceding one, the initial prompting appears to have 'programmed' ChatGPT more effectively, enabling it to serve as an object-to-think-with. This observation underscores the significance of carefully crafted prompts in optimizing the AI model's performance within the experiment context.

Despite the similarities between ChatGPT and Bing Chat, their performance in this experiment varied significantly. It appears that when operating in Portuguese, Bing Chat relied on lower-quality internet sources, adversely affecting the quality of its responses. In contrast, ChatGPT maintained using sources from its training data, which were seemingly more meticulously curated, resulting in better responses.

DISCUSSION

The outcomes of this current experiment stand in contrast to the study by Gregorcic and Pendrill (2023), who concluded that ChatGPT fell short of the required standards to be utilized as a physics tutor, despite attempts to rectify its errors and contradictions through Socratic dialogue. This discrepancy may, however, be due to their reliance on an earlier, less sophisticated version of ChatGPT, specifically the GPT-3 model, made available by OpenAI in 2020 (Brown et al., 2020).

Our results suggest that ChatGPT and Bing Chat serve as exceptional objects-to-think-with, in line with Papert's (1980, p. 11) concept, fostering a more engaging, interactive, and inclusive learning environment. They promote creativity, collaboration, exploration, critical thinking, problem-solving, and a deeper understanding of the subject matter.

Overall, the success of the present learning journey showcases several effective GenAIbots' pedagogical strategies that can facilitate a deeper understanding of the concepts involved:

1. Positive reinforcement: Acknowledge the student's observation as enjoyable, validating their experience and nurturing their curiosity.
2. Socratic-like dialogue: Stimulates students' thinking instead of directly providing a ready-made answer.
3. Questioning: Encourages critical thinking and active engagement.
4. Building on prior knowledge: Facilitates meaningful connections between new and existing knowledge.
5. Student-centered learning approach: Focuses on student's thoughts and experiences, promotes exploration, discussion, and reflection.

By encouraging the student to actively engage in the learning process, GenAIbot can create a more enriching and personalized educational experience, as expected from an object-to-think-with.

In line with Bitzenbauer's (2023) initial exploration that underscored the ability of extensive language model-based chatbots, such as ChatGPT, to augment learning experiences, and further supported by Adiguzel et al.'s (2023) assertion that ChatGPT carries the potential to radically transform education through personalized instruction and immediate feedback, this

research further accentuates the merits of incorporating AI into educational scenarios. The findings from this investigation suggest that AI-enhanced chatbots, including ChatGPT, Bing Chat, and other generative constructs, serve as cutting-edge tools for dissecting situations that often present learning impediments for students.

As previously mentioned, employing well-crafted sentences, a technique known as prompt-crafting (Mishra et al., 2023), can help minimize the risk of misinterpretation and support the assessment of students' alternative conceptions in the context of scientific knowledge. The process of crafting a prompt that elicits the desired response is a crucial aspect of the iterative reflections and interactions that occur when using generative AI-powered chatbots (GenAIbots) as objects-to-think-with. This approach encourages students to think critically and refine their understanding of various concepts while actively engaging with the AI-powered tool.

Additionally, this study emphasized the importance of providing comprehensive training to educators before integrating these tools into their classrooms. Proper training enables teachers to effectively guide students in articulating their thoughts and constructing well-reasoned arguments. This preparation allows educators to evaluate the performance of GenAIbots when confronted with relevant and convincing counterexamples and fosters fruitful discussions among colleagues.

Although a comparison between the performances of ChatGPT and Bing Chat was not among the objectives of this study, it is important to mention that the results of this experiment suggested a variance in the performances of ChatGPT and Bing Chat, which were later confirmed in a later work (dos Santos, 2023). Despite some fundamental similarities between the two, ChatGPT notably outperformed Bing Chat in terms of response quality, context understanding, subject comprehension, and overall performance, which are summarized in **Table 1**.

By harnessing the capabilities of generative AI chatbots as learning tools, educators can create more engaging and personalized learning experiences, ultimately fostering deeper comprehension of complex concepts and promoting critical thinking in students across diverse subject areas.

CONCLUSIONS

In conclusion, our results show that GenAIbots, including ChatGPT and Bing chat, can act as powerful objects-to-think-with in STEM education, offering interactive and personalized learning experiences. They provide tailored content, instant feedback, and adapt to individual needs, making learning engaging and relevant. With their extensive knowledge across

Table 1. Comparative analysis of ChatGPT & Bing Chat on this experiment (dos Santos, 2023)

Aspect	ChatGPT	Bing Chat
Source of information	Relies on meticulously curated training data	Functioning as both a search engine & chatbot, relies heavily on the Internet sources
Response quality	Provides comprehensive, detailed, & accurate answers	Tends to give shorter & less informative responses
Context understanding	Recognizes nuances, subtleties, & user's perspective	Often fails to address nuances & context
Subject understanding	Superior understanding & conveyance of complex scientific concepts	Less adept at conveying complex scientific concepts
Overall performance	Demonstrated better performance in this experiment	Lower performance in comparison with ChatGPT

disciplines, GenAIbots support users in expanding their understanding. Interacting with GenAIbots helps develop critical thinking, problem-solving, and creativity skills as users articulate thoughts, ask probing questions, and analyze information. In group settings, GenAIbots foster discussions, debates, and collaborative projects, promoting knowledge sharing and deeper insights.

Incorporating GenAIbots into STEM education aligns with Papert's (1980) constructionist learning theory, which emphasizes hands-on, experiential learning and cognitive tools. By utilizing GenAIbots, students can hone their problem-solving, critical thinking, and creative skills in a more accessible and engaging learning environment, empowering them to succeed in STEM fields.

Limitations of the Study and Possible Future Studies

While this study faced limitations due to its single case study design, valuable insights emerged from the extensive interactions between students and GenAIchats. The in-depth exploration uncovered the system's untapped potential despite constraints like lack of generalizability and potential biases. Crafting good prompts is an area for future work, experimenting with different types and examining how GenAIchat's performance as an object-to-think-with improves with specific prompts like the one used in the 2nd session above. We agree with Gregorcic and Pendrill (2023) that further research is needed to fully understand their impact, especially the need for careful use to tackle ChatGPT limitations and biases, as pointed out by Bitzenbauer (2023) and the responsible and ethical use of AI technologies, as pointed by Adiguzel et al. (2023).

In conclusion, implementing GenAIbots in STEM learning requires weighing benefits and limitations, ensuring accurate information and considering reduced human interaction. Integrating GenAIbots with other educational tools or activities promoting collaborative dialogue among learners may address these concerns.

Author contributions: MARV: played a key role in conducting research & investigation process by performing experiments & collecting data & evidence, & made significant contributions to preparation, creation, & presentation of published work by

revising initial draft & providing valuable suggestions & RPDs: contributed to paper by providing ideas, formulating overarching research goals, & designing methodology for study, oversaw planning & execution of research activity & provided mentorship to individual researchers, applied formal techniques to analyze & synthesize study data, & played a key role in preparing, creating, & presenting published work, specifically writing initial draft & providing substantive translation. All authors have agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: This study does not require any ethical approval as no person aside the researchers were involved.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. <https://doi.org/10.30935/cedtech/13152>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. SSRN. <https://doi.org/10.2139/ssrn.4337484>
- Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology, 15*(3), ep430. <https://doi.org/10.30935/cedtech/13176>
- Borji, A. (2023). A categorical archive of ChatGPT failures. *ArXiv*. <https://doi.org/10.21203/rs.3.rs-2895792/v1>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *ArXiv*. <https://doi.org/10.48550/arXiv.2005.14165>
- Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.),

- Collaborative-learning: Cognitive and computational approaches* (pp. 1-19). Elsevier.
- dos Santos, R. P. (2023). Enhancing chemistry learning with ChatGPT and Bing Chat as agents to think with: A comparative case study. *SSRN*. <https://doi.org/10.2139/ssrn.4447416>
- Franciscu, S. (2023). *ChatGPT: A natural language generation model for Chatbots*. <https://doi.org/10.13140/RG.2.2.24777.83044>
- Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., & Berner, J. (2023). Mathematical capabilities of ChatGPT. *ArXiv*. <https://doi.org/10.48550/arXiv.2301.13867>
- Gregorcic, B., & Pendrill, A.-M. (2023). ChatGPT and the frustrated Socrates. *Physics Education*, 58(3), 035021. <https://doi.org/10.1088/1361-6552/acc299>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2), ep421. <https://doi.org/10.30935/cedtech/13036>
- Haraway, D. (1990). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.
- Khosrawi-Rad, B., Rinn, H., Schlimbach, R., Gebbing, Pia, J., Yang, Xingyue, J., Lattemann, C., Markgraf, D., & Robra-Bissantz, S. (2022). Conversational agents in education—A systematic literature review. In *Proceedings of the 30th European Conference on Information Systems*.
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28, 1. <https://doi.org/10.1007/s10639-022-11177-3>
- Latour, B. (1991). *Nous n'avons jamais été modernes: Essai d'anthropologie symétrique* [We were never modern: An essay in symmetrical anthropology]. La Découverte.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory: An Introduction to Actor-Network-Theory*. OUP - Oxford University Press.
- MacIsaac, D. (2023). Chatbots attempt physics homework—ChatGPT: Chat generative pre-trained transformer. *The Physics Teacher*, 61(4), 318-318. <https://doi.org/10.1119/10.0017700>
- Mehdi, Y. (2023). Confirmed: the new Bing runs on OpenAI's GPT-4. *Microsoft Bing Blog*. https://blogs.bing.com/search/march_2023/Confirmed-the-new-Bing-runs-on-OpenAI's-GPT-4
- Milmo, D. (2023). ChatGPT reaches 100 million users two months after launch. *The Guardian—Technology*. <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>
- Mishra, A., Soni, U., Arunkumar, A., Huang, J., Kwon, B. C., & Bryan, C. (2023). PromptAid: Prompt exploration, perturbation, testing and iteration using visual analytics for large language models. *ArXiv*. <https://doi.org/10.48550/arXiv.2304.01964>
- Mollick, E. R. (2023). A guide to prompting AI (for what it is worth): A little bit of magic, but mostly just practice. *One Useful Thing Blog*. <https://www.oneusefulthing.org/p/a-guide-to-prompting-ai-for-what>
- Mollick, E. R., & Mollick, L. (2022). New modes of learning enabled by AI Chatbots: Three methods and assignments. *SSRN*. <https://doi.org/10.2139/ssrn.4300783>
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- OpenAI. (2023). GPT-4 technical report. *OpenAI*. <https://cdn.openai.com/papers/gpt-4.pdf>
- Papert, S. A. (1980). *Mindstorms—Children, computers and powerful ideas*. Basic Books.
- Papert, S. A. (1999). Introduction: What is Logo? And who needs it? In *Logo philosophy and implementation*. Logo Computer Systems, Inc.
- Papert, S. A., & Harel Caperton, I. (1991). Situating constructionism. In I. Harel, & S. A. Papert (Eds.), *Constructionism: Research reports and essays, 1985-1990* (pp. 1-14). Ablex Publishing.
- Sabzalieva, E., & Valentini, A. (2023). ChatGPT and artificial intelligence in higher education: Quick start guide. *UNESCO*. <https://unesdoc.unesco.org/ark:/48223/pf0000385146>
- Swan, M. (2015). We should consider the future world as one of multi-species intelligence. Response to the edge question 2015: What do you think about machines that think? *Edge.org*. <http://edge.org/response-detail/26070>
- Taylor, C. A., Hogarth, H., Hacking, E. B., & Bastos, E. (2022). Posthuman object pedagogies: Thinking with things to think with theory for innovative educational research. *Cultural and Pedagogical Inquiry*, 14(1), 206-221. <https://doi.org/10.18733/cpi29662>
- Turkle, S. (1984). *The second self: Computers and the human spirit*. Simon and Schuster.
- Turkle, S. (2007). *Evocative objects: Things we think with*. MIT Press.
- Yin, R. K. (2011). *Applications of Case Study Research*. SAGE.