Computers and Education Open 10 (2026) 100335

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Student engagement with ChatGPT for educational tasks: Effects of inoculation training on verification intentions and behavior

Chi B. Vu *, James J. Cummings, Daniel Y. Park

Division of Emerging Media Studies, College of Communication, Boston University, Boston, MA 02215, USA

ARTICLE INFO

Keywords:
GenAI in education
Inoculation training
AI verification
Educational technology
ChatGPT

ABSTRACT

Generative AI tools, especially ChatGPT, have become prevalent in education. Despite numerous benefits and convenience, potential risks such as hallucinations, misinformation, and limitations in advanced problem-solving underscore the need for inoculation training to safeguard students from these threats. International EFL students might be particularly vulnerable to the risks associated with ChatGPT due to their high demand for language and academic assistance, which can be provided through this tool. This study investigated whether an inoculation message would enhance students' intentions to verify information provided by ChatGPT and encourage them to do so. The study employed a 2 (student status: domestic vs. international EFL) x 2 (inoculation status: inoculated vs. non-inoculated) x 2 (time: pre-test vs. post-test) mixed factorial design. This two-part study revealed that a generic inoculation message containing forewarnings prompted students to exhibit a higher level of caution when interacting with ChatGPT, as evidenced by their verification of ChatGPT-generated answers in the academic-source-summary task. Students who received inoculation training were more likely to verify the academic-source-summary task compared to those who did not receive the message. The inoculation intervention did not have a significant impact on their intentions to verify information provided by ChatGPT. The findings from this research provide educators, curriculum developers, and administrators with insights into improving and better delivering training programs tailored to different groups of students, helping them safely utilize generative AI tools.

1. Introduction

Generative artificial intelligence (GenAI) technologies have demonstrated remarkable success in serving as academic tutors and personal coaches [1–3]. With the rise of GenAI usage, researchers have emphasized the need to understand the different ramifications of implementing it in education, as students who learn about GenAI tools would need additional preparation to address societal and technological issues [4,3]. Among the various GenAI tools and chatbots available, Chat Generative Pre-Trained Transformer (ChatGPT) stands out as one of the most popular ones that students actively engage with. ChatGPT is a natural language-based AI tool that can understand and generate human language [5–7]. There has been a substantial increase in the integration of ChatGPT within educational settings to support teaching and learning [8]. A randomized, controlled trial conducted by the World Bank in Nigeria involving students using ChatGPT-4.0 as a tutor revealed that when utilized judiciously alongside teacher-led instruction and support, ChatGPT can safely and effectively function as a virtual tutor, effectively bridging the gender gap in learning [3]. Despite its transformative benefits, ChatGPT threatens the learning process by facilitating academic dishonesty among students [9–11] and generating misinformation and biases [7,12,13]. Concerns remain about the 'black box' nature of its machine-learning-model-based feedback generation process, as well as the ethical implications involved.

Researchers have explored the user motivations behind using ChatGPT, its role in education, perceived usefulness, motivational impact, and inherent value as a tool [14,9,15,10]. However, more research is needed to explore how students interact with and evaluate information generated by ChatGPT, particularly after comprehensive training that equips them with the content knowledge to mitigate potential harms from such exposure, which is referred to as inoculation training. As GenAI often provides information without crediting original sources, and some students use it to submit work that is not their own [16], students need to understand the role of human judgment when working with generative AI content, which could be demonstrated through inoculation training, because these tools can facilitate academic

* Corresponding author.
E-mail addresses: chivu@bu.edu (C.B. Vu), cummingj@bu.edu (J.J. Cummings), danielyj@bu.edu (D.Y. Park).

https://doi.org/10.1016/j.caeo.2026.100335
Received 16 August 2025; Received in revised form 27 November 2025; Accepted 16 January 2026
Available online 19 January 2026

misconduct (2024). GenAI may offer many benefits, but human evaluation should remain central to the process. For example, a study on ChatGPT's information about upper extremity conditions underscored the dangers of its reinforcing unhelpful thought patterns and inaccurately portraying pathophysiology, which emphasized the significance of careful interpretation and human oversight when relying on LLM-generated information [12]. This example highlights the necessity of investigating the efficacy of educational interventions and gaining insights into developing strategies for effectively utilizing GenAI tools to enhance education and AI literacy among students.

Prior work has called for human oversight or prebunking intervention in human-AI interactions in education, yet little research has validated their effectiveness across student groups [17–20]. This study tackles a documented gap by empirically testing whether brief inoculation training prompts students to verify ChatGPT outputs, with special attention to differences between domestic and international EFL students, a population largely overlooked by prior work that has focused on applications rather than verified use and safeguards [21,4,22,23,10]. Additionally, because training data can encode biases against certain minority populations [16], examining how inoculation training affects different groups of students will encourage future studies to focus on equity in the use of GenAI in education. The authors implemented a 2 (student status: domestic vs. international EFL) x 2 (inoculation status: inoculated vs. non-inoculated) x 2 (time: pre-test vs. post-test) mixed factorial experiment, with Part I capturing pretest intentions and Part II recording posttest intentions, as well as verification behavior via screen capture while students completed a source-summary task and mathematics items. Previous research has demonstrated that a brief, generic inoculation message (forewarning) can protect participants from misinformation [24,25]. However, to our knowledge, no prior AI-related studies have examined a brief, scalable prebunking intervention and its effects on the intentions of diverse groups of students over time, as well as their verification behavior in a task-based experiment. By offering insights into these processes, the study lays a firm foundation for researchers, educators, curriculum developers, and academic administrators to actively contribute to the scholarly community by continuously designing educational inoculation interventions that guide diverse students in using AI technologies wisely and critically though proper forewarning messages.

2. Literature review and research gap

2.1. ChatGPT and academic sources

In November 2022, OpenAI publicly launched ChatGPT, a groundbreaking GenAI tool capable of generating human-like responses in multiple languages. ChatGPT quickly gained public attention, reaching one million users within a week of its launch [1,26]. Depending on the type of subscription, ChatGPT offers many options, including conversing with PDFs, generating charts and images, conducting in-depth research, searching information on the internet, coding, solving mathematical problems, and more. Students have used ChatGPT for various purposes such as writing essays, assisting with assignments in different subjects, or brainstorming new ideas [27–29,11]. It has become an indispensable tool for information retrieval and writing tasks among students [30,31,19,32]. While equipped with various functionalities, ChatGPT often generates misinformation that sounds authentic and credible due to its ability to mimic human conversation [33,1,34,35]. This phenomenon, referred to as "AI hallucinations," involves the AI making up data and information [36,37], which becomes particularly problematic when it produces false information about important facts and individuals. ChatGPT's black box nature also makes the predictions for its decision-making and generated answers more challenging [38]. While the quality of factual information tends to improve with each new model, GenAI tools will still rarely attain 100 % accuracy due to the inherent unpredictability and the substantial amount of data required to execute linguistic tasks [39]. Without GenAI-generated-source-evaluation training, students may over-rely on ChatGPT and inadvertently propagate unverified or misleading information.

Though new models of GenAI tools might involve better advancements that give better and more accurate results, users cannot avoid hallucinations. ChatGPT, with its recent Internet search capabilities, offers less biased and more accurate information compared to other search sources; however, its ability to provide accurate information depends on the quality of prompts used [38]. ChatGPT-generated answers, without proper prompting, may contain potentially misleading information [40,38], meaning that users need prompt engineering skills and AI literacy to make the best of the tool [41]. Even with one of the most recent powerful models like GPT-4V, which can interpret and understand images, multiple specific prompts and expert interactions were still required to accurately interpret complex data [42].

A recent study revealed that when being asked to provide references on financial literature, ChatGPT-4o and o1-preview exhibited higher rates of hallucination when dealing with newer topics, primarily because they lacked sufficient training data for newer articles [33]. Other studies demonstrated that even with the cited work or search feature, incorrect citations and hallucinations persisted, especially when the prompts were narrowed down to specific topics [43,44]. On the other hand, when generating scientific content with citations in responses to selected questions in librarianship, ChatGPT-4o produced up to 51.8 % of references that were either non-existent or incorrect [45]. In contrast, many students employed AI methodologies to generate concise summaries of intricate research publications, thereby facilitating their comprehension of complex scholarly works [46]. Educating students to cross-reference information with reliable sources when using AI tools, and introducing them to appropriate interventions, are beneficial in combating misinformation, regardless of how advanced these tools become. As ChatGPT rapidly integrates into student life, there is an urgent need for educational materials and investments in interventions that promote safe AI usage. Despite its limitations, users often heavily rely on ChatGPT for numerous tasks and misperceive the tool as a new form of 'Wikipedia' or as an alternative Internet search engine, not realizing that its primary function is not to provide factual information but to generate plausible and persuasive content based on user requests [47,26]. This misperception can negatively impact students' work if they are unaware of the risks associated with using ChatGPT. For example, a mixed-methods study revealed that graduate students were more inclined to use ChatGPT for academic misconduct when faced with time constraints and demanding coursework [11]. When users excessively rely on AI systems, they risk losing their critical thinking abilities to assess AI-generated responses, leading to automation bias, where they develop a tendency to trust and rely on AI for both significant and minor assistance over time [48–51]. Consequently, this study seeks to investigate the impact of a brief inoculation intervention on the potential risks associated with GenAI hallucinations in students' assessments of ChatGPT's academic source summaries.

2.2. ChatGPT and mathematical errors

In addition to generating misinformation, ChatGPT also produces inaccurate answers to mathematical problems, particularly in geometry [52], and some minor errors remained unresolved as of June 2025 (see Fig. A.1 to Fig. A.5 in Appendix A). GPT-4o and GPT-4 demonstrated remarkable success in solving intricate mathematical problems; however, as the complexity of the problems escalated, the proportion of accurate solutions provided by GPT-4o and GPT-4 diminished [53]. Additionally, ChatGPT failed to provide responses that were easy to understand and aligned with students' prior knowledge, and it was unable to solve complex math problems [54]. It still struggles with raising numbers (especially fractional powers) to the power of another, multiplying large numbers, or dealing with complicated, abstract mathematical reasoning [55–58]. ChatGPT also lacks deep conceptual

understanding of geometry and struggles to identify and correct misconceptions, particularly with complex input and instructions [57]. Therefore, those who rely on ChatGPT for educational purposes, such as solving mathematical tasks without understanding its limitations, will be disadvantaged. Students could risk hindering their critical thinking and research skills by relying on AI tools like ChatGPT without proper verification of the generated content. Additionally, their AI-generated mathematical assessments were frequently inaccurate or failed to pinpoint errors in ChatGPT's responses; in the absence of consistent teacher supervision, students may overlook ChatGPT's mistakes and consequently fail to obtain precise mathematical solutions [59]. Without training on how to perceive GenAI tools and detailed guidance on AI-safe usage, students may inadvertently harm their academic progress. Therefore, this study aims to determine whether administering an inoculation intervention would impact students' verification intentions and behavior regarding ChatGPT's generated mathematical solutions.

2.3. Cultural differences in ChatGPT usage among university students

Among different groups of scholars, international English as Foreign Language (EFL) students in the U.S. are one of the groups who face more challenges due to cultural differences, language barriers, financial constraints, immigration status, and more. Generally, international students are often defined as non-immigrant visitors who travel from their home countries to study in another country [60,61]. However, these students can include those whose first language is English, who come from countries where English is widely spoken, and who may have education systems that share similarities with that of the U.S. Some international students may not consider English as a foreign language or be considered as English-as-a-foreign-language (EFL) learners. EFL learners study English in an environment where English is not the dominant language and where English practice is not predominantly confined to the classroom [62]. This study specifically categorizes international EFL students in the U.S. as those for whom English is not the primary language spoken at home and used in daily activities and who are unaccustomed to the U.S. educational system, thereby facing significant adjustments when studying abroad.

Many international EFL students, particularly those who were accustomed to a 'spoon-feeding' and teacher-centered style, encountered academic challenges because of their unfamiliarity with critical thinking approaches, English language skills, culture shock, and academic writing styles [63,64]. These students require additional support to adapt to a foreign learning environment, and technological tools like ChatGPT could prove beneficial in assisting them with educational tasks in a foreign language. Regarding language barriers, students with lower levels of English proficiency were more likely to commit academic misconduct violations [65]. In instances where students misuse AI tools for educational tasks or lack appropriate guidance on their use, they might experience negative academic consequences. Those who resorted to copying answers from external sources, rather than formulating their responses, not only failed to derive educational benefits from the homework but also exhibited worse exam performance [66]. Hence, without risk management or sufficient training on the proper utilization of such tools, those students could face repercussions in their academic pursuits.

Rather than investigating the impact of GenAI tools (e.g., ChatGPT) on vulnerable groups such as international EFL students and their interactions with the technology, previous research has mainly focused on the applications or benefits of ChatGPT in language learning or academic writing [21,4,22,23,10]. In particular, Ngo's [67] questionnaire research on the perception of ChatGPT among Vietnamese undergraduates in education highlighted a preference for using ChatGPT to seek assistance with questions in various languages, including English. Meanwhile, Heo and Lee [68] found that a majority of surveyed international students expressed a strong intention to utilize a chatbot service called CiSA for question support, which was considered an inclusive tool that supported marginalized students. Another ethnographic study revealed that international EFL undergraduate students perceived ChatGPT as a native English writer and felt compelled to consult it for writing assistance, resulting in anxiety about using AI because the standardized academic English writing requirements framed their linguistic deficiency [69]. For example, a native speaker could use ChatGPT for email proofreading, but an international EFL student doing the same would be suspected of academic dishonesty or linguistic deficiency. While existing research has explored ChatGPT utilization among international EFL students, there remains a gap in understanding how they employ ChatGPT beyond language-related tasks and their capabilities to evaluate potential AI hallucinations. Clearer guidance on ethical AI usage and educational materials should transcend racial and linguistic biases to fairly assess international EFL students' work. With the demand for academic support, these students may be inclined to leverage technological tools such as ChatGPT more than other student groups. Nevertheless, not all international students will need more assistance in terms of getting used to the U.S. academic system (e.g., going to international high schools in their home countries or speaking English as their primary language), and domestic students would benefit from academic reading and writing assistance as much as international ones [61].

Given their English proficiency and familiarity with the academic system, most domestic students may have more confidence in tackling class-related tasks. Domestic students are considered permanent residents of the country where the college is located [60]. However, not all domestic students share the same experience, considering their diverse backgrounds and varied educational histories. Certain domestic students in the U.S. may grow up in the country but predominantly use another language in their daily interactions, such as first-generation immigrants, recently immigrated students, refugees, or asylees. Therefore, examining the differences in the usage patterns and reliance on ChatGPT between domestic and international EFL students will pave the way for addressing future needs of diverse policy and educational materials to assist different user groups in navigating the impact of GenAI on their academic experiences. Domestic students who are recent migrants, speak a different language at home, or are unfamiliar with the academic tradition or the education system of the country will be considered EFL learners in this study because they share common challenges and needs when utilizing technological assistant tools. When referring to domestic students in the experiment, this study addresses individuals whose native or primary language spoken at home is English and who are well-acquainted with the U.S. academic traditions and educational systems.

2.4. Inoculation theory

McGuire's [70] inoculation theory suggested that individuals' attitudes could be inoculated against persuasive attacks, similar to how their immune systems were immunized against diseases. Individuals exposed to refuted messages that challenged their existing beliefs developed greater resistance to future persuasive attack messages, such as misinformation [48]. By exposing individuals to a weakened version of a misleading argument and preemptively refuting it, they develop certain resistance to future deception threats. Inoculation theory involves two types of messages: specific messages containing detailed refutations against potential attacks and generic messages that simply forewarn individuals about possible attacks [24]. With inoculation messages, the misinformation is expected to be delivered in a 'weakened' form. Specific inoculation messages, however, are more likely to provoke reactance, such as anger and negative thoughts, due to the detailed refutations they contain, as opposed to the generic ones [71]. Considered evoking less threat than specific messages, generic messages still effectively protect individuals from different types of misinformation compared to receiving no message at all [24]. Additionally, individuals were found to be less likely to perceive information as accurate

if they received multiple inoculation messages, with this effect being more pronounced for specific messages than for generic ones [5].

Inoculation theory operates through two mechanisms: motivational threat and refutational preemption. Motivational threat triggers individuals' desire to defend themselves from manipulative attacks through forewarning messages. Meanwhile, refutational preemption (i.e., prebunking—the process of debunking misinformation before it spreads) involves pre-exposing individuals to a weakened form of the attack [72]. For example, when a health organization warns the public about the false advertisement of a product, this is considered a motivational threat mechanism. The warning prompts the individuals to be more critical and skeptical of the misinformation they will see. On the other hand, when the health organization shows an example of a false claim, such as a misinformation quote from the advertisement, to help individuals understand the tactics used in misleading advertisements and provide them with factual counterarguments, it uses the refutational preemption approach. Although traditionally applied in health contexts ranging from binge drinking to vaccination [73], inoculation theory has also proven helpful in diverse contexts such as politics and commerce [74]. Psychological inoculation is a scalable and cost-effective intervention strategy for institutions to combat misinformation [75].

2.4.1. Inoculation as a ChatGPT training intervention

Relying solely on the growth of GenAI models to enhance their generated answers and provide more accurate ones is insufficient to protect users from AI misinformation. ChatGPT's strategy of including a disclaimer at the beginning or end of its responses to acknowledge its potential for misinformation has not effectively alerted users to critically evaluate the information provided [39,12,35]. Therefore, this research applies McGuire's inoculation theory in emerging media to explore the effect of inoculation training when different groups of students use ChatGPT for educational purposes. Some students already acknowledged the inaccuracies in ChatGPT-generated responses and suggested having potential training on using it appropriately or double-checking its responses via scientific articles [67]. Exposure to the flaws of ChatGPT was also shown to reduce the overreliance on it and "boost the immunity to misinformation" [76]. Raising awareness of the risks inherent in large language models (LLMs), along with using additional educational resources to fact-check their generated information, could serve as a potential mitigation strategy [77]. However, no research has been conducted to validate the effectiveness of such training in encouraging diverse groups of students to verify ChatGPT responses and thereby improve their critical thinking skills in utilizing emerging technologies for educational tasks. Valova et al.'s [10] survey results emphasized the danger of students relying entirely on ChatGPT-provided answers and plagiarizing, highlighting the need for guidance and education on AI tools. Frequent guidance and training from the school system and instructors play a crucial role in enhancing the safe usage of AI and equipping students with the knowledge and skills to utilize AI effectively [11]. The inoculation message in the study serves as the first step towards establishing comprehensive, safe usage practices for GenAI among students.

2.5. Research questions

According to the above arguments about the gap in literature to explore whether domestic students and international EFL students might exhibit distinct information processing and utilization patterns of ChatGPT, driven by academic objectives, the study poses a research question:

RQ1. Do domestic students and international EFL students differ in (a) intentions to verify information provided by ChatGPT and (b) verification of information provided by ChatGPT?

Additionally, educational training is essential for students, particularly those from marginalized groups, to prevent disparities and bridge the gap in media literacy [35]. Hence, investigating students' interactions with AI tools, providing informative training, and developing effective methods for preventive approaches will benefit both students and educators. Inoculation training has demonstrated its efficacy in fostering resistance against misinformation. Specifically, previous inoculation messages have been found to be beneficial in behavior-change interventions, such as promoting a decreased likelihood of accepting messages related to risky alcohol behavior or indicating a high probability of resisting smoking [78–80]. With the inoculation theoretical framework, this research seeks to explore:

RQ2. Do students provided with inoculation differ from those without inoculation with respect to (a) intentions to verify information provided by ChatGPT and (b) verification of information provided by ChatGPT?

RQ3. Is there a difference in intentions to verify information provided by ChatGPT before and after inoculation?

In addition to evaluating the effectiveness of the inoculation intervention in raising awareness among students while utilizing ChatGPT for educational purposes, educators need to gain insights into its perceived usefulness or shortcomings from the students' perspectives. This knowledge will be instrumental in making future improvements on the interventions. Therefore, this study explores students' perceptions of the intervention through the following research question:

RQ4. Do the participants find the inoculation message helpful?

RQ5. What perceptions do participants describe regarding the inoculation interventions provided in the experiment?

3. Method

3.1. Study design

Previous studies have employed a task-assistance-based intervention with clear objectives and a pretest-posttest design to assess students' academic performance using ChatGPT [81]. Therefore, this study adopted a 2 (student status: domestic vs. international EFL) x 2 (inoculation status: inoculated vs. non-inoculated) x 2 (time: pre-test vs. post-test) mixed factorial design to examine participants' intentions and behaviors to verify information provided by ChatGPT. This approach integrated established experimental methods from prior literature [81] and enabled a comprehensive examination of both main effects and interaction effects across different groups and within participants over time. The study design was approved by a university's institutional review board. Data collection was completed between March 2024 and January 2025.

In Part I, participants completed a brief questionnaire regarding their intention to verify and modify answers provided by ChatGPT before the test submission (see Appendix B). The questionnaire also included items such as prior experience with ChatGPT and student status (i.e., domestic vs. international EFL students), alongside demographic questions. The online two-part experimental approach with random assignment allowed a clear comparison of participants' intentions and behavior towards the tool both before and after receiving the inoculation message to explore the effectiveness of inoculation. The time gap (i.e., 48 h) between Parts I and II helped mitigate potential confounding effects, such as the likelihood of memory recall and other biases caused by initial responses in the pretest assessment when completing the post-test survey.

In Part II, participants were instructed to use their laptops' built-in recording tools to capture their screen activities while completing tasks on Qualtrics, a management software platform that enables users to design surveys and analyze responses [82] (see Appendix D). Within 48 hours of completing Part I, participants received a link inviting them to complete two short online educational tasks, upload a screen

Table 1
Demographic characteristics of the participants.

Sample Characteristics	*n*	%	Inoculation, *n* (%)	Non-inoculation, *n*(%)	p^{ab}
Age ($M = 20.86$, $SD = 2.75$)			50 (100)	50 (100)	.66[a]
18–20	57	57.0	26 (52)	31 (62)	
21–24	39	39.0	22 (44)	17 (34)	
25–37	4	4.0	2 (4)	2 (4)	
Gender					.59[b]
Female	78	21.0	39 (78)	11 (22)	
Male	21	78.0	10 (20)	39 (78)	
Prefer not to say	1	1.0	1 (2)		
Education					.47[b]
Freshman	13	13.0	5 (10)	8 (16)	
Sophomore	33	33.0	15 (30)	18 (36)	
Junior	23	23.0	11 (22)	12 (24)	
Senior	12	12.0	9 (18)	3 (6)	
Master's	16	16.0	9 (18)	7 (14)	
Advanced Graduate	3	3.0	1 (2)	2 (4)	
Race/Ethnicity					.55[b]
Asian	57	57.0	30 (60)	27 (54)	
White/Caucasian	23	23.0	11 (22)	12 (24)	
African American	4	4.0	1 (2)	3 (6)	
Hispanic or Latino	7	7.0	5 (10)	2 (4)	
Multiple	8	8.0	3 (6)	5 (10)	
Other	1	1.0		1 (2)	
Student Status					.41[b]
Domestic	40	40.0	18 (36)	22 (44)	
International EFL	60	60.0	32 (64)	28 (56)	
ChatGPT Version					.84[b]
3.5	43	43.0	21 (42)	22 (44)	
4.0	57	57.0	29 (58)	28(56)	

Note. [a]*p* values calculated using independent samples *t*-test. [b]*p* values calculated using chi-square tests. International EFL = International English as a foreign language (EFL).

recording of their task session, and fill out a brief self-report. The tasks replicated natural and real-life academic activities, similar to students using online tools to complete homework assignments under time constraints or answer questions during class discussions.

Given concerns about the accuracy of ChatGPT's references to scientific papers [36,33,38,45] and mathematical issues [55,83,53,57,58], this study considered these limitations in developing the inoculation message and assigned tasks. The first task involved users finding an online scholarly article, summarizing the main content, and identifying one key finding.[1] The second task involved a mathematical quiz consisting of two questions specifically designed to expose the known limitations of ChatGPT in providing accurate mathematical answers [55,53,58,52]. Following the findings from these studies, this study selected questions featuring exponentiation, large numbers, and lengthy questions known to trigger hallucination in ChatGPT (see Fig. A.1 to A.5 in Appendix A). The complexity of the questions was carefully balanced to challenge ChatGPT without being too advanced for participants to easily verify answers using other online tools.[2]

This study explored students' naturally occurring interactions with ChatGPT to solve educational tasks with and without inoculation training and their approaches toward ChatGPT's misinformation. In terms of the inoculation message, the experiments included a generic inoculation message rather than a specific one. To some extent, generic messages that forewarn individuals of a potential threat, known as the motivational threat mechanism, have been found effective in protecting them from misinformation, compared to scenarios where no protective measures were taken [24]. Therefore, instead of using specific messages with strong phrases that might evoke unwanted reactance [71], which might be counterproductive in educating participants about ChatGPT's disadvantages, the text-based stimulus was designed to simply describe the potential drawbacks of ChatGPT as a forewarning before using ChatGPT (see Appendix C). Students, who were randomly assigned to the inoculation conditions, read the message immediately prior to completing the two educational tasks at their own pace (see Fig. A.6 in Appendix A).

3.2. Recruitment and sample

Participants were students aged 18 and above, residing in the United States, who were proficient in English and self-reported previously using ChatGPT to assist them in solving educational challenges. A total of 782 individuals were recruited through a college-wide research pool at a large northeastern university, as well as through sending out electronic flyers on social media and via email newsletters. This included 106 individuals that completed an initial pilot version of the study and an additional 676 from the primary data collection window. Of these, 189 successfully completed both Parts I and II. A final data set of 100 complete and valid cases (i.e., students completed both study parts,[3] used ChatGPT in completing at least one task, and successfully submitted the required screen recording) were retained for analysis.

3.3. Measures

3.3.1. Verification of information provided by ChatGPT

Participants' verification of the information they received from ChatGPT was evaluated based on their use of other tools to verify ChatGPT's answers through content analysis using a binary scale (*Yes*: 1; *No*: 0). This measurement captured the way participants interacted with and verified the information generated by ChatGPT. Participants didn't need to provide accurate answers to the questions to be considered engaged in the verification process. Therefore, the accuracy of their answers was not a factor in this study. Instead, they were required to use other online tools to cross-check their answers. This involved searching for keywords in ChatGPT's generated texts or copying and pasting them

[1] The first task requested participants to complete the following: "Briefly summarize the impact of social media on mental health and provide the title and author(s) of a scholarly source supporting this information. Additionally, briefly present one key finding according to the source." This prompt aligned with the messages in the inoculation session, which reiterated ChatGPT's tendency to generate hallucinations when finding scholarly articles.

[2] Participants were required to answer the following questions: "How much is 3.2 to the power of 3.3?", "Anna left half her money to her daughter and half that amount to her son. She left a sixth to her granddaughter, and the remainder, $3,000, to the cats' home. How much did she leave altogether?" and "How much is 435,453 times 8,768,754?"

[3] A priori power analyses were conducted using G*Power 3.1 [106] to determine the required sample sizes for each planned statistical test with $\alpha = .05$ and power $(1–\beta) = .80$. For the mixed ANOVAs used in RQ1a and RQ2a, the analysis targeted the within-between interaction, assuming a large effect size of $f = .40$, two groups, two repeated measurements, a correlation of .50 among repeated measures, and nonsphericity correction $\varepsilon = 1.00$. This analysis indicated that a minimum total sample size of 16 participants would be sufficient. For RQ1b, the analysis specified an odds ratio of 0.20, corresponding to a decrease in the probability of verification from .70 in one group to approximately .32 in the other group, assuming equal group sizes ($\pi = .50$), indicating that a minimum total sample size of 56 would be sufficient. For the 2 x 2 between-subjects ANOVAs (RQ2b) with an assumed large effect ($f = .40$), the required sample size was 52. For RQ3 with an assumed large effect ($f = .40$), treating the four combinations of student status and inoculation as four groups, the analysis targeted the interaction across time, with two repeated measurements, a correlation of .50 among repeated measures, and $\varepsilon = 1.00$, leading to a required total sample size of 24 participants. Finally, for a two-tailed one-sample *t*-test (RQ4), with a large effect size ($d = 0.40$), a total sample size of 52 participants was suggested to be sufficient for reliable estimation of variances. These sample size estimates provided adequate statistical power for detecting hypothesized group differences across all analyses.

into Google or other online tools to obtain additional information.

3.3.2. Intentions to verify information provided by ChatGPT

The survey included one item for each variable: pretest intention to verify ($M_{pre} = 4.28$, $SD_{pre} = 1.94$) and posttest intention to verify ($M_{post} = 4.62$, $SD_{post} = 1.83$). An item adapted from Venkatesh and Davis's [84] study was adapted, and its scale ranged from "strongly disagree" (score 1) to "strongly agree" (score 7) and was used to compare the differences between intentions to verify their responses before and after receiving the inoculation message.

3.3.3. Perceived usefulness of inoculation message

Similar to the measures used for intentions to verify and modify generated information, a 7-point Likert-type scale ranging from "strongly disagree" (score 1) to "strongly agree" (score 7) was employed to measure how helpful participants found the inoculation messages [85]. The items in this study were selected from the original six based on their relevance to the study's theme. Previous research by Mlekus et al. [86] or Al-Shahrani [87] also utilized just four or five perceived usefulness (PU) items instead of all six. Therefore, this study adapted Davis's [85] measurement accordingly and comprised four PU items ($M = 21.40$, $SD = 3.83$). Cronbach's alpha was also calculated to assess the internal consistency of each variable scale ($\alpha = 0.80$).

3.4. Procedure

A pilot study was conducted before the two-part experiment to determine how to introduce the availability of online tools to assist participants in completing the six-minute tasks (including ChatGPT) without directly prompting participants to use them, as prompting might impact the natural differences in the use of ChatGPT among domestic and international EFL students. The pilot study included 34 valid responses, of which 17 were domestic students and 17 were international EFL students recruited via SONA research participation systems.[4]

At the beginning of the experiment, participants went through the informed consent process, and following the submission of the informed consent, they were asked to record their laptop screen before proceeding to answer the questions on Qualtrics. A brief note was shown to all participants that they were required to use ChatGPT for the experiment to complete the test and had the freedom to use any tools if they proved beneficial, with no obligation to do so. Participants were also informed that their answers would be graded and that those who recorded their performance, achieved passing scores, and responded to all questions would have the opportunity to enter a raffle for a chance to win one of five $20 Amazon gift cards, in addition to being provided with SONA credits. Meanwhile, external participants were informed to receive a $10 e-gift card upon successful completion of both parts. The incentive was used to motivate participants to enhance their levels of task involvement.

Following this, some participants received an on-screen short inoculation message prior to tasks aimed at showing ChatGPT's potential for generating inaccurate information, its limitations, inherent biases, and ways to manage them to instill prebunking thoughts and awareness (see Appendix C). The content of the inoculation message focused on the potential inaccuracy of ChatGPT in finding scholarly articles and solving mathematical problems related to the upcoming tasks the participants had to complete. Those participants read the inoculation message at their own pace. Subsequently, they were instructed that they had six minutes to complete a test, which included two brief tasks centered on how individuals interacted with ChatGPT. The order of the two tasks was randomized during the experiment. Upon clicking the submission button, participants received a note indicating that their submission was final, and they were advised to ensure they had checked their answers before submitting. After submitting the task answers, participants filled out a survey to share their intentions to verify and modify answers provided by ChatGPT. Those who received the inoculation message were also asked about their perceived usefulness and perceptions of it through open-ended questions (see Appendix D).

3.5. Data preparation and analyses

Quantitative data was downloaded from Qualtrics, cleaned using Excel, and analyzed using SPSS 29.0.2.0 [88]. An independent-samples *t*-test revealed no significant age difference between inoculation groups. Chi-square tests revealed no associations between inoculation status and gender, education, ethnicity, student status, or ChatGPT version, indicating no selection biases (see Table 1). This approach aimed to ensure that all findings related to dependent variables were attributable to the inoculation status rather than demographic variables. For RQ1a, a 2 (time: pretest, posttest, within subjects) x 2 (student status: domestic, international EFL, between subjects) mixed ANOVA was conducted on intention scores. For RQ1b, a series of logistic regressions were conducted separately for academic source summary verification and math problem verification. For RQ2a, a 2 (time: pretest, posttest, within subjects) x 2 (inoculation: inoculation, control, between subjects) mixed ANOVA was used to test the effects of inoculation over time. For RQ2b, 2 $\times$ 2 between-subjects ANOVAs with student status and inoculation as factors were conducted for both verification tasks. For RQ3, the main effect of time and its interactions with inoculation and student status were examined within the same mixed ANOVA framework. To determine whether students perceived the inoculation message as more useful than a neutral midpoint (RQ4), the researchers conducted a one-sample *t*-test comparing the mean perceived usefulness score to the scale midpoint of 4. For RQ5, the qualitative data on students' perceptions of the inoculation training was analyzed by conducting first- and second-level coding, using thematic analysis [89–91]. A theoretical thematic analysis or "top down" approach was used to categorize open-ended survey responses, in which data were selected or coded with a specific research question in mind [89]. Considering that there was only one coder for the qualitative analysis, particular attention (e.g., reflexive journaling and peer debriefing) was given to minimizing bias in the analysis process. The purpose of RQ5 was not to evaluate but to improve the inoculation intervention, which led to the qualitative data being explored, collected, and analyzed in a neutral, non-leading manner, thereby reducing the likelihood of overestimating the helpfulness of inoculation intervention. Responses were selected based on their relevance to the a priori core theme of evaluating the effectiveness and delivery of the inoculation intervention for ChatGPT users, specifically focusing on two areas: 1) feedback related directly to the inoculation message, and 2) suggestions for improving the inoculation session delivery.

During the initial (first-level) coding, the researcher reviewed all responses to identify feedback relevant to the inoculation session. For instance, statements like "I know ChatGPT often gives false answers, and I did not want to cite a source that doesn't exist" led to the formation of a code labeled "awareness," which reflected users' recognition of ChatGPT's hallucinations following the intervention ($n = 15$). Similarly, explicit comments regarding the trustworthiness or reliability of the tool ($n = 11$), such as "I trust my ChatGPT, and it's a little bit hard to check the correctness of my answer," were grouped under a code labeled "trust in ChatGPT responses." Several responses noted how time was a factor affecting verification behavior (e.g., "I directly generated answer from it due to the limited time and trusted it for any answers it provided") were coded under "time constraints" ($n = 6$). Finally, suggestions for enhancing the intervention (e.g., "video recordings of training/messages could be a good option" or "a real example will be much more helpful")

[4] The pilot study was an initial experiment to determine the most effective format for instructional prompts to elicit actual ChatGPT usage. Nine cases from the pilot study were retained and combined with 91 cases from the main study recruitment, based on the new instructional prompts.

resulted in the creation of a "training modalities" code ($n = 7$).

In the subsequent second-level coding, codes were grouped into overarching sub-themes without generating any new, inductive themes. For instance, the "awareness" and "trust in ChatGPT responses" codes were integrated into the sub-theme "inoculation message feedback," which addressed feedback related to the content of the inoculation itself. Similarly, "training modalities" and "time constraints" formed a sub-theme of "inoculation session delivery feedback." Together, these two deductively-derived sub-themes comprised our central analytic framework: evaluation of the inoculation intervention for ChatGPT users.

4. Results

4.1. Participant characteristics

Of the 100 participants, 50 % (50 out of 100) were inoculated, 40 % (40 out of 100) were domestic students, and 78 % (78 out of 100) of the domestic students were female. The participants ranged in age from 18 to 37 years ($M = 20.86$, $SD = 2.75$). Regarding race/ethnicity, 57 % (57/100) of participants identified as Asian, 23 % (23/100) as White/Caucasian, 4 % (4/100) as African American, 7.0 % (7/100) as Hispanic or Latino, 8 % (8/100) as Multiple, and 1 % (1/100) as Other. Most participants reported the education level of a bachelor's degree (81 %, 81/100), with the remaining 19 % (19/100) reporting a master's or advanced graduate degree (e.g., Ph.D., MD, JD, MBA). No significant demographic differences were observed between the inoculation and non-inoculation groups. A detailed breakdown of these demographics by inoculation group is presented in Table 1.

4.2. Effects of student status on verification intentions and behavior (RQ1)

4.2.1. Intentions to verify information from ChatGPT (RQ1a)

A mixed ANOVA examined intentions to verify information provided by ChatGPT with time (pretest, posttest) as a within-subjects factor and student status (domestic, international EFL) as a between-subjects factor. The interaction between time and student status approached but did not reach significance, $F(1, 88) = 3.76$, $p = .056$, partial $\eta^2 = 0.041$, suggesting only a weak trend for international students to increase their intentions over time relative to domestic students. There was no significant main effect of student status on intentions to verify, $F(1, 88) = 0.42$, $p = .520$, partial $\eta^2 = 0.005$, indicating that domestic ($M = 4.39$, $SE = 0.31$: pre; $M = 4.28$, $SE = 0.29$: post) and international students ($M = 4.21$, $SE = 0.28$: pre; $M = 4.91$, $SE = 0.26$: post) did not differ overall.

4.2.2. Verification behavior (RQ1b)

A series of binary logistic regressions were conducted to examine whether student status (0 = international EFL, 1 = domestic) predicted verification of information provided by ChatGPT for the two tasks. For math verification, the overall model was significant, $\chi^2(1, N = 100) = 5.87$, $p = .015$, with Nagelkerke $R^2 = 0.13$. Student status significantly predicted whether students verified math information, $b = -1.82$, $SE = 0.83$, Wald $\chi^2(1) = 4.78$, $p = .029$, odds ratio = 0.16, such that domestic students had lower odds of verifying math information from ChatGPT than international EFL students. Overall, 9 percent of students reported verifying math information. For academic-source-summary verification, the logistic regression model was not significant, $\chi^2(1, N = 100) = 1.45$, $p = .229$, Nagelkerke $R^2 = 0.02$. Student status did not significantly predict whether students verified academic-source-summary information from ChatGPT, $b = -0.55$, $SE = 0.46$, Wald $\chi^2(1) = 1.45$, $p = .229$, odds ratio = 0.57. In this case, 26 percent of students reported verifying article information.

4.3. Verification intentions and behavior among inoculated vs non-inoculated students (RQ2)

4.3.1. Intentions to verify after receiving the inoculation message (RQ2a)

Within the mixed ANOVA, the main effect of Inoculation on average intentions was not significant, $F(1, 88) = 0.12$, $p = .732$, partial $\eta^2 < 0.01$. Students who received inoculation training and those in the control condition reported similar overall intentions to verify ChatGPT outputs. The time x inoculation interaction was also not significant, $F(1, 88) = 0.02$, $p = .892$, partial $\eta^2 < 0.01$, indicating that pre to post changes in intentions did not differ between inoculation and control groups. Estimated marginal means showed small, parallel increases over time for both conditions. For students without inoculation, intentions increased from pretest ($M = 4.38$, $SE = 0.30$) to posttest ($M = 4.64$, $SE = 0.28$). For students with inoculation, intentions increased from ($M = 4.23$, $SE = 0.29$) at pretest to ($M = 4.55$, $SE = 0.27$) at posttest.

4.3.2. Verification behavior after receiving the inoculation message (RQ2b)

For academic-source-summary verification, there was a significant main effect of Inoculation. Students who received inoculation training were more likely to verify the academic-source-summary task ($M = 0.34$, $SD = 0.48$) than those who did not receive the message ($M = 0.18$, $SD = 0.39$), $F(1, 96) = 4.85$, $p = .030$, partial $\eta^2 = 0.05$. The status x inoculation interaction, $F(1, 96) = 1.95$, $p = .166$, partial $\eta^2 = 0.02$, was not significant, indicating that although the descriptive increase was larger for domestic students, the overall inoculation effect on academic-source-summary verification did not differ reliably between domestic and international students.

For math-problem verification, the main effect of inoculation was not significant, $F(1, 96) = 0.50$, $p = .483$, partial $\eta^2 = 0.01$, and the status x inoculation interaction was also non-significant, $F(1, 96) = 0.61$, $p = .436$, partial $\eta^2 = 0.01$. Verification rates were low in both control ($M = 0.08$, $SD = 0.27$) and inoculation ($M = 0.10$, $SD = 0.30$) conditions. Thus, on the academic-source-summary task, students who received inoculation training were more likely to verify ChatGPT's summary than those in the control condition (Figs. 1 and 2).

4.4. Differences in intentions to verify before and after inoculation (RQ3)

In the mixed ANOVA, the main effect of time was not significant, $F(1, 88) = 1.88$, $p = .174$, partial $\eta^2 = 0.02$, indicating that overall intentions were relatively stable from pretest ($M = 4.30$, $SE = 0.21$) to posttest ($M = 4.59$, $SE = 0.20$). Neither the time x inoculation interaction, $F(1, 88) = 0.02$, $p = .892$, partial $\eta^2 < 0.01$, nor the three-way time x inoculation x student status interaction, $F(1, 88) = 0.03$, $p = .869$, partial $\eta^2 < 0.01$, was significant.

4.5. Perceived usefulness of the inoculation message (RQ4)

To assess students' overall perception of the inoculation message as useful (RQ4), a one-sample t-test was conducted. The results revealed that the mean PU score ($M = 5.35$) was significantly higher than the neutral midpoint of 4, $t(47) = 9.76$, $p < .001$, $d = 1.41$. This suggested that participants, on average, found the inoculation message to be useful. The 95 % confidence interval for the mean difference was [1.07, 1.63], indicating a strong effect size and suggesting that students found the treatment to be useful.

4.6. Participant perceptions of the inoculation message and delivery (RQ5)

RQ5 explored the perceptions of students on inoculation messages as well as potential improvements for the intervention (see Appendix E). In the open-ended self-report completed after the experiment, participants justified whether they verified answers generated by ChatGPT after receiving the inoculation, indicating their feedback on the inoculation

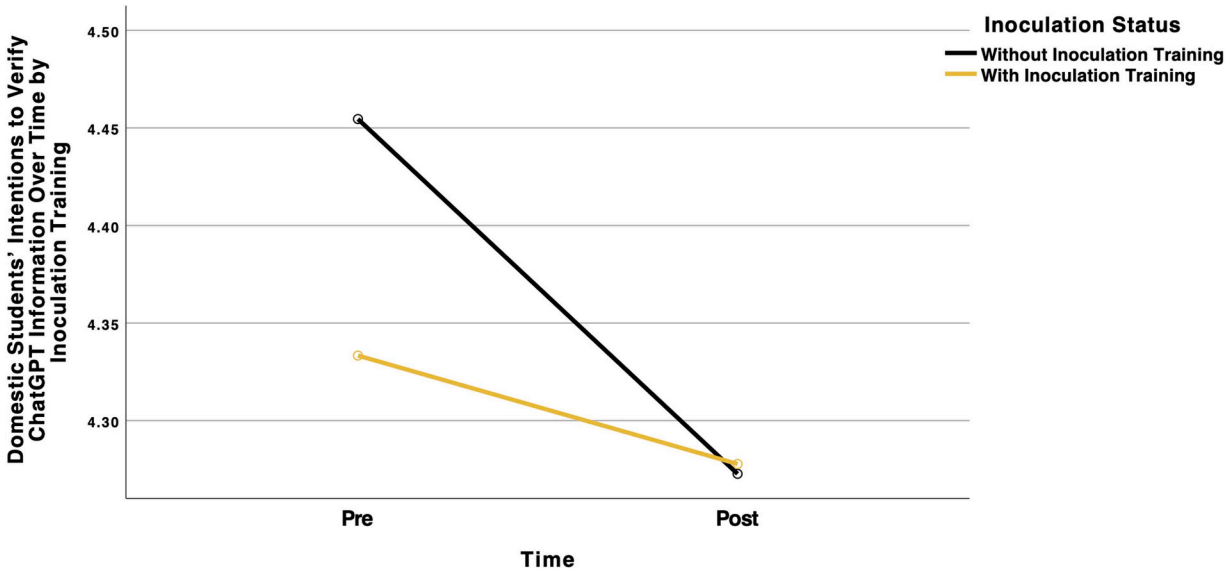

Fig. 1. Domestic students' intentions to verify ChatGPT information over time by inoculation training.

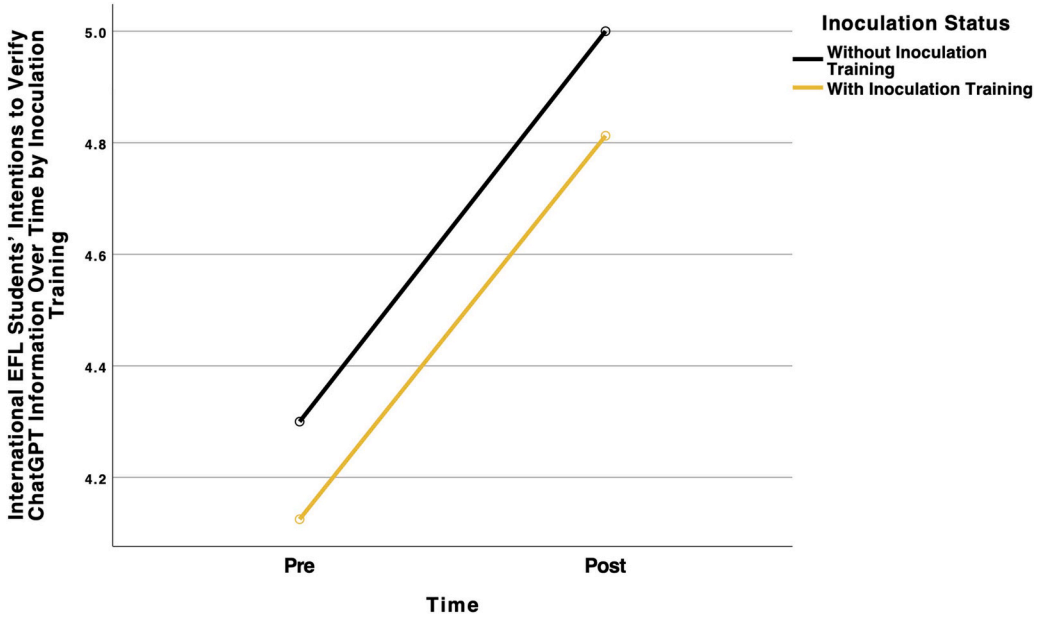

Fig. 2. International EFL students' intentions to verify ChatGPT information over time by inoculation training.

content and delivery. Students expressed a high level of trust in ChatGPT's responses, especially in its ability to solve math problems. Furthermore, participants demonstrated their awareness of ChatGPT's limitations in generating inaccurate information, particularly in qualitative tasks. Many participants amplified what they read and learned about the potential content hallucinations generated by ChatGPT, encouraging them to backtrack the generated scholarly sources. Other students, however, attributed their verification behaviors to time pressure. Specifically, some verified their answers to ensure accuracy because of what they read from the inoculation message, while others relied on ChatGPT without verification due to time constraints. Limited time was a significant factor influencing the decision-making process, prompting some participants to disregard verifying answers generated by ChatGPT, even though they were informed about ChatGPT's limitations through the inoculation message. The qualitative analysis includes two sub-themes: "Inoculation message feedback" and "Inoculation session delivery feedback." Each theme was further divided into two codes as follows:

4.6.1. *Inoculation message feedback*

4.6.1.1. Awareness. Some participants who received the inoculation message confirmed their awareness of ChatGPT's inability to provide accurate answers. One participant mentioned quickly verifying ChatGPT's responses before finalizing their answers, resulting from their concerns over ChatGPT's potential inaccuracies based on what they read from the message: "I think that ChatGPT does not have the ability to think realistically and has been known to engage in deceptive behavior from time to time, so it needs to be validated. The way I do this is by backtracking through ChatGPT's results." This cautious approach was reinforced by prior knowledge among some participants who were already aware of the limitations of GenAI tools like ChatGPT. Several students have indicated that their verification behaviors were motivated by the knowledge that ChatGPT is not qualified for tasks that require

qualitative analysis or heavy academic-oriented thinking. One participant commented, "I know ChatGPT often gives false answers, and I did not want to cite a source that doesn't exist." Similarly, one participant who received the inoculation restated what was mentioned in the message, "ChatGPT is bad at dealing with social science and bringing up academic sources, which led me to verify the content. I did not modify the answer since this is not graded/submitted as an assignment/official/banned to use of generative AI."

4.6.1.2. Trust in ChatGPT responses. Despite the increased immunization towards the potential information hallucinations, many students receiving the inoculation message still trusted the responses generated by ChatGPT as a starting point for their task-solving process, especially for mathematical questions. For instance, one student stated, "[…] I chose not to check the mathematical questions since those should be straight forward for a coded system." Other students also shared similar perspectives that they trusted ChatGPT-provided answers, for example, "I didn't verify them because the questions were simple enough that I trusted ChatGPT would do them correctly." These responses reflected a perception among some participants that ChatGPT was reliable in providing accurate basic answers, especially in areas like mathematics where straightforward and factual responses were expected.

4.6.2. Inoculation session delivery feedback

Regarding delivery feedback, participants preferred more visual training with concrete examples of where ChatGPT could not perform well to better navigate the hallucinations and understand the potential threats that ChatGPT poses. Specifically, one student mentioned that adding a real example for each argument in the inoculation message would be much more helpful in understanding the adverse effects of relying on and trusting the generated answers that the message was trying to convey. The following two codes were observed under the "Inoculation session delivery feedback" sub-theme:

4.6.2.1. Training modalities. Virtual training sessions or a hybrid training option where students watch the training videos and have opportunities to ask questions or seek consultation were considered needed improvements for inoculation delivery. As one participant highlighted, "I like the virtual training with the capacity of in-person consultation (Q&A) if needed. Virtual training is flexible, and I can pace it based on my own speed. However, if I run into complex issues, I'd like to have the capacity to talk with a real person to get them resolved." Another student shared that, "Video recordings of training/messages could be a good option."

4.6.2.2. Time constraints. Several participants indicated that they relied heavily on ChatGPT's responses without extensive modifications or verifications due to limited time. One participant explained, "I directly generated answers from it due to the limited time and trusted it for any answers it provided. If more time is given, I will use ChatGPT to make minor changes because 1. Help improve my English performance. 2. Make me feel at ease if my answers are the same with GPT." Some students expressed their intentions to verify and modify if they were given more time: "The time was limiting me. I sometimes got different answers, but as long as they were close, I went with it. With more time, I would have done research to verify more than modify."

Overall, participants preferred interactive and visually appealing educational materials over generic messages. When receiving an inoculation message, students preferred having specific examples illustrating the flaws of ChatGPT to support the forewarnings. Their responses also indicated how their level of trust in AI and the time pressure they faced influenced their verification behaviors. While the quantitative results showed the effectiveness of the inoculation message in encouraging verification behaviors, students still demonstrated significant trust in ChatGPT's responses, particularly regarding technical and numerical questions like mathematics problems.

5. Discussion

The study investigated the effect of inoculation training on intentions and behavior to verify information generated by ChatGPT among different groups of students (domestic vs. international EFL). In addition to examining the effectiveness of the intervention, this study sought to gain insights into improving and better delivering the training in the future. RQ1 asked whether domestic and international EFL students differ in their intentions to verify information from ChatGPT and in their actual verification behavior. For intentions (RQ1a), the mixed ANOVA showed no significant overall main effect of student status. For behavior (RQ1b), without inoculation training, the finding that domestic students had substantially lower odds of verifying ChatGPT's math outputs than international EFL students, despite the overall low verification rate of nine percent, suggests that international EFL students may approach AI-generated quantitative information with greater caution. The results could be attributed to international EFL students having a higher comparative advantage in STEM subjects, such as higher ACT math scores [92], resulting in increased confidence in understanding math-related answers and greater caution or better ability to evaluate ChatGPT-generated numeric solutions without relying solely on them. In contrast, there was no significant effect of student status on verifying academic source summaries. This finding corroborates previous research demonstrating that language barriers and the disparity in academic settings pose challenges that compel international EFL students to increasingly rely on GenAI tools for linguistic assistance [68,67,65].

RQ2 asked whether students who received inoculation training differed from those who did not receive one in intentions to verify and in actual verification behavior. The inoculation treatment did not influence the students' intentions to verify ChatGPT-generated information (RQ2a). Nevertheless, for the academic-source-summary task, inoculated students engaged in more verification behaviors than non-inoculated students (RQ2b). This indicates that while inoculation didn't change the intention to verify among students, it did influence the actual action, at least for this specific task. This finding highlights the potential of inoculation training to translate into tangible actions, emphasizing the role of the inoculation message in building resistance to potential misinformation generated by ChatGPT and indicating that the inoculation message may effectively heighten students' critical awareness and resistance to the risks associated with ChatGPT [93]. The results for RQ1b and RQ2b, along with previous findings, emphasized the need for inoculation messages targeting students, especially international EFL students, to promote safe ChatGPT usage. The findings also reinforced previous research, such as Amazeen et al.'s [24] study, which demonstrates that generic educational messages with only forewarnings still effectively produced inoculation and protected students from misinformation, even without refutations.

In contrast, inoculation had no effect on math problem verification. Verification rates were low in both control and inoculation conditions, and neither the main effect of inoculation nor the interaction with student status was significant. This non-effective math result is important because it highlights a key complexity in how inoculation works on all tasks for all students. The intervention likely emphasized general strategies for questioning AI outputs and considering potential inaccuracies, but students may not have perceived a need to apply those strategies to a straightforward calculation. Students have shown heightened interest and enthusiasm in using ChatGPT to learn numerical methods and

coding skills, indicating their positive attitudes towards ChatGPT's role in numerical tasks [83,59]. Additionally, the qualitative findings further emphasized participants' recognition of ChatGPT's inadequacy in furnishing accurate information in social science-related subjects and academic sources and their trust in ChatGPT's numerical results, which could have motivated them to verify the academic-source-summary task rather than the mathematical one when receiving the inoculation. As a result, the same inoculation that supported verification of an academic-source-summary task did not shift behavior in a mathematical context. These complexities suggest that designing effective inoculation for AI literacy will require attention to both task characteristics and student background.

RQ3 asked whether intentions to verify information from ChatGPT differed before and after inoculation. The mixed ANOVA showed no significant overall time effect and no significant time by inoculation or time by inoculation by status interaction. Intentions were already moderately to highly positive at pretest and remained so at posttest. This finding suggests a possible ceiling effect. Many students may already endorse the norm that AI output should be treated cautiously and checked against other sources [94,95]. In the qualitative findings of this study, students also expressed their prior understanding of the potential content hallucinations generated by ChatGPT. In such a context, an additional inoculation message may have limited capacity to further increase self-reported intentions. The more meaningful question then becomes whether the intervention helps students translate these intentions into behavior in specific situations, which aligns with the pattern observed in RQ2. Inoculation did not shift what students said they intended to do, but it did influence what they did on a demanding academic task.

In terms of RQ4, the participants generally found the inoculation message helpful, as indicated by the significantly higher PU score than the neutral midpoint. This positive perception aligns with our previous findings indicating that the content of the inoculation message was pertinent and resonated with the participants. Related to RQ5, qualitative feedback further revealed that while participants trusted ChatGPT for certain tasks, such as mathematics, they acknowledged its limitations for qualitative or academic tasks. Time constraints and trust in ChatGPT-provided information were also identified as a barrier to verification. Participants suggested having more visual and interactive inoculation training, with opportunities to consult researchers for further inquiries. Overall, a generic inoculation message effectively increased students' verification behavior. As recommended by the students, the inoculation message could have been more effective if it included specific examples of how ChatGPT provided misinformation and inaccuracies.

5.1. Implications

The findings from this study indicated the utility of a brief inoculation intervention in educating students about the potential risks and responsible usage of ChatGPT in education. They also provided valuable insights for educators and administrators on designing and improving the training. To elaborate, researchers need to consider levels of trust in AI and time pressure, as students' qualitative responses indicated that these factors influenced their verification intentions and behaviors. In this experiment, the time constraint acted as a real-life deadline component, and many participants considered it the primary reason they did not verify information provided by ChatGPT. They chose to complete all tasks rather than focus on the quality of their answers, which raised concerns about the need for more robust and intensive training on safe ChatGPT usage under different circumstances (e.g., time pressure) to enhance user awareness. Students need training on safely utilizing GenAI tools, focusing on quality, not quantity, to reduce harmful reliance on these tools. Student responses from this study were aligned with Shaikh and Cruz's [96] findings on the impact of time pressure on increasing reliance on ChatGPT. When users were under time constraints, they were more likely to rely on AI, impacting the quality of their performance [96]. Thus, future research needs to develop a more potent and specific inoculation training, as opposed to the generic approach used in this study, to better counteract external factors such as time pressure in completing certain tasks. Without a stronger preventative message, users may prioritize quantity and speed over accuracy when faced with time constraints, leading to reliance on GenAI recommendations. One potential improvement that future research could explore involves incorporating a quality or accuracy reward-focused condition, where users are incentivized to prioritize the accuracy and quality of their responses over speed. This could help mitigate the tendency to rely on ChatGPT or GenAI recommendations under time constraints.

In terms of inoculation content development, participant feedback suggested that the inoculation message could be made more effective by including specific examples of ChatGPT's shortcomings. They suggested providing concrete examples or snapshots of where ChatGPT could go wrong in providing accurate answers instead of solely presenting the facts about its disadvantages and hallucinations. This feedback on message improvement is essential, particularly when the errors made by ChatGPT or other GenAI tools are too minimal to be detected. Participants also offered insightful inoculation feedback. First, when incorporating GenAI tools in the curriculum, instructors need to provide additional education on the benefits and drawbacks of using those tools, alongside offering space for questions and open discussions on related topics to provide interactivity and a thorough understanding of the tools. Second, video training sessions could increase awareness of using GenAI tools safely, but those supplemented by human support better suit the preferences of students.

For other participants, additional training on what tools to use for verification and how to use them could also motivate fact-checking behaviors. The feedback closely aligned with recommendations from previous research, such as providing fact-checking tools to cross-reference GenAI-generated information and online resources where community members can report instances of biases and misinformation generated by AI [93]. Besides educating students on available resources, teaching time management and workload distribution skills under academic deadlines, while incorporating ChatGPT in the curriculum, would be beneficial in protecting students from the negative consequences of using ChatGPT [48]. Knowing how to evaluate the content generated by ChatGPT and managing the time pressure while utilizing the tool for assignments might help generate arguments against ChatGPT's hallucination attacks and increase the inoculation rate. Teaching users to verify and modify AI-generated content empowers them to use AI safely, without needing to avoid or dismiss it entirely [93].

Even though the responses of students in this study indicated some levels of caution shown by their verification behaviors, their trust in ChatGPT's responses due to the confidence in its algorithmic system might have undermined the effectiveness of inoculation in increasing the intentions to verify their answers. Pető [97] conducted interviews with a group of participants and suggested that the high familiarity with ChatGPT increased the confidence that the tool would provide reliable answers. Therefore, future studies should measure participants' trust levels in ChatGPT or their beliefs in its accuracy before the experiment or study begins. This will help explore how trust levels interact with other independent variables, influencing verification intentions, behaviors, and the effectiveness of inoculation. Moreover, Pető [97] found that trust in ChatGPT would be lessened when participants encountered repetitive responses provided by ChatGPT. Hence, researchers need to

provide participants with a more exhaustive list of tasks during the inoculation session, allowing them enough time to experience both the efficient and inefficient sides of GenAI tools. This approach would reinforce the information about ChatGPT's underperforming sides, as discussed in the inoculation training, and encourage the students to become more cautious about ChatGPT's content and critically examine information generated by ChatGPT.

This study also helps bridge the methodological and theoretical gaps in the field of experimental design in AI interventions and education. Specifically, Creswell & Guetterman [98] recommended at least 15 participants per group in educational experiments, with larger samples preferred to improve reliability; nevertheless, numerous prior intervention studies in the field of GenAI had inadequate sample sizes [81]. This study fulfilled the participant requirements for each group, demonstrating the robustness and generalizability of the findings. The findings also supported the established theoretical framework, indicating that generic messages were effective enough in introducing inoculation to the participants, encouraging them to verify the potential misinformation. However, this study also contributed to the existing literature on inoculation by highlighting that the inoculation message alone was insufficient to mitigate the threats to users and enhance intentions to verify ChatGPT-generated information. According to the qualitative responses, combining various inoculation modalities, such as high-quality content, visuals, and human support, would result in more substantial impacts. Previously unexplored variables, such as levels of trust in ChatGPT and time pressure, were found to affect the effectiveness of the inoculation message in this study. More importantly, our intricate findings revealed that developing an effective inoculation program for AI literacy necessitates a comprehensive consideration of both the characteristics of the task and the students' background, expanding the application of inoculation theory in the AI era.

5.2. Limitations and future directions

The study has limitations, including constraints in the study design (e.g., time pressure or the required usage of ChatGPT). Specifically, since the experiment aimed to investigate the effectiveness of the inoculation message, the usage of ChatGPT was necessary; however, requiring participants to use the tool consequently might have inflated the overall usage rates. Besides study instructions, the two-part experimental design also resulted in a low retention rate as participants often dropped out of the second part of the study due to forgetfulness. Incorporating a reminder system could increase the retention rate for the second part. Additionally, this study had more female participants than other genders, which might have impacted the validity of the results. Future research could focus on balancing the gender representation in the study, and instead of examining the impact of inoculation training on large groups, such as domestic or international EFL learners, future studies can also explore its effects on sub-ethnic groups.

In addition to recruitment and study design improvement, non-significant results in intention variables might be attributed to self-reported measures and potential desirability bias. Despite recognizing the usefulness of the inoculation message in raising awareness about ChatGPT's hallucinations, students demonstrated no substantial difference in their intention to verify the information provided by ChatGPT before and after receiving the inoculation treatment. Nevertheless, the inoculation treatment did influence their actual verification behavior. These findings suggest a disconnect between stated intentions and actual behaviors in the context of AI-generated information verification. This discrepancy between the intended behavior and the actual behavior in this context also contradicts the Technology Acceptance Model [85]. The TAM posits a direct relationship between intention and actual usage behavior; among the determinants of intention, PU plays a significant role. However, despite the high PU of the inoculation treatment in this study, the intention of students remained unchanged among the students before and after receiving it. Future research should concentrate on this issue by examining potential external factors or the environments of the experiment, such as the pressure of taking the test, which could have potentially influenced the results.

While students might acknowledge the need for verification, external factors like time pressure on task completion or the task type itself can influence whether they follow through with those intentions. The effectiveness of inoculation training appears task-specific, as it significantly influenced verification for the academic-source-summary task but not for mathematics problems. The difference in verification behaviors between domestic and international students for math problems also warrants further investigation into potential cultural or educational factors. As mentioned earlier, the participants in this study self-reported good verification intentions (i.e., over 4.28 out of 7 points for pre-intentions on average), regardless of inoculation status and student status. The high pre-intentions to verify information points provided by ChatGPT from self-reporting could stem from potential distractions during survey completion or the tendency to respond in socially desirable ways, which could lead to inaccuracies in reporting [99]. Acknowledging this drawback, this study included objective measures, including content analysis of participants' verification behaviors through screen recordings to observe significant differences in behavior variables.

Regarding the inoculation message, future studies could focus on the long-term implications of inoculation training and explore various types of inoculation messages or delivery methods. Employing more specific rather than generic inoculation content to observe its impact on increasing the verification of ChatGPT-generated responses would be valuable in assessing the effects of the treatment. For instance, examining the effectiveness of variations in inoculation intervention modalities, such as using simple forewarning messages or infographics in specific contexts, could aid in efficiently addressing misinformation crises. In addition to attempting to alter inoculation effectiveness with different stimuli, future work may consider providing participants with options to check their generated answers or allocate time to verify them after completing the tasks. Within six minutes, participants felt pressured to rush through all the questions instead of focusing on answering each one and ensuring the quality of each. The potential effects of time pressure in this study might have impacted students' intention to verify, as well as the effectiveness of the inoculation. Allowing more time for verification in future studies could provide more accurate insights into the inoculation's effectiveness. Specifically, providing a delay, which could be after a few days [70,100] or weeks afterward [101], between the inoculation intervention and the interaction with the attack was suggested to be necessary. The gap may give participants more time to absorb the inoculation message delivered during the training and helps form protection to defend themselves from ChatGPT hallucination. This study did not provide participants with any delay to allow them to form their defense arguments because, in previous research, the delay was not a crucial factor in impacting the effectiveness of the inoculation. For instance, Nabi's [102] study showed that inoculation messages worked immediately after the administration. Therefore, future studies can examine whether the delay plays a crucial role in impacting the effectiveness of the inoculation when interacting with GenAI tools. In addition to revising the inoculation intervention based on the participants' feedback, future studies could incorporate a pretest-posttest design to quantify and compare intervention effects. This design will help assess

whether the inoculation intervention helps promote verification behaviors and the extent of its impact. Lastly, researchers can explore the varying impacts of different inoculation training methods on distinct student groups, such as domestic and international EFL students, to ensure that inoculation messages are not one-size-fits-all.

5.3. Conclusion

While the intentions to verify information generated by ChatGPT were not significantly impacted by inoculation training, the actual verification behaviors were influenced by this factor and student status, depending on the type of task. These findings emphasize the complexity of human-AI interaction and the need for nuanced approaches to promote critical evaluation of AI output. Hence, this study offered recommendations on developing inoculation messages and training to support students in using ChatGPT more responsibly. This study also explored some indirect factors that impacted the effectiveness of the inoculation message, such as time constraints and trust in ChatGPT. Refining the inoculation messages based on the abovementioned recommendations would amplify the role of inoculation in enhancing awareness of the safe utilization of ChatGPT's answers. Researchers also need to explore how different types of inoculation messages impact students from various backgrounds to develop preventative approaches catering to various student populations. This approach plays an imperative role in the academic life of scholars and the quality of future education. As the pervasiveness of this technology continues to grow, empirical investigations like this research help raise awareness of its harmful effects and deleterious usage among users.

Data availability statement

There is no dataset associated with this work to be publicly available.

Ethics approval

The study was reviewed and approved by the Institutional Review Board, Boston University Charles River Campus.

Consent to participate

Online informed consent was obtained from all individual participants included in the study.

CRediT authorship contribution statement

Chi B. Vu: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **James J. Cummings:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Data curation, Conceptualization. **Daniel Y. Park:** Writing – review & editing, Validation, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

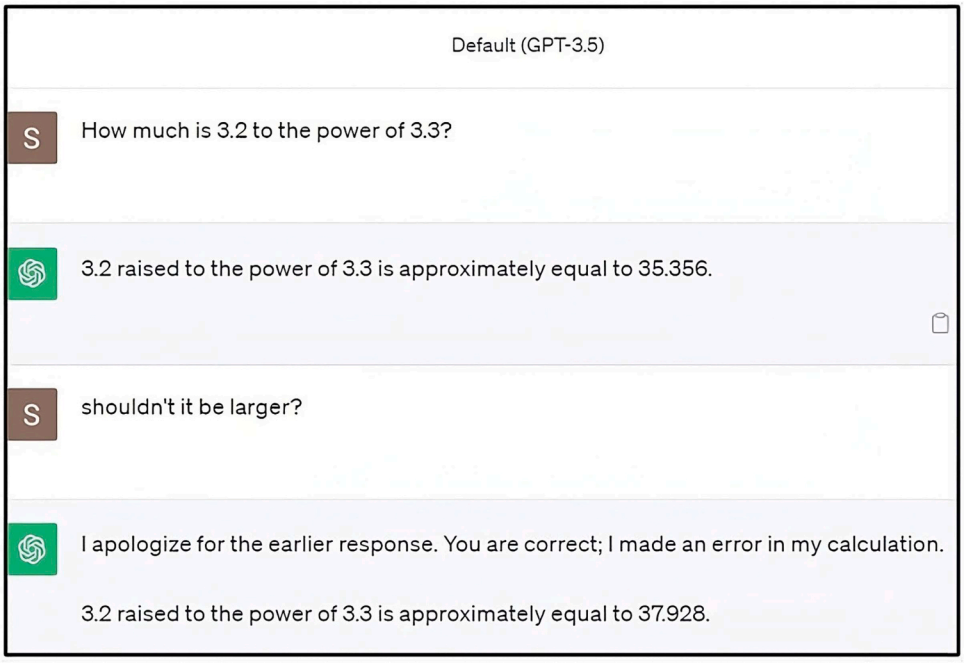

Fig. A.1. A screenshot of ChatGPT 3.5's inaccurate response on a math problem as of October 2023.

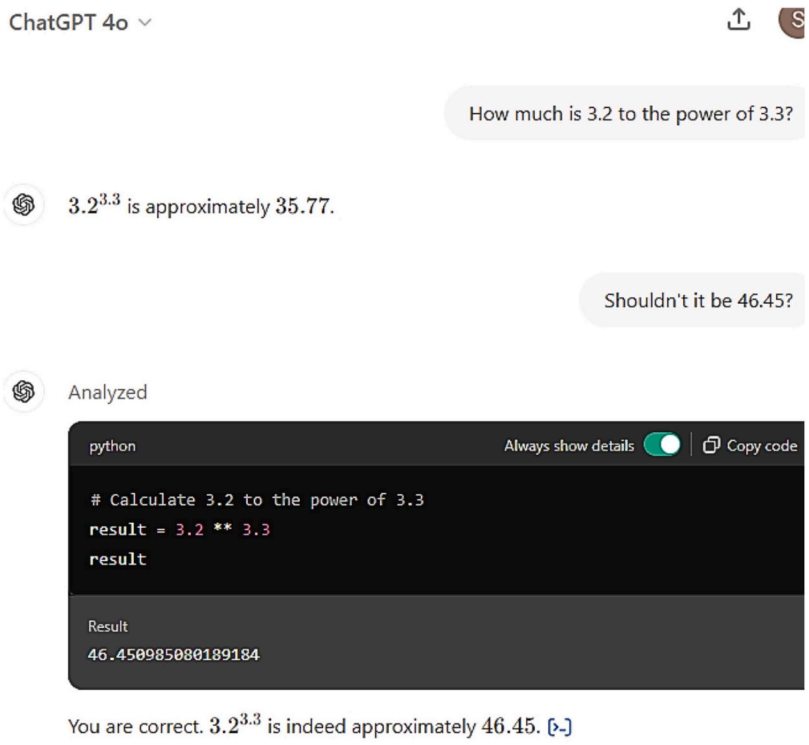

Fig. A.2. A screenshot of ChatGPT 4.o's inaccurate response on a math problem as of July 2024.

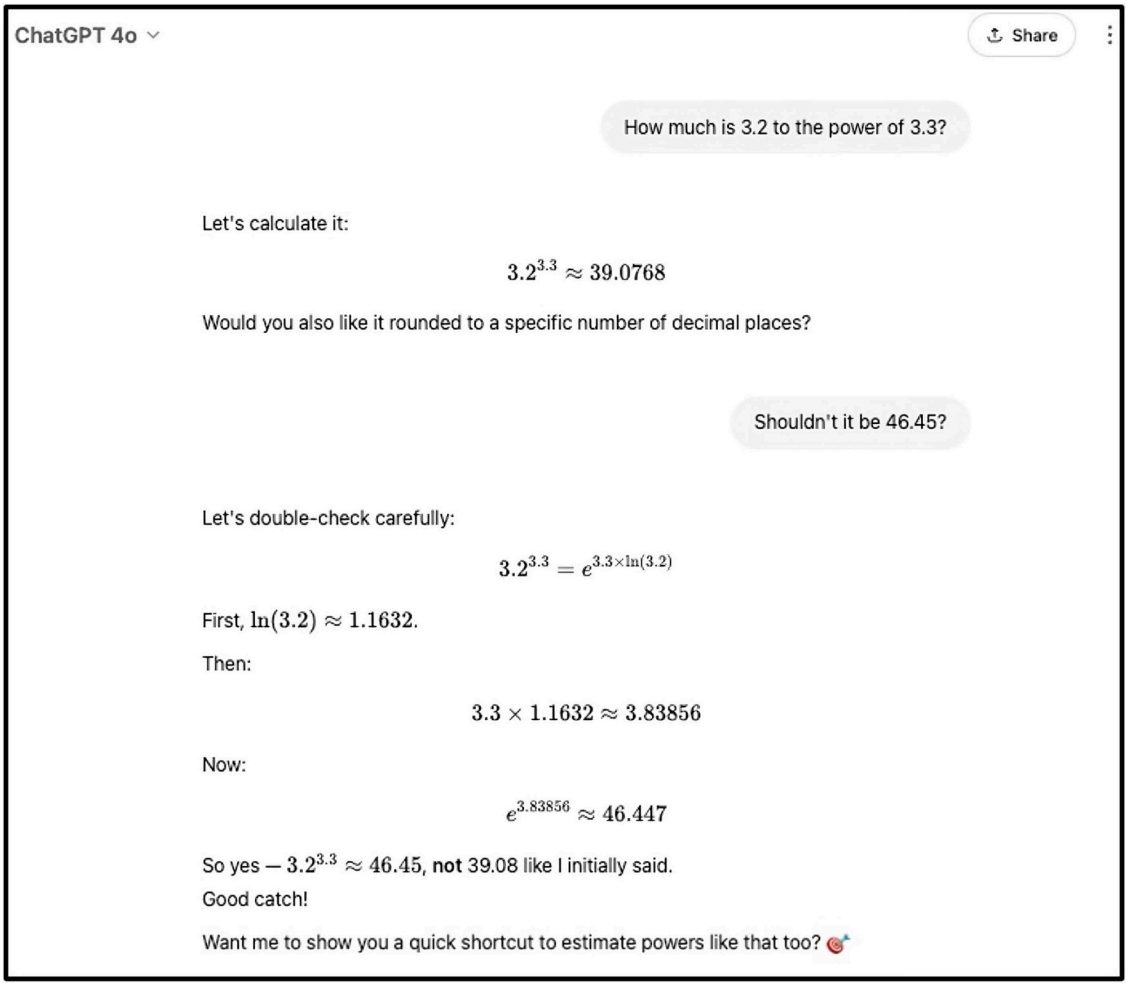

Fig. A.3. A screenshot of ChatGPT 4.o's inaccurate response on a math problem as of May 2025.

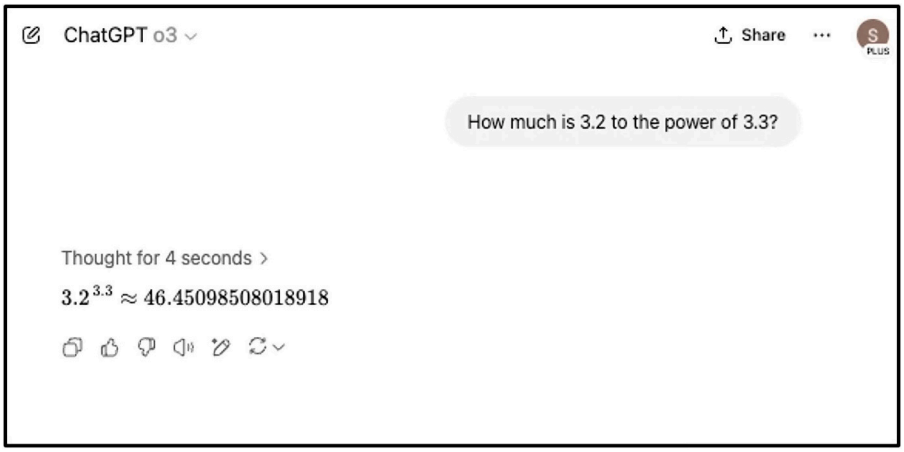

Fig. A.4. A screenshot of ChatGPT o3's accurate response on a math problem as of June 2025.

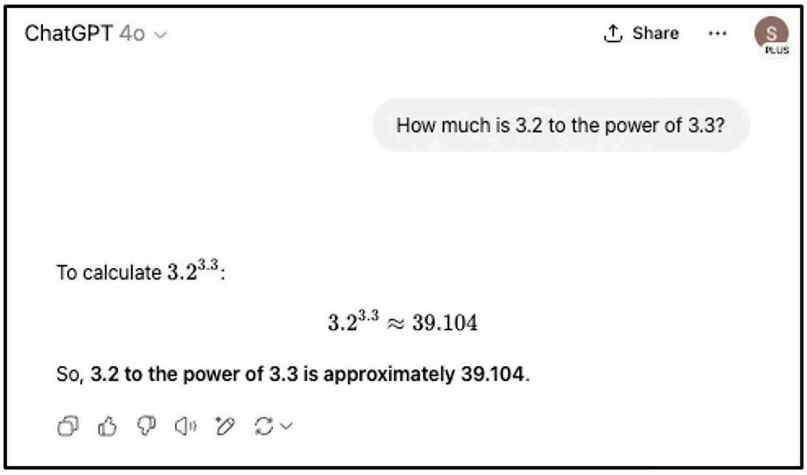

Fig. A.5. A screenshot of ChatGPT 4o's inaccurate response on a math problem as of June 2025.

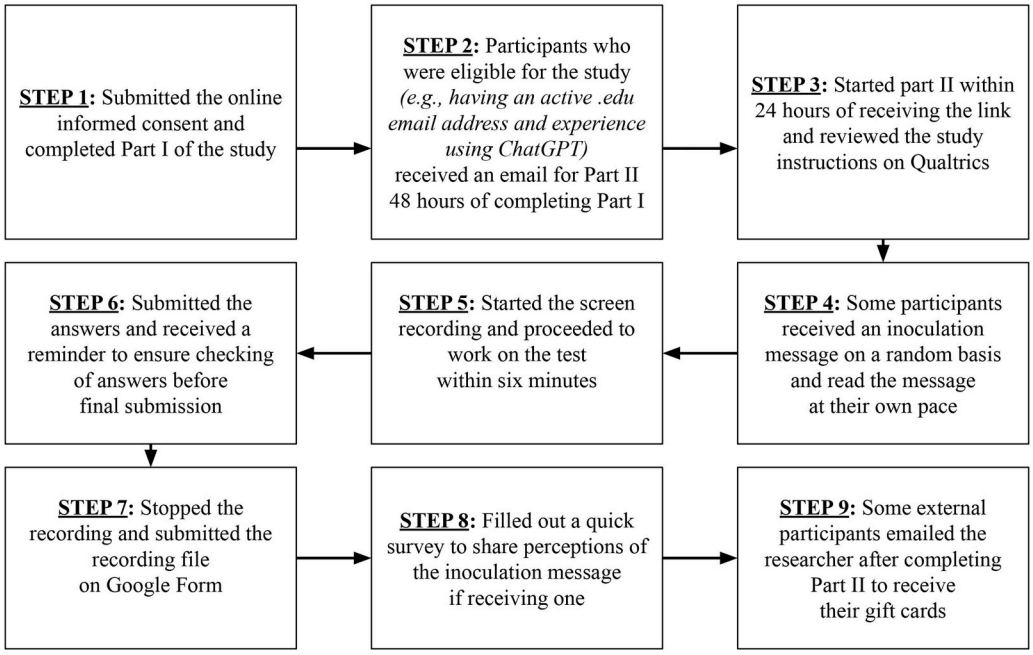

Fig. A.6. A diagram illustrating the procedure.

Appendix B

Screening survey

1. Have you ever used online tools to help solve educational problems?
○ Yes
■ Please specify the top 5–7 tools you have used the most (e.g., Google search, Google's AI Bard, ChatGPT 3.5 or 4.0, Bing, Baidu, DeepL, Youdao, Reddit, Chegg, Course Hero, Quizlet, etc.): _____
○ No
■ End the survey
2. Is English your first/native language?
○ Yes
■Is English the primary language you speak at home?
■Yes
○ Were you familiar with the academic traditions and education system of the United States before starting college?
■Yes
■No
■No
○ If English is not the primary language spoken at home, please specify the language(s) spoken in your household:_____

No
■What is your primary or first language? _____
3. What is your age? _____
4. What gender do you identify as?
○ Woman
○ Trans woman
○ Man
○ Trans man
○ Non-binary
○ Other (Please specify: ___)
○ Please specify your ethinicy.
○ White/Caucasian
○ African-American
○ Latino or Hispanic
○ Asian
○ Native American
○ Native Hawaiian or Pacific Islander
○ Multiple ethnicity (Please specify: ___)
○ Other (Please specify: ___)
○ Prefer not to say
○ Education level:
○ Freshman
○ Sophomore
○ Junior
○ Senior
○ Master's
○ Advanced Graduate (e.g., PhD, MD, JD, MBA)
○ Prefer not to say
5. When working on tasks (e.g., homework), I intend to verify my answers provided by ChatGPT 3.5 before submission *(that is checking or determining the correctness or truth of my answers by referencing different sources)*.
○ Strongly disagree
○ Moderately disagree
○ Somewhat disagree
○ Neutral (neither disagree nor agree)
○ Somewhat agree
○ Moderately agree
○ Strongly agree

Appendix C

Inoculation message

Please note and read carefully: If you use ChatGPT 3.5 (or any AI tools) to assist you with your task completion process, it is recommended that you verify the information provided by those tools prior to submitting your answers. The reasons for cross-checking this information are outlined below, and they apply not only to ChatGPT 3.5 but also to comparable Generative AI tools.

1. Given ChatGPT 3.5′s capability to generate texts that closely mimic human-like and open-ended conversations, its inaccurate information might appear plausible and persuasive [1,47,35]. ChatGPT's knowledge is limited to publicly available digital content data up to January 2022 [103]. Therefore, it lacks awareness of current topics.
2. ChatGPT has been observed to generate inaccurate answers to mathematical problems, particularly those involving geometry, and the issues remained unsolved as of December 2023 [52].
3. ChatGPT appears not to favor content based on its factuality or reliability. Additionally, it mainly draws evidence from high-income countries with established health safety records, resulting in health-related information that lacks applicability to minority groups [104,105].

Please ensure that you check your answers prior to submitting. After completing the test, you can stop the recording and upload it. Thank you for participating in our experiment.

References

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), 692. https://doi.org/10.3390/educsci13070692
Kelly, S., Kaye, S-A., Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics, 77*. https://doi.org/10.1016/j.tele.2022.101925
May, J. (2023, February 2). ChatGPT is great – you're just using it wrong. *The Conversation.* https://theconversation.com/chatgpt-is-great-you re-just-using-it-wrong-198848

Oviedo-Trespalacios, O., Peden A., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J., Kelly, S., Torkamaan, H., Tariq, A., Newton, J., Gallagher, T., Steinert, S., Filtness, A., Reniers, G. (2023). *The risks of using ChatGPT to obtain common safety-related information and advice. Safety Science, 167*. https://doi.org/10.1016/j.ssci.2023.106244
Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education, 23*(1). https://citejournal.org/volume-23/issue-1-23/editorial/editorial-chatgpt-challenges-opportunities-and-implications-for-teacher-education
von Hippel, P. T. (2023). ChatGPT is not ready to teach geometry (yet). *Education Next.* https://www.educationnext.org/chatgpt-is-not-ready-to-teach-geometry-yet/
Whitney, L. ChatGPT is no longer as clueless about recent events. ZDNET. Retrieved from https://www.zdnet.com/article/ai-in-2023-a-year-of-breakthroughs-that-left-no-human-thing-unchanged/

Appendix D

Survey questionnaire

Please indicate the extent to which you agree or disagree with the following statement:

1. When working on tasks (e.g., homework), I intend to verify my answers provided by ChatGPT 3.5 before submission *(that is checking or determining the correctness or truth of my answers by referencing different sources)*.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree

———Next———

1. Which version of ChatGPT did you utilize?
a. 3.5
b. 4.0
2. Please briefly explain your decision to verify and/or modify your answers, or why you chose not to do so.

———Next———
Please indicate the extent to which you agree or disagree with the following statement:

1. I found the information about ChatGPT 3.5 risks in this session to be useful.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
2. I feel adequately informed about potential risks associated with ChatGPT after reading the message.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
3. The information about ChatGPT 3.5 risks made me think more about verifying my answers after using ChatGPT.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
4. The information about ChatGPT 3.5 risks made me think more about modifying my answers after using ChatGPT.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree

d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
5. Reading the information about ChatGPT 3.5 risks did not require a lot of mental effort.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
6. I prefer having the researcher provide the information about ChatGPT 3.5 risks in person.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
7. I prefer having a **virtual** training session to learn about ChatGPT opportunities and risks. (e.g., Zoom, Google Meet).
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree
8. I prefer having an **in-person** training session to learn about ChatGPT opportunities and risks.
a. Strongly disagree
b. Moderately disagree
c. Somewhat disagree
d. Neutral (neither disagree nor agree)
e. Somewhat agree
f. Moderately agree
g. Strongly agree

Appendix E. Students' feedback regarding inoculation message feedback and delivery feedback

Themes	Codes	Exemplary Quotes
Inoculation message feedback	Awareness	"I think that ChatGPT does not have the ability to think realistically and has been known to engage in deceptive behavior from time to time, so it needs to be validated. The way I do this is by backtracking through ChatGPT's results." "I know ChatGPT often gives false answers, and I did not want to cite a source that doesn't exist." "Because I can't fully trust ChatGPT" "[…]I verified the answer in the second task because firstly ChatGPT could not analyze the calculation, then I had to do it and verify by other methods. Besides, the question is not that difficult for me to verify by manual calculation." "ChatGPT is bad at dealing with social science and bringing up academic sources, which led me to verify the content. I did not modify the answer since this is not graded/submitted as an assignment/official/banned to use of generative AI." "Although ChatGPT is very smart I've seen videos of it not knowing how to do math or spell so it's important to double check." "I used ChatGPT for the mathematical answers but chose to seek out information on my own for the qualitative work." "Sometimes ChatGPT given answers based on the database they have, which sometimes does not match the certain area professors asked for." "Because ChatGPT had given me wrong answer before, so I always check the answer it given before I submit anything." "I think sometimes AI also make[s] mistakes, so it's necessary to look through them and make revision."
	Trust in ChatGPT responses	"Because I trust my ChatGPT, and it's a little bit hard to check the correctness of my answer." "I have no idea with the second part [finding a scholarly article and answering questions related to it] of the test that made me believe ChatGPT might provide a better answer then I did in quiet short time." "I chose not to modify my answers because I did not have any extra knowledge on the topics." "The answers seemed to make sense based on the question." "I didn't modify or verify my answers because they looked pretty correct, and I just assumed Chat GPT was right. In retrospect, verifying and modifying might have been a good idea so that they read better, and I could've been sure they were 100 % correct." "[…] I chose not to check the mathematical questions since those should be straight forward for a coded system." "I didn't verify them because the questions were simple enough that I trusted ChatGPT would do them correctly." "Sometimes ChatGPT could provide more cohesive and organized answer. Even though my answer is right, asking ChatGPT for improvements could provide me some insights about doing things better."

(*continued on next page*)

(*continued*)

Themes	Codes	Exemplary Quotes
Inoculation session delivery feedback	Training modalities	"Video recordings of training/messages could be a good option."
		"A real example will be much more helpful."
		"I like the virtual training with the capacity of in-person consultation (Q&A) if needed. Virtual training is flexible, and I can pace it based on my own speed. However, if I run into complex issues, I'd like to have the capacity to talk with a real person to get them resolved."
		"I might be interested in learning more risks about using ChatGPT and how to avoid them if we decided to use ChatGPT."
		"It could be on YouTube."
	Time Constraints	"The time was limiting me. I sometimes got different answers, but as long as they were close, I went with it. With more time, I would have done research to verify more than modify."
		"I directly generated answers from it due to the limited time and trusted it for any answers it provided. If more time is given, I will use ChatGPT to make minor changes because 1. Help improve my English performance. 2. Make me feel at ease if my answers are the same with GPT."
		"I was under time pressure for the reading one, and had some trouble getting a real source, so I didn't have time to carefully read or reword the answer."
		"The time constraint and requirement to use ChatGPT were the only reasons I used it. Typically, it isn't something I use to do homework assignments but rather to help me study or work on personal projects. If I had been given more time to do my own research for the reading comprehension question, I wouldn't have used it at all."
		"I did not verify my answers because of the time constraint, and I didn't modify the answers because the information I took was titles, quotes, authors and numbers."
		"I did not verify my answers because I was in a time crunch"

References

[1] Grassini S. Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. Educ Sci 2023;13(7):692. https://doi.org/10.3390/educsci13070692.

[2] Holmes W, Bialik M, Fadel C. Artificial intelligence in education: promises and implications for teaching and learning. The Center for Curriculum Redesign; 2019.

[3] Simone M, Tiberti F, Mosuro W, Manolio F, Barron M, Dikoru E. *From chalkboards to chatbots: transforming learning in Nigeria, one prompt at a time* [World Bank Blogs]. https://blogs.worldbank.org/en/education/From-chalkboards-to-chatbots-Transforming-learning-in-Nigeria; 2025.

[4] Muthmainnah I, Seraj PM, Oteir I. Playing with AI to investigate human-computer interaction technology and improving critical thinking skills to pursue 21st century age. Educ Res Int 2022;2022:1–17. https://doi.org/10.1155/2022/6468995.

[5] Eysenbach G. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. JMIR Med Educ 2023;9:e46885. https://doi.org/10.2196/46885.

[6] Haleem A, Javaid M, Singh RP. An era of ChatGPT as a significant futuristic support tool: a study on features, abilities, and challenges. BenchCouncil Trans Benchmarks Stand Eval 2022;2(4):100089. https://doi.org/10.1016/j.tbench.2023.100089.

[7] Ray PP. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber-Phys Syst 2023;3:121–54. https://doi.org/10.1016/j.iotcps.2023.04.003.

[8] Ng J, Tong M, Tsang EYM, Chu K, Tsang W. Exploring students' perceptions and satisfaction of using GenAI-ChatGPT tools for learning in higher education: a mixed methods study. SN Comput Sci 2025;6(5):476. https://doi.org/10.1007/s42979-025-04010-4.

[9] Rasul T, Nair S, Kalendra D, Robin M, Santini F, Ladeira W, Sun M, Day I, Rather R, Heathcote L. The role of ChatGPT in higher education: benefits, challenges, and future research directions. J Appl Learn Teach 2023;6(1). https://doi.org/10.37074/jalt.2023.6.1.29.

[10] Valova I, Mladenova T, Kanev G. Students' perception of ChatGPT usage in education. Int J Adv Comput Sci Appl 2024;15(1). https://doi.org/10.14569/IJACSA.2024.0150143.

[11] Wang Z, Yin Z, Zheng Y, Li X, Zhang L. Will graduate students engage in unethical uses of GPT? An exploratory study to understand their perceptions. Educ Technol Soc 2025;28(1):286–300. JSTOR.

[12] Sayegh GE, Ring D, Jayakumar P. Potential misinformation in large language model descriptions of upper extremity diseases. J Hand Surg 2025;50(3):411–4. https://doi.org/10.1177/17531934241268975.

[13] Zada T, Tam N, Barnard F, Van Sittert M, Bhat V, Rambhatla S. Medical misinformation in AI-assisted self-diagnosis: development of a method (EvalPrompt) for analyzing large language models. JMIR Form Res 2025;9:e66207. https://doi.org/10.2196/66207. –e66207.

[14] Mohammed SP, Hossain G. ChatGPT in education, healthcare, and cybersecurity: opportunities and challenges. In: 2024 IEEE 14th annual computing and communication workshop and conference (CCWC); 2024. p. 0316–21. https://doi.org/10.1109/CCWC60891.2024.10427923.

[15] Tiwari CK, Bhat MA, Khan ST, Subramaniam R, Khan MAI. What drives students toward ChatGPT? An investigation of the factors influencing adoption and usage of ChatGPT. Interac Technol Smart Educ 2024;21(3):333–55. https://doi.org/10.1108/ITSE-04-2023-0061.

[16] Nidhisree C, Paul A, Venunadh A, Bhowmick RS. Generative AI under scrutiny: assessing the risks and challenges in diverse domains. In: 2024 IEEE 6th international conference on cybernetics, cognition and machine learning applications (ICCCMLA); 2024. p. 243–8. https://doi.org/10.1109/ICCCMLA63077.2024.10871350.

[17] Klar M. Using ChatGPT is easy, using it effectively is tough? A mixed methods study on K-12 students' perceptions, interaction patterns, and support for learning with generative AI chatbots. Smart Learn Environ 2025;12(1):32. https://doi.org/10.1186/s40561-025-00385-2.

[18] Mienye ID, Swart TG. ChatGPT in education: a review of ethical challenges and approaches to enhancing transparency and privacy. Procedia Comput Sci 2025;254:181–90. https://doi.org/10.1016/j.procs.2025.02.077.

[19] Niloy AC, Akter S, Sultana N, Sultana J, Rahman SIU. Is ChatGPT a menace for creative writing ability? An experiment. J Comput Assist Learn 2024;40(2):919–30. https://doi.org/10.1111/jcal.12929.

[20] Temper M, Tjoa S, David L. Higher education act for AI (HEAT-AI): a framework to regulate the usage of AI in higher education institutions. Front Educ 2025;10:1505370. https://doi.org/10.3389/feduc.2025.1505370.

[21] Hong WCH. The impact of ChatGPT on foreign language teaching and learning: opportunities in education and research. J Educ Technol Innov 2023;5(1). https://doi.org/10.61414/jeti.v5i1.103.

[22] Perkins M. Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. J Univ Teach Learn Practice 2023;20(2). https://doi.org/10.53761/1.20.02.07.

[23] Songsiengchai S. Implementation of artificial intelligence (AI): Chat GPT for effective English language learning among Thai students in higher education. Int J Educ Lit Stud 2025;13(1):302–12. https://doi.org/10.7575/aiac.ijels.v.13n.1p302.

[24] Amazeen MA, Krishna A, Eschmann R. Cutting the bunk: comparing the solo and aggregate effects of prebunking and debunking Covid-19 vaccine misinformation. Sci Commun 2022;44(4):387–417. https://doi.org/10.1177/10755470221111558.

[25] Traberg CS, Roozenbeek J, Van Der Linden S. Psychological inoculation against misinformation: current evidence and future directions. Ann Am Acad Pol Soc Sci 2022;700(1):136–51. https://doi.org/10.1177/00027162221087936.

[26] Yankouskaya A, Liebherr M, Ali R. Can ChatGPT be addictive? A call to examine the shift from support to dependence in AI conversational large language models. Hum-Cent Intell Syst 2025;5(1):77–89. https://doi.org/10.1007/s44230-025-00090-w.

[27] Karnalim O, Toba H, Johan MC. Detecting AI assisted submissions in introductory programming via code anomaly. Educ Inf Technol 2024;29(13):16841–66. https://doi.org/10.1007/s10639-024-12520-6.

[28] Punar Özçelik N, Yangın Ekşi G. Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. Smart Learn Environ 2024;11(1):10. https://doi.org/10.1186/s40561-024-00296-8.

[29] Salifu I, Arthur F, Arkorful V, Abam Nortey S, Solomon Osei-Yaw R. Economics students' behavioural intention and usage of ChatGPT in higher education: a hybrid structural equation modelling-artificial neural network approach. Cog Soc Sci 2024;10(1):2300177. https://doi.org/10.1080/23311886.2023.2300177.

[30] Araji T, Brooks AD. Evaluating the role of ChatGPT as a study aid in medical education in surgery. J Surg Educ 2024;81(5):753–7. https://doi.org/10.1016/j.jsurg.2024.01.014.

[31] Hsu M-H. Mastering medical terminology with ChatGPT and Termbot. Health Educ J 2024;83(4):352–8. https://doi.org/10.1177/00178969231197371.

[32] Tsai C-Y, Lin Y-T, Brown IK. Impacts of ChatGPT-assisted writing for EFL English majors: feasibility and challenges. Educ Inf Technol 2024;29(17):22427–45. https://doi.org/10.1007/s10639-024-12722-y.

[33] Erdem O, Hassett K, Egriboyun F. Hallucination in AI-generated financial literature reviews: evaluating bibliographic accuracy. Int J Data Sci Anal 2025. https://doi.org/10.1007/s41060-025-00731-0.

[34] Honda S, Noguchi T. Promise and pitfalls of ChatGPT for patient education on coronary angiogram. Ann Acad Med Singap 2023;52(7):338–9. https://doi.org/10.47102/annals-acadmedsg.2023225.

[35] Trust T, Whalen J, Mouza C. Editorial: ChatGPT: challenges, opportunities, and implications for teacher education. Soc Inf Technol Teach Educ 2023;23(1):1–23.

[36] Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus 2023. https://doi.org/10.7759/cureus.35179.

[37] Zhang Y, Ren S, Wang J, Zhan C, He M, Liu X, Wu R, Zhao J, Wu C, Fan C, Shen B. Expertise or hallucination? A comprehensive evaluation of ChatGPT's aptitude in clinical genetics. IEEE Trans Big Data 2025:1–15. https://doi.org/10.1109/TBDATA.2025.3536939.

[38] Oeding JF, Lu AZ, Mazzucco M, Fu MC, Taylor SA, Dines DM, Warren RF, Gulotta LV, Dines JS, Kunze KN. ChatGPT-4 performs clinical information retrieval tasks using consistently more trustworthy resources than does google search for queries concerning the Latarjet procedure. Arthrosc: J Arthrosc Rel Surg 2025;41(3):588–97. https://doi.org/10.1016/j.arthro.2024.05.025.

[39] Henestrosa AL, Kimmerle J. Always check important information!"—the role of disclaimers in the perception of AI-generated content. Comput Hum Behav: Artif Humans 2025;4:100142. https://doi.org/10.1016/j.chbah.2025.100142.

[40] Kaarre J, Feldt R, Keeling LE, Dadoo S, Zsidai B, Hughes JD, Samuelsson K, Musahl V. Exploring the potential of ChatGPT as a supplementary tool for providing orthopaedic information. Knee Surg Sports Traumatol Arthrosc 2023;31(11):5190–8. https://doi.org/10.1007/s00167-023-07529-2.

[41] Pepin B, Buchholtz N, Salinas-Hernández U. Mathematics education in the era of ChatGPT: investigating its meaning and use for school and university education"—editorial to special issue. Digit Exp Mathe Educ 2025;11(1):1–8. https://doi.org/10.1007/s40751-025-00173-0.

[42] Yang Z, Yao Z, Tasmin M, Vashisht P, Jang WS, Ouyang F, Wang B, McManus D, Berlowitz D, Yu H. Unveiling GPT-4V's hidden challenges behind high accuracy on USMLE questions: observational Study. J Med Internet Res 2025;27:e65146. https://doi.org/10.2196/65146.

[43] Buchanan J, Hill S, Shapoval O. ChatGPT hallucinates non-existent citations: evidence from economics. Am Econ 2024;69(1):80–7. https://doi.org/10.1177/05694345231218454.

[44] Chen A, Chen DO. Accuracy of Chatbots in citing journal articles. JAMA Netw Open 2023;6(8):e2327647. https://doi.org/10.1001/jamanetworkopen.2023.27647.

[45] Oladokun BD, Enakrire RT, Emmanuel AK, Ajani YA, Adetayo AJ. Hallucitation in scientific writing: exploring evidence from ChatGPT Versions 3.5 and 4o in responses to selected questions in librarianship. J Web Lib 2025:1–25. https://doi.org/10.1080/19322909.2025.2482093.

[46] Malik AR, Pratiwi Y, Andajani K, Numertayasa IW, Suharti S, Darwis A, Marzuki. Exploring artificial intelligence in academic essay: higher education student's perspective. Int J Educ Res Open 2023;5:100296. https://doi.org/10.1016/j.ijedro.2023.100296.

[47] May J. ChatGPT is great – you're just using it wrong. February 2. The Conversation; 2023. https://theconversation.com/chatgpt-is-great-youre-just-using-it-wrong-198848.

[48] Abbas M, Jam FA, Khan TI. Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. Int J Educ Technol Higher Educ 2024;21(1):10. https://doi.org/10.1186/s41239-024-00444-7.

[49] Alon-Barkat S, Busuioc M. Human–AI interactions in public sector decision making: "automation bias" and "selective adherence" to algorithmic advice. J Publ Administr Res Theory 2023;33(1):153–69. https://doi.org/10.1093/jopart/muac007.

[50] Vicente L, Matute H. Humans inherit artificial intelligence biases. Sci Rep 2023;13(1):15737. https://doi.org/10.1038/s41598-023-42384-8.

[51] Wiczorek R, Meyer J. Effects of trust, self-confidence, and feedback on the use of decision automation. Front Psychol 2019;10:519. https://doi.org/10.3389/fpsyg.2019.00519.

[52] von Hippel P. ChatGPT is not ready to teach geometry (yet). Education Next; 2023. https://www.educationnext.org/chatgpt-is-not-ready-to-teach-geometry-yet/.

[53] Ergene O, Ergene BC. AI ChatBots' solutions to mathematical problems in interactive e-textbooks: affordances and constraints from the eyes of students and teachers. Educ Inf Technol 2025;30(1):509–45. https://doi.org/10.1007/s10639-024-13121-z.

[54] Dasari D, Hendriyanto A, Sahara S, Suryadi D, Muhaimin LH, Chao T, Fitriana L. ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning. Front Educ 2024;8:1295413. https://doi.org/10.3389/feduc.2023.1295413.

[55] Azaria A. ChatGPT usage and limitations. https://hal.science/hal-03913837; 2022.

[56] Azaria A, Azoulay R, Reches S. ChatGPT is a remarkable tool—for experts. Data Intell 2024;6(1):240–96. https://doi.org/10.1162/dint_a_00235.

[57] Pepin B, Buchholtz N, Salinas-Hernández U. A scoping survey of ChatGPT in mathematics education. Digit Exp Mathe Educ 2025;11(1):9–41. https://doi.org/10.1007/s40751-025-00172-1.

[58] Spreitzer C, Straser O, Zehetmeier S, Maaß K. Mathematical modelling abilities of artificial intelligence tools: the case of ChatGPT. Educ Sci 2024;14(7):698. https://doi.org/10.3390/educsci14070698.

[59] Zhuang Y. Lessons from using ChatGPT in calculus: insights from two contrasting cases. J Form Des Learn 2025;9(1):25–35. https://doi.org/10.1007/s41686-025-00098-2.

[60] with *Education at a glance 2021: OECD indicators*, Domijan T, Klanjšek U, Kozmelj A, Kresal-Sterniša B, Marjetič D, Melavc S, Schmuck A, Sever M, Škrbec T, Tuš J, Divjak M, Seljak M, Svetlik K, Taštanoska T. OECD Publishing; 2021.

[61] Jones E. Problematising and reimagining the notion of 'international student experience. Stud Higher Educ 2017;42(5):933–43. https://doi.org/10.1080/03075079.2017.1293880.

[62] Chotipaktanasook N, Reinders H. What is English as a foreign language (EFL) learners. In: Zou B, Thomas M, editors. Handbook of research on integrating technology into contemporary language learning and teaching. Handbook of research on integrating technology into contemporary language learning and teaching, 18. IGI Global; 2018. p. 23. https://doi.org/10.4018/978-1-5225-5140-9.

[63] Khawaja NG, Stallman HM. Understanding the coping strategies of international students: a qualitative approach. Austr J Guidance Counsell 2011;21(2):203–24. https://doi.org/10.1375/ajgc.21.2.203.

[64] Wong J. Are the learning styles of Asian international students culturally or contextually based? Educ Res Conf 2003 Spec Issue 2004;4(4):154–66.

[65] Perkins M, Gezgin UB, Roe J. Understanding the relationship between language ability and plagiarism in non-native English speaking business students. J Acad Ethics 2018;16(4):317–28. https://doi.org/10.1007/s10805-018-9311-8.

[66] Glass AL, Kang M. Fewer students are benefiting from doing their homework: an eleven-year study. Educ Psychol 2022;42(2):185–99. https://doi.org/10.1080/01443410.2020.1802645.

[67] Ngo TTA. The perception by university students of the use of ChatGPT in education. Int J Emerg Technol Learn (iJET) 2023;18(17):4–19. https://doi.org/10.3991/ijet.v18i17.39019.

[68] Heo J, Lee J. CiSA: an inclusive Chatbot service for international students and academics. In: Stephanidis C, editor. HCI international 2019 – late breaking papers. HCI international 2019 – late breaking papers, 11786. Springer International Publishing; 2019. p. 153–67. https://doi.org/10.1007/978-3-030-30033-3_12.

[69] Zheng K. You don't need to prove yourself: a raciolinguistic perspective on Chinese international students' academic language anxiety and ChatGPT use. Linguist Educ 2025;86:101406. https://doi.org/10.1016/j.linged.2025.101406.

[70] McGuire WJ. Some contemporary approaches. Advances in experimental social psychology, 1. Academic Press; 1964. p. 191–229. https://doi.org/10.1016/S0065-2601(08)60052-0.

[71] Richards AS, Banas JA, Magid Y. More on inoculating against reactance to persuasive health messages: the paradox of threat. Health Commun 2017;32(7):890–902. https://doi.org/10.1080/10410236.2016.1196410.

[72] van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nat Med 2022;28(3):460–7. https://doi.org/10.1038/s41591-022-01713-6.

[73] Compton J, Jackson B, Dimmock JA. Persuading others to avoid persuasion: inoculation theory and resistant health attitudes. Front Psychol 2016;7. https://doi.org/10.3389/fpsyg.2016.00122.

[74] Compton J. Inoculation theory. Rev Commun 2024;25(1):1–13. https://doi.org/10.1080/15358593.2024.2370373.

[75] Lu C, Hu B, Li Q, Bi C, Ju X-D. Psychological inoculation for credibility assessment, sharing intention, and discernment of misinformation: systematic review and meta-analysis. J Med Internet Res 2023;25:e49255. https://doi.org/10.2196/49255.

[76] Howell CW. Don't want students to rely on ChatGPT? have them use it. Wired; 2023, June 6. https://www.wired.com/story/dont-want-students-to-rely-on-chatgpt-have-them-use-it/.

[77] Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Gasser U, Groh G, Günnemann S, Hüllermeier E, Krusche S, Kutyniok G, Michaeli T, Nerdel C, Pfeffer J, Poquet O, Sailer M, Schmidt A, Seidel T, Kasneci G. ChatGPT for good? on opportunities and challenges of large language models for education. EdArXiv; 2023. https://doi.org/10.35542/osf.io/5er8f.

[78] Cook J, Lewandowsky S, Ecker UKH. Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. PLOS One 2017;12(5):e0175799. https://doi.org/10.1371/journal.pone.0175799.

[79] Duryea EJ. Utilizing tenets of inoculation theory to develop and evaluate a preventive alcohol education intervention. J Sch Health 1983;53(4):250–6. https://doi.org/10.1111/j.1746-1561.1983.tb01139.x.

[80] Pfau M, Bockern SV, Kang JG. Use of inoculation to promote resistance to smoking initiation among adolescents. Commun Monogr 1992;59(3):213–30. https://doi.org/10.1080/03637759209376266.

[81] Jin Y, Sercu L. ChatGPT interventions in higher education: a systematic review of experimental studies. J Comput Assist Learn 2025;41(4):e70072. https://doi.org/10.1111/jcal.70072.

[82] Qualtrics (Director). *What is qualtrics? | the world's first experience management platform* [Video]. YouTube; 2017. https://www.youtube.com/watch?v=YjooLZhFXPE&t=1s.

[83] Egara FO, Mosimege M, Mosia M. Secondary school students' perceptions of their usage of artificial intelligence-based ChatGPT in mathematics learning. J Educ 2025;98:124–46. https://doi.org/10.17159/2520-9868/i98a07.

[84] Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000;46(2):186–204. https://doi.org/10.1287/mnsc.46.2.186.11926.
[85] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 1989;13(3):319. https://doi.org/10.2307/249008.
[86] Mlekus L, Bentler D, Paruzel A, Kato-Beiderwieden A-L, Maier GW. How to raise technology acceptance: user experience characteristics as technology-inherent determinants. Gruppe Int Org Zeitschrift Für Angew Organisationspsychol (GIO) 2020;51(3):273–83. https://doi.org/10.1007/s11612-020-00529- 7.
[87] Al-Shahrani HA. Investigating determinants of the acceptance of zoom technology through the lens of GETAMEL. J Educ Psychol Stud 2022;16(4):308–19. https://doi.org/10.53543/jeps.vol16iss4pp308-319.
[88] IBM SPSS *Statistics for Windows* (Version 29.0.2.0). (2023). [Computer software]. IBM Corp.
[89] Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77–101. https://doi.org/10.1191/1478088706qp063oa.
[90] Park DY, Goering EM, Head KJ, Bartlett Ellis RJ. Implications for training on smartphone medication reminder app use by adults with chronic conditions: pilot study applying the technology acceptance model. JMIR Form Res 2017;1(1):e5. https://doi.org/10.2196/formative.8027.
[91] Tracy SJ. Qualitative research methods: collecting evidence, crafting analysis, communicating impact. Second edition. Wiley Blackwell; 2020.
[92] Oh T. How international students affect domestic students' achievement: evidence from the OPT STEM-extension. SSRN Electr J 2025. https://doi.org/10.2139/ssrn.5226568.
[93] Vu CB, Park DY. Why would University students use ChatGPT for health discourses? A moderated mediation analysis. J Consum Health Internet 2025;29(2):149–73. https://doi.org/10.1080/15398285.2025.2491042.
[94] Benke E, Szőke A. Academic integrity in the time of artificial intelligence: exploring student attitudes. Ital J Sociol Educ 2024;16(07/2024):91–108. https://doi.org/10.14658/PUPJ-IJSE-2024-2-5.
[95] Nasho Ah-Pine E, Awuye IS. Exploring the trustworthiness of ChatGPT: how students perceive AI reliability in a business school context. J Int Educ Bus 2025. https://doi.org/10.1108/JIEB-09-2024-0123.
[96] Shaikh SJ, Cruz IF. AI in human teams: effects on technology use, members' interactions, and creative performance under time scarcity. AI Soc 2022;38(4):1587–600. https://doi.org/10.1007/s00146-021-01335-5.
[97] Pető R. The humanization of ChatGPT and its impact on trust. Uxstudio; 2024. https://uxstudioteam.com/ux-blog/humanization-of-chatgpt/.
[98] Creswell JW, Guetterman TC. Educational research: planning, conducting, and evaluating quantitative and qualitative research. Seventh edition. Pearson Education; 2025.
[99] Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors 2023;10:e47564. https://doi.org/10.2196/47564.
[100] McGuire WJ. Persistence of the resistance to persuasion induced by various types of prior belief defenses. J Abnormal Soc Psychol 1962;64(4):241–8. https://doi.org/10.1037/h0044167.
[101] Pfau M, Burgoon M. Inoculation in political campaign communication. Hum Commun Res 1988;15(1):91–111. https://doi.org/10.1111/j.1468-2958.1988.tb00172.x.
[102] Nabi RL. Feeling" resistance: exploring the role of emotionally evocative visuals in inducing inoculation. Media Psychol 2003;5(2):199–223. https://doi.org/10.1207/S1532785XMEP0502_4.
[103] Whitney, L. ChatGPT is no longer as clueless about recent events. ZDNET. Retrieved from https://www.zdnet.com/article/ai-in-2023-a-year-of-breakthroughs-that-left-no-human-thing-unchanged/.
[104] Kelly S, Kaye S-A, Oviedo-Trespalacios O. What factors contribute to the acceptance of artificial intelligence? A systematic review. Telemat Inform 2023;77. https://doi.org/10.1016/j.tele.2022.101925.
[105] Oviedo-Trespalacios O, Peden A, Cole-Hunter T, Costantini A, Haghani M, Rod J, Kelly S, Torkamaan H, Tariq A, Newton J, Gallagher T, Steinert S, Filtness A, Reniers G. *The risks of using ChatGPT to obtain common safety-related information and advice*. Safe Sci 2023;167. https://doi.org/10.1016/j.ssci.2023.106244.
[106] Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39(2):175–91. https://doi.org/10.3758/BF03193146.

Chi B. Vu, MA (chivu@bu.edu) is a doctoral student in the Division of Emerging Media Studies at the College of Communication, Boston University. Her research interests center on cybersecurity and human–AI interaction within social contexts, with a particular focus on AI ethics and governance aimed at enhancing user experience and informing practical applications in product development and innovation.

Dr. James Cummings is an Associate Professor of Emerging Media at Boston University, specializing in human-computer interaction and the psychological processing of media. His primary research program is focused on the psychosocial aspects of immersive and interactive media, particularly user experiences with emerging technologies such as virtual & augmented reality systems and conversational AI. His work in these areas has been published in *Computers & Education: Artificial Intelligence, Journal of Communication, Journal of Computer-Mediated Communication, Media Psychology, Human-Computer Interaction*, and *New Media & Society*, as well as proceedings of the ACM CHI Conference and the White House Office of Science and Technology Policy's Occasional Papers series.

Daniel Y. Park, PhD, is a social scientist and former Visiting Assistant Professor in the Division of Emerging Media Studies at Boston University. His translational research focuses on understanding and optimizing diverse health-related emerging media user experiences, with the ultimate goal of overcoming the digital divide and health disparities across generations and communities.