

# Testing ChatGPT in International Relations Classrooms: Potentialities, Limitations, and What's Next

Paola Rivetti, *Dublin City University, Ireland*

Rituparna Banerjee, *Dublin City University, Ireland*

David O'Mullane, *Dublin City University, Ireland*

## ABSTRACT

Whereas significant attention has been devoted to online/blended teaching and related tools, open GenAI chatbots and large language models and writing programs have received comparatively less attention as instruments that impact our teaching and assessment methods. The pedagogies of political science and international relations somehow trail behind in understanding and addressing them. For those who are teaching these subjects, it is of great importance to come to terms with AI and its impact on how we should assess students. This article describes an eight-week laboratory during which we experimented with and discussed ChatGPT's utilizations with our students. The goal of the laboratory was to revise which type of learning objectives our teaching should have and which type of assessment methods are best suited for an environment in which GenAI is present and used regularly. Should we ban it or should we instead focus on teaching students how to use it responsibly and ethically, as well as developing a shared understanding of what GenAI is good for (if anything at all)? Consequently, which learning objectives and assessment methods should we adopt?


**T**his article explores the potential and the limitations of using ChatGPT (Generative Pre-trained Transformer) in international relations (IR) classrooms by drawing on an eight-week laboratory experiment conducted between October and December of 2023.

The experiment involved first-year students of BA programs in social and political sciences at Dublin City University (DCU) in Ireland. The three authors co-led the laboratory (Author 1 is an associate professor and coordinator of the module, "Introduction to International Relations and Security" at DCU. Authors 2 and


3 currently are PhD candidates at DCU). They contributed in an equal and balanced manner to the development of the overall project and the content for each laboratory session. Supported by DCU and the Irish Higher Education Authority, the laboratory expanded on the observation that whereas many political scientists and IR scholars have embraced ChatGPT and other Generative Artificial Intelligence (Gen-AI) tools for writing and publishing purposes, the same enthusiasm for testing applied uses of Gen-AI has not taken place in the teaching and learning realm.

## BACKGROUND

Within DCU and, more generally, higher-education institutions, faculty meetings in 2023 and 2024 revealed a mix of worries, including concerns about plagiarism, over the perceived blandness of AI-generated content, and even fears of Gen-AI replacing educators altogether. Teaching a diverse set of modules at both

**Corresponding author: Paola Rivetti**  is associate professor of law and government at Dublin City University. She can be reached at [paola.rivetti@dcu.ie](mailto:paola.rivetti@dcu.ie).

**Rituparna Banerjee** is a PhD candidate in communications at Dublin City University. She can be reached at [rituparna.banerjee3@mail.dcu.ie](mailto:rituparna.banerjee3@mail.dcu.ie).

**David O'Mullane**  is PhD candidate in applied language and intercultural studies at Dublin City University. He can be reached at [david.omullane2@mail.dcu.ie](mailto:david.omullane2@mail.dcu.ie).

© The Author(s), 2025. Published by Cambridge University Press on behalf of American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

the BA and MA levels, Author 1 had mixed feelings about the possibility of Gen-AI replacing educators and lecturers. However, her primary concern was the detrimental impact of an uninformed and naïve use of Gen-AI tools on her students' learning process.

In particular, Author 1 observed that students' essays increasingly lacked accurate references and presented a superficial engagement with scholarly sources. Furthermore, assigned essays and other written work often included unreferenced factual claims that, even when correct, lacked proper context and resembled a chatbot-generated text. Proving plagiarism from Gen-AI tools, however, is challenging because they produce different narratives each time. To circumvent instances of plagiarism altogether, colleagues suggested avoiding essays and other forms of written assessments in favor of multiple-choice quizzes, exams, and oral examinations. Although Author 1 had used these forms of assessment previously, she remained convinced of the importance of essays for developing writing, argumentation, and critical-thinking skills during university years. Because Gen-AI chatbots seemed to hinder these skills, Author 1 decided to explore their role in teaching and learning practices. This led to the creation of a laboratory for first-year students, where educators and students could jointly explore and learn how to use ChatGPT. ChatGPT's free version (3.5) was chosen due to its popularity.

ChatGPT is a deep learning-based language model that can respond intelligently and conversationally to a plain-language prompt (Floridi and Chiriatti 2020). Launched in November 2022, the AI chatbot had 100 million users within only two months (Milmo 2023). As Halaweh (2023) noted in their paper on the responsible implementation of ChatGPT in educational settings, the accessibility of the tool combined with these staggering early-user figures, suggests that—much like calculators, laptops, and mobile phones—this form of technology will become an integral part of life. As Holmes and Tuomi (2022, 543) noted, the burgeoning need for “AI literacy” places teachers and students at the forefront of a rapidly developing technological landscape. There undoubtedly is an onus on third-level institutions to proactively incorporate these AI tools into their approaches to teaching and assessing students. Several institutions in the United Kingdom, such as The Russell Group and the Quality Assurance Agency for Higher Education, have issued guidance on how to best incorporate AI into teaching and learning, being cognizant of the technology's strengths and weaknesses (Centre for Teaching and Learning 2023; The Russell Group 2024; Quality Assurance Agency for Higher Education 2023). Concerning ChatGPT's weaknesses, OpenAI—the company responsible for developing ChatGPT—noted that one limitation of their language model was that it could create “plausible-sounding but incorrect or nonsensical answers” (OpenAI 2022). This was particularly relevant for these laboratories because students could receive responses from the chatbot riddled with non-existent book titles, fabricated direct quotes, and correct-sounding but entirely false reference information. OpenAI (2023) admitted that the implicit structure of the language model makes correcting this issue extremely challenging. As the Centre for Teaching and Learning (2023) noted in its report on AI in academic practice, competitor language models also have struggled to rectify the fabrication of information. Being cognizant of this inherent tendency toward confabulation, the authors encouraged students to be “the human in the loop” (Mollick and Mollick 2023), casting a critical eye on

ChatGPT's outputs and cross referencing its academic sources through a trusted site (e.g., Google Scholar).

### The Laboratory

The laboratory consisted of eight weekly sessions offered to first-year BA students. The students voluntarily opted in to the laboratory, which received approval by the University's Research Ethics Committee (Ref. No. DCU-FHSS-2023-039). They were enrolled in the module, “Introduction to IR and Security” (Module Code: LG1170), which is a core module for five BA programs. In 2023–2024, the class consisted of 203 students. For the laboratories, the class was divided into seven groups of about 30 students each. This was done randomly for logistical reasons, although it enhanced intraclass relations because the group consisted of first-year students who did not know one another. Because completing the laboratories was not mandatory, students' attendance inevitably decreased after the week-long midterm break in November 2023. However, an average of 10 students attended each session, resulting in approximately 70 students participating each week, excluding the final week of the semester when low attendance necessitated canceling some sessions (Rivetti, Banerjee, and O'Mullane 2024).

The laboratory sessions took place in a computer-equipped room, where each student could work on one computer or team up with two or three other students and work together in a group on one computer. In this context, the students autonomously decided to work individually or in groups. In the latter scenario, they independently selected the other group members. Each session consisted of three parts: practical exercises with ChatGPT, class discussion about the results, and student feedback. To enhance students' voices, the authors asked them to participate in an end-of-the-laboratory survey. Seventy-two students participated in the survey, which was reflective of the number of students who opted to attend all eight of the laboratory sessions.

It is important to note that the laboratory offered two key benefits for students in addition to only using ChatGPT. First, IR students on our campus typically lack access to computers in their classroom environment. This is especially the case for first-year students, many of whom also may struggle with off-campus computer access. When asked how many had tried ChatGPT in the past, few replied positively. Of the 72 students taking part in the final survey, approximately 64% had prior experience in using an AI chatbot. However, discussing students' experiences with Gen-AI language models during the introductory sessions revealed that these interactions tended to be centered on exploratory usage and basic questioning. Therefore, the laboratory bridged a digital gap for our students, giving them the skills to write prompts that yielded more-targeted and higher-quality responses relevant to IR studies.

The second benefit stemmed from the small class size during lab sessions. This allowed the authors to closely supervise students, facilitating discussions that ranged from ChatGPT to course content, essay writing, source mining, and referencing. Such in-depth exchanges were difficult during regular classes due to size limitations (i.e., more than 200 students) and budget constraints on tutorials. The laboratory therefore was an opportunity for the students to learn about ChatGPT and to access useful information about the module's content and assessment methods, which otherwise would have remained more difficult to access.

---

## PLANNING THE LABORATORY: LEARNING GOALS AND PRACTICAL EXERCISES

The three authors met several times to structure the laboratory and, in particular, the content of each session. During this preparatory phase, Maria Clara Menezes, a PhD student at the same university who had an interest in the project because of her utilization of ChatGPT in her research practice, joined them. Together, the authors and Menezes identified two key sets of primary learning objectives for the laboratory. The first was to cultivate a shared understanding of ChatGPT's functionalities, fostering a common ground for both instructors and students. Related secondary learning objectives included crafting effective prompts, interacting with the chatbot, and critically evaluating the information and sources that it provides. Additionally, the laboratory aimed to equip students with the ability to assess ChatGPT's reliability compared to other sources, such as textbooks and Google Scholar. Developing the crucial skill of sourcing reliable information online was another primary objective. This is a transferable skill and a foundational steppingstone for first-year students, especially considering the growing challenge of discerning trustworthy sources—not only for students but for educators as well.

The second set of primary learning objectives focused on the intersection of essay writing and ChatGPT. The laboratory concept emerged from concerns about students' weak essay-writing skills and the difficulty of detecting plagiarism from ChatGPT, which caused a growing belief among colleagues that written assessments were outdated (Grove 2024; Marche 2022; Stokel-Walker 2022). OpenAI discontinuing its low-accuracy AI detector in 2023 and the complications of effectively prohibiting the use of ChatGPT fueled this sentiment. Instead, the laboratory took a different approach and discussed which tasks (if any) related to essay writing that ChatGPT could complete. Another related secondary learning objective included identifying the tasks that are instrumental for developing an essay. This became an opportunity for students to discuss essay structure and learn how to use ChatGPT for those tasks. Examples of tasks that we used ChatGPT for included brainstorming for topics and arguments, structuring an essay, compiling and controlling lists of references, and finding and double-checking information about possible case studies.

During the preparatory phase, the three authors and Menezes collaboratively refined the learning objectives, ensuring that the exercises in each laboratory session served the identified objectives. To find relevant exercises and pair them with the learning objectives, the authors and Menezes conducted a literature review about the use of ChatGPT in the classroom, as well as the challenges and opportunities it poses to pedagogy in the social sciences. Particular attention was given to research that discussed practical uses of ChatGPT in the classroom (Atlas 2023; Centre for Teaching and Learning 2023; Goodlad and Stoerger 2023; Herft 2023; Mollick and Mollick 2023; Theophilou et al. 2023). Building on the literature and their experience, the four researchers structured each laboratory session with exercises that they paired to specific learning objectives. The exercises used in class are discussed in detail in the online appendix, where the learning objectives of each exercise and the questions that guided our interactions with the students pre- and post-exercise are listed. Students were encouraged to use topics covered in their IR lectures when undertaking exercises and designing prompts. For example, when prompting the chatbot to act as a tutor (see the online

appendix), a student could request that the AI gave particular attention to theoretical perspectives (e.g., realist, liberal, feminist, and post-colonial) relevant to studying IR. Whereas the content expressed by the AI chatbot was related to IR, the prompt work the students engaged in could be applied to other subjects and disciplines.

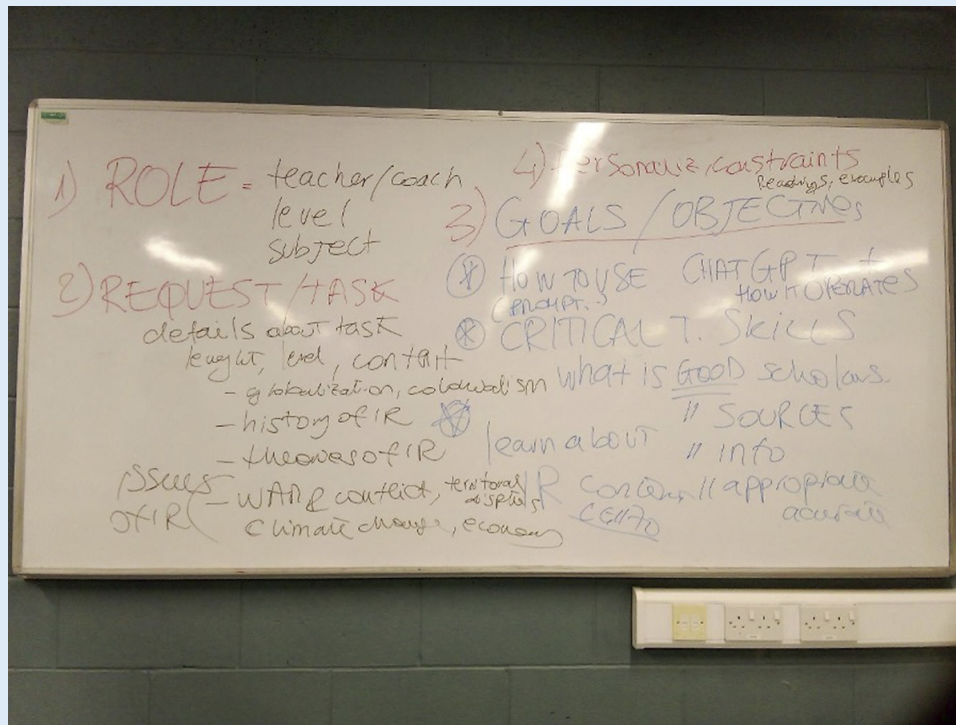
## CHATGPT IN THE CLASSROOM: POTENTIALITIES AND LIMITATIONS

Running the exercises in the classroom proved beneficial to the students as well as for the authors' learning process about how ChatGPT works and how it can be used to learn about a topic and develop an essay. However, we also appreciated its limitations. Furthermore, we encountered technical difficulties. For example, we discovered that there is a maximum of 20 messages that ChatGPT can receive and react to per hour. This meant that the students could not interact freely with the chatbot and had to incorporate several questions and messages in a single entry to avoid maxing out the messages. Therefore, the students and the researchers had to rethink their interaction strategy. Before starting the laboratory, the researchers created a series of Gmail and ChatGPT accounts for the students. The goal was to have their interactions stored and accessible for reflection and analysis post-laboratory. Although the accounts usually worked, they occasionally crashed and the researchers had to ask students to work in larger groups or create and use backup accounts. As is well known, ChatGPT is trained on data from before September 2021. Depending on the pedagogical finalities of the exercise, the authors used this knowledge cutoff date and asked students to refrain from referring to post-2021 events or, conversely, to focus on post-2021 events to enhance their autonomy in finding sources and information.

Exercise 1 presented an excellent opportunity to discuss ChatGPT with the students. To begin, we asked them what the objectives of the laboratory should be and what they wanted to get out of it. Responses included learning how to use ChatGPT for IR-related content, which allowed us to discuss how to discern good scholarship and to reflect on the activities that we should prioritize to achieve our goals. After the brainstorming, we discussed how to write a prompt, building on the objectives that we had identified. The authors prepared a list of characteristics that good prompts should include (Picture 1) and asked the students to use them to write their prompt. Although Exercise 1 was well received by the students, the researchers reflected on the necessity to enhance the critical approach to interactions with ChatGPT. In fact, the students' first impression was that ChatGPT offered high-quality information. It is interesting that they changed their opinion in the course of the laboratory and after sustained interactions with ChatGPT.

Consequently, the authors placed more emphasis on discussing the quality of results. Exercises 2, 3, and 5 offered good opportunities to do so. Exercise 2 included inputting a pre-edited prompt and interacting with the machine. The prompt was prepared by building on Exercise 1 and was discussed with the students and then input. In Exercise 3, the students were required to draft their own prompt autonomously, based on Exercises 1 and 2. As ChatGPT began to generate text, the goal was to interact and push the machine to generate better information by challenging it—for example, by asking for more details, fact checking the information given, and requesting more examples. The students' observations became increasingly critical. They noticed that ChatGPT “reuses the prompt

Picture 1  
Exercise 1



over and over” with “no original content, just repeating the prompt”—although this might be related to the fact that the prompts gave ChatGPT very detailed instructions to follow, especially in [Exercise 2](#). However, after running their prompt several times and interacting with ChatGPT, the students noticed that the machine derailed from it. For example, in some cases, it faked a conversation between “a coach” and “a student” ([Exercise 2](#)) instead of interacting with the student as a coach. In other cases, it decided independently to move across the requests made in the prompt, jumping from one command to another without respecting the order it was given ([Exercise 3](#) and [5](#)).

In introducing ChatGPT, OpenAI (2022) warned that the model “is often excessively verbose and overuses certain phrases” because its training data privileged lengthy, wordy answers that appeared “more comprehensive.” Furthermore, when the students asked for more case studies and readings, ChatGPT often became increasingly vague and unreliable. We reflected on the fact that ChatGPT’s responses are based on the quality of the data used to train it and therefore biases in training data could lead to inaccurate responses. In particular, ChatGPT’s use to provide guidance on sensitive and complex issues (e.g., international conflicts) was a particular concern voiced by both the students and the authors.

In the feedback sessions, the students noticed that using ChatGPT can be time-consuming, considering that the information must be double-checked constantly. This was a good opportunity to discuss the advantages and disadvantages of using Google Scholar and the textbook instead of ChatGPT to search for information and readings. The students noticed that it is more difficult to control the reliability of ChatGPT’s narrative because, generally, there are no

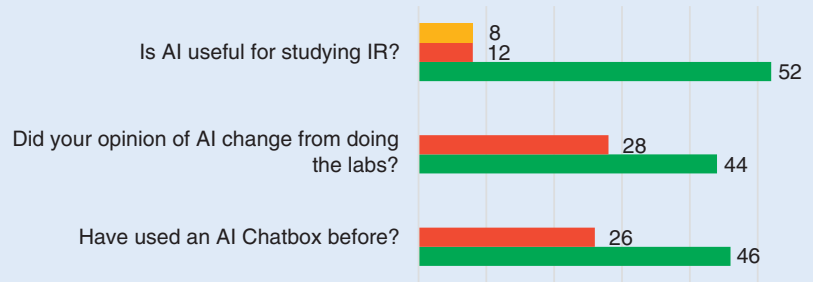
sources and the risk of confabulation is higher. They also reflected that whereas ChatGPT may suggest relevant readings, they were not able to understand the rationale for its suggestions. In fact, ChatGPT assigns readings based on probabilistic codes and words, not pedagogical considerations. It follows that Gen-AI chatbots’ reasoning process is unknowable, with significant implications for the users’ learning process.

The students also noticed ChatGPT’s standardized writing style, which included typical lists of items and vague explanations. The narrative is repetitive, they observed, which makes it possible for lecturers to identify a ChatGPT-generated text. [Exercises 2](#) and [3](#) helped the entire class to reflect on the usefulness of ChatGPT. We established that it can be useful for developing an initial plan for an essay and formulating initial thoughts about how to structure it, including how to organize it by sections. The narrative, however, needs to be rewritten and reframed. At this stage, students no longer seemed to trust ChatGPT as a knowledge base; due to the inherent risk of confabulation, they viewed it more as a useful “jumping-off point” for examining a topic or structuring an exercise.

[Exercises 4](#) and [5](#) provided the opportunity to upgrade the class practice to testing the uses of ChatGPT for the module’s assignment, especially relative to the final essay. [Exercise 4](#) provided the opportunity to discuss what is a good research question. In one-on-one conversations with the authors, first-year students could reflect on the issues of applicability and practicality of research questions, including the availability of sources and case studies to answer them. In [Exercise 5](#), students could test their knowledge of ChatGPT as well as the module’s IR-related content while they structured and worked on an essay. This presented the opportunity to discuss essay

Figure 1

## Student Relationship to ChatGPT



	Have used an AI Chatbox before?	Did your opinion of AI change from doing the labs?	Is AI useful for studying IR?
■ Sometimes			8
■ No	26	28	12
■ Yes	46	44	52

writing; an essay’s structure; and the specific function of each section, from the literature review to case-studies analysis. Confabulation, fake facts, inaccurate references, and reading suggestions were common, and the students identified them. Both exercises enhanced the students’ critical approach to ChatGPT.

### THE SURVEY: STUDENTS’ REFLECTIONS ON THE LABORATORY

Following the completion of the laboratories, students were asked to participate in a short survey reflecting on the strengths and weaknesses of ChatGPT for studying IR and the likelihood of their using it in the future as a study aid. Seventy-two students completed the survey.

As shown in figure 1, almost two thirds of respondents had prior experience using an AI chatbot. However, developing prompts and interacting with the chatbot tended to alter students’ perceptions of

the applicability of AI in their learning. A majority of respondents (52 students) agreed that ChatGPT was a useful tool in studying IR; eight students believed it was somewhat useful; and 12 thought it not useful at all due to its tendency to confabulate material. A student who had an affirmative perspective on the use of ChatGPT stated that AI made it easier to study IR because it allowed the student to easily “compare different theories.” Another student had a more ambivalent perspective, stating that the tool was useful “as a starting point to further explore IR” but, because the chatbot had a “tendency to give wrong information, [they] will not be relying much on this in the future and will stick to [their] own research skills.”

Students also were asked if they would incorporate ChatGPT going forward in their studies. Of the respondents, 78% (56 students) stated that they would use ChatGPT in the future and 22% (16 students) confirmed that they would not use it as a tool for their IR studies. As shown in figure 2, 50% of the group

Figure 2

## Student Post-Laboratory Incorporation of ChatGPT

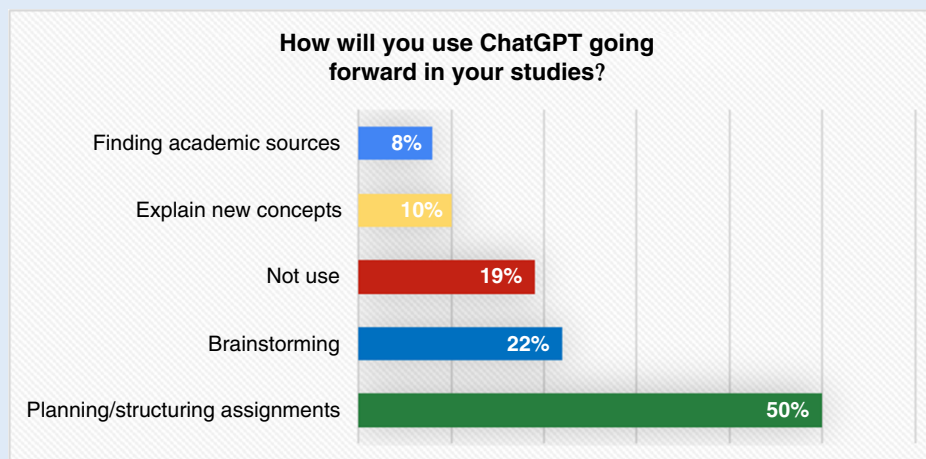
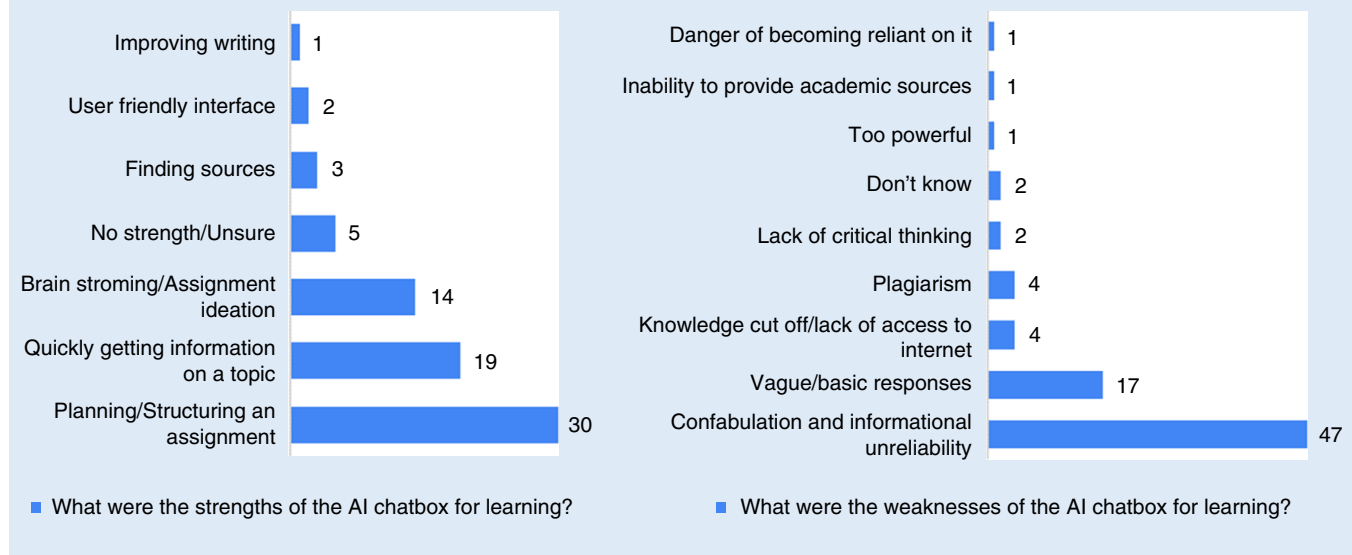


Figure 3  
Student Post-Laboratory Reflections on ChatGPT



stated that they would use ChatGPT to help with planning and structuring a future assignment and slightly less than 25% viewed ChatGPT as a useful tool for ideation when preparing for an essay. Seven students (10%) believed that ChatGPT was a resource for explaining new or difficult concepts arising in their studies. Six students (8%) stated that they would use the AI chatbot as a source of information in their studies. This reflected the lessons learned during the laboratories on the inherent risk of confabulation.

Finally, students were asked to reflect on the strengths and weaknesses of using ChatGPT in an academic setting (figure 3). It is interesting that only four students mentioned plagiarism in their response, opting instead to concentrate on the language model’s tendency to confabulate or give circuitous and vague responses. Students seemed to believe that the chatbot’s core strength was as an organizational tool or as an initial entry point into an assignment rather than a *bonafide* trusted source of information. This implied a pragmatic approach to the use of AI in learning, with students casting a critical eye on ChatGPT’s outputs while simultaneously drawing on AI’s ability to stimulate ideas, forge potent lines of academic inquiry, and provide structure to their work.

### CONCLUSION

Before presenting the key conclusions derived from this study, we address some of the limitations of this study. First, the number of students who participated in the laboratories and undertook the survey was limited. In this respect, the results of the survey relied on perceptions of those students who elected to participate in the laboratory and, therefore, cannot be compared to others outside of this group who did not attend the laboratory. It follows that we are not able to offer a definitive conclusion about the impact of the laboratory on the students’ learning experience because we cannot compare the two groups. Our study, however, reports on a practical experience and offers pedagogical reflections about incorporating Gen-AI chatbots in the classrooms through a straightforward description of the exercises that we conducted during the laboratory.

Despite the fact that this limits our study’s validity to an extent, the study offers insights into a learning experience in which other colleagues may be interested.

Second, our intentional use of Open AI’s ChatGPT, given that it is a widely used and popular chatbot in academic settings, poses another limitation. Our study, in fact, does not provide insights into the technological and pedagogical experience of using other, even comparable chatbots that students might be using. Nonetheless, this limitation also affords new directions and group-control mechanisms for future research.

A third limitation arises from how the question of plagiarism was only partially addressed in our survey (see figure 3). The impact of introducing and applying Gen-AI in the classroom and in comparable academic settings directly prompts further inquiry into what type of malpractice could be considered plagiarism. Such an inquiry would invoke wider questions on the legal definition of plagiarism, including but not limited to academic administration—particularly how they are communicated to and understood by students. We fully recognize that such questions will need further investigation in future research.

Such limitations aside, a major conclusion emerging from this study concerns the students’ pedagogical experience in the laboratory setting. As the end-of-laboratory survey and observations throughout the sessions revealed, a learning curve with the use of ChatGPT became evident. The teaching and learning experience fostered a critical approach to Gen-AI, apparent in students reporting how they became less trusting and more discerning in their evaluation of its responses.

A second conclusion that can be drawn from the study pertains to the many ways in which the students benefited from the laboratory. They gained practical knowledge in how to deploy a tool that is now part of their learning landscape. Accordingly, they built an understanding of when and how ChatGPT is useful, enhancing their critical-thinking skills more generally and confirming a trend observed by other scholars (Guo and Lee 2023; Messeri 2023; Yu 2023).

---

A third conclusion concerns essay writing and assignments. As evident in the survey's results and our observations, although students agreed that ChatGPT proved useful to plan an essay

In addition to the core benefits of student learning, the laboratory offered several secondary advantages. It partially addressed challenges associated with a shortened semester and a large class

*As the end-of-laboratory survey and observations during the lab revealed, a learning curve with the use of ChatGPT became evident. The experience fostered a critical approach to Gen-AI, apparent in students reporting how they became less trusting of its responses.*

and organize its content, they acknowledged the risk of confabulation and the machine's inability to think critically—which, in turn, precluded its usability for producing an essay in full. This finding could assuage some concerns concerning AI and plagiar-

size, effectively bridging a technology gap. Most significant, the authors learned that they and educators in general are non-replaceable. However, concerns remain regarding the risk of language uniformity and standardization.

*ChatGPT's performance was useful for enhancing writing instruction by motivating students with immediate feedback in the early stages of the writing process. The key takeaway is the importance of students developing critical skills to evaluate and integrate feedback, especially when coming from an automated source. Our study therefore also focused on the necessity of enhancing students' critical-thinking skills.*

ism in higher education, particularly because of how a positive correlation appeared between the frequency and depth of students' engagement with the chatbot and their criticism about its ability to provide academically rigorous responses. However, this conclusion returns us to the previous point related to the need for further inquiry about the forms and definition of plagiarism and how Gen-AI has been transforming it.

A fourth key finding from the laboratory emerged in how the learning experience enhanced students' critical-thinking skills in more independent terms. This finding aligns with a recent study

Finally, and crucially, as many survey answers suggested, students likely will continue using ChatGPT, even when they are aware of its many limitations. This should encourage institutions to dedicate more resources to educate the educators about ChatGPT—which, unfortunately, few universities do. It is crucial that educators learn to use ChatGPT so that they can educate students about AI-generated content and feedback. The goal is to remind students that they, not the chatbots, have the ultimate say in applying their learning and expressing their ideas.

*Students likely will continue using ChatGPT, even when they are aware of its many limitations. This should encourage institutions to dedicate more resources to educate the educators about ChatGPT—which, unfortunately, few universities do. The goal is to remind students that they, not the chatbots, have the ultimate say in applying their learning and expressing their ideas.*

by Steiss et al. (2024) that compared human and AI (ChatGPT) feedback on student essays. Whereas human evaluators provided better overall feedback, particularly in areas such as clarity and accuracy, ChatGPT's performance was useful only for enhancing writing instruction by motivating students with immediate feedback in the early stages of the writing process. The same study also highlighted the importance of educators being aware of ChatGPT's limitations due to its reliance on algorithms, which tends to produce oversights and inconsistencies. The key takeaway is the importance of students developing critical skills to evaluate and integrate feedback, especially when coming from an automated source. Our study therefore also focused on the necessity of enhancing students' critical-thinking skills.

#### **SUPPLEMENTARY MATERIAL**

To view supplementary material for this article, please visit <http://doi.org/10.1017/S1049096524000817>.

#### **DATA AVAILABILITY STATEMENT**

Research documentation and data that support the findings of this study are openly available at the *PS: Political Science & Politics* Harvard Dataverse at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZB1S4P&version=1.0>.

#### **CONFLICTS OF INTEREST**

The authors declare that there are no ethical issues or conflicts of interest in this research. ■

### REFERENCES

- Atlas, Stephen. 2023. "ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI." College of Business Faculty Publications. South Kingstown: University of Rhode Island. [https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1547&context=cba\\_facpubs](https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1547&context=cba_facpubs).
- Centre for Teaching and Learning. 2023. "Beyond ChatGPT: The State of Generative AI in Academic Practice for Autumn 2023." Oxford, UK: University of Oxford. [wwwctl.ox.ac.uk/beyond-chatgpt](http://wwwctl.ox.ac.uk/beyond-chatgpt).
- Floridi, Luciano, and Massimo Chiriatti. 2020. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines* 30 (4): 681–94. <https://doi.org/10.1007/s11023-020-09548-1>.
- Goodlad, Lauren M. E., and Sharon Stoerger. 2023. "Teaching Critical AI Literacies: Advice for the New Semester." New Brunswick, NJ: Rutgers University. <https://otear.rutgers.edu/teaching/aiadvice#:~:text=Develop%20rubrics%20that%20emphasize%20critical,as%20to%20encourage%20intrinsic%20motivation>.
- Grove, Jack. 2024. "Academics Despair as ChatGPT-Written Essays Swamp Marking Season." June 17. [www.timeshighereducation.com/news/academics-despair-chatgpt-written-essays-swamp-marking-season](http://www.timeshighereducation.com/news/academics-despair-chatgpt-written-essays-swamp-marking-season).
- Guo, Ying, and Daniel Lee. 2023. "Leveraging ChatGPT for Enhancing Critical-Thinking Skills." *Journal of Chemical Education* 100 (12): 4876–83. <https://doi.org/10.1021/acs.jchemed.3c00505>.
- Halaweh, Mohamad. 2023. "ChatGPT in Education: Strategies for Responsible Implementation." *Contemporary Educational Technology* 15 (2). <https://doi.org/10.30935/cedtech/13036>.
- Herft, Andrew. 2023. "A Teacher's Prompt Guide to ChatGPT Aligned with What Works Best Guide." <https://usergeneratededucation.files.wordpress.com/2023/01/a-teachers-prompt-guide-to-chatgpt-aligned-with-what-works-best.pdf>.
- Holmes, Wayne, and Ilkka Tuomi. 2022. "State of the Art and Practice in AI in Education." *European Journal of Education* 57 (4): 542–70. <https://doi.org/10.1111/ejed.12533>.
- Marche, Stephen. 2022. "The College Essay Is Dead. Nobody Is Prepared for How AI Will Transform Academia." December 6. [www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371](http://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371).
- Messeri, Lisa. 2023. "Teaching with ChatGPT: Critiquing Generative Artificial Intelligence from the Classroom." *Anthropology Now* 15 (1): 84–92. <https://doi.org/10.1080/19428200.2023.2230098>.
- Milmo, Dan. 2023. "ChatGPT Reaches 100 Million Users Two Months after Launch." *The Guardian*, February 2. [www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app](http://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app).
- Mollick, Ethan, and Lilach Mollick. 2023. "Assigning AI: Seven Approaches for Students, with Prompts." Ithaca, NY: Cornell University. <https://arxiv.org/ftp/arxiv/papers/2306/2306.10052.pdf>.
- OpenAI. 2022. "Introducing ChatGPT." November 20. <https://openai.com/blog/chatgpt>.
- OpenAI. 2023. "New AI Classifier for Indicating AI-Written Text." January 31. <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>.
- Quality Assurance Agency for Higher Education. 2023. "Reconsidering Assessment for the ChatGPT Era: New QAA Advice Published." July 31. [www.qaa.ac.uk/news-events/news/reconsidering-assessment-for-the-chatgpt-era-new-qaa-advice-published#](http://www.qaa.ac.uk/news-events/news/reconsidering-assessment-for-the-chatgpt-era-new-qaa-advice-published#).
- Rivetti, Paola, Rituparna Banerjee, and David O'Mullane. 2024. "Replication Data for 'Testing ChatGPT in IR Classrooms: Potentialities, Limitations, and What's Next.'" *PS: Political Science & Politics*. DOI: 10.7910/DVN/ZB1S4P&version=1.0.
- Steiss, Jacob, Tamara Tate, Steve Graham, Jazmin Cruz, Michael Hebert, Jiali Wang, et al. 2024. "Comparing the Quality of Human and ChatGPT Feedback of Students' Writing." *Learning and Instruction* 91 (1): 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>.
- Stokel-Walker, Chris. 2022. "AI Bot ChatGPT Writes Smart Essays: Should Professors Worry?" December 9. [www.nature.com/articles/d41586-022-04397-7](http://www.nature.com/articles/d41586-022-04397-7).
- The Russell Group. 2024. "New Principles on Use of AI in Education." Washington, DC, March 7. <https://russellgroup.ac.uk/news/new-principles-on-use-of-ai-in-education>.
- Theophilou, Emily, Cansu Koyuturk, Mona Yavari, Sathya Bursic, Gregor Donabauer, Alessia Telari, et al. 2023. "Learning to Prompt in the Classroom to Understand AI Limits: A Pilot Study." *AIXIA 2023: Advances in Artificial Intelligence. XXIInd International Conference of the Italian Association for Artificial Intelligence*. Rome, November 6–9. In *Proceedings*, ed. Roberto Basili, Domenico Lembo, Carla Limongelli, and Andrea Orlandini, 481–96. <https://doi.org/10.48550/arXiv.2307.01540>.
- Yu, Hao. 2023. "Reflection on Whether Chat GPT Should Be Banned by Academia from the Perspective of Education and Teaching." *Frontiers in Psychology* 14:1–12. <https://doi.org/10.3389/fpsyg.2023.1181712>.