

Is ChatGPT good or bad for education?

Alisa-Isabelle Gubanova

Abstract

The newest advances in Artificial Intelligence (AI), and specifically, ChatGPT, resulted in a lot of excitement and confusion. Some people believe that it will destroy our education while others claim that it will be used as a powerful tool making it better. There are many concerns about accuracy of the answers, ethical aspects including plagiarism, and slightly outdated information. In this study we decided to focus on the use of AI models in the education of high school students with the goal to understand the risks and the benefits. We conducted a survey asking high school students about their use of ChatGPT and other AI models. We learned that many students use AI for their studies, though it is not a very reliable tool. Our conclusion is that it does not make accurate calculations and cannot generate meaningful essays but can be used as an intelligent supportive tool for grammar checking and search for the formulas on a given topic.

Keywords: Large Language Models, ChatGPT, Education

Introduction

The use of ChatGPT in education has been a common topic of conversation among K-12 teachers and college faculty (Baidoo-Anu, 2023, Adeshola, 2023, Lo, 2023, İpek, 2023, Imran, 2023). Many view it as a threat to education and envision many ethical issues, while others view it as a helpful tool (Rahman, 2023, Mhlanga, 2023, Zhai, 2022, Arif, 2023, Sok, 2023). The positive properties of AI-based education include personalized learning experiences and efficient content generation, support for educators and enhanced learning experiences. The drawbacks are degradation of critical thinking and independent problem-solving skills, concerns related to data privacy, biases, hallucinations. However, it can be expected that the actual situation requires much more detailed understanding. The idea behind this research is to fill this knowledge gap and find out in which classes students have used ChatGPT or other AI models and what exactly they use it for and evaluate it on case-by-case basis.

Methods: To conduct our study, we used several digital technologies. First, to collect statistics about the use of ChatGPT by high school students, we created a survey that was administered via a Google Form. The survey was advertised on Instagram and administered in person, with a total of 24 responses. The study was conducted anonymously. Based on our study we found the most popular AI model used by students and identified the most popular use cases. Since ChatGPT was the most popular AI model, we conducted a set of tests to evaluate the quality of the results generated by ChatGPT 3.5 in response to various questions from different fields of knowledge. We present and analyze our findings in the Results and Conclusion sections.

Procedure

The very first step was to administer the survey through the Instagram platform. Some of the questions had multiple choice answers, while others were open-ended. The users could answer N/A if they had no answer for a particular question. The list of the questions is the following:

1. Have you ever used ChatGPT or any other AI models for your studies?
 - Yes
 - No
 - I have never heard about ChatGPT or AI
 - I have heard about ChatGPT and AI, but have never tried them
2. In which classes/subjects did you use ChatGPT or any other AI model?
3. Have you used ChatGPT, Quillbot, Google AI, or AI Quizlet? Are there any other models that you have used?
4. How often do you use ChatGPT or other AI models?
 - Every now and then
 - 1-2 times a month
 - Several times a month
 - Several times a week
 - Every day
5. What kind of questions did ChatGPT or other AI answer right?
6. What kind of questions did ChatGPT or other AI answer wrong?
7. What is the best thing about ChatGPT and/or other AI models?
8. What is the worst thing about ChatGPT and/or other AI models?
9. What grade you are in?

Data

Based on our survey, 12 respondents answered that they used ChatGPT or other AI models for their studies, while the other 12 said "No". None of them chose an option that "they hadn't heard of ChatGPT or other AI models". Out of 12 who stated that they've used AI models, 8 said that they used ChatGPT, 3 - QuillBot, 3 - AI Quizlet, 2 - Google AI, and 1 used Bing AI. Several of them mentioned multiple AI systems that they used for studies. Figure 2 demonstrates that the most used AI model is ChatGPT, and some students were using multiple models (Figure 1).

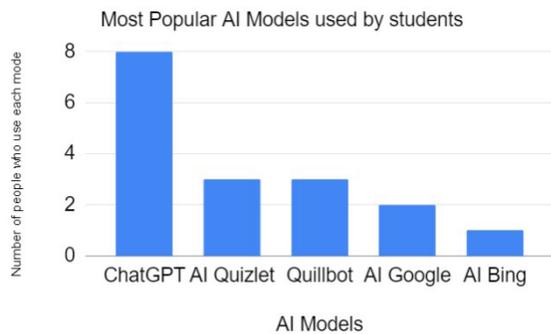


Fig. 1. Number of respondents using each of the AI models.

Figure 2 demonstrates that AI models were more popular among older high school students (classes 10-12), than 9th grade or the last year of middle school (8).

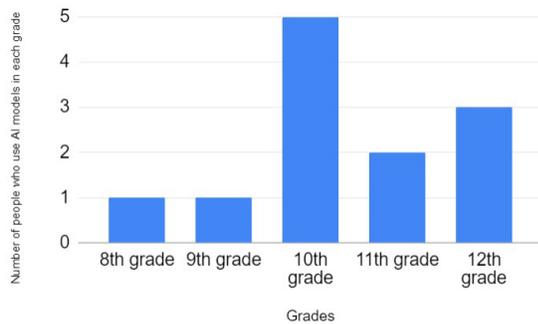


Fig. 2. Grades of the respondents that answered positively about using AI models.

Figure 3 demonstrates different classes where students use their AI models, followed by Figure 4 demonstrating specific applications.

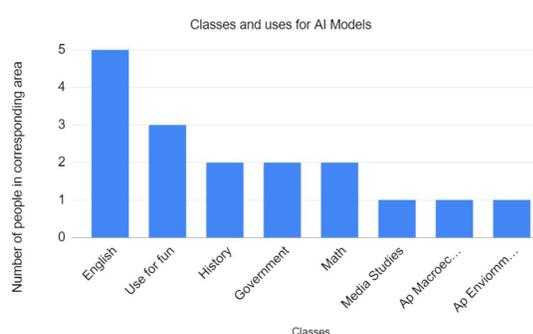


Fig. 3. Classes where students use AI models for education

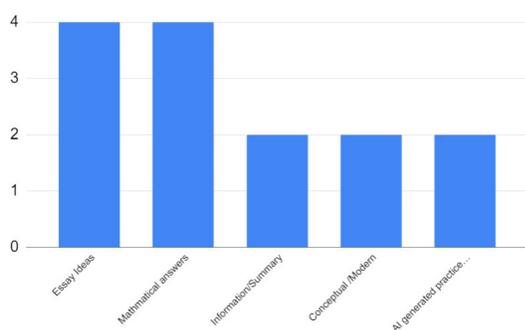


Fig. 4. Specific use-cases where students use AI models, mostly ChatGPT

Based on the survey results we decided to 1) focus on ChatGPT, since it is the most popular AI model. 2) Evaluate its performance on a) essay writing, b) mathematical problems, c) recent information, d) science problems. The results of the evaluation are the following.

Results

a) Essay writing by ChatGPT.

Figure 5 shows a fragment from a ChatGPT generated essay. While each individual sentence looks normal, reading of the whole text demonstrates that it lacks any structure, progression, or development of thought. It is impossible to extract any moral, or main ideas, or conclusions. Overall, it looks like a short, polished summary lacking any analysis or creative thought.

Recommendation: ChatGPT should not be used for essay writing, not only because of ethical issues, but also due to very poor quality of the analytical part of the essay. However, other tasks, such as summarization, grammar check, finding synonyms, and other fine-tuning of a human-written document can be done with this tool. Under those circumstances it will benefit students' education by helping them to expand their vocabulary and enriching their sentence structure.

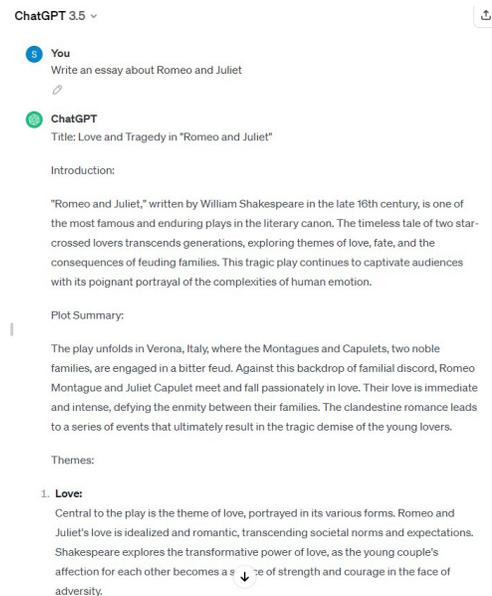


Fig. 5. ChatGPT generated essay (fragment).

b) Mathematical problem solving with ChatGPT

Our experiments demonstrated that ChatGPT fails even simple math questions (Figure 6). Both Google and calculator showed the correct answer, while ChatGPT was off by 4.

Recommendation: Not recommended for math education.

c) Recent information

We evaluated ChatGPT 3.5's knowledge of recent events. Since it was trained on data up to January 2022, it did not know anything recent. Also, often it just did not want to give a specific answer. For example, Figure 7 demonstrates that it uses general statements, such as "several studio albums" with "a number of songs". Googling this question resulted in answer "274", (as of 08/29/24).

Recommendation: for more reliable information we would still prefer Google over the free version of ChatGPT 3.5.

ChatGPT

To find the sum of the numbers 1, 15, 79, 19, 94, 86, 75, 199, 316, 319, 22, add them together:

$$1 + 15 + 79 + 19 + 94 + 86 + 75 + 199 + 316 + 319 + 22 = 1231$$

Therefore, the sum is 1231.

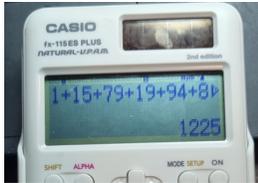
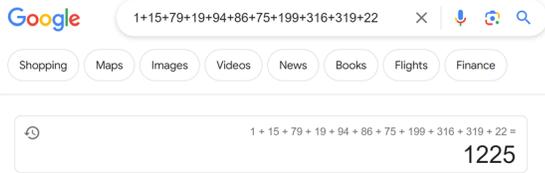


Fig. 6. Simple math problem solved by ChatGPT, Google and a calculator

You
how many songs does Taylor Swift have

ChatGPT
As of my knowledge cutoff date in January 2022, Taylor Swift has released numerous songs across her career. The exact number may vary depending on how one counts, including singles, album tracks, and collaborations. Swift has released several studio albums, and each album typically contains a number of songs.

For the most accurate and up-to-date information on the number of songs Taylor Swift has, I recommend checking her official website, discography, or reputable music databases. Additionally, Taylor Swift may have released new music since my last update in January 2022.



Fig. 7. ChatGPT answer about the number of Taylor Swift songs.

d) Science problems.

We decided to evaluate ChatGPT on a physics problem to see if it can give us an answer. Figure 8 shows that it keeps explaining things, but never manages to produce and answer. It is somewhat helpful because it mentions the right formulas and a human might be able to find an answer, but it cannot solve the problem for you.

Recommendations: This might not be too bad for learning physics, since it can help a student who is completely lost, while not doing all the work. Overall: recommended (cautiously).

You
There are 2 planets. Planet A revolves around the Earth. Planet B revolves around Jupiter. The distance from planet A and Earth is 238,900 miles. The distance of Planet B and Jupiter is 3 times that. The mass of planet A is $7.34767309 \times 10^{22}$ kilograms, and planet B is 5 times that mass. What is the difference in their gravitational pull?

ChatGPT
The gravitational force between two objects is given by Newton's law of gravitation:

$$F = \frac{Gm_1m_2}{r^2}$$

where:

- F is the gravitational force,
- G is the gravitational constant ($6.67430 \times 10^{-11} \text{ N m}^2/\text{kg}^2$),
- m_1 and m_2 are the masses of the two objects, and
- r is the distance between the centers of the two masses.

For the difference in gravitational pull, you would need to calculate the gravitational force for each planet and then find the difference.

Let's denote:

- F_A as the gravitational force between Planet A and Earth,
- F_B as the gravitational force between Planet B and Jupiter.

Given that the distance from Planet A and Earth (r_A) is 238,900 miles, and the distance of Planet B and Jupiter (r_B) is 3 times that ($r_B = 3 \times 238,900$), you can calculate F_A and F_B .



- F_A as the gravitational force between Planet A and Earth,
- F_B as the gravitational force between Planet B and Jupiter.

Given that the distance from Planet A and Earth (r_A) is 238,900 miles, and the distance of Planet B and Jupiter (r_B) is 3 times that ($r_B = 3 \times 238,900$), you can calculate F_A and F_B .

$$F_A = \frac{Gm_Am_E}{r_A^2}$$

$$F_B = \frac{Gm_Bm_J}{r_B^2}$$

Then, find the difference $\Delta F = F_B - F_A$.

Given that the mass of Planet A is $7.34767309 \times 10^{22}$ kilograms and the mass of Planet B is 5 times that, $m_B = 5 \times m_A$, substitute these values into the formulas to find the gravitational forces and their difference.

You
give me an answer

ChatGPT
Let's calculate the gravitational forces and their difference:

Given:

- Gravitational constant $G = 6.67430 \times 10^{-11} \text{ N m}^2/\text{kg}^2$
- Mass of Planet A (m_A) = $7.34767309 \times 10^{22}$ kilograms
- Mass of Planet B (m_B) = $5 \times 7.34767309 \times 10^{22}$ kilograms
- Distance from Planet A to Earth (r_A) = 238,900 miles
- Distance from Planet B to Jupiter (r_B) = 3 times r_A

Fig. 8. ChatGPT is trying to solve physics problem.

Conclusion

To summarize, the main contribution of this paper is the detailed analysis of how ChatGPT can help students to improve their education. Specifically, we believe that ChatGPT can support the development and refinement of writing skills. While it cannot produce a high-quality essay, it may still be beneficial, as it discourages the unethical use of AI and ensures that students do not entirely bypass the writing process. ChatGPT's inability to solve physics problems does not preclude it from being a valuable supporting tool. It generates useful hints and formulas that can guide students who feel completely lost, without doing all the work for them. It can also serve as an intelligent tutor, providing guidance and support to help students navigate challenging concepts while encouraging independent problem-solving. We do not recommend ChatGPT for mathematical applications. If it cannot do basic addition, we cannot trust any other operations that it would do. Finally, the model is not very up to date with respect to recent events, and it fails questions that require analysis or counting of objects, even in a simple form. Overall, we think that AI is at its infancy, and it will continue to produce more interesting outcomes, but even at its current state it can enrich high school education with the new personalized support that has not been available previously. If it is used ethically and critically, it can be very beneficial.

References

- Adeshola, I., & Adepoju, A. P. (2024). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 32(10), 6159-72. <https://doi.org/10.1080/10494820.2023.2253858>
- Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is ChatGPT a blessing or blight in disguise? *Medical Education Online*, 28(1), 2181052. <https://doi.org/10.1080/10872981.2023.2181052>.
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>.
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>.

- İpek, Z. H., Gözümlü, A. I. C., Papadakis, S., & Kallogiannakis, M. (2023). Educational applications of the ChatGPT AI system: A systematic review research. *Educational Process: International Journal*, 12(3), 26-55. <https://doi.org/10.30935/cedtech/13605>.
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>.
- Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. In *FinTech and artificial intelligence for sustainable development: The role of smart technologies in achieving development goals* (pp. 387-409). Cham: Springer Nature Switzerland. <https://dx.doi.org/10.2139/ssrn.4354422>.
- Rahman, M. M., & Watanobe, Y. (2023). "ChatGPT for education and research: Opportunities, threats, and strategies." *Applied Sciences*, 13(9), 5783.
- Sok, S., & Heng, K. (2023). ChatGPT for education and research: a review of benefits and risks. *Cambodian Journal of Educational Research*, 3(1), 110-121. <https://doi.org/10.3390/educsci13040410>.
- Zhai, X. (2022). "ChatGPT user experience: Implications for education." Available at SSRN 4312418.