

Applicability of ChatGPT for generating and calculating partial derivatives of multivariable functions

Mária Bakó¹

INFO

Received 02/01/2025

Accepted 18/01/2025

Available on-line 04/02/2025

Responsible Editor: László

Várallyai

Keywords:

ChatGPT, prmathematics,
education.

ABSTRACT

In this article, we present an experiment conducted with first-year students majoring in Commerce and Marketing at the Faculty of Economics at the University of Debrecen, along with its results.

The aim of the experiment was to make the teaching of Mathematics more engaging for the students and to familiarize them with the limitations of using ChatGPT 3.5.

We tasked ChatGPT with generating a solved problem for each student, focused on calculating the first and second-order partial derivatives of multivariable functions. Subsequently, the students were asked to solve the assigned problems themselves and verify the solutions provided by ChatGPT. This approach was intended to help students recognize the limitations of the tool and develop a critical perspective on its use.

1. Introduction

Researchers have been studying the application of artificial intelligence (AI) in education for over three decades (O’Shea and Self, 1986; Luckin et al., 2016; Hwang et al., 2020). Progress in this area—like in many others—has been exponential. The launch of ChatGPT in 2022 made AI’s capabilities evident to the general public, shocking many with the realization that this algorithm could pass significant exams with good results. Naturally, this raised concerns about the potential for cheating, presenting educational institutions with a major dilemma. Should its use be allowed or restricted? How can we identify whether a student used ChatGPT to complete their homework (Burkhart, 2023, Davis, and Kumar, 2023)? Should we ban it or integrate it into education? Less than two months after its launch, research revealed that up to one-fifth of students were already using AI programs for assessment tasks (Cassidy, 2023). However, the actual usage rate may be significantly higher. A January 2023 survey of over 1,000 university students found that more than one-third were utilizing ChatGPT for writing assessments. Among these users, 75% acknowledged that it constituted cheating but chose to use it regardless (Intelligent, 2023). These behaviors have prompted some universities to prohibit the use of ChatGPT and led certain academics to label such tools as a “threat” and a “plague on education” (Sawahel, 2023; Weissman, 2023). Although the release of ChatGPT has sparked significant controversy regarding its implications for higher education, it also presents clear opportunities to improve student learning and accessibility (Sullivan et al., 2023).

What is ChatGPT?

ChatGPT is a large language model (LLM) developed by OpenAI, based on the GPT (Generative Pre-trained Transformer) architecture. It generates human-like text responses in conversational contexts and understands various topics and writing styles. By employing deep learning techniques, ChatGPT learns patterns and structures from its training data, enabling it to generate coherent and contextually relevant responses to user inputs. ChatGPT can engage in dialogue, answer questions, provide explanations, generate creative text, and assist in various linguistic tasks.

Numerous scientific publications have highlighted its strengths and weaknesses of ChatGPT (Mohammadreza et al., 2023, Lesson plan, 2023) The literature clearly indicates that ChatGPT should not be banned from higher education but rather integrated into teaching and learning processes (Gentile et al., 2023; Oliveira, 2023; Borji, 2023, Bakó, 2024). The potential of ChatGPT to enhance education,

¹ Mária Bakó

University of Debrecen

e-mail: bako.maria@econ.unideb.hu

solve mathematical problems, and support student learning is immense. For instance, it can aid teachers and educators in creating personalized and relevant educational materials for students (Guo et al., 2023). This can result in improved student engagement, increased motivation, and better academic performance (Wardart 2023). Additionally, the ChatGPT model can be a valuable resource for educational assessments and evaluations. Teachers can use it to efficiently assess student assignments and provide feedback, as highlighted in studies by Crust (2023) and De Winter (2023).

In the case of mathematics, several studies have examined ChatGPT's capabilities in generating and solving both basic and more complex problems. Azaria, in (Azaria, 2022), points out arithmetic errors made by ChatGPT when multiplying large numbers, calculating roots and powers of a number, or performing addition and subtraction of irrational numbers. A qualitative study described in (Wardart, 2023) demonstrated that ChatGPT is a useful educational tool, but caution is required when using it, and guidelines must be developed for safe usage. In fact, the article describes several incorrect answers to fairly simple mathematical questions.

As the above makes clear, while ChatGPT possesses impressive capabilities, it occasionally provides incorrect or nonsensical answers. This is because ChatGPT primarily focuses on understanding and generating language, not solving mathematical and logical problems (Liu et al, 2023, Wardart, 2023). Mathematics involves precise calculations, logic, and formal reasoning, which fundamentally differ from linguistic tasks. Moreover, solving mathematical problems often requires a deeper understanding of concepts and a step-by-step reasoning process to achieve accurate solutions. While ChatGPT excels at generating plausible responses, its lack of formal understanding and mechanisms for executing mathematical computations can result in difficulties in producing accurate mathematical outcomes.

Based on this, it is evident that despite its many merits, using ChatGPT in education presents significant challenges and potential risks. One major risk lies in the ChatGPT algorithm's reliance on the data used during training, which can easily amplify biases present in the data (Cotton et al., 2023). Additionally, there is a danger that students might become overly reliant on ChatGPT, hindering the development of critical thinking and independent problem-solving skills. Cruz's paper examines the applicability of ChatGPT in teaching Calculus (Cruz et al., 2023). The results showed that ChatGPT produced numerous inaccurate answers and problematic questions in arithmetic and symbolic computations. However, despite these apparent limitations, ChatGPT, with proper guidance, can sometimes arrive at accurate solutions to such problems, highlighting the importance of teacher supervision when students use this large language model. According to their view, introducing ChatGPT in teaching sessions could be advantageous if we reverse the narrative: instead of accepting ChatGPT's answers and thus impeding the development of independent problem-solving skills, the answers could be debated. This could achieve the opposite effect, as meaningful discussions between students and teachers might enhance students' critical thinking in a simple and friendly environment. Moreover, with careful use, ChatGPT can offer a personalized approach to learning, as its dialog is friendly, dynamic, and adaptable—a fast way to acquire theoretical knowledge. These aspects are crucial because the current generation of young people in higher education was born and raised in the digital age and is accustomed to quickly adopting, interacting with, and customizing new software. For this generation, it feels as natural as any other aspect of their lives, and they expect to find this in their educational training as well. The greatest challenge for today's educators will be to integrate these tools into their teaching practices.

Despite being aware of all these limitations of ChatGPT in solving mathematical problems, we attempted to involve it in teaching the course Economic Mathematics as an experiment. The goal was to make the course more diverse and to familiarize our students with the limitations of using ChatGPT 3.5. We asked ChatGPT to generate unique problems—with solutions—for each student, aligned with the topic of the given lesson. Students were then tasked with solving their assigned problems and verifying ChatGPT's solutions. Our goal was to help students recognize the tool's limitations and develop a critical approach to using ChatGPT in the future.

In this article, we present the tasks generated by ChatGPT for calculating partial derivatives of multivariable functions, along with the results and insights gained from this experiment. We discuss the

types of problems ChatGPT generated, their variety, difficulty level, the accuracy of its solutions, and how well our students managed to solve these problems.

2. Methodology

The study involved 110 first-year students majoring in Commerce and Marketing at the Faculty of Economics, University of Debrecen. The experiment was conducted during the second semester of the 2023/24 academic year.

Our primary goal with the experiment was to make the teaching of Mathematics more engaging for our students. Secondly, we aimed to assess ChatGPT's performance in generating and solving partial derivatives of multivariable functions, as well as its potential use by educators in creating supplementary materials. We evaluated the model's ability to generate diverse questions within this topic and its capability to solve the problems it generated. These aspects are detailed in this article.

Each student was assigned a number, and every week they were required to solve the problem corresponding to their assigned number and upload their solution to the university's Moodle system. Students were allowed to use any tools to solve the problems. Besides providing additional practice, our goal was to help students familiarize themselves with ChatGPT's limitations and discourage uncritical reliance on the tool in the future.

We typically requested ChatGPT to generate 20–30 problems at once; however, in many cases, it only produced 10–20 problems. It often required repeated prompts to obtain the necessary number of tasks.

The following prompt was used to generate the tasks: *Generate 20 two-variable functions and provide their first- and second-order partial derivatives.*

3. Results

3.1. The problems

We asked ChatGPT to generate 110 problems, along with their solutions, for determining the first- and second-order partial derivatives of multivariable functions. Generating 110 problems for a specific topic is a relatively large request, but unfortunately, the generated problems lacked significant variety. In terms of difficulty, the problems were computationally intensive and prone to many potential errors.

Overall, we can conclude that while the generated problems were generally good, there was a significant repetition of problem types. Essentially, the first 50 problems were diverse, but after that, it seemed as though the model got "stuck," and the subsequent problems were variations of the same types. For example, functions like

$$f(x, y) = \frac{\cos(xy)}{y^2} + e^{x-y^2}$$

frequently appeared.

The first 50 problems were notably diverse and highly useful for practicing the topic. However, the latter part of the problem set became monotonous, with the problems being more challenging and containing more errors in the provided solutions. It was interesting to observe that, even in the beginning, when solving simpler tasks, ChatGPT 3.5 often made mistakes when calculating the second-order partial derivatives. Below, we can see the solution to Problem 5:

$$\begin{aligned} &5. \text{ Function: } f(x, y) = \sin(xy) \\ &\text{First-order partial derivatives:} \\ &\frac{\partial f}{\partial x} = y \cos(xy) \\ &\frac{\partial f}{\partial y} = x \cos(xy) \\ &\text{Second-order partial derivatives:} \\ &\frac{\partial^2 f}{\partial x^2} = -y^2 \sin(xy) \\ &\frac{\partial^2 f}{\partial y^2} = -x^2 \sin(xy) \\ &\frac{\partial^2 f}{\partial x \partial y} = (1 - xy^2) \cos(xy) \end{aligned}$$

We can observe that the solution is correct up until the last step, where the mixed partial derivative is completely incorrect. The correct solution would be as follows:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \cos(xy) - xy \sin(xy)$$

Furthermore, a very common error was observed in the solution of Problem 48. In the case of the pure second-order derivative with respect to y , we can see that ChatGPT performs well at the beginning but fails to complete the process—it simply forgets to differentiate $\sin(y)$.

48. Function: $f(x, y) = \frac{\sin(x)}{y} + \cos(y)$
 First-order partial derivatives:
 $\frac{\partial f}{\partial x} = \frac{\cos(x)}{y}$
 $\frac{\partial f}{\partial y} = -\frac{\sin(x)}{y^2} - \sin(y)$
 Second-order partial derivatives:
 $\frac{\partial^2 f}{\partial x^2} = -\frac{\sin(x)}{y}$
 $\frac{\partial^2 f}{\partial y^2} = \frac{2 \sin(x)}{y^3} + \sin(y)$
 $\frac{\partial^2 f}{\partial x \partial y} = -\frac{\cos(x)}{y^2}$

3.2. First-order partial derivatives

Regarding ChatGPT's solutions, we concluded that it is not reliable for solving partial derivatives. It can be observed that even with very simple problems, there are mistakes in calculating the first-order partial derivative. If we examine problem 37, we can notice that in the second part of the problem, the first-order partial derivative of e^{x+y} with respect to x and y is similarly miscalculated, resulting xe^{x+y} and ye^{x+y} , whereas the correct solution would be e^{x+y} .

37. Function: $f(x, y) = \ln(xy) + e^{x+y}$
 First-order partial derivatives:
 $\frac{\partial f}{\partial x} = \frac{1}{x} + ye^{x+y}$
 $\frac{\partial f}{\partial y} = \frac{1}{y} + xe^{x+y}$

As shown in Figure 1a, ChatGPT solved 38% of the problems incorrectly, making errors in 42 out of 110 problems. On the other hand, our students performed slightly better, as shown in Figure 1b, with 70% of them solving the problems correctly.

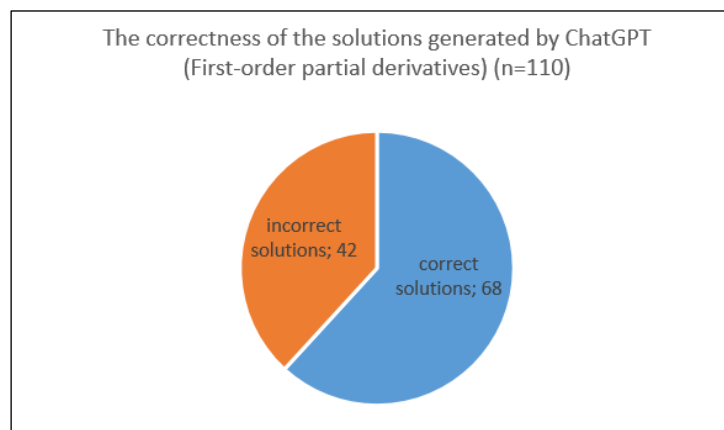


Figure 1a. The correctness of the solutions generated by ChatGPT (First-order partial derivatives).

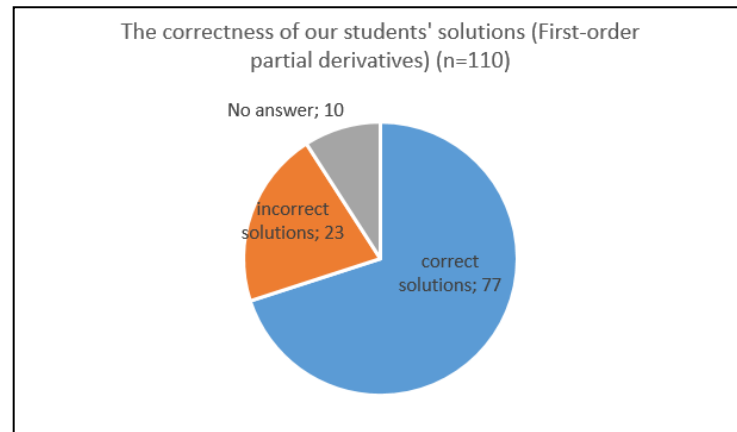


Figure 1b. The correctness of our students' solutions(First-order partial derivatives).

The most interesting insight is presented in Figure 2, which shows the proportion of tasks solved correctly and incorrectly by ChatGPT and how well our students performed on those tasks. We can conclude that students managed to correctly solve nearly half of the problems ChatGPT had solved incorrectly. On the other hand, while 14 students made errors on problems ChatGPT had solved correctly, 28 students successfully corrected ChatGPT's incorrect solutions.

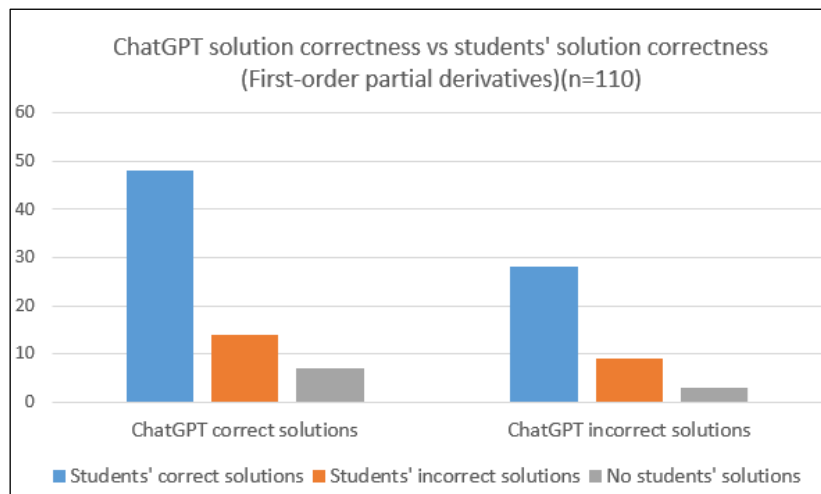


Figure 2. ChatGPT solution correctness vs students solution correctness (First order partial derivatives)

3.3. Second-order partial derivatives

As for the second-order partial derivatives, the results were significantly worse. Naturally, if the first-order partial derivative was incorrect, the second-order derivative could not be correct either. However, when comparing *Figure 1a* with *Figure 3a*, we can observe that the number of incorrect solutions generated by ChatGPT doubled. While the first-order partial derivatives were incorrect in 42 cases, the second-order partial derivatives were incorrect in 84 cases, accounting for errors in 76% of the solutions.

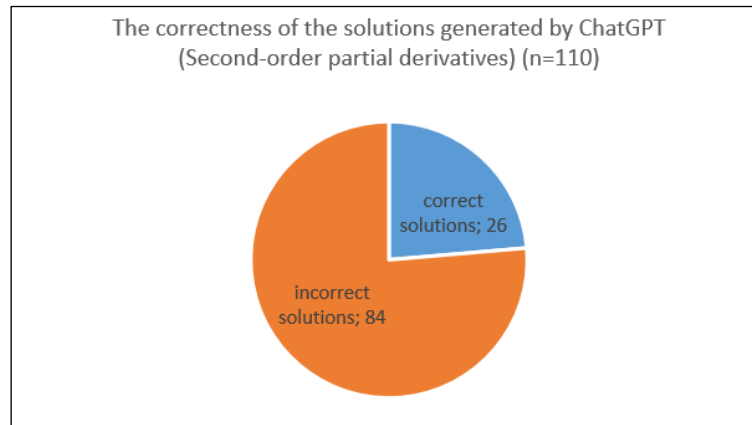


Figure 3a. The correctness of the solutions generated by ChatGPT (Second–order partial derivatives).

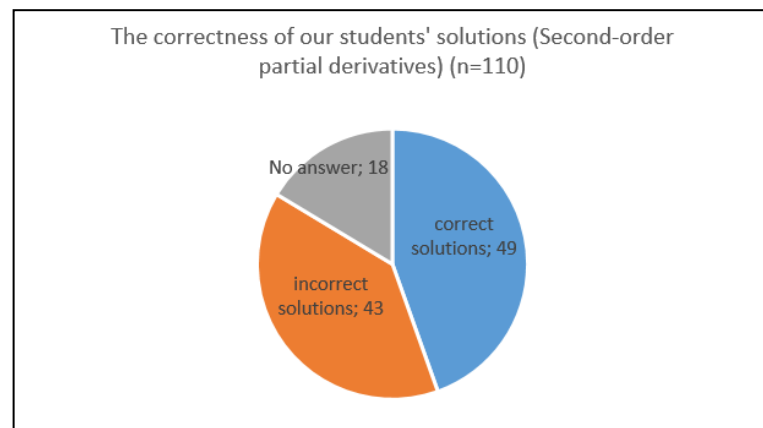


Figure 3b. The correctness of our sutents' solutions (Second –order partial derivatives).

Looking at our students' responses, we can conclude that they performed somewhat better than ChatGPT, solving 45% of the tasks correctly, compared to ChatGPT's 25%. However, while the first-order derivatives were incorrect in only 23 cases, the second-order partial derivatives were incorrect in 42 cases, representing 38% of the total solutions.

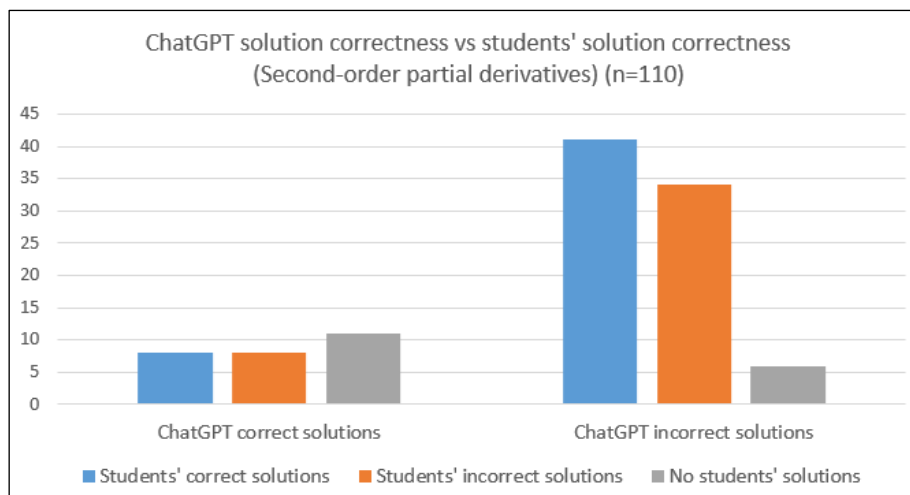


Figure 4. ChatGPT solution corectness vs students solution corectness (Second order partial derivatives)

In Figure 4, we can see that out of the 84 incorrect solutions provided by ChatGPT, our students corrected 41 of them. However, they made errors on 8 of the problems that ChatGPT had solved correctly.

4. Conclusions

The objective of our experiment was to make the teaching of *Economic Mathematics* more engaging, evaluate how diverse the problems generated by ChatGPT were, assess the reliability of its solutions, provide students with additional practice opportunities, and familiarize them with the limitations of using ChatGPT 3.5. ChatGPT produced a range of problems with varying levels of difficulty, and about half of them were suitable for students to use as practice. However, when it came to solving these problems, the AI didn't perform as well as we hoped. ChatGPT 3.5 had a relatively high error rate, even for first-order partial derivatives (38%), which increased significantly to 76% for second-order partial derivatives. This performance was significantly lower than that of our students, who demonstrated an error rate of 21% for first-order partial derivatives and 39% for second-order partial derivatives. Despite these shortcomings, ChatGPT added value by exposing students to a new way of learning and prompting discussions about the role of AI in education. It gave them extra practice material while also teaching them to approach AI-generated content critically, recognizing its strengths and weaknesses. While ChatGPT isn't ready to replace traditional methods or human expertise, it can still be a useful tool for enhancing the learning experience when used thoughtfully.

References

- Azaria, A. (2022). ChatGPT Usage and Limitations <https://doi.org/10.31219/osf.io/5ue7n>
- Bakó, M., (2024). Applicability of ChatGPT for generating and solving of exercises in basic probability theory, ICERI2024 Proceedings, Pages 4348-4354, ISBN: 978-84-09-63010-3, doi: 10.21125/iceri.2024.1091
- Borji, A. (2023). A Categorical Archive of ChatGPT Failures, <tps://arxiv.org/abs/2302.03494>
- Burkhart, M. (2023). AI program ChatGPT presents useful resource to students, educators warn against cheating. The Technician. <https://www.proquest.com/wire-feeds/ai-program-chatgpt-presents-useful-resource/docview/2771844930/se-2?accountid=10675>
- Cassidy, C. (2023, January 17). Lecturer detects bot-use in one fifth of assessments as concerns mount over AI in exams. The Guardian. <https://www.theguardian.com/australianews/2023/jan/17/lecturer-detects-bot-use-in-one-fifth-of-assessments-as-concerns-mount-over-ai-in-exams>
- Cotton, D. R. E., Cotton, P. A., Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT, *Innovations in Education and Teaching International*, DOI: 10.1080/14703297.2023.219014
- Crust, G. (2023). ChatGPT employability study skills and curriculum development. <https://doi.org/10.13140/RG.2.2.35643.28960>
- Cruz, J. P., de Sousa Oliveira, M.P., Cação, I. (2023). Teaching and learning calculus with ChatGPT: Benefits and limitations, ICERI2023 Proceedings, Sevilla 2023, Pages:456-463 ISBN: 978-84-09-55942-8, doi: 10.21125/iceri.2023.0175
- Davis, M., & Kumar, A. (2023). 'Pay attention to it, ignore it or push back on it': Brown professors discuss AI's impact on academic integrity. The Brown Daily Herald. <https://www.proquest.com/wire-feeds/pay-attentionignore-push-back-on-brown/docview/2774393345/se2?accountid=10675>
- De Winter, J. (2023). Can ChatGPT pass high school exams on English language comprehension? https://www.researchgate.net/publication/366659237_Can_ChatGPT_pass_high_school_exams_on_English_Language_Comprehension
- Gentile, M., Citta, G, Perna, S., Allegra, M. (2023). Do we still need teachers? Navigating the paradigm shift of the teacher's role in the AI era, *Front. Educ.* :1161777. doi:10.3389/educ.2023.1161777
- Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J., & Wu, Y. (2023). How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection. arXiv. <https://doi.org/10.48550/arXiv.2301.07597>
- Hwang, G-J., Xie, H., Wah, B. W., Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education, *Computers and Education: Artificial Intelligence*, Volume 1, 2020, <https://doi.org/10.1016/j.caeai.2020.100001>
- Intelligent. (2023). Nearly 1/3 college students have used ChatGPT on written assessments. <https://www.intelligent.com/nearly-1-in-3-collegestudents-have-used-chatgpt-on-written-assignments/>
- Lesson Plan (2023): Teaching and Learning in the Era of ChatGPT - The New York Times ([nytimes.com](https://www.nytimes.com))

- Liu, H., Ning, R., Teng, Z., Liu, J., Zhou, Q., Zhang, Y.: (2023): Evaluating the Logical Reasoning Ability of ChatGPT and GPT-4, 2023, arXiv preprint arXiv:2304.03439.
- Luckin, R., Holmes, W., Griffiths, M., Forcier, L. B. (2016). *Intelligence Unleashed, An argument for AI in Education*, London: Pearson 2016, ISBN: 9780992424886, https://www.researchgate.net/profile/Rosemary-Luckin/publication/299561597_Intelligence_Unleashed_An_argument_for_AI_in_Education/links/5812782608ae1f5510c2aa4d/Intelligence-Unleashed-An-argument-for-AI-in-Education.pdf
- Mohammadreza, F., Seyyed, K. B., Omid, N., Arjen, W. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research, *Innovations in Education and Teaching International*, DOI: 10.1080/14703297.2023.2195846
- Oliveira, A.(2023). Talk at University of Aveiro, <https://www.youtube.com/watch?v=PICpMwv9n4o&t=4443s>
- O’Shea, T., Self, J. (1986). *Learning and teaching with computers: The artificial intelligence revolution*. Prentice Hall Professional Technical Reference.
- Sawahel, W. (2023). Embrace it or reject it? Academics disagree about ChatGPT. *University World News*. <https://www.universityworldnews.com/post.php?story=20230207160059558>
- Sullivan, M., Kelly, A. and McLaughlan, P. (2023) ‘ChatGPT in higher education: Considerations for academic integrity and student learning’, *Journal of Applied Learning and Teaching*, 6(1). Available at: <https://doi.org/10.37074/jalt.2023.6.1.17>.
- Wardart, Y, Tashtoush, M. A., Alali, R., Jarrah, A.(2023). ChatGPT: A Revolutionary Tool for Teaching and Learning Mathematics, May 2023, *Eurasia Journal of Mathematics, Science and Technology Education* 19(7):1-18, DOI:10.29333/ejmste/13272
- Weissman, J. (2023). ChatGPT is a plague upon education. *Inside Higher Ed*. <https://www.insidehighered.com/views/2023/02/09/chatgpt-plague-upon-educationopinion>