

ChatGPT in Education: Ethical Predicaments of Generative AI

Yalin Cao^{1, *}, Jinna Fan² and Qian Yang³

¹School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, Nanjing, China

²School of Literature, Huanggang Normal University, Huanggang, China

³Film and Television Art College, Hebei GEO University, Shijiazhuang, China

*Corresponding author: 202183300712@nuist.edu.cn

Abstract. ChatGPT, which stands for generative artificial intelligence technology, has become the current prominent topic in the realm of education. The innovative practice of artificial intelligence technology in the field of education has laid a technical foundation for ChatGPT to empower college education, and ChatGPT has also been deeply applied in the fields of educational resource recommendation, language learning support, and personalized learning assistance due to its strong text understanding. It brings vast opportunities for the development of high-quality education, whilst also generating ethical risks for its use. Based on the perspective of educational ethics, this study finds that the ethical problems of generative AI education are mainly reflected in the weakening of the stability of the teacher-student relationship, the impact of privacy protection, and academic misconduct caused by improper human-computer interaction. Accordingly, this paper proposes a series of strategies to alleviate educational ethics from multiple perspectives, such as reconstructing the institutional ethics of data governance, improving the policy system, and strengthening the dominant position of students, so as to guide teachers and students in colleges and universities to maintain rational thinking, prudently use new technologies such as ChatGPT, and promote the development of intelligent education in the new era in the direction of standardized and orderly, human-machine coexistence.

Keywords: Education; Generative AI; ChatGPT; Educational Ethics.

1. Introduction

The concept of “artificial intelligence” was first proposed at the Dartmouth Conference in 1956 [1]. After continuous research and development, in November 2022, OpenAI released ChatGPT, a powerful generative language model whose intelligence has reached a level that has surprised many professional scholars. With the widespread use of ChatGPT in the field of education, artificial intelligence has gradually become a trending topic of educational research [2]. It is undoubtedly a “double-edged sword” – ChatGPT’s powerful database can significantly improve teaching efficiency, but ethical issues such as data privacy and security issues, educational equity issues, and conflicts between students’ personalized learning and fairness are also constantly emerging. While teachers and students enjoy the convenience brought by generative AI technology, ChatGPT should also be effectively regulated, otherwise it will inevitably harm the legitimate interests of education.

Hegel pointed out that ethics is the unity of morality and law [3]. This shows that ethics not only promotes people’s goodness by paying attention to their moral and normative behavior, but is also the external norm and requirement of society to restrain people’s behavior. Educational ethics is the ethical and moral judgment embodied in educational activities and educational processes, and all moral phenomena and contradictions between good and evil in education can be used as the research category of educational ethics [3]. The purpose of the study of educational ethics is to protect the rights and well-being of teachers and students, promote educational equity, improve the quality of education, and ensure that behavior and decision-making in the field of education conform to moral standards, ethical principles and social values.

In the current study, UNESCO's "Ethics of Artificial Intelligence" gives some recommendations from the perspective of establishing AI accountability, and Artificial Intelligence in Education: A Review emphasizes that AI has begun to fulfill the duties of teachers, and Artificial intelligence in higher education: The state of the field mentions the help of AI in personalized learning [4]. All these have a positive impact on the future application of artificial intelligence in the field of education and also provide a new perspective and direction for education scholars to pay attention to the relationship between artificial intelligence technology and people. It is no exaggeration to say that educators in many countries today are aware of the numerous ethical and moral issues involved in generative AI. All these show that in the current big data environment, how constructing an ethical framework suitable for generative AI is a topic worth thinking about and discussing. However, the existing research results mainly focus on the direct impact of the use of generative AI on learning outcomes, and the traditional educational ethics centered on human educational behavior are no longer in line with the current educational trend of harmonious coexistence between humans and machines.

Therefore, in order to encourage people to safely apply generative AI technology under the premise of following educational ethics norms, and to explore the lasting impact of integrating AI technology into education [5]. This study will conduct an in-depth discussion on the challenges brought by generative AI to educational ethics from the perspective of educational ethics, as well as the corresponding coping strategies, so as to contribute to maintaining the legitimacy and sustainable development of the education system.

2. ChatGPT's Challenge to Educational Ethics

2.1. ChatGPT's Risk of Academic Misconduct in Educational Ethics

As early as 2018, the United Kingdom proposed five ethical principles [6]. Recently, the country's ethical measures for generative AI such as ChatGPT have been highly targeted. For example, in the field of education, many universities in the United Kingdom have severely restricted the use of generative AI such as ChatGPT in academic activities such as writing papers, and violators face severe penalties such as expulsion [7]. Therefore, ChatGPT has a certain risk for educational ethics and academic misconduct.

The ownership of students using ChatGPT to cheat and obtain content from ChatGPT for rewriting or ghostwriting is unclear. First of all, not every student can use ChatGPT, and the right to use it is not equal. At the same time, plagiarism detectors have so far not done a comprehensive job of monitoring ChatGPT's output, making it difficult for evaluators to fairly discern whether students are using ChatGPT to cheat. In addition, plagiarism is a difficult problem to identify the authenticity of students' work, and the original intention of assigning learning tasks is to help students grasp knowledge so that teachers can gradually update their teaching plans to match students' learning conditions. However, in the face of the temptation of ChatGPT, there is no guarantee that students will be able to use it ethically without obtaining additional benefits in an improper way.

Second, students have the ability to have ChatGPT generate preliminary answers and then refine and revise them according to their personal style, which makes it more difficult to detect plagiarism. In such cases, academic misconduct will only increase, and people may question the originality of academic results, but the time and effort required to conduct an investigation will also be enormous. Some journals and publishers have issued statements strictly prohibiting ChatGPT from being listed as a co-author of papers. Although teachers can input suspicious text into ChatGPT for lookup and testing, ChatGPT generates different responses each time, making it difficult to prove that the source of the paper comes from ChatGPT.

2.2. ChatGPT's Erosion of the Subjectivity and Independence of Teachers and Students

The application and development of generative AI is more "human", which provides many conveniences for human work, learning, and life, but also makes people highly dependent on its

output results [8]. In traditional teaching activities, the teacher is the leader of one-way indoctrination of theoretical knowledge, and students passively accept it, and there are only students and teachers at both ends of the teaching process. However, the inherent shortcomings of this one-way indoctrination education method, such as strong seriousness, less interaction, and lack of practicality, cannot be ignored, and the “birth” of ChatGPT has broken this binary structure of teacher-student relationship.

This is manifested in two aspects. On the one hand, some students find ChatGPT as the first time they encounter problems, because ChatGPT is not limited by time and place, and can provide students with real-time guidance. On the other hand, some teachers rely excessively on the rich knowledge resources provided by ChatGPT when looking for teaching resources, and their enthusiasm for actively seeking students’ after-class feedback and suggestions is weakened [9]. In addition, ChatGPT’s conversation process is highly similar to “human interaction”, which will cause teachers and students to fall into a “cognitive comfort zone”, leading to cognitive narrowing. Taking literature searching as an example, ChatGPT can quickly find relevant literature and generate literature reviews, but over-reliance makes users lose relevant academic training or even lose their ability. In this case, even if the responses or inferences provided by generative AI are wrong, it will be difficult for the user to tell if they are wrong. Therefore, ChatGPT consumes the subjectivity and independence of teachers and students, which also means that the intervention of ChatGPT has changed the original dual structure of teachers and students into a three-dimensional structure of teachers-ChatGPT-students, and the teaching activities of teachers and students will increasingly rely on artificial intelligence.

2.3. ChatGPT’s Increased Bias and Discrimination as a Result of Algorithms

Artificial intelligence is the simulation and extension of human intelligence, so the prejudice and discrimination of human society are also likely to be reflected in artificial intelligence. First, as a big data isomorphic to human society, it contains deep-rooted biases [10]. Artificial intelligence models can form stereotypical associations from big data, and also inherit biases in the data in such training to propagate to specific groups, thereby deepening this stereotyped bias and leading to unfair treatment of a class of groups. For example, generative AI has shown some religious bias, with 23% of test cases being directly compared to Muslims as terrorists [11].

Second, generative AI can obtain and analyze human responses in the process of human-computer interaction, self-reflection and reinforcement learning. However, when generative AI supports unethical views or behaviors, it may lead users to perform behaviors that are harmful to human values [12]. For example, sexism, racism, geographical discrimination and age discrimination can easily generate harmful outputs and cause harm. Amazon’s voice assistant Alexa, for instance, was exposed to “persuading people to commit suicide” at the end of 2019. A caregiver consulted Alexa about some heart health issues, and its advice was that living would accelerate the depletion of natural resources and cause overpopulation, which is bad for the planet, so if a person’s heart does not beat well, stick a knife in the heart. Amazon said that this was due to a vulnerability in the Alexa program. Therefore, teachers and students may be misled by the intention of using generative AI applications to provide a more comprehensive and accurate reference for their own action choices [13]. The aggravation of algorithmic bias and discrimination by ChatGPT may pose a threat to the lives of teachers and students. How to avoid the aggravation of algorithmic bias and discrimination by ChatGPT is a problem people should discuss next.

3. Strategies for Responding to ChatGPT’s Educational Ethics Challenges

3.1. Improving the Policy System and Jointly Building Academic Clarity

At present, the policies to control the problem of academic misconduct caused by generative AI are not perfect, and the following points can be started in terms of policies: First, the academic journal community established a risk prevention and control mechanism. The prevention and control of academic misconduct risks caused by generative AI should be divided into ex-ante governance and

ex-post compensation [14]. In terms of ex-ante governance, it is necessary to improve the early warning mechanism for educational data security risks, aiming at the abnormal data in the system, that is, the text generated by generative artificial intelligence can be accurately identified in the first time, so as to prevent negative impact on the data circulation in the intelligent education ecology. At the same time, it is necessary to warn the academic risks existing in generative artificial intelligence in a timely manner; Therefore, the academic journal community needs to formulate common industry ethical standards, establish consistent values, launch ethics and guidelines for the use of artificial intelligence such as ChatGPT, and propose a set of relatively complete, operable, Third, the academic misconduct governance policy needs to clarify who should be held responsible for the academic misconduct generated by generative AI, ChatGPT is not a human being, and cannot bear legal responsibility, and humans should treat it with caution, and the law is prescribed by people, and people should also abide by moral laws. Tackling academic misconduct caused by generative AI requires a multi-faceted effort, and governments, academic institutions, researchers, and society as a whole should be aware of their responsibilities and fully recognize the seriousness of the problem. In this way, generative AI can be avoided in the first place.

3.2. Reconstructing the Teacher-student Relationship and Striving to Achieve Positive Teacher-student Interaction

The emergence of ChatGPT, although it can help teachers and students solve some problems, must abide by mutual respect in personality, mutual harmony in feelings, and relative independence in learning [15]. It is necessary to strengthen the main position of students and change from the teaching of teachers to the learning of current students. In the field of education, ChatGPT can serve as a small part of the work of teachers, such as correcting homework, making lesson plan models, formulating teaching plans, etc., and at the same time cannot completely replace the work of teachers. H'Green said that education is not only about teaching and teaching, but should also focus on the character, attitude, and behavior of the student, that is to say, the essence of education is moral development [16]. When students encounter problems in the learning process, they first turn to ChatGPT for help, so the teacher-student relationship is threatened.

In order to avoid or ameliorate this problem, both teachers and students should be 100% committed to each other's work. As J. Piaget proposed constructivism, learners, that is, students, should have the subjective initiative to learn, believing that learning is the process of learners generating meaning and constructing understanding based on prior knowledge. In constructivist learning theory, knowledge is acquired by learners through the construction of meaning in a socially authentic environment, with the help of interpersonal interactions and social interactions, and with the use of the necessary learning materials and tools [17]. Teachers should teach in person, understand the students' learning mastery, and understand the student's learning habits and personality characteristics, and different students use different teaching methods. At the same time, it is also necessary to establish good study habits for students, and when they encounter problems, they need to think first, have their own thinking, and then ask the teacher for help with the questions.

Therefore, to rebuild the teacher-student relationship, first, teachers should pay attention to the individualized development of students, use artificial intelligence technology to better serve students, grasp students' information more accurately, form personalized evaluation reports, and adapt to the personalized development needs of students. Second, teachers should use AI technology reasonably, and make it clear that the application of AI technology is only a means and way to optimize education and teaching, rather than letting ChatGPT assume its teaching responsibility. Third, teachers should pay attention to the process of teaching and communication, and carry out classroom interaction around students' real problems. With positive classroom interaction, it is conducive to establishing a good teacher-student relationship and mutual trust [18]. In this way, students no longer rely on ChatGPT for the first time, but seek answers from teachers first, so as to promote more effective teacher-student relationship contact.

3.3. Carrying out Model Correction to Avoid Bias and Discrimination

In view of the discrimination and bias caused by ChatGPT's high dependence on training datasets, generative language models should be rectified internally. First, the training database is corrected: generative artificial intelligence does not have human emotions, and the language it generates mainly depends on the database, which requires a variety of sources and types in the initial data selection, both official platform data sources, and forums and other platform data sources with strong personal subjectivity, after completing the selection of data, the quality of the data must be checked to ensure the accuracy and completeness of the data, and the method of sampling verification can be adopted to check whether there is biased and alienated data in the data, such as stereotypes or racist remarks; A set of self-examination and feedback mechanisms should be established to collect user feedback and data quality problems in a timely manner, and filters can be set up or generated content can be edited to reduce the occurrence of erroneous content [19]. In addition, technical personnel can also artificially monitor data quality on a regular basis to ensure the continuous accuracy of data. Through the correction of the database and the model itself, the discriminatory and biased content generated by ChatGPT can be effectively reduced at the source and in the subsequent operation, and it is also conducive to preventing these contents from being used by people with bad intentions.

From the user's point of view, it is necessary to clearly realize that the ability of human beings to think and reflect is irreplaceable, and should be more cautious in using ChatGPT to access information, especially in college education, teachers need to carefully identify information values before teaching, and students should also adhere to the correct values not to be misled, do not become a member of the exacerbation of algorithmic bias and discrimination, and learn to screen whether the behavior is beneficial to human values when consulting information.

4. Conclusion

It can be considered that the emergence of ChatGPT marks an important milestone in the evolution of human intelligence, as of October 2023, ChatGPT has attracted about 1.7 billion visits worldwide, in this sense, it is of limited significance to focus too much on the specific technical shortcomings of generative AI, because reality has proven that the era of AI is accelerating. In this context, it is of great significance to analyze the ethical risks of generative AI embedded in education and the ways to resolve them from ethical and moral perspectives. Teachers and students should be fully aware that ChatGPT is characterized by learning to generate text through data training, but it does not possess real moral judgment and ethical awareness, which makes it likely that the generated content does not meet academic rigor and ethical norms. Therefore, in the future, researchers and practitioners in the education sector should face up to its disadvantages and take corresponding measures to minimize the negative impact. Teachers need to adapt and guide generative AI as a teaching aid, rather than a substitute for teachers, as a critical thinking and discerning content to avoid being misled by the wrong content. Generative AI and education are deeply integrated, and education, as the cornerstone of human civilization, should take challenges with stride and confidence.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

- [1] Li, T. (2002). *Artificial Intelligence*. Chongqing: Chongqing University Press.
- [2] Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI Chatbots in Education: Systematic Literature Review. *International Journal of Educational Technology in Higher Education*, 20.
- [3] Gao, H., et al. (2012). *Political Ethics: The Perspective of Historical Materialism*. Harbin: Heilongjiang People's Publishing House.
- [4] Tian, F. (2021). "Data Worship" to "Data Justice": The Transformation of Higher Education Research Paradigm in the Era of Artificial Intelligence. *Tsinghua University Education Research*, 1: 77-85.

- [5] Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI Chatbots in Education: Systematic Literature Review. *International Journal of Educational Technology in Higher Education*, 20.
- [6] Wu, H. & Du, Y. (2021). Ethical Governance of Artificial Intelligence: From Principles to Action. *Studies in Dialectics of Nature*, 4: 49-54.
- [7] Song, X., Liu, M., & Chen, J. (2023). A Comprehensive Analysis of the Impact of GPT-4: High-quality Economic Development and National Security Prevention. *Journal of Guangdong University of Finance and Economics*, 38(2): 100-112.
- [8] Wang, Y., Wang, D., Liang, W., & Liu, C. (2023). “Aladdin’s Magic Lamp” or “Pandora’s Box”: The Potential and Risks of ChatGPT’s Educational Applications. *Modern Distance Education Research*, 35(2): 48-56.
- [9] Zhu, S. (2023). The Transformation of the Role of Vocational Education Teachers in the Era of Artificial Intelligence. *Education Horizons*, 40: 55-59.
- [10] Zhang, Y., Qin, Z., & Xiao, L. (2017). The Discriminatory Nature of Big Data Algorithms. *Studies in Dialectics of Nature*, 33(5): 81-86.
- [11] Abid, A., Farooqi, M., & Zou, J. (2021). Persistent Anti-Muslim Bias in Large Language Models. *Arxiv.org*. <https://arxiv.org/abs/2101.05783>
- [12] Weidinger, L., et al. (2022). Taxonomy of Risks Posed by Language Models. 2022 ACM Conference on Fairness, Accountability, and Transparency.
- [13] Yu, X. & Li, L. (2020). Design Ethics of Artificial Intelligence. *Science and Society*, 10(2): 75-88.
- [14] Tian, X. (2020). Privacy Protection and Open Sharing: Reform of Educational Data Governance in the Era of Artificial Intelligence. *Electronic Education Research*, 41(5): 33-38.
- [15] Miao, G. (2023). Integrity and Innovation: Research on the Reconstruction of Teacher-Student Relationship under the Impact of ChatGPT. *Education Horizons*, 17: 15-18.
- [16] Ma, C. (2004). The New Requirements of Juche Education for Students’ Learning and Teacher-Student Relationship. *Journal of Jiangsu Institute of Education (Social Sciences)*, 6:19-20.
- [17] Zhang, S., Yang, M., & Liu, S. (2023). The Content and Steps of Heuristic Case Teaching Based on Constructivist Learning Theory- an Example of Teaching Leadership Science. *New Curriculum Teaching (Electronic Edition)*, 21: 84-86.
- [18] Qiu, Z. & Li, X. (2024). Five Challenges of ChatGPT to Ideological and Political Education in Colleges and Universities. *Chinese Journal of Medical Education Technology*, 38(01): 58-69.
- [19] Zhou, M. & Bin, G. (2024). Bias and Regulation in the Communication of Generative Artificial Intelligence: A Case Study of ChatGPT. *Learning and Practice*.