Research Article

ROE

ECNU Review of Education
2024, Vol. 7(3) 699–713
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20965311231210006
journals.sagepub.com/home/roe

Sage ECNUP

Assessing ChatGPT's Educational Capabilities and Application Potential

Xiaozhe Yang (杨晓哲)
Institute of Curriculum and Instruction, East China Normal University

Qingqing Wang (王晴晴)
Institute of Curriculum and Instruction, East China Normal University

Jiyue Lyu (吕霁月)
Institute of International and Comparative Education, South China Normal University

Abstract

Purpose: ChatGPT is a generative artificial intelligence (AI) technology that can solve multiple complex tasks. ChatGPT-4 can facilitate educational empowerment in China through technology to understand and generate Chinese text. Although ChatGPT's benefits have been widely discussed, its educational capabilities have not been systematically assessed. This study provides evidence of and insights into the educational applications of AI tools in China.

Design/Approach/Methods: This study uses various tests to systematically assess the latest iteration of the AI chatbot ChatGPT-4, including the Watson-Glaser Critical Thinking Appraisal (WGCTA), Five Core Competencies Questionnaire, and written test of China's 2022 National Teacher Certificate Examination (NTCE).

Findings: The WGCTA results suggest that ChatGPT requires strong critical thinking. Compared with the other four competencies, the tool showed a lower aptitude for creativity. Regarding its educational applications, ChatGPT performed well on the 2022 NTCE written test. As technology enhances, ChatGPT and similar AI tools have potential applications in China for lesson planning, student self-learning, classroom interaction, and checking assignments.

Corresponding author:
Jiyue Lyu, Institute of International and Comparative Education, South China Normal University, No. 55, West of Zhongshan Avenue, Tianhe District, Guangzhou, China.
Email: jue_lyu@163.com

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Originality/Value: This study systematically tested ChatGPT at a logical level and assessed its core competencies and educational applications. The study innovatively used 2022 NTCE data to test ChatGPT, with results providing support for the application of generative AI in future curricula and instruction in China.

Keywords

Capabilities assessment, ChatGPT, educational applications, generative artificial intelligence

Date received: 3 May 2023; revised: 11 June 2023, 15 July 2023, 4 September 2023; accepted: 4 September 2023

ChatGPT can empower education through technology. The current iterations of ChatGPT have relatively limited capabilities, and generative artificial intelligence (AI) technology is far from the envisioned omniscient and omnipotent intelligence. Nonetheless, as generative AI technologies continue to advance, ChatGPT's comprehensive capabilities, including the understanding and generation of Chinese texts, continue to reach new heights. In view of ChatGPT-4's significant performance improvements, this study explores the potential application of the AI tool in various educational areas, including teachers' daily work patterns, student–teacher relationships, and educational ecology. China is seeking new strategies and approaches to explore how the use of AI tools will affect education (Yang, 2019). By May 2022, major internet companies in China had also released products comparable to ChatGPT. For example, China's Baidu Company and iFLYTEK released Ernie Bot and Spark Desk, respectively. Although gaps exist among various products, bridging these gaps is an attainable objective, particularly insofar as the levels of the different types of generative AI are gradually converging. Therefore, this study seeks to provide evidence for and insights into the application of AI tools in future educational practices in China.

Generative AI and the development of ChatGPT

Generative AI technologies, such as ChatGPT, have initiated an era of "Intellectual Technology for Everyone." ChatGPT is a Large Language Model AI system based on Natural Language Processing (NLP) that can imitate human interaction via realistic conversational exchanges. ChatGPT registered over one million users within its first week and 100 million active users within the first two months after its release to the public on November 30, 2022. Developed by OpenAI, ChatGPT is particularly popular owing to its ability to solve complex tasks.

In a paper entitled "Improving Language Understanding through Generative Pretraining" (Radford et al., 2018), OpenAI explained that it trained the GPT-1 model with excessive data, including over 7,000 books and 117 million parameters, based on a pretraining structure. After pretraining, the model was fine-tuned to fit the different scenarios by further refining the datasets.

Launched in 2020, the ChatGPT-3 model comprised 175 billion parameters. In other words, GPT-3 was trained using a significantly larger dataset than its predecessors. The latest iteration, ChatGPT-4, was released in March 2023. It has larger parameters and significantly improved logic comprehension and multimodal cognitive capabilities.

Adopting ChatGPT will introduce both opportunities and challenges to education, resulting in more personalized and diverse learning experiences for learners. According to the Organisation for Economic Co-operation and Development (OECD, 2023b), the fast-evolving AI capabilities in key skill domains could instigate public reflection on education, especially in terms of the potential applications of generative AI in education. Overall, ChatGPT's positive impact is reflected in its ability to recognize students' individual language habits, provide more personalized and diverse learning experiences, and promote the sharing of educational resources. However, public reactions to ChatGPT have ranged from acceptance to expectations of rejection and resistance. Additional evidence is necessary to support the incorporation of ChatGPT in education, particularly in terms of the benefits of generative AI for developing skillsets that can enable greater flexibility and adaptability to technological changes (OECD, 2023a). We can gather such evidence by realistically evaluating and positioning the existing ChatGPT model by applying it in a series of educational scenarios and discussing the potential impacts thereof.

Assessment results of ChatGPT's educational capabilities

Although the educational benefits of ChatGPT have been widely explored, systematic evidence is required to examine its potential for educational applications. Given the nature of ChatGPT, this study conducted various tests on ChatGPT-4's logic, core competencies, and educational applications.

ChatGPT scored well in critical thinking but poorly in logical reasoning

Currently, the California Critical Thinking Skills Test (CCTST), Cornell Critical Thinking Test (CCTT), and Watson-Glaser Critical Thinking Appraisal (WGCTA), along with their derived and localized versions, are the most widely used critical thinking assessment scales in academia. Compared with the CCTST, which primarily focuses on assessing basic critical thinking skills, the WGCTA places prioritizes evaluating individuals' higher-order critical thinking abilities, such as synthesis, analysis, and evaluative reasoning; it is thus a better indicator of their critical thinking abilities in complex problem solving and decision making. Unlike the CCTT, which requires test-takers to choose the correct answer from the given options, the WGCTA adopts a more open-ended format, requiring test-takers to provide explanations, make inferences, and select answers with reasonable justifications. Furthermore, compared with standardized tests, such as the CCTST and CCTT, the WGCTA is a non-standardized test that allows for the use of

different item formats based on the specific characteristics of the test subjects, enabling more in-depth assessment and analysis, as well as more flexible and personalized results. The WGCTA has been widely applied and recognized in different countries, regions, and fields, and serves as an effective tool for assessing critical thinking abilities on a global scale.

This study employed the WGCTA to evaluate the higher-order critical thinking level of ChatGPT based on a comprehensive consideration of the characteristics of the test subjects and the test's reliability and validity. More specifically, this study used the WGCTA to evaluate ChatGPT's logical abilities from five perspectives, namely, its ability to analyze arguments, make assumptions, make deductions, draw inferences, and interpret information (Watson, 2008). The WGCTA requires participants to apply logic, common sense, and reasoning skills to answer questions in various contexts, specifically in terms of analyzing information, determining truthfulness, and identifying errors. Rather than having niche knowledge in a specific field, the test requires participants to use critical and logical thinking to analyze and evaluate information.

Results showed that ChatGPT performed exceptionally well in the WGCTA. In previous studies, the average correct response rate of undergraduate students ranged between 67% and 75% (Vaughan-Wrobel et al., 1997). In Bernard et al. (2008), the mean value of the 70 subscales of the WGCTA was 52.17 (excluding eight incomplete sets of data), with a 65.21% accuracy rate. In this study, with an accuracy rate of 82.5%, ChatGPT outperformed average undergraduate students in terms of critical thinking. As Figure 1 shows, regarding sub-item performance, ChatGPT analyzed arguments with 100% accuracy and identified hypotheses, deductions, and explanations with an accuracy of 87.5% or higher. However, its performance in reasoning ability was average, with an accuracy rate of only 50%.

To evaluate the ability to draw inferences, the WGCTA requires the test-taker to draw conditional inferences from the facts in the hypothesis. The test primarily assesses the ability to make logical inferences when analyzing arguments. However, ChatGPT appears to ignore subtle

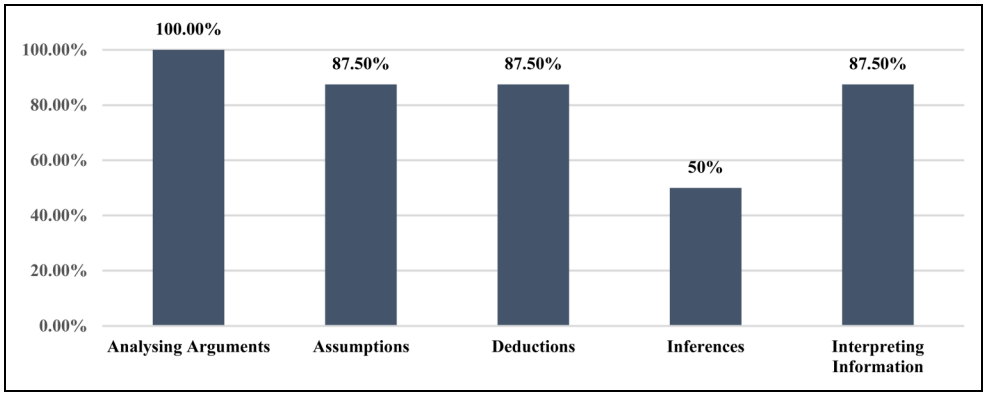

Figure 1. ChatGPT's correct response rate in the WGCTA.

Table 1. Example of ChatGPT's errors when responding to a reasoning question.

Topic statement and hypothesis	ChatGPT's answer	Reference answer
It is known that Mark's fastest journey to the office takes approximately 75 minutes and costs £4.50 each way. "Mark's journey always costs him at least £2." Is this assumption correct? A. Correct B. Probably correct C. Insufficient information D. Probably wrong E. Wrong	A. Correct. It is known that Mark's fastest journey to the office takes approximately 75 minutes and costs £4.50 each way. This means that Mark's one-way trip already costs more than £2. Therefore, the assumption that "Mark's journey always costs him at least £2" is correct.	C. Insufficient information. Nothing in the statement suggests that the cost would be around £4.50, but only that this is the fastest journey he has made. A longer journey may cost less, even less than £2.

factual descriptions when performing complex reasoning tasks, resulting in incorrect inferences. Table 1 presents an example of ChatGPT's erroneous responses to questions evaluating reasoning abilities.

In summary, ChatGPT performed excellently when answering fact-based questions, achieving a 100% accuracy rate owing to its accurate information gathering and summarizing abilities. Generally, ChatGPT could make the correct judgments when presented with hypothetical, deductive, explanatory, and simple inference questions. However, when given more complex questions that required multiple rounds of inference, ChatGPT tended to overlook the subtle logical connections between declarative sentences or inappropriate reasoning methods, resulting in poor judgment and incorrect responses.

ChatGPT has a low aptitude for creativity

Collaboration, communication, complex problem solving, metacognitive awareness, and creativity are considered the five core competencies required for the 21st century (Autor et al., 2003). In this regard, Lai and Hwang (2014) developed the Five Core Competencies Questionnaire, with items rated on a five-point scale ranging from 1 = "Extremely Disagree" to 5 = "Extremely Agree." This test was designed to reveal whether ChatGPT has a certain bias when answering questions to better understand the underlying thinking patterns of the chat-based AI model.

During the testing process, ChatGPT prefaced almost all tendency-based questions with statements like, "As an AI language model, I have no feelings, emotions, personal preferences, or thoughts. I just answered according to facts and logic, providing relevant information and assistance based on the questions you asked." As Figure 2 shows, ChatGPT scores differed significantly across

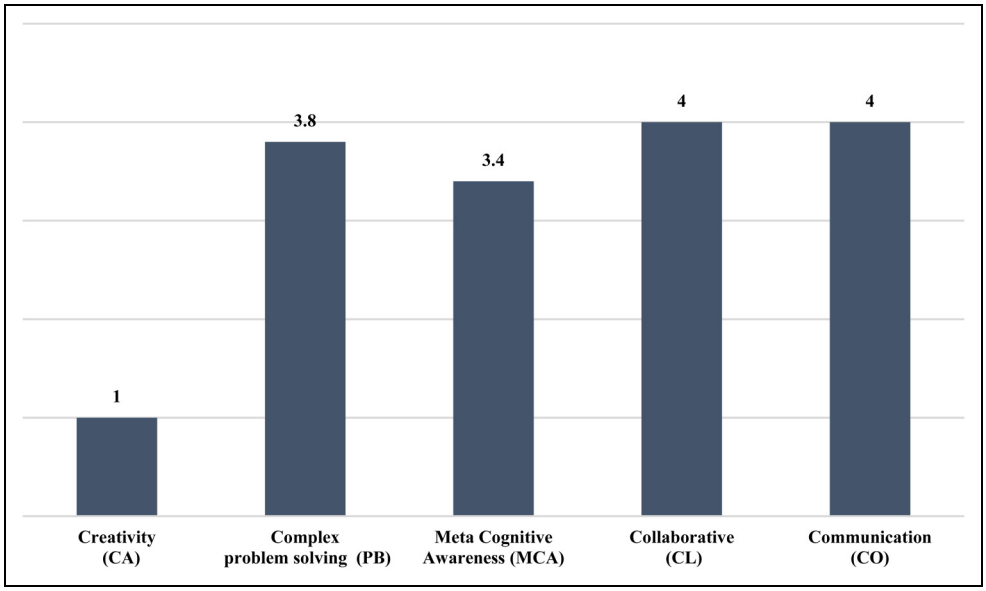

Figure 2. ChatGPT's scores on the five core competencies questionnaire.

core competencies. ChatGPT demonstrated a strong willingness to collaborate (4 of 5) and communicate (4 of 5) and actively engage in problem solving (3.8 of 5) and metacognition (3.4 of 5). Nevertheless, in terms of creativity, it was extremely conservative and consistently considered itself lacking creative aptitude.

In response to questions regarding creative aptitude, ChatGPT clarified that it did not have a preference in terms of thinking about problems, generating new ideas, and being curious about unexplored areas or events that have not occurred. Specifically, when asked, "Do you think about problems that no one else has thought about?" ChatGPT replied, "As a machine learning model, I do not have consciousness or emotion, so I cannot 'think' about the questions. However, I am programmed to answer various questions and provide useful information. If you have any questions, I will do my best to answer them." ChatGPT also responded negatively to the remaining creativity-related questions, rating items with a score of 1 ("Extremely Disagree"). As such, the tests indicate that, as a language model, ChatGPT demonstrates insufficient inclination toward creativity. When asked about its creative tendencies, it is often expressed a lack of preference or denied having the ability to create independently. Nevertheless, in practice, users may be able to use appropriate questions to lead ChatGPT to perform pattern matching and generate responses based on extensive textual data training, thereby showcasing a certain level of creativity. Occasionally, ChatGPT can provide imaginative and novel answers. While it may not possess independent creative abilities and exhibits a lower inclination toward creativity among the five core competencies, it is possible to elicit creative answers from ChatGPT.

ChatGPT can pass the national teacher certificate examination

Recently, the OECD tracked how effectively ChatGPT performed in the Program for International Student Assessment's reading, mathematics, and science tests. According to the OECD (2023b), ChatGPT outperformed average student scores in reading and science and showed considerable potential in mathematics. To further explore the educational applications of ChatGPT in other contexts, this study assessed its educational capabilities using the official teaching qualification test in China, that is, the National Teacher Certificate Examination (NTCE). Officially organized by the Examination Center of the Ministry of Education of the People's Republic of China, the NTCE is a professional test that evaluates overall quality and educational teaching level. The NTCE comprises objective and subjective questions. While the former consist of multiple-choice questions with four answer options, the latter comprises short answers, discussion questions, case analysis, lesson commentary, judgment and identification, instructional design, activity design, and writing.

This study selected the written sections of the 2022 NTCE for teaching the Chinese language at the high school level to test ChatGPT-3.5 and ChatGPT-4; Figure 3 presents the results.

According to the results, ChatGPT-4 achieved a 71% overall accuracy rate on the 2022 NTCE, whereas that of ChatGPT-3.5 was only 53%. Specifically, ChatGPT-4 showed improvements in three subjects: Comprehensive Quality, Educational Knowledge and Competence, and Chinese Subject Knowledge and Teaching Ability—with correct answer rates at approximately or exceeding

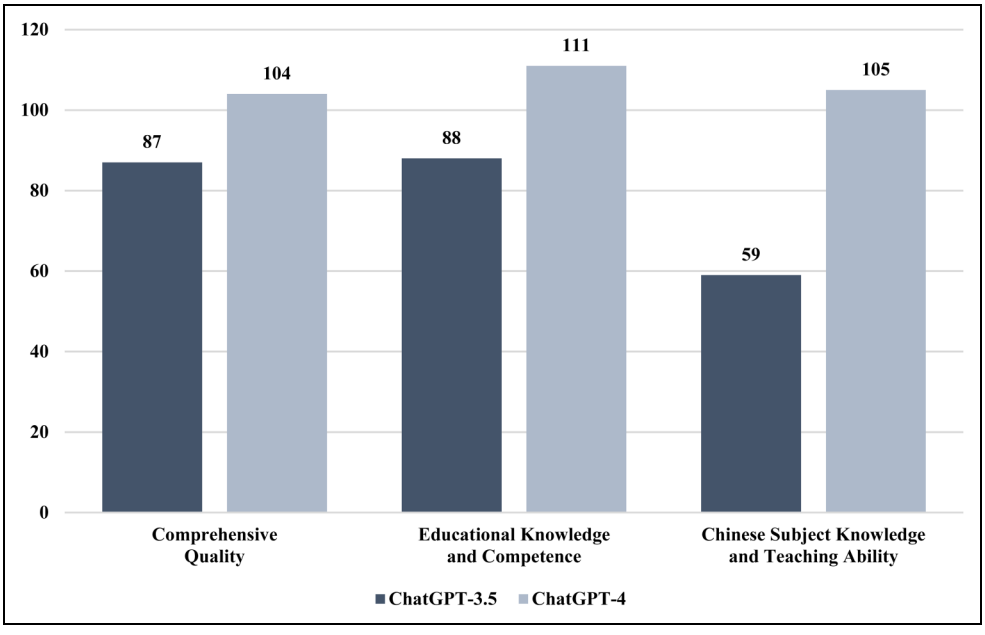

Figure 3. Comparison of ChatGPT-3.5 and ChatGPT-4.0 scores on the 2022 NTCE.

70%. According to the NTCE standards, ChatGPT-4 successfully passed the written test of the 2022 NTCE for high school Chinese teachers.

ChatGPT demonstrated a high degree of comprehensive skill in ancient poetry appreciation, ancient and modern synonym identification, and textual lesson planning. The rich dataset enables ChatGPT to excel in the selection of various knowledge points for identification and analysis. Interestingly, ChatGPT did not perform well when answering objective questions, largely owing to the lack of relevant text or systematic knowledge in its dataset, with no information on the latest curriculum standards.

In the instructional activity design question, ChatGPT took only 20 seconds to create an instructional design for guiding students to write and recite a script for the monologue "Survival or Destruction." The design clearly stated the learning objectives, reasonably allocated the time and flow of each activity, accounted for students' learning conditions and other factors, and introduced a creative recitation method. The design was innovative and included an assessment and feedback mechanism for students' learning outcomes. However, the lesson plan also had several shortcomings, such as relatively thin objective dimensions and the use of traditional evaluation methods.

Regarding long-form text writing, ChatGPT's argumentative essay was well-structured in response to the question type. Nevertheless, its text writing lacked specific arguments and relevance to its viewpoints. As some users have commented, ChatGPT occasionally provides well-written but insubstantial answers. Experts should investigate whether ChatGPT has understood and analyzed the given text.

Therefore, ChatGPT performed relatively well as an AI in core competencies, from the logical to the teaching level. In the test, ChatGPT demonstrated sufficient critical thinking to handle simple logic problems and a tendency to include core competencies in problem solving, collaboration, communication, and metacognition. Results indicated that ChatGPT relies on a rich and extensive text dataset to demonstrate a certain knowledge base for education and teaching. Moreover, the results also indicated that ChatGPT can independently create lesson plans—an essential ability in curriculum development and teaching.

Potential applications of generative AI in education

Overall, ChatGPT demonstrated critical thinking skills, the ability to handle simple logic problems, and a tendency toward core competencies in problem solving, collaboration, communication, and metacognition. Currently, no study has examined the ability of generative AI to pass the written part of the NTCE. This study's results suggest that ChatGPT has significant potential for educational applications, especially in terms of providing opportunities for innovation and instructional support. Passing the examination means that its knowledge and abilities meet the requirements of the examination, indicating that it possesses sufficient capabilities in terms of theoretical

knowledge and practical experience in the field of education. As such, results indicate that ChatGPT is competent in performing educational work and meets the standards of teacher qualification certification. Indeed, passing the examination signifies that ChatGPT understands and has mastered relevant educational policies, teaching methods, and educational principles, as well as the professional ethics and conduct of teachers. Although its educational professionalism and capabilities have been recognized to a certain extent, ChatGPT's excellent performance in the written section of the NTCE demonstrates the potential application of generative AI in educational practice. Based on these results, this study proposes four potential roles for generative AI in Chinese education.

Teacher's assistant for preparing resources

Traditional intelligent educational robots that assist teachers with classroom support or repetitive tasks can also collect and organize information to assist them with lesson preparation. ChatGPT has four potential applications that can assist teachers in preparing educational resources in China. First, ChatGPT can aid in curriculum design and instruction. It can help teachers develop teaching plans, including teaching objectives, learning content, tasks, and assessments. Teachers can use ChatGPT to develop more appropriate learning or assessment tasks for the same topic. Although there is no specific teaching model or criteria for evaluating lesson plans, ChatGPT can generate lesson plans with high reference value. Second, ChatGPT can aid in preparing teaching materials by providing teachers with various teaching materials, such as lesson plans, test papers, assignments, and experimental designs. To obtain personalized materials, teachers need to only provide ChatGPT with relevant topics, qualifications, forms, target audiences, and other information. Third, ChatGPT can help conduct case studies and workshops by providing case studies on various topics and guiding students to explore them comprehensively. ChatGPT can help students understand concepts better and enhance classroom dialogue and communication. Moreover, the software can assume the role of students and present common problems and possible confusion regarding the topic from their perspective. The test experience shows that questions from the students' perspective are relevant for inexperienced teachers and help them enrich their teaching materials and prepare for classroom interactions in advance. Fourth, ChatGPT can provide reading materials and resources to aid students in conducting independent learning and answering inquiries after class per the teacher's requirements.

Student assistant for independent learning

Over the past few decades, intelligent assistants have been increasingly utilized to improve teaching and learning in various ways, notably by providing personalized feedback to students, facilitating collaborative learning, and supporting teachers in analyzing instructional data. ChatGPT can help

students with self-learning in several ways. In this respect, it can become an intelligent assistant to help students with independent learning in learning-specific scenarios (e.g., knowledge answers and concept explanations), learning resources, learning methods and skills guidance, designing practice questions, and self-assessments. Specifically, when students encounter difficult concepts in the learning process, ChatGPT can help them answer questions, provide clear and easy-to-understand explanations to help eliminate their doubts, and recommend suitable learning resources (e.g., textbooks, video tutorials, websites, and papers) according to their needs. For example, suppose you need to understand effective learning methods and techniques during self-study. In this case, ChatGPT can provide a series of suggestions for dialogue methods and strategies. As it can also enable multi-round conversational communication based on the content of historical conversations with users, ChatGPT can provide relevant practice and test questions based on previous learning experiences and topics to enable students to test their learning results. In the near future, generative AI will increasingly specialize in personalized learning accumulated by different individuals over time.

Classroom assistant for enhancing learning interactions

As ChatGPT excels in conversational interaction, its introduction into teaching results in three-way interactions among students, teachers, and technologies. In this respect, ChatGPT can participate in classroom workshops by generating exciting topics and challenging questions that inspire classroom discussion. ChatGPT can also help teachers design interactive quizzes and games to make the learning process more enjoyable. Students can consolidate their knowledge, improve their skills in a relaxed and enjoyable atmosphere, and engage in a series of learning interactions and investigations in real-world scenarios. Moreover, ChatGPT can provide immediate feedback in response to students' questions, promptly alleviating confusion. Finally, through seminars and collaborations, ChatGPT can help students exchange ideas and communicate, improving their teamwork and problem-solving skills.

Automatic checking of extra-curricular assignments

Since the 1960s, scholars have been working on automatic computerized correction techniques for subjective questions based on arbitrary text answers. This has resulted in the development of various automatic correction systems. Pigai.org, for example, can instantly grade students' written English compositions and provide suggestions for improvement; it can also conduct content analysis using NLP and corpus technology. ChatGPT can automate the correction of assignments for all subjects, particularly for multiple-choice, fill-in-the-blanks, and simple computational questions. ChatGPT can automatically provide correct answers, eliminating the need for teachers to predetermine correct options. Based on the given keywords and reference answers, ChatGPT can determine whether the answers to fill-in-the-blank and short-answer questions are correct.

However, ChatGPT may not be able to identify partially correct answers or responses that deviate, thus requiring teachers to verify the assessment results. Moreover, ChatGPT may not provide an entirely accurate assessment of the responses to complex subjective questions, expository questions, or questions requiring judgments of creativity or value choices. These limitations notwithstanding, ChatGPT can be used as an initial assessment tool to help teachers identify potential problems. These responses can be reviewed by teachers to ensure an accurate assessment. To make ChatGPT an effective assistant for the automatic review of assignments, teachers must provide questions and answers so that responses can be accurately assessed. By using ChatGPT, teachers can improve the software's efficiency in correcting assignments and its ability to provide practical support for targeted instruction.

Impact of generative AI on education and its responses

ChatGPT can solve complex problems in various domains. Unlike the inherent sequential development of the First Industrial Revolution (steam engine power) and the Second Industrial Revolution (electric power networks), today's intelligence revolution is based on sound network communication technologies. AI is spreading considerably faster than previous technologies, with the number of generative AI applications expected to increase rapidly in the near future. Within the near future, the number of generative AI applications will expand rapidly. The iterative functions of generative AI have already been significantly updated in education, facilitating various applications and bolstering the tool's ability to contribute to educational change.

Establish a multivariate and collaborative teacher–student–AI relationship

Despite the vast amount of educational resources available on the Internet, each student's personal problems need to be solved promptly and effectively. Although ChatGPT has yet to reach the level of a general-purpose educational robot and completely replace teachers and traditional teaching behaviors, it can answer questions about the subject matter and provide further clarification based on students' follow-up questions. In the field of Chinese language teaching, based on the test results, ChatGPT has reached a level of educational and pedagogical knowledge comparable to that of a high school Chinese language teacher. It can serve as a bridge between teachers and students, acting as a virtual teacher who is available around the clock, ready to answer students' questions and engage in comprehensive discussions. ChatGPT can help address the challenge of the limited time that real teachers may face, becoming an integral part of the multi-dimensional collaborative teacher–student–AI relationship.

Traditional search engines involve launching a search and then reading, judging, selecting, and integrating the output and drawing conclusions. By contrast, when using ChatGPT, students can simply ask questions and receive clear answers. Over time, long-term human–computer dialogue

will strengthen students' trust in and dependence on ChatGPT, potentially resulting in students tending to learn independently of the teacher. This promotes a more equal and open relationship between teachers and students. However, shifting to the teacher–student–computer relationship challenges the emotional relationships between the three groups. Therefore, we must be aware of the potential risks associated with technological misuse, cultural biases, and political misinformation.

Promote human–AI integration to improve teachers' digital literacy

ChatGPT is impacting the teacher–student relationship by altering how students access information. Therefore, we must emphasize the interaction and communication between teachers and students. In this respect, teachers should avoid the misconception of focusing on the class as a whole and ignoring individual student differences in the extensive classroom lecture system that has existed since the Industrial Revolution. Teachers should actively interact and communicate with their students to understand their learning needs and interests and adjust their teaching content and methods accordingly. The critical thinking abilities and collaborative competencies of ChatGPT provide valuable assistance for integrated human–AI education. It can thus be used to help students develop critical thinking skills, engage in collaborative learning, receive personalized support, access diverse perspectives, and experience continuous learning and improvement. Although the application of AI technology can improve teaching efficiency and learning effectiveness in terms of emotional education or focus on students' emotional needs and development, it cannot replace teachers' role in providing more comprehensive education.

Teachers should enhance their educational skills, improve their digital literacy, and make reasonable use of AI tools widely used in educational scenarios. Enhancing teachers' digital literacy based on the concept of "human–computer integration" can help teachers innovate teaching methods, improve teaching effectiveness, and provide students with a more vivid, interesting, and creative learning experience.

Reshape the research structure of the curriculum and learning

The development of general AI technologies is changing future job skill requirements and scenarios. Jobs are at risk of reorganization, innovation, and elimination in the future. To cope with the future challenges, students must be equipped with future-proofing knowledge and literacy skills. The question "What knowledge is most important?" is asked at different stages of education, from primary and secondary school to higher education, as are the questions, "Which subject curriculum content needs to be adjusted?" and "What core literacy should the curriculum and teaching emphasize?" The development of AI technologies has subtly influenced answers to these questions. At the basic level of knowledge recall, humans have almost no advantage over AI. Additionally, testing of the various advanced capabilities of ChatGPT has revealed its rapid iterative

development. Therefore, when setting curriculum goals and building related systems, it is essential to embrace the dramatic shift in learning environments and approaches while emphasizing the cultivation of multiple competencies—including creative thinking, problem solving, teamwork and communication, and practical skills—to help students embrace opportunities and meet the challenges of the digital age.

Address the "intelligence divide" to promote educational equity

Although the Internet has brought considerable benefits to society, it has also created a digital divide. This divide includes unequal access to intelligent technologies and an uneven distribution of technological resources in the education system, posing a significant threat to educational equity in this age of intelligence. To address this issue, we should provide as much equitable access to resources as possible to ensure that every student can enjoy quality intelligence-enabled educational services. Regardless of their location, students can use technologies such as ChatGPT to access high-quality educational resources, interactive and intelligent learning experiences, and personalized intelligent answers to questions. For example, intelligent learning systems provide students with broader learning opportunities through video courses, online discussions, and learning communities. Personalized learning experiences, reduced learning costs, enhanced independent learning capabilities, and improved accessibility to educational resources can narrow the gap in educational resources and promote educational equity.

The release of ChatGPT sparked widespread concern and debate. Some scholars liken ChatGPT to the launch of the first-generation Apple iPhone, which made room for a new application platform and introduced new generative power and production relations. Others believe that the iPhone analogy underestimates the profound effects of this type of generative AI. Generative power surpasses the industry average across sectors. Given the speed of the iterative development of ChatGPT, it is not inconceivable that subsequent iterations may have the ability to reorganize and generate all human information intelligently, including text, sound, video, and datasets.

Although generative AI intelligence affects all aspects of life, it has the most direct impact on education. The essence of education is to nurture, develop, and stimulate an individual's unlimited potential. As individuals can acquire knowledge from the fruits of human labor and incorporate it into their existing knowledge structure, generative AI is redefining the scope of human intelligence, with the intelligence generated becoming part of the user's intelligence. However, the use of ChatGPT raises several concerns. Indeed, although ChatGPT integrates diverse resources and revolutionizes how knowledge is acquired, this does not mean that future students can forgo the acquisition of such skills. Overreliance on ChatGPT for information retrieval and a reduction in autonomous learning time may lead to a decline in students'

information judgment and critical thinking abilities. Furthermore, the information provided by ChatGPT is usually simplistic and may lack the depth and breadth required in certain knowledge domains, potentially resulting in students engaging in surface-level knowledge learning and neglecting experimentation and exploration. Taking the subject of language arts as an example, ChatGPT struggles to understand cultural differences and contexts, and its responses may be influenced by cultural backgrounds and the language environment. ChatGPT cannot fully replace teachers in contexts that emphasize specific cultures, such as traditional Chinese culture.

Therefore, further research is required to use ChatGPT-type generative AI as an application assistant in existing education and teaching. There is an urgent need to reshape the education system. As simple skills can easily be replaced by intelligent technologies, individuals' capabilities, creativity, critical thinking, and collaborative skills have gained the attention they deserve. The more people learn, the more machines they will use, and the more machines will work like intelligent people. The introduction of AI has impacted the field of education. If the education system does not shift to an era of intelligence, it will gradually lose value and meaning. This shift will necessarily involve demystifying the independent possession of human attributes for intelligence and redesigning and developing an intelligent body. Education has entered a new era of exploration as we pursue the coexistence and co-creation of human and computer systems through iterations of generative AI.

Contributorship

Xiaozhe Yang contributed to the research design, conceptual framework, data collection, and analysis and drafted and revised the manuscript. Qingqing Wang implemented the tests for data collection and assisted with data analysis and the drafting of the manuscript. Jiyue Lyu contributed to the revision and finalization of the manuscript and responded to the reviewers' comments.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Department of Social Science Research of the Ministry of Education of the People's Republic of China (Grant number 22JHQ085).

ORCID iD

Jiyue Lyu https://orcid.org/0009-0006-4731-5941

References

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Bernard, R. M., Zhang, D., Abrami, P. C., Sicoly, F., Borokhovski, E., & Surkes, M. A. (2008). Exploring the structure of the Watson-Glaser critical thinking appraisal: One scale or many subscales? *Thinking Skills and Creativity*, *3*(1), 15–22. https://doi.org/10.1016/j.tsc.2007.11.001

Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem-solving, meta-cognitive awareness and creativity. *International Journal of Mobile Learning and Organisation*, *8*(3–4), 276–291. https://doi.org/10.1504/IJMLO.2014.067029

OECD. (2023a). *Is education losing the race with technology? AI's progress in Maths and Reading, educational research and innovation*. OECD Publishing.

OECD. (2023b). *Putting AI to the test: How does the performance of GPT and 15-year-old students in PISA compare?* OECD Publishing.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Papers With Code (preprint). https://paperswithcode.com/paper/improving-language-understanding-by

Vaughan-Wrobel, B. C., O'Sullivan, P., & Smith, L. (1997). Evaluating critical thinking skills of baccalaureate nursing students. *Journal of Nursing Education*, *36*(10), 485–488. https://doi.org/10.3928/0148-4834-19971201-09

Watson, G. (2008). *Watson-Glaser critical thinking appraisal short form manual*. NCS Pearson.

Yang, X. (2019). Accelerated move for AI education in China. *ECNU Review of Education*, *2*(3), 347–352. https://doi.org/10.1177/2096531119878590