Journal of Contemporary Educational Research, 2025, Volume 9, Issue 1
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

New Opportunities, Challenges, and Strategies for Educational Evaluation Reform in the Era of Artificial Intelligence

Li Jin[1]*, Hua Zhang[2]

[1]Chongqing Energy College, Chongqing 402260, China
[2]Chongqing ChinaSoft Excellence Information Technology Co., Ltd., Chongqing 400000, China

Corresponding author: Li Jin, 153102378@qq.com

Abstract: Under the rapid impetus of artificial intelligence (AI) technology, human society is stepping into the age of intelligence at an unprecedented speed. A new generation of information technology such as AI is not only a new engine of economic development, but also a gas pedal of social development, which has had a profound impact on the field of education. In the face of the opportunities and challenges of the AI era, it is particularly urgent to build a scientific and reasonable education evaluation system. This paper combines the context of the times with the new technology of AI to discuss the opportunities, challenges, and implementation strategies of educational evaluation reform in the era of AI, with a view to providing references for the construction of the educational evaluation system and the development of high-quality education in the new era.

Keywords: Artificial intelligence; Education evaluation reform; New policy; Formative evaluation; Value-added evaluation; Multiple evaluation

Online publication: February 10, 2025

1. Introduction

Driven by the wave of artificial intelligence, the field of education is experiencing unprecedented changes. From educational philosophy to talent cultivation goals, from educational content to educational resources and environment, and to educational methods and talent cultivation methods, each link deeply reflects the impact of intelligence. Educational evaluation, as a key link in navigating the course of education, must keep pace with the times and adapt to the needs of the new era. In the face of the opportunities and challenges of the artificial intelligence era, it is particularly urgent to build a scientific and reasonable education evaluation system. In October 2020, the Overall Program for Deepening the Reform of Educational Evaluation in the New Era (hereinafter referred to as the "Overall Program") issued by the Central Committee of the Communist Party of China (CPC) and the State Council, explicitly proposes that we should make full use of information technology

to improve the scientific, professional, and objective nature of educational evaluation [1], encouraging and supporting the intelligence and innovation of evaluation tools, means and methods. This requires us not only to make full use of the new generation of information technology such as artificial intelligence to optimize the evaluation means and enrich the evaluation content, but also to ensure the fairness of the evaluation process and the effectiveness of the results, so as to guide education in the direction of a more personalized, comprehensive, and quality-oriented development. This is not only an important issue in the current education reform, but also a necessary way to realize the high-quality development of education.

2. Opportunities for artificial intelligence-driven educational evaluation reforms

2.1. The enabling role of artificial intelligence technology: Reshaping the precision and efficiency of educational evaluation

The rapid development of artificial intelligence technology, especially the breakthroughs in cutting-edge fields such as big data analysis, machine learning, natural language processing, etc., has brought revolutionary changes to the field of education evaluation. These technologies not only make it possible to collect, process, and analyze huge amounts of educational data, but also realize real-time monitoring and in-depth insight into students' learning behavior, cognitive development, and even psychological changes. By building an intelligent education evaluation system, teachers can grasp the learning dynamics of students in real time, identify problems in a timely manner, and provide personalized feedback and suggestions, effectively improving the accuracy and timeliness of education evaluation.

In addition, artificial intelligence can intelligently push learning resources that meet the needs of students based on their historical learning data and current learning status, realizing the personalized customization of learning paths. This intelligent push mechanism not only meets the personalized learning needs of students, but also promotes the optimal allocation and efficient use of educational resources. Empowered by artificial intelligence technology, education evaluation is no longer just a summation, but a dynamic assessment and guidance throughout the entire learning process, providing strong support for the comprehensive development of each student.

2.2. A profound shift in educational philosophy: From knowledge transfer to core literacy development

The popularization and application of artificial intelligence technology have profoundly affected the reshaping of the concept of education. Traditionally, education focuses on the transmission and accumulation of knowledge, but in the era of artificial intelligence, this concept is gradually shifting to the cultivation of core literacy. Core literacy includes independent learning ability, critical thinking skills, innovation ability, and interpersonal communication skills, which are crucial for students' survival and development in the future society. Educational evaluation, as an important means of testing the effectiveness of education, must keep up with the shift in educational philosophy, expanding from single knowledge evaluation to comprehensive quality evaluation.

Through the support of artificial intelligence technology, educational evaluation can more comprehensively and objectively assess the development of students' core literacy. For example, analyzing students' homework and discussion records using natural language processing technology can assess their critical thinking and expression skills; monitoring students' online learning behavior through data analysis technology can shed light on their independent learning ability and time management skills. This multi-dimensional and all-encompassing

assessment helps to more accurately grasp students' individual differences and developmental needs, providing a more scientific basis for tailoring teaching to students' needs.

2.3. Guidance and support of new policies: Escorting the reform of education evaluation

The introduction of the Overall Program provides clear policy guidance and support for education evaluation reform in the new era. The program emphasizes making full use of information technology means to improve the scientific, professional, and objective nature of education evaluation, and encourages and supports the intelligence and innovation of evaluation tools, means, and methods. This policy background provides a broad development space and policy guarantee for the application of artificial intelligence technology in education evaluation.

Led by the policy, local education departments have increased investment to promote the intelligent upgrade of the education evaluation system. On the one hand, the research and development and application promotion of artificial intelligence technology have been strengthened to enhance the technical content and intelligent level of education evaluation; on the other hand, the relevant policies, regulations, and standard systems have been improved to guarantee the safe, reliable, and effective application of artificial intelligence technology in education evaluation. These initiatives provide a solid institutional guarantee and a favorable development environment for artificial intelligence-driven education evaluation reform.

To summarize, the enabling role of artificial intelligence technology, the profound transformation of educational concepts, and the guidance and support of new policies together constitute three major opportunities for artificial intelligence-driven education evaluation reform. Seizing these opportunities and actively promoting the intelligent upgrading and innovative development of the education evaluation system is of great significance for improving the quality of education and promoting the overall development of students.

3. Challenges to educational evaluation in the age of artificial intelligence

In the current education system, the evaluation method still mainly relies on standardized quantitative evaluation, which is easy to operate and compare, but its limitations are also becoming more and more prominent. Quantitative evaluation overemphasizes test scores, often taking scores as the only or main criterion for measuring students' strengths and weaknesses, thus ignoring students' performance in non-cognitive areas such as moral character, emotional attitude, innovative spirit, practical ability, and so on. This type of evaluation is particularly inadequate in the age of artificial intelligence because it cannot comprehensively and accurately reflect students' comprehensive qualities and abilities.

The introduction of artificial intelligence technology provides the possibility of constructing a diversified and multidimensional evaluation system. Through the comprehensive use of qualitative and quantitative evaluation, combined with big data analysis, machine learning, and other advanced technological means, it is possible to gain a more comprehensive and in-depth understanding of students' learning status, interest preferences, abilities, and specialties, and then form a more scientific and reasonable evaluation conclusion. However, to achieve this goal, it is still necessary for educators and researchers to continue exploring and innovating and to develop a new type of evaluation that is more in line with the needs of the times.

3.1. Traditional and outdated means of evaluation: Limiting the efficiency and depth of evaluation

Despite the rapid progress of science and technology, the means of educational evaluation are still relatively traditional and backward. In many regions and schools, educational evaluation still relies mainly on the

traditional paper-and-pencil test and manual grading. This method is not only inefficient but also susceptible to the influence of human factors, which makes it difficult to ensure the objectivity and accuracy of evaluation results. The introduction of artificial intelligence technology provides strong support for the intelligent upgrading of educational evaluation tools. Through the development of intelligent evaluation tools and systems, real-time monitoring and data analysis of the students' learning process can be realized to improve the efficiency and accuracy of evaluation [2]. At the same time, intelligent evaluation tools can also automatically adjust evaluation strategies and standards according to students' learning behaviors and characteristics to achieve more personalized and precise evaluation. However, to realize this goal, it is necessary to overcome the challenges of technology, funding, talent, and other aspects.

3.2. One-sidedness of evaluation content: Neglecting the assessment of non-cognitive domains

The current content of educational evaluation mostly focuses on the mastery of subject knowledge and skills, while the assessment of non-cognitive areas such as students' emotions, attitudes, and values is relatively insufficient. This one-sidedness of evaluation content not only limits the ability to assess students' overall development but also affects the realization of educational goals. In the era of artificial intelligence, the content of educational assessment needs to be more comprehensive and diversified. In addition to the assessment of disciplinary knowledge and skills, the assessment of non-cognitive domains such as core literacy, innovation ability, teamwork ability, etc. should also be incorporated into the evaluation system. Through the comprehensive use of a variety of evaluation methods and tools, students' comprehensive quality and ability should be comprehensively assessed, so as to provide a more scientific basis for teaching students according to their aptitude. At the same time, it is also necessary to strengthen the assessment of students' learning processes and methods and pay attention to their performance in terms of learning attitudes, methods, and strategies.

3.3. The unitary nature of evaluation standards: Difficult to adapt to the needs of individualized development

Existing education evaluation standards are often based on scores or promotion rates, and this single evaluation standard is difficult to fully reflect students' individual differences and potential. In the era of artificial intelligence, with the diversified development of social demand for talent, the single evaluation standard can no longer meet the needs of personalized education. In order to meet the needs of personalized education, a multi-dimensional evaluation standard system needs to be established. This includes focusing on the application of process evaluation and value-added evaluation to comprehensively assess students' growth progress and development potential. It is also necessary to develop differentiated evaluation standards and methods according to the characteristics of different school years, subjects, and students, so as to achieve the goals of tailor-made education and personalized evaluation. In addition, there is a need to strengthen the publicity and promotion of the assessment standards and raise awareness and acceptance of the multidimensional assessment standards among all sectors of society.

3.4. Limitations of evaluation subjects: Lack of multi-party participation and synergy

The current subject of education evaluation is relatively narrow, mainly teachers and students, lacking the participation of parents, society, and other parties. This limitation affects the objectivity and fairness of evaluation to a certain extent and is not conducive to the formation of educational synergy. In the era of artificial intelligence, it is necessary to build an evaluation system in which multiple evaluation subjects participate

together. This includes strengthening the participation and voice of parents, communities, enterprises, and other social sectors in education evaluation, and forming a synergy in education evaluation. The transparency and credibility of education evaluation should be enhanced through the introduction of a third-party evaluation mechanism and a social supervision mechanism, among other means. At the same time, it is also necessary to strengthen the cultivation and guidance of students' self-evaluation ability, encourage students to actively participate in activities such as self-evaluation and peer evaluation, and improve their sense of subjectivity and self-reflection ability. The implementation of these measures can promote the development of educational evaluation in the direction of greater democratization, scientification, and humanization.

4. Strategies for using artificial intelligence to help educational evaluation system innovation

4.1. Promoting the diversification of evaluation methods to achieve comprehensive evaluation

In the era of artificial intelligence, the traditional single, standardized quantitative evaluation method has been difficult to meet the evaluation needs of students' comprehensive quality. Therefore, promoting the diversification of evaluation methods and realizing the organic combination of process evaluation, formative evaluation, and summative evaluation have become an important direction of educational evaluation reform. By introducing artificial intelligence technology, an intelligent evaluation system can be constructed to monitor students' learning processes and behavioral data in real time. Such a system can record every attempt and interaction of students in the learning process, thus comprehensively reflecting students' learning attitudes, methods, effectiveness, and other dimensions. For example, analyzing students' classroom discussion records using natural language processing technology can assess their critical thinking and oral expression abilities; tracking students' learning paths and interaction behaviors on online platforms through data analysis technology can understand their independent learning abilities and problem-solving strategies [3]. This diversified evaluation method can reveal students' comprehensive quality and ability level more comprehensively and provide richer information support for educational decision-making.

4.2. Developing intelligent evaluation tools to improve evaluation efficiency and accuracy

With the rapid development of big data, cloud computing, and other technologies, the development of intelligent evaluation tools has become an important way to improve the efficiency and accuracy of educational evaluation. These tools can automate a large amount of repetitive work and reduce the burden on teachers, while improving the objectivity and scientificity of evaluation. Specifically, an intelligent paper marking system can be developed, which utilizes technologies such as image recognition and natural language processing to quickly and accurately complete the correction of test papers and improve the efficiency of paper marking. At the same time, a learning situation analysis system can be developed to monitor students' learning status and progress in real time and provide teachers with timely teaching feedback. In addition, an intelligent push system can be developed to recommend personalized learning resources for students according to their learning needs and progress, so as to realize tailor-made teaching. The application of these intelligent assessment tools can not only improve assessment efficiency but also provide students with more personalized learning support.

4.3. Expanding evaluation content and focusing on comprehensive quality evaluation

In the era of artificial intelligence, educational evaluation should focus on the comprehensive evaluation of

students' comprehensive quality. In addition to traditional subject knowledge and skills, non-cognitive factors such as students' emotions, attitudes, and values should be included in the evaluation system. These non-cognitive factors have an important impact on students' lifelong development and social adaptability. The use of artificial intelligence technology to collect and analyze students' multidimensional data, such as learning behavior data, psychological characteristics data, social interaction data, etc., can provide a scientific basis for comprehensive quality evaluation [4]. For example, by analyzing students' learning behavior data, we can understand their learning attitude, self-discipline, and time management ability; through the analysis of psychological characteristics data, we can assess their psychological quality such as emotional stability and stress resistance; through the analysis of social interaction data, we can evaluate their social literacy such as teamwork ability, communication skills, and leadership. The comprehensive analysis of these data can reflect the comprehensive quality and ability level of students.

4.4. Establishing multi-dimensional evaluation standards and focusing on process evaluation and value-added evaluation

The construction of a multi-dimensional evaluation standard system is one of the keys to education evaluation reform. In the era of artificial intelligence, more attention should be paid to process evaluation and value-added evaluation in order to comprehensively and objectively assess students' growth and development. Process evaluation focuses on students' performance and efforts in the learning process, and evaluates their learning attitudes, methods, and effectiveness by recording and analyzing every attempt and every bit of progress in the learning process. This type of assessment helps teachers identify problems, adjust teaching strategies, and provide students with personalized learning suggestions. Value-added assessment focuses on assessing students' growth and progress over a period of time and evaluating their development potential and effectiveness by comparing the changes in their initial and final states. This type of assessment helps encourage students to actively participate in the learning process and personalize their development.

The use of artificial intelligence technology to track students' learning progress and development trajectory enables the precise implementation of process and value-added evaluation. For example, by constructing a student growth file system, every attempt, progress, and feedback of students in the learning process is recorded, forming a complete map of students' development trajectory. At the same time, using data analysis technology to conduct in-depth mining and analysis of students' growth trajectory, we can find out students' growth patterns and development bottlenecks, providing a scientific basis for educational decision-making.

4.5. Promoting the diversification of evaluation subjects and forming a multi-party joint evaluation mechanism

In the era of artificial intelligence, educational evaluation should encourage the participation of parents, communities, third-party organizations, and other parties in the evaluation system. This multi-party co-evaluation mechanism can help improve the objectivity and fairness of evaluation and promote the democratization and scientization of educational evaluation. The sharing and exchange of evaluation information through artificial intelligence technology can break the limitations of closed information and one-way transmission in traditional evaluation. For example, a home-school co-education platform can be constructed to allow parents to learn about students' performance and learning progress at school in real time; a community education evaluation system can be developed to allow community members to participate in evaluating students' social practices and volunteer services, etc.; and third-party organizations can be introduced

to carry out independent evaluations to ensure that the results of the evaluations are objective and fair. The implementation of these measures will help to form an open, inclusive, and pluralistic education evaluation ecosystem [5].

At the same time, attention should be paid to cultivating students' self-evaluation ability. By guiding students to use intelligent evaluation tools for self-reflection and self-evaluation, their self-knowledge and self-improvement can be promoted. For example, a student self-reflection system can be developed to allow students to summarize and reflect on themselves after completing learning tasks; a peer evaluation mechanism can be introduced to allow students to discover their strengths and weaknesses in mutual evaluation. The implementation of these measures will help to develop students' independent learning ability and lifelong learning ability.

5. Conclusion and prospects

The arrival of the era of artificial intelligence provides unprecedented opportunities and challenges for the reform and innovation of the educational evaluation system. With the rapid development of the new generation of information technology, educational evaluation is no longer limited to the traditional standardized quantitative evaluation, but is gradually moving towards diversification, intelligence, and comprehensiveness. By making full use of advanced technologies such as big data, cloud computing, and artificial intelligence, the education evaluation system can more objectively and scientifically reflect students' comprehensive quality and ability. Under the guidance of the Overall Program for Deepening the Reform of Educational Evaluation in the New Era, we need to actively promote the diversification of evaluation methods, focusing not only on outcome evaluation, but also on process evaluation and qualitative evaluation. Through intelligent evaluation systems, we can monitor students' learning process in real time and collect and analyze a large amount of behavioral data, so as to achieve more accurate and personalized evaluation. At the same time, the development of intelligent evaluation tools is also an important direction of educational evaluation innovation. Taking advantage of artificial intelligence technology, more efficient and accurate evaluation tools can be developed to reduce the burden on teachers and improve the efficiency of evaluation. These tools can not only help teachers teach accurately but also provide personalized learning suggestions for students and promote their overall development. In addition, expanding the content of evaluation, establishing multi-dimensional evaluation standards, and promoting the diversification of evaluation subjects are also critical tasks in education evaluation reform. We need to pay attention to the all-round development of students, evaluating not only their mastery of knowledge but also their non-cognitive abilities such as thinking ability, innovation, and teamwork. At the same time, students, parents, teachers, and the community are encouraged to participate in the evaluation process, forming a pluralistic and co-governance education evaluation system.

Looking ahead, with the continuous development and application of artificial intelligence technology, the education evaluation system will be continuously improved and optimized. By building an educational evaluation system that meets the requirements of the new era, we can better cultivate high-quality talents with innovative spirit and practical skills, and provide strong talent support for the country's economic and social development. We believe that with the help of artificial intelligence technology, the education evaluation system will usher in a better tomorrow.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Central Committee of the Communist Party of China, 2020, Overall Program for Deepening the Reform of Educational Evaluation in the New Era, viewed December 10, 2024, https://www.gov.cn/zhengce/2020-10/13/content_5551032.htm

[2] Gao X, 2024, Research on Optimization Strategies of Artificial Intelligence-Enabled Education Evaluation Reform, Master's thesis, Hubei Normal University.

[3] China Big Data Industry Ecological Alliance, 2021, China Big Data Industry Development White Paper, viewed December 10, 2024, https://www.smartcity.team/reports/2021%E4%B8%AD%E5%9B%BD%E5%A4%A7%E6%95%B0%E6%8D%AE%E4%BA%A7%E4%B8%9A%E5%8F%91%E5%B1%95%E7%99%BD%E7%9A%AE%E4%B9%A6/

[4] Xu J, Shen X, 2023, Reshaping the Learner-Centered Education Evaluation Ecology—A Global Observation Based on the Intelligent Development of Education Evaluation. Open Education Research, 29(03): 40–46.

[5] Long H, 2021, Educational Assessment Reform in AI Era: Opportunities, Challenges and Path Selection. Journal of China Examinations, (11): 10–18 + 34.