A Study on Multimodal AI-Powered Outcome-Based Teaching Assessment System

Chaomin Gao[1,2], Fang Liu[1,*]
[1]*School of Business Administration, Baise University, Baise, Guangxi, China*
[2]*The Revitalization and Development of Old Revolutionary Areas in Guangxi, Baise, Guangxi, China*
**Corresponding Author*

Abstract: To address the core problems existing in the current Outcome-Based Education (OBE) teaching assessment in universities, such as simplistic assessment dimensions, delayed feedback, and ambiguous goal achievement levels, this study proposes a dynamic teaching assessment system driven by multimodal AI. By integrating multimodal data, the system establishes a dynamic mapping mechanism between teaching objectives and assessment content, realizing a closed-loop process from goal input to feedback improvement. Specifically, the system employs semantic analysis technology to decompose vague and abstract curriculum objectives into measurable, implementable, and evaluable behavioral indicators, and generates an intelligent question bank covering different cognitive levels based on the principle of "reverse design and forward implementation". Meanwhile, through a human-machine collaborative evaluation mechanism, it combines AI's automatic scoring of objective questions and structured answers with teachers' review and arbitration of open-ended answers, thereby improving assessment efficiency and reliability. Finally, an implementation path is proposed, starting from pilot applications in standardized courses, gradually expanding to practical courses, and ultimately establishing a university-wide teaching objective database, so as to provide infrastructure support for the digital transformation of education.

Keywords: Multimodal Artificial Intelligence; OBE Teaching Assessment; Human-Machine Collaborative Evaluation; Digital Transformation of Education

1. Introduction

At present, the OBE teaching assessment in universities generally faces three major challenges. Firstly, the assessment dimensions are monotonous. The first aspect is the simplification of assessment dimensions. Traditional student evaluations heavily rely on written tests and standardized tests, with the assessment content mainly focusing on the memorization of theoretical knowledge and the solution of fixed question types. The evaluation methods are rather rigid. Although this model is convenient for unified scoring and organization, it is difficult to comprehensively measure students' comprehensive abilities. In particular, it is unable to effectively assess high-level qualities such as practical operation, teamwork, innovative thinking, and problem-solving. This often leads to one-sided evaluation results, which is not conducive to the all-round development of students [1,2]. Secondly, the teaching feedback is lagging. The traditional teaching evaluation system has a significant problem of feedback lag. Its assessment mechanism relying on the single evaluation method of the final exam leads to the lack of timely and effective remedial measures even when it is found that students' academic performance is poor. This lag is contrary to the educational philosophy of "centering on student development" and is difficult to achieve the fundamental goal of comprehensively improving the quality of talent cultivation [3,4]. Finally, the degree of goal achievement is ambiguous. Although the OBE concept always emphasizes the high consistency between teaching objectives and assessment evaluations, in actual operation, the assessment indicators designed and formulated by humans often deviate from or even disconnect from the preset goals of the course, making it difficult to precisely correspond to various ability requirements. This makes it impossible for teaching, learning and evaluation to effectively connect, and the complete closed loop of "teaching - learning - evaluation" cannot be truly formed. As a result, it is impossible to objectively determine whether students have achieved the expected learning outcomes [5,6].

The rapid development of artificial intelligence technology has provided new opportunities for the reform of educational evaluation. The collection and analysis of multimodal data (text, images, charts, sounds, etc.) make it possible to conduct comprehensive and dynamic assessment of students' abilities [7]. For instance, research based on multimodal learning analysis indicates that through a cross-modal, cross-spatial, cross-data, and cross-analysis system, learning experiences can be optimized more precisely [8]. Moreover, the application of artificial intelligence technology not only enhances the efficiency and accuracy of educational evaluation but also enables the implementation of the entire-process evaluation of students' abilities through intelligent assessment systems [9]. For example, artificial intelligence can process high-dimensional and massive educational data, clearly describe the knowledge map that students have mastered, and provide personalized teaching suggestions for teachers [10,11].

This research aims to design and validate a multimodal-driven OBE teaching assessment system with three core assessment objectives. First, establish a dynamic mapping mechanism between teaching objectives and assessment contents to ensure that the assessment indicators comprehensively cover the course objectives. Second, enhance assessment efficiency through human-computer collaborative evaluation, liberating teachers from mechanical tasks and focusing on the optimization of teaching design. Third, achieve real-time feedback and a closed loop for teaching improvement, facilitating personalized learning for students and precise teaching for teachers [12]. Through the implementation of this system, it can effectively solve the problems existing in traditional OBE teaching assessment, promoting the development of educational evaluation reform towards a more efficient and precise direction.

2. Literature Review

The concept of OBE (Outcome-Based Education) emphasizes that the final learning results of students are the starting point and end point of teaching design [13]. Its core concept is mainly reflected in two aspects. On the one hand, based on the principle of reverse design, from the perspective of social needs and training objectives, it is disassembled layer by layer into curriculum objectives, teaching activities and assessment standards [14]. On the other hand, a continuous improvement mechanism is established to diagnose teaching problems through assessment data and dynamically adjust teaching content [15]. However, in the implementation of traditional OBE, the assessment indicators often rely on teachers' experience and lack quantitative basis. For example, the expression of "cultivating critical thinking" in the course objective is difficult to be directly translated into measurable assessment questions [16].

In recent years, artificial intelligence technology has gradually permeated the field of educational assessment. For instance, automatic grading systems have achieved rapid scoring for objective questions such as multiple-choice and fill-in-the-blank questions, but open-ended answers , such as essay questions, still rely on manual evaluation [17]; learning behavior analysis predicts learning outcomes through data from online learning platforms, such as click streams and video viewing durations, but single-modal data (such as pure text) is difficult to fully reflect abilities [18]; personalized feedback tools push learning resources based on error records, but the resource matching accuracy is limited by the coverage of the preset knowledge base [12].

At present, existing research has two major limitations: first, the static association between goals and assessment, with most systems merely attaching teaching goals as labels to the question bank, lacking a dynamic adjustment mechanism [19]; second, the absence of human-machine collaboration mechanism, excessive reliance on machine scoring may lead to evaluation biases, but no rules for the division of labor and collaboration between teachers and AI have been established [20].

In the context of artificial intelligence empowering educational evaluation, scholars have proposed various solutions. For instance, the artificial intelligence-based classroom evaluation framework constructs four application scenarios - classroom language analysis, classroom behavior analysis, classroom emotion analysis, and classroom evaluation system - through a systematic architecture consisting of four layers: the object layer, data layer, technology layer, and application layer [21]. Moreover, the application of artificial intelligence in educational evaluation also involves the construction of the learning occurrence knowledge-action model. This model, by correlating knowledge-action with observable language stimuli data, makes the interaction between internal cognitive activities and the external teaching environment an object that artificial intelligence can analyze [22].

However, the application of artificial intelligence in educational evaluation also faces many challenges. For example, the application of educational big data may ignore human development, resulting in one-sided evaluation results [23], and the "preference", "discrimination" and "preset" attributes of artificial intelligence technology may affect the development of human-computer collaboration [24]. Therefore, future research needs to conduct in-depth exploration in the following three aspects. The first is to improve the theoretical basis of artificial intelligence enabling education evaluation and strengthen the leading role of new concepts of education evaluation [25]. The second is to expand the function of artificial intelligence education assessment and promote the development of artificial intelligence assessment technology [26]. Third, consolidate the supporting system of artificial intelligence education evaluation and strengthen the construction of intelligent education evaluation infrastructure.

The goal of this study is to build a dynamic assessment system driven by teaching objectives and improve the evaluation reliability through human-machine collaboration. This goal not only conforms to the core requirements of the OBE concept, but also conforms to the latest development trend of artificial intelligence technology in the field of education evaluation. The research focuses on how to deeply integrate artificial intelligence technology and OBE concept, realize the dynamic adjustment of teaching objectives and assessment standards, and how to improve the scientificity and fairness of educational evaluation through human-machine collaboration mechanism.

3. System Design

3.1 Core Logic of the System

This system starts from the teaching objectives and builds a closed-loop process of "target input → assessment generation → learning support → intelligent assessment → feedback improvement". First, teachers upload the teaching syllabus that links the course objectives with the graduation requirements. The system automatically analyzes the key words of the course objectives and converts them into quantifiable assessment indicators, and matches them with the corresponding question bank. For example, "mastering market research methods" can correspond to a practical question like "design a consumer questionnaire". Subsequently, students conduct targeted review under the guidance of artificial intelligence, and the system dynamically pushes learning resources based on the analysis of incorrect question records (such as automatically generating explanatory videos for the knowledge points with high error rates). The assessment stage is conducted in an online format, where the system automatically grades objective questions and structured answers, and teachers mainly review open-ended answers. Both parties collaborate to complete the scoring confirmation. Finally, the system generates a course report containing data such as the overall students' achievement rate of the objectives and the distribution of frequently missed knowledge points, and synchronizes it to the educational administration system, forming a complete "teaching - learning - assessment" data chain [27].

3.2 Key Functional Components

The core functions of the system revolve around the deep integration of teaching objectives and multimodal data. The teaching objective analysis module, through semantic analysis, breaks down vague course objectives (such as "cultivating critical thinking") into specific behavioral indicators (such as "able to identify logical flaws in arguments") and establishes association rules with question types. The intelligent question bank generation module, based on the subject content in the knowledge base and historical test data, automatically generates questions covering different cognitive levels (memory, application, innovation), and teachers can adjust the proportion of questions to ensure balanced assessment. In the learning guidance phase, students communicate with the AI through voice questions, image-text interaction, etc. For example, in a programming course, after students upload their code, the system not only points out grammatical errors but also demonstrates the correct debugging steps through video. In the assessment stage, a human-machine collaborative mechanism is adopted: objective questions are scored directly by the system, while subjective questions (such as experimental reports) are initially scored through keyword matching and structural analysis, and teachers review disputed answers. If the difference between human and machine scores exceeds 15%, the system automatically retrieves the student's answer process data (such as the trajectory of answering time) to assist in arbitration. The final teaching improvement report is presented in visual charts, such as using a heat map to compare the target achievement rates of different classes, providing teachers with a basis for adjusting teaching strategies [18].

4. Conclusions and Suggestions

4.1 Conclusions

The multimodal OBE teaching assessment system constructed in this study can achieve the seamless integration of teaching objectives, assessment implementation, and teaching improvement through artificial intelligence technology. The system is centered on the "outcomes-based" concept of OBE and transforms the traditional static assessment based on teachers' experience into data-driven dynamic evaluation. The system effectively resolves three major contradictions. Firstly, by automatically correlating teaching objectives with assessment questions, it solves the problem of "examining what is not learned". Secondly, it uses multi-modal data (such as voice questions and operation videos) to achieve comprehensive ability assessment, filling the gap in traditional written tests. Thirdly, through real-time feedback and visual reports, it shortens the response cycle of "assessment - improvement", enabling teachers to quickly identify teaching weak points.

The three major pain points of traditional OBE assessment can be solved through technology empowerment. For teachers, the system reduces the time spent on mechanical tasks such as question generation, paper assembly, and grading. For students, the real-time feedback mechanism shortens the response time of assessment results, and students can precisely review the weak knowledge points pushed by the system. At the teaching management level, the cross-course data analysis dashboard generated by the system can help managers quickly identify the shortcomings in the training plan. In addition, the system can seamlessly connect with the educational administration system through API, avoiding duplicate data entry, and providing infrastructure for the long-term tracking and quality certification of educational big data.

4.2 Practical Suggestions

To promote the implementation of the system, it is suggested that it be carried out in three stages.

The first stage is the initial stage: conduct pilot tests in public foundational courses with high standardization (such as university English and computer basics), and accumulate question banks and multi-modal evaluation rules.

The second stage is the expansion stage: extend to practical courses (such as experimental operations and graduation projects), and develop specialized assessment modules for video action recognition, code logic analysis, etc.

The final stage is the deepening stage: establish a university-wide teaching objective database, and through cross-course data mining, identify systematic deficiencies in the training plan (such as the widespread lack of data analysis skills among students in a certain major).

At the same time, it is necessary to strengthen teacher training to help them understand the system logic (such as the method of breaking down teaching objectives), and avoid the occurrence of "technological dependence". For example, in the evaluation of open-ended answers, teachers still need to lead the scoring of the creative dimension to prevent excessive intervention by machines.

4.3 Future Outlook

The current limitations of the system mainly lie in two aspects. Firstly, the processing of multimodal data requires high performance of campus networks and computing devices, and some remote institutions may encounter hardware bottlenecks. Secondly, the student privacy protection mechanism still needs to be improved. For example, the storage and use of experimental operation videos need to have clear ethical norms. In the future, lightweight model deployment solutions can be explored, and federated learning technology can be introduced to achieve cross-university experience sharing while protecting data privacy. Moreover, with the continuous evolution of educational large models, the system is expected to further expand its application scenarios. For instance, by simulating virtual working environments (such as e-commerce platforms), it can assess students' professional qualities and truly achieve the connection between "learning" and "employment".

Acknowledgments

This paper is funded by the Key Research Base of Humanities and Social Sciences of Universities in Guangxi Zhuang Autonomous Region "The Research Base for the Revitalization and Development of Old Revolutionary Areas in Guangxi", and also supported by the 2025 Guangxi Higher Education Undergraduate Teaching Reform Project "Research on the Construction of a Closed Loop for OBE Course Evaluation Driven by Artificial Intelligence" (2025JGA347).

References

[1] Wang Huiya. Structuring the Multi-dimensional Teaching System of Environmental Impact Assessment Based on OBE Concept. China Modern Educational Equipment, 2021(21): 107-110.
[2] Chen Jiajun, Zhao Jiangtao. Exploration of integrating virtual and real welding formation practice teaching under the Outcome-Based Education concept. MW Metal Forming, 2026, (03): 32-37.
[3] Zhang Zhimei, Cheng Liying, Li Liu, et al. Reform of "Signals and Systems" curriculum based on OBE education concept. Journal of Shenyang Normal University (Natural Science Edition), 2019, 37(3): 284-288.
[4] Wang Qiufen, Wang Yongxin. Teaching Reform of Operating System Principles Based on OBE. Journal of Higher Education, 2018(22): 124-126.
[5] Zhao Cunyou, Xu Peng, Chen Guojing. Discussion on the Design of Student Performance Assessment System Based on the OBE Concept. Heilongjiang Education (Theory & Practice), 2020(10): 66-67.
[6] Guo Jiawen, Zhou Xiaojie, Li Xinheng, et al. Implementation and Continuous Improvement of the OBE-Based Teaching Model. China Metallurgical Education, 2026, (01): 54-59.
[7] Chen X. Multimodal learning analytics in education: A systematic review. Journal of Educational Technology & Society, 2023, 26(1): 1-13.
[8] Mu Zhijia. Multimodal Learning Analytics: New Growth Points of Learning Analytics Studies. e-Education Research, 2020(5): 27-32, 51.
[9] Lu Xingnan, Gao Xuewei. On the AI Enabling Education Evaluation Reform: Development Trend, Risk Inspection and Countermeasures. Journal of the Chinese Society of Education, 2023(2): 48-54.
[10]Guo Zinan, Wang Wanding, Liu Jia. Artificial Intelligence Facilitating Educational Evaluation Reform. AI-View, 2022(2): 126-132.
[11]Zou Qin, Ma Xiaoyan, Sun Lichao, et al. Construction and teaching practice of an AI- and OBE-integrated smart course Genetic Engineering. Chinese Journal of Biotechnology, 2026, 42(03): 1413-1423.
[12]Hwang G J. Vision, challenges, roles and research issues of Artificial Intelligence in education. Computers & Education: Artificial Intelligence, 2020, 1: 100001.
[13]Spady W G. Outcome-Based Education: Critical Issues and Answers. Arlington, VA: American Association of School Administrators, 1994.
[14]Zeng Fanju, Tan Yongqian. Teaching Reform of Digital Electronic Technology Based on OBE Teaching Mode. University, 2021(39): 122-124.
[15]Li Juan Liang. Research on the Reform of Electronic Testing and Experiment Courses Based on the OBE Concept. China High and New Technology, 2021(3): 143-144.
[16]Killen R. Outcomes-based education: Principles and possibilities. Unpublished manuscript, University of Newcastle, faculty of education, 2000: 1-24.
[17]Zawacki-Richter O. Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 2019, 16(1): 1-27.
[18]Chen X. Multimodal learning analytics in education: A systematic review. Journal of Educational Technology & Society, 2023, 26(1): 1-13.
[19]Li L. AI in education: A review of the literature. Journal of Educational Technology Development and Exchange, 2021, 14(1): 1-16.
[20]Williamson B. The social life of AI in

education. Postdigital Science and Education, 2023, 5(1): 1-20.

[21]Wu Libao, Cao Yanan, Cao Yiming. Reform and Practical Paths of Classroom Teaching Evaluation under Artificial Intelligence. China Educational Technology, 2021(5): 94-101.

[22]Wang Jianhua, Xu Shan, Wu Jinling. Exploration of Online Open Course Construction Based on the OBE Model in the Context of Artificial Intelligence - Taking the "Nutrition Diet Design" Course as an Example. University, 2025, (32): 15-18.

[23]Mao Gang, Zhou Yueliang, He Wentao. Development Trend of Teaching Evaluation Theory under Background of Educational Big Data. e-Education Research, 2020, 41(10): 22-28.

[24]Chen Qianqian, Zhang Lixin. Ethical Thinking on Educational Artificial Intelligence: Phenomenon Analysis and Vision Construction: Based on the Analysis Perspective of “Human-Machine Collaboration. Journal of Distance Education, 2023, 41(3): 104-112.

[25]Li Xiaoyan, Ma Hanbo, Wang Mingchun, et al. Innovative Exploration of Integrating Artificial Intelligence into the OBE Practical Teaching Model. Modern Food, 2025, (19): 121-123.

[26]Biggs J, Tang C. Teaching for quality learning at university. Maidenhead: McGraw-Hill Education, 2011.

[27]Shao Sicheng, Hou Jifei, Tang Jun, et al. New Paths for Enabling the "Clean Production" Course's Ideological and Political Education through Artificial Intelligence in the New Era. Journal of Wuxi University, 2026, 42(01): 53-59.

