Evaluation of Learning Efficiency in MOOC Teaching Systems Based on Kirkpatrick's Model and AHP-fuzzy Comprehensive Evaluation Method

Yangbo Wu [1,*], Ng Giap Weng [1], Beng Yu Leau [1], and Ying Lin [2]

[1] Faculty of Computing and Informatics, University Malaysia Sabah, Sabah, Malaysia
[2] Faculty of Psychology and Education, University Malaysia Sabah, Sabah, Malaysia
Email: wuyangbo@126.com (Y.W.); nggiapweng@ums.edu.my (N.G.W.); lybeng@ums.edu.my (B.Y.L.); linying0813@126.com (Y.L.)
*Corresponding author

Manuscript received March 13, 2026; revised April 9, 2026; accepted May 20, 2026; published August 21, 2026

***Abstract*—Conventional single-dimensional scoring methods for Massive Open Online Course (MOOC) teaching systems often fail to capture holistic learning efficiency and suffer from low evaluation accuracy. To address these limitations, this paper proposes a novel learning efficiency evaluation method integrating Kirkpatrick's four-level model with the Analytic Hierarchy Process (AHP) and fuzzy comprehensive evaluation. First, a multidimensional indicator system is constructed across Reaction, Learning, Behavior, and Results levels. The AHP is then employed to calculate the weight vectors of each indicator, ensuring rational weight distribution through consistency testing. Subsequently, a fuzzy relation matrix is established based on membership degrees derived from expert assessments. By synthesizing the AHP-derived weights with the fuzzy matrix via weighted average operators, the method defuzzifies the results to yield a comprehensive efficiency score. Experimental validation involving 200 learners demonstrates that the proposed method outperforms existing approaches, achieving a Root Mean Squared Error (RMSE) of 0.45, Mean Absolute Error (MAE) of 0.35, and an accuracy rate of 89.5%. The results indicate that this method maintains stable, high-confidence evaluation performance across all levels, effectively handling the fuzziness and uncertainty inherent in learning efficiency assessment, thereby providing a reliable basis for optimizing MOOC teaching systems.**

***Keywords*—Kirkpatrick's model, Analytic Hierarchy Process (AHP) method, fuzzy comprehensive evaluation method, Massive Open Online Course (MOOC) teaching system, learning efficiency evaluation**

I. Introduction

Online education is currently experiencing rapid growth, with Massive Open Online Courses (MOOCs) playing a vital role in the educational landscape. Scientifically and effectively evaluating the learning efficiency of MOOC teaching systems is crucial for enhancing teaching quality and effectiveness, yet existing research still has limitations in this assessment.

In related research fields, Huaren and Chunyan [1] evaluated the combat effectiveness of composite brigade equipment using a combined grey entropy weighting method. However, this approach has limited adaptability for handling complex problems involving fuzziness and uncertainty. Hossain *et al.* [2] employed machine learning algorithms to evaluate urban resource efficiency and management. However, this approach struggles to comprehensively cover multi-level soft indicators, compromising reliability. Luki [3] applied the Multi-Attributive Border Approximation Area Comparison (MABAC) method to assess Serbian bank efficiency. Yet, its heavy reliance on pre-set weights results in low reliability. Zbaltan [4] employed machine learning on audit data to enhance fraud detection and risk assessment capabilities. However, its limited adaptability to complex problems compromised reliability. Tong and Zhan [5] analyzed the online learning behavior data of MOOC learners and established an evaluation model using multiple linear regression, multilayer perceptron, and classification regression tree to assess the learning performance of MOOCs. However, the evaluation accuracy of this method is limited when applied to uncertain learner features.

To address these methodological limitations, this paper proposes a MOOC teaching system learning efficiency evaluation method based on the Kirkpatrick's Model and Analytic Hierarchy Process (AHP) fuzzy comprehensive evaluation. The current mainstream methods mostly focus on learning behavior data analysis or single satisfaction surveys, making it difficult to make an overall evaluation of complex ability transformation. This method achieves multi-level and coherent inference from observable response and learning data, internalized knowledge and skills, and potential behavioral intentions and long-term outcomes. This method can handle the incompleteness and fuzziness of evaluation information, unify scattered, accurate, or fuzzy feedback into a comprehensive fuzzy diagnosis of overall teaching effectiveness, and output evaluation results with confidence information. This transcendence is reflected in the shift from discrete indicator measurement to comprehensive ability diagnosis, and from static result reporting to dynamic teaching insights. This method is based on the Kirkpatrick four level evaluation model and constructs a multidimensional ability evaluation system, providing a unified analytical framework for the evaluation of dispersed and usually single level learning outcomes in MOOC environments. At the level of method integration, this article proposes a phased and nested hybrid method flow. In response to the complex interrelationships among internal indicators at different levels of the Kirkpatrick model, the AHP method is used to determine the relative weights of each indicator, thereby solving the problem of simple weighting in traditional evaluations; In response to the inherent subjectivity and fuzziness of interpreting ability assessment, this article deeply integrates fuzzy comprehensive evaluation into each evaluation level, and applies the weight vectors determined by AHP to fuzzy comprehensive operations, achieving structured coupling between quantitative data, qualitative judgment, and hierarchical models.

The research structure of this article is as follows: Firstly, based on the Kirkpatrick model, a four-dimensional index evaluation system is constructed; secondly, a hierarchical structure model of indicators was established using the AHP method, and the weight vectors of each indicator were calculated; finally, calculate the membership degree and construct a fuzzy relationship matrix; Implementing deblurring processing through weight vector synthesis operation to achieve fuzzy comprehensive evaluation of MOOC teaching system learning efficiency. Evaluate the effectiveness of the research method in this article through experimental testing.

II. Evaluation Method for Learning Efficiency of MOOC Teaching System

A. Evaluation Index System for Learning Efficiency Based on Kirkpatrick Model

To establish a scientifically sound and comprehensive learning efficiency evaluation system tailored to the characteristics of MOOC teaching systems, enabling a thorough and objective assessment of learner outcomes, this paper constructs an evaluation indicator system based on the Kirkpatrick's Model.

The Kirkpatrick's Model comprises four progressive levels—Reaction, Learning, Behavior, and Results—enabling a comprehensive and systematic evaluation of training effectiveness. Drawing upon the principles of the Kirkpatrick's Model, this paper constructs a learning efficiency evaluation indicator system for MOOC teaching systems across these four levels.

Response-level indicators focus on learner satisfaction with the MOOC teaching system, covering course design, instructor competence, and platform functionality. Course design rationality depends on the alignment of course content with learner needs and cognitive levels, as well as the system's logical coherence. Instructor teaching proficiency is evaluated based on their subject expertise, pedagogical skills, and communication abilities—assessing whether knowledge is conveyed clearly and engagingly. Platform usability examines whether the interface and operational processes are simple and intuitive.

Learning-level indicators assess mastery of knowledge, skills, and attitudes. Knowledge acquisition is evaluated through tests and assignments; skill enhancement is measured by practical skill proficiency; and learning attitude shifts are observed by tracking pre- and post-learning attitude changes and initiative.

Behavior-level indicators evaluate changes in learners' knowledge application behaviors, covering knowledge application frequency, problem-solving ability, and teamwork skills.

Outcome-level indicators evaluate the impact of the MOOC teaching system on individual learners and organizational business results. This includes personal performance improvement, comparing work performance before and after learning to assess the system's facilitating effect; and organizational benefit growth, analyzing the system's influence on overall organizational benefits such as sales revenue and cost changes.

To more clearly illustrate the MOOC teaching system's learning efficiency evaluation indicator framework, this paper presents an indicator framework diagram, as shown in Fig. 1.

Additionally, indicators across all levels are organized into Table 1.

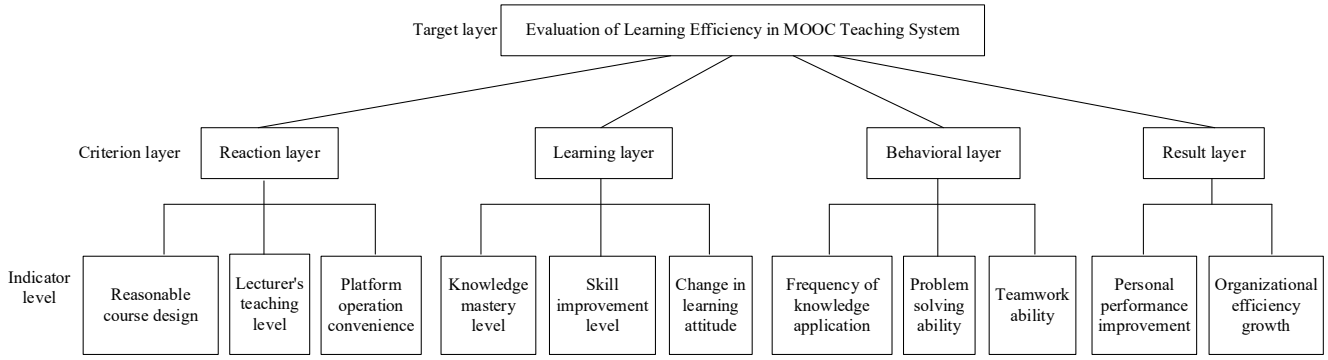

Fig. 1. MOOC teaching system learning efficiency evaluation indicator framework diagram.

Table 1. MOOC teaching system learning efficiency evaluation indicator system table

Level	Indicator name	Indicator description
Reaction layer	Reasonable course design	Does the course content meet the needs and cognitive level of learners, and is it systematic and logical
	Lecturer's teaching level	The lecturer's professional knowledge, teaching skills, and communication abilities
	Platform operation convenience	Is the interface design and operation process of the MOOC platform simple and easy to understand
Learning layer	Knowledge mastery level	Learners' understanding and memory of course knowledge
	Skill improvement level	Have learners mastered the relevant practical skills
	Change in learning attitude	Changes in learners' learning attitudes before and after learning
Behavioral layer	Frequency of knowledge application	The number of times learners apply their learned knowledge in practical work or life
	Problem solving ability	The ability of learners to apply their learned knowledge to solve practical problems
	Teamwork ability	Learners' ability to collaborate with others in team projects
Result layer	Personal performance improvement	Comparison of learners' job performance before and after learning
	Organizational efficiency growth	The impact of MOOC teaching system on overall organizational efficiency

Together, these indicators constitute the MOOC teaching system learning efficiency evaluation indicator system

presented in this paper.

B. Weight of Evaluation Indicators Based on AHP Method

To accurately reflect the relative importance of different evaluation indicators in assessing learning efficiency within MOOC teaching systems, thereby enhancing the persuasiveness and credibility of evaluation results, this paper employs the AHP method to determine the weighting of learning efficiency evaluation indicators.

The AHP method is a systematic analysis approach combining qualitative and quantitative elements. It decomposes complex decision-making problems into multiple hierarchical levels, determines the relative importance of factors through pairwise comparisons, and subsequently calculates the weights of each factor [6, 7]. This paper employs the AHP method to determine the weights of the MOOC teaching system learning efficiency evaluation indicators. Based on the previously constructed indicator system, a hierarchical structure model is established. The objective level is "MOOC Teaching System Learning Efficiency Evaluation". The criterion level consists of four hierarchical levels (Response Level, Learning Level, Behavior Level, and Outcome Level), while the indicator level comprises specific indicators corresponding to each level. Education experts were invited to conduct pairwise comparisons of indicators across levels to determine their relative importance to the criterion level above [8, 9]. Comparisons used a 1–9 scale as shown in Table 2.

Table 2. Meaning of the 1–9 importance scale

Serial number	Scale value	Meaning
(1)	1	Indicates that two factors are equally important
(2)	3	Indicating that one factor is slightly more important than another
(3)	5	Indicating that one factor is significantly more important than another
(4)	7	Indicating that one factor is stronger and more important than another
(5)	9	Indicating that one factor is extremely more important than another
(6)	2, 4, 6, 8	Indicate the median value of the adjacent judgments mentioned above

Based on the expert comparison results, a pairwise comparison matrix was constructed. For example, for the four factors at the criterion level, the pairwise comparison matrix A is constructed as follows:

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & a_{14} \\ a_{21} & 1 & a_{23} & a_{24} \\ a_{31} & a_{32} & 1 & a_{34} \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} \quad (1)$$

In Eq. (1), a_{ij} represents the importance of the ith factor relative to the jth factor in evaluating the learning efficiency of MOOC teaching systems.

Next, this paper employs the eigenvalue method to compute the weight vector. Solving the eigenvalue equation $M_e - \lambda_l = 0$ of the pairwise comparison matrix M_e yields the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Identifying the maximum eigenvalue M_{max}, solving the equation $(M_e - \lambda_{max}())$ yields the eigenvector $X = (x_1, x_2, \cdots, x_n)^T$. After normalization, the weight vector w_i is obtained:

$$w_i = \frac{x_i \|A\|}{\sum_{i=1}^{n} x_i}, (i = 1,2, \cdots, n) \quad (2)$$

In Eq. (2), x_i represents the weights of each factor in the learning efficiency evaluation of the MOOC teaching system.

The elements w_i in the weight vector represent the weights of each factor. To ensure consistency of the pairwise comparison matrix, a consistency test is performed in this paper. The consistency index C_I is calculated:

$$C_I = \frac{\lambda_{max}}{n-1} \quad (3)$$

In Eq. (3), C_I denotes the consistency index of the pairwise comparison matrix in the learning efficiency assessment of MOOC teaching systems.

Next, calculate the consistency ratio CR:

$$CR = \frac{C_I w_i}{\frac{1}{n}\sum_{i=1}^{n} \frac{1}{i}} \quad (4)$$

In Eq. (4), CR denotes the consistency ratio of the pairwise comparison matrix in MOOC teaching system learning efficiency evaluation.

When $CR < 0.1$, the pairwise comparison matrix is deemed to have satisfactory consistency, and the weight vector w_i can be accepted; otherwise, the pairwise comparison matrix requires adjustment, followed by recalculation of the weight vector [10, 11].

After calculating the weights for each factor in the criterion layer through the above steps, the composite weights for each factor in the indicator layer relative to the target layer are then computed. The composite weight can be expressed by the following Eq. (5) [12, 13]:

$$w_i^* = w_j \times \left(CR \sum_{i=1}^{n} w_{kj} \cdot w_{ij}\right) \quad (5)$$

In Eq. (5), w_i^* denotes the composite weight of the ith indicator in the learning efficiency assessment of the MOOC teaching system, w_j represents the weight of the jth criterion layer factor, and w_{kj} indicates the weight of the kth indicator relative to the jth criterion layer factor.

By applying the AHP method to determine the weights of each evaluation indicator, an objective weighting basis is provided for subsequent fuzzy comprehensive evaluation.

C. Fuzzy Comprehensive Evaluation Method for Learning Efficiency of MOOC Teaching System

Since the learning efficiency of MOOC teaching systems is influenced by multiple factors, it exhibits fuzziness and uncertainty. Traditional evaluation methods struggle to effectively address these characteristics [14–16]. Therefore, this paper employs a fuzzy comprehensive evaluation approach to assess the learning efficiency of MOOC teaching systems, thereby more accurately reflecting the true state of learning efficiency.

The factor set U_e comprises all indicators from the previously constructed indicator system, i.e., $U_e = \{u_1, u_2, \cdots, u_n\}$, where u_i denotes the ith indicator in the MOOC teaching system learning efficiency assessment [17, 18]. The comment set V_e consists of various evaluation results provided by assessors for the evaluated

object. This paper adopts a five-level system: $V_e = \{v_1, v_2, v_3, v_4, v_5\}$, corresponding to Excellent, Good, Average, Poor, and Very Poor, respectively.

This paper determines membership degrees by counting the frequency of statistical factors across each comment grade and normalizing the results, thereby forming a fuzzy relationship matrix R_{efgh} and the membership distribution of behavioral factors [19, 20]. This can be expressed by the following formula:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \quad (6)$$

In Eq. (6), r_{nm} represents the membership degree of the n th factor for the m th evaluation level in the learning efficiency assessment of the MOOC teaching system.

The aggregation of expert judgments is achieved by collecting evaluations from multiple independent experts on each factor belonging to each language category. For each factor, experts select multiple categories from the preset language categories and count the number of times each category is selected. The mapping from frequency to membership value is normalized by dividing the count frequency of each language category by the total number of evaluations, and the resulting ratio is directly used as the membership value. The membership function used is a frequency based linear membership function without any additional transformations. The operation of language categories is completed by defining a set of fuzzy sets, each language category corresponds to a fuzzy set, and its membership degree is calculated by statistical frequency methods. Let the factor set be X equal to x_1 to x_n, and the language category set be L equal to L_1 to L_m. For each factor x_i, there are k experts evaluating it. Let Z_{ij} be the number of experts who classify factor x_i as language category L_j. The formula for calculating the membership degree μ of factor x_i to language category L_j is:

$$\mu = \frac{Z_{ij} w_i^*}{k} \quad (7)$$

Assuming the evaluation factor is F1, the language categories are low, medium, and high, and three experts participate in the evaluation. Expert 1 classified F1 as low, Expert 2 as medium, and Expert 3 as high. The statistics show that the low category count is 1, the medium category count is 1, the high category count is 1, and the total number of evaluations is 3. Calculate the membership degree, F1 has a membership degree of 0.333 for low categories, 0.333 for medium categories, and 0.333 for high categories. Represent the factors in the rows of the fuzzy relationship matrix sample, list the language categories, and assign membership values to the elements. The matrix is constructed based on three factors and three language categories, with specific values shown in Table 3.

Table 3. Sample of the fuzzy relation matrix

Factor	Low	Middle	High
F1	0.333	0.333	0.333
F2	0.000	0.667	0.333
F3	0.333	0.333	0.333

In the matrix, the membership degree of factor F2 is generated from the assumed data: among the three experts, the zero position is classified as low, the two positions as medium, and the single position as high. The calculated membership degrees are 0.000 for the low category, 0.667 for the medium category, and 0.333 for the high category.

The weight vector w_i determined earlier via the AHP method is combined with the fuzzy relationship matrix R through specific operators to obtain the comprehensive evaluation result vector B. Four common fuzzy synthesis operators exist; this paper employs the weighted average operator, whose calculation formula is as follows:

$$b_j = \sum_{i=1}^{n} (w_i \cdot B)^n = min\left\{\mu, \sum_{i=1}^{n} (w_i \cdot R) \cdot \frac{1}{\sum_{i=1}^{n} w_i \cdot \sum_{i=1}^{n} r_{ij}}\right\} \quad (8)$$

In Eq. (8), b_j represents the j th element in the comprehensive evaluation result vector B of the MOOC teaching system.

De-fuzzification is applied to the comprehensive evaluation result vector B to obtain the final score or grade. This paper employs the weighted average method to calculate the final score using the following formula:

$$S_e = \sum_{j=1}^{n} v_j \cdot \left(\frac{\sum_{j=1}^{n} b_j}{b_j}\right) \quad (9)$$

In Eq. (9), S_e represents the final score for learning efficiency in the MOOC teaching system, while v_j denotes the score corresponding to the jth grade in the comment set.

Finally, the score calculated via Eq. (9) constitutes the assessment result for the MOOC teaching system's learning efficiency. Thus, this paper completes the research on evaluating MOOC teaching system learning efficiency based on the Kirkpatrick's Model and AHP fuzzy comprehensive evaluation method. When the MOOC environment lacks longitudinal tracking data between the behavior layer and the outcome layer, this study adopts a statistical validation strategy based on cross-sectional data for empirical exploration rather than assuming structural relationships between different levels. Conduct statistical analysis using real-time data provided by the same group of learners after the course, including response level satisfaction, learning level knowledge and skill test scores, behavior level application intention scale scores, and outcome level long-term benefit perception scores. Firstly, the correlation coefficient between the evaluation results of adjacent levels was calculated, and it was found that there was a moderate positive correlation between the evaluation values of the reaction layer and the learning layer, and a significant positive correlation between the intention evaluation values of the learning layer and the behavior layer. This preliminarily supports the hypothesis that low-level outcomes have a predictive effect on high-level outcomes. Secondly, the path relationship between levels was tested using a structural equation model. The model considers the reaction and learning layers as exogenous variables, and the behavioral intention and outcome perception as endogenous variables. The analysis showed that the standardized path coefficients from the reaction layer to the learning layer, and

then from the learning layer to the behavioral layer, all reached a statistically significant level. This indicates that there is an expected positive influence path between levels, rather than an independent parallel relationship.

III. Experimental Setup

A. Environment

To conduct an experiment evaluating learning efficiency in MOOC teaching systems, this paper establishes an integrated experimental environment featuring data collection, analysis processing, and result visualization capabilities.

The open-source MOOC platform Moodle was selected as the experimental vehicle, deployed on an Alibaba Cloud ECS server. It was configured with Linux (Ubuntu 20.04), Apache, MySQL, and PHP to form a LAMP architecture. Using the Moodle backend, an experimental course was created, uploading resources such as videos, documents, and test questions, while setting up chapter structures, assignment submission modules, and discussion forums. The platform database was extended with a user behavior log table to record structured data like login times. JavaScript scripts embedded event listener code in the frontend to capture unstructured behavioral data, which was saved to the log file system.

This paper uses Python 3.8 to build the environment, installing libraries like NumPy for matrix weight calculations, Scikit-learn for constructing fuzzy membership functions, and Matplotlib with Seaborn for visualization. The AHP module incorporates expert input to compute eigenvalues and CR values, constructing a fuzzy comprehensive evaluation module to generate scores. An API integrates the MOOC platform with the data analysis system. Flask develops middleware to periodically extract learner data from MySQL and convert it to Pandas storage. Post-experiment, the program runs to produce scores and radar charts.

B. Data Source

The experiment design recruited 200 participants with diverse learning backgrounds and knowledge levels to engage in the experimental course. To minimize the impact of individual differences on outcomes, recruitment aimed for balanced distribution across age, major, and learning experience. No additional grouping was performed; all participants studied the same experimental course to ensure consistency in the experimental environment.

The experimental course selected covered multiple knowledge domains, featured practical content, and had a well-structured chapter layout, incorporating rich videos, documents, quizzes, and assignments. A 12-week learning cycle was established, during which participants completed scheduled tasks including video viewing, document reading, discussion participation, assignment submission, and test-taking. Structured data such as login events were recorded using a user behavior log table extended from the platform database. Unstructured data like mouse clicks were captured via front-end embedded JavaScript scripts and stored in a log file system.

C. Expert Solutions

Prior to the experiment, education experts were consulted to determine metric weightings. A hierarchical model was constructed based on the Kirkpatrick's Model indicator system. Experts were organized to perform pairwise comparisons of indicators across levels, determining importance on a 1–9 scale to form a pairwise comparison matrix. Eigenvalue methods were used to calculate weight vectors, which were tested and adjusted if unsatisfactory, yielding combined weights for the criterion and indicator layers.

The experiment invited 5 experts in the field of educational technology to form an evaluation team, all of whom hold associate professor or above titles and have been engaged in online education research for over 8 years. Experts first construct an indicator system based on the four levels of the Kirkpatrick model, and independently score each indicator at each level to determine the degree of membership. The specific operation is as follows: for each learner's underlying indicator, experts compare the five level comment set to give a rating judgment of excellent, good, moderate, poor, and very poor. Calculate the frequency of judgment for each indicator and level by 5 experts, and normalize to obtain the membership values in the fuzzy relationship matrix. Taking the rationality index of course design in the reaction layer as an example, if 3 out of 5 experts are judged as excellent and 2 as good, the membership vector will be 0.6 excellent, 0.4 good, 0.0 moderate, 0.0 poor, and 0.0 very poor.

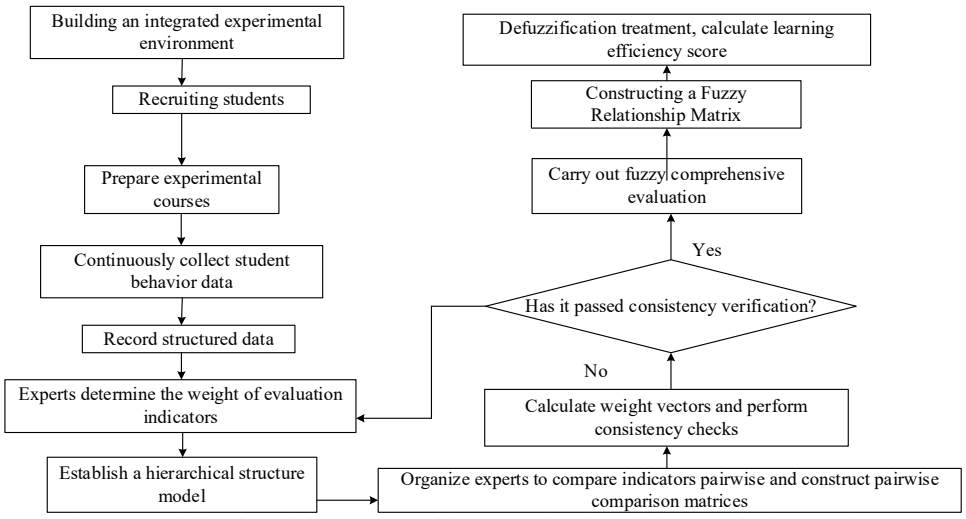

Fig. 2. Overall flowchart of the MOOC teaching system learning efficiency evaluation experiment.

After the experiment, fuzzy comprehensive evaluation was conducted using the constructed indicator system as the factor set and a five-level rating scale as the comment set. Expert scoring determined the membership degrees of each factor across evaluation grades, forming a fuzzy relationship matrix. The weight vector from AHP was combined with this matrix using a weighted average operator to yield the comprehensive evaluation result vector. Finally, weighted averaging was applied for defuzzification, calculating individual learner and overall course learning efficiency scores. The overall experimental workflow is shown in Fig. 2.

By thoroughly analyzing the distribution map of learner efficiency scores, we observe the score distribution, identify the predominant evaluation levels among learners, and analyze the differences in characteristics between learners at different levels. Utilizing the course evaluation radar chart across all levels, we visually examine the performance of the four layers—including the reaction layer—in the overall course evaluation. This validates the effectiveness of the assessment method based on the Kirkpatrick's Model and AHP fuzzy comprehensive evaluation in practical teaching applications.

In the specific implementation process of this research method, experts in the field of education are invited to compare pairwise the indicators at each level in the evaluation index system of the learning efficiency of MOOC teaching system. Compare using a 1–9 importance scale. Based on the comparison results of experts, judgment matrices for the criterion layer and indicator layer were constructed, and eigenvalues, eigenvectors, consistency indicators, and consistency ratios were calculated sequentially to obtain the combined weights of each indicator. The criterion layer consists of four factors: the response layer, the learning layer, the behavior layer, and the outcome layer. The judgment matrix criteria constructed by experts after comparison are shown in Table 4.

Table 4. Criterion layer judgment matrix

Factor	Conversion layer	Learning layer	Behavior Layer	Result layer
Conversion layer	1	1/3	1/4	1/3
Learning layer	3	1	2	3
Behavior Layer	4	1/2	1	2
Result layer	3	1/3	1/2	1

Use the eigenvalue method to calculate the weight vector. Firstly, calculate the maximum eigenvalue $\lambda_{\max}$ and corresponding eigenvector of the judgment matrix. After calculation, the maximum eigenvalue $\lambda_{\max} = 4.1558$ is obtained. After normalizing the eigenvectors, the weight vector w_m of each factor in the criterion layer is obtained:

$$w_m = [0.091, 0.435, 0.294, 0.179]^T \tag{10}$$

Perform consistency checks. The consistency index C_I is calculated:

$$C_I = \frac{\lambda_{max}}{n-1\frac{4.1558-4}{4-1}519} \tag{11}$$

The reaction layer includes three indicators: the rationality of course design, the teaching level of instructors, and the convenience of platform operation. The judgment matrix constructed by expert comparison is shown in Table 5.

Table 5. Judgment matrix for response layer indicators

Index	Rationality of Course Design	Lecturer teaching proficiency	Easy platform operation
Rationality of Course Design	1	1/2	2
Lecturer teaching proficiency	2	1	3
Easy platform operation	1/2	1/3	1

The maximum eigenvalue $\lambda_{\max} = 3.0092$ is calculated. After normalizing the eigenvectors, the weight vector w_f for each factor in the criterion layer is obtained:

$$w_f = [0.297, 0.539, 0.164]^T \tag{12}$$

Perform consistency checks. The consistency index C_I is calculated:

$$C_I = \frac{\lambda_{max}}{n-1\frac{3.0092-3}{3-1}046} \tag{13}$$

The result calculated by Eq. (13) is 0.0046 < 0.1, which passes the consistency test.

The learning layer includes three indicators: degree of knowledge mastery, degree of skill improvement, and change in learning attitude. The judgment matrix constructed by expert comparison is shown in Table 6.

Table 6. Learning layer indicator judgment matrix

Index	Knowledge mastery level	Skill level improvement	Changes in learning attitude
Knowledge mastery level	1	1	3
Skill level improvement	1	1	3
Changes in learning attitude	1/3	1/3	1

The maximum eigenvalue $\lambda_{\max} = 3.0000$ was calculated. After normalizing the eigenvectors, the weight vector w_g for each factor in the criterion layer was obtained:

$$w_g = [0.4286, 0.4286, 0.1429]^T \tag{14}$$

Perform consistency checks. The calculation result of consistency index C_I is:

$$C_I = \frac{\lambda_{max}}{n-1\frac{3.0000-3}{3-1}} \tag{15}$$

The result calculated by Eq. (15) is 0 < 0.1, passing the consistency test. The behavior layer comprises three indicators: knowledge application frequency, problem-solving capability, and teamwork proficiency. The judgment matrix constructed by experts is presented in Table 7.

The maximum eigenvalue $\lambda_{\max} = 3.0092$ was calculated. After normalizing the eigenvectors, the weight vector w_k for each factor in the criterion layer was obtained:

$$w_k = [0.297, 0.539, 0.164]^T \tag{16}$$

Perform consistency checks. The calculation result of consistency index C_I is:

$$C_I = \frac{\lambda_{max}}{n-1\frac{3.0092-3}{3-1}46} \quad (17)$$

The result calculated by Eq. (17) is 0 < 0.1, passing the consistency test. The outcome layer comprises two indicators: individual performance improvement and organizational benefit growth. The judgment matrix constructed by the experts is presented in Table 8.

Table 7. Behavior layer indicator judgment matrix

Index	Frequency of knowledge application	Problem-solving ability	Team collaboration skills
Frequency of knowledge application	1	1/2	2
Problem-solving ability	2	1	3
Team collaboration skills	1/2	1/3	1

Table 8. Result layer indicator judgment matrix

Index	Personal Performance Improvement	Organizational efficiency growth
Personal Performance Improvement	1	2
Organizational efficiency growth	1/2	1

The maximum eigenvalue $\lambda_{\max} = 2.0000$ was calculated. After normalizing the eigenvectors, the weight vector w_s for each factor in the criterion layer was obtained:

$$w_s = [0.6667, 0.3333]^T \quad (18)$$

Perform consistency checks. The calculation result of consistency index C_I is:

$$C_I = \frac{\lambda_{max}}{n-1\frac{2.0000-2}{2-1}} \quad (19)$$

The result calculated by Eq. (19) is 0 < 0.1, which passes the consistency test.

Calculate the combined weights of each indicator relative to the target layer using Eq. (5). The results of the combination weight calculation are shown in Table 9.

The sum of weights for all indicator combinations equals 1.0000, satisfying the normalization condition. To reduce the potential systemic bias caused by a single disciplinary background in the expert group, this study established an interdisciplinary balance selection mechanism when constructing the Analytic Hierarchy Process expert group. The selection of expert group members follows the principle of domain coverage, and the final five invited experts come from five different professional fields: interpretation teaching, online education, educational measurement and evaluation, applied linguistics, and curriculum design. This ensures that when judging the relative importance of indicators, multidimensional perspectives such as teaching practice, online learning characteristics, evaluation effectiveness, language learning rules, and curriculum structure can be comprehensively considered. Each expert independently conducts pairwise comparisons and judgments, and then summarizes and calculates the weights of each indicator. To verify the effectiveness of this mechanism, we analyzed the consistency ratio of expert group judgments and calculated a CI value below the acceptable standard of 0.1. This indicates that experts from different disciplines have a high degree of consistency in their overall understanding of the relative importance of indicators, and interdisciplinary combinations do not lead to irreconcilable differences, but rather enhance the comprehensiveness and stability of weight results.

Table 9. Combination weights of evaluation indicators for learning efficiency in MOOC teaching systems

Top class	Name of index	Policy Layer Weight	Indicator layer weight	Combination weight
Conversion layer	The rationality of the course design	0.091	0.297	0.0270
	The instructor's teaching proficiency	0.091	0.539	0.0490
	Easy platform operation	0.091	0.164	0.0149
Learning layer	Knowledge mastery level	0.435	0.4286	0.1864
	Level of skill improvement	0.435	0.4286	0.1864
	Change in learning attitude	0.435	0.1429	0.0622
Behavior Layer	Frequency of knowledge application	0.294	0.297	0.0873
	Problem-solving ability	0.294	0.539	0.1585
	Team collaboration skills	0.294	0.164	0.0482
Result layer	Personal Performance Improvement	0.179	0.6667	0.1193
	Organizational efficiency growth	0.179	0.3333	0.0597

IV. Experimental Results and Analysis

According to the experimental design, the proposed evaluation method based on the Kirkpatrick's Model and AHP fuzzy comprehensive evaluation was applied to assess the learning efficiency of the MOOC teaching system. Using a simple random sampling method, use Python's random sample function to extract from all 200 samples without replacement, with a sampling seed set to 42 to ensure reproducibility. The purpose of sampling is to visually present the specific scores of learners with different levels of learning efficiency in the form of individual cases, while maintaining the statistical characteristics of the entire sample. The demographic distribution of the 200 learners in the entire sample is as follows. In terms of age distribution, 88 learners aged 18 to 25 accounted for 44.0%, 62 learners aged 26 to 30 accounted for 31.0%, 36 learners aged 31 to 40 accounted for 18.0%, and 14 learners aged 40 and above accounted for 7.0%. In terms of professional background, 72 students majoring in computer science account for 36.0%, 50 students majoring in engineering account for 25.0%, 38 students majoring in business account for 19.0%, 28 students majoring in humanities and social sciences account for 14.0%, and 12 students majoring in other fields account for 6.0%. In terms of online learning experience, 144 people have online learning experience, accounting for 72.0%, and 56 people have no online learning experience, accounting for 28.0%. In terms of gender distribution, 112 males account for 56.0% and 88 females account for 44.0%. In terms of educational

level, 98 undergraduate students account for 49.0%, 74 master's students account for 37.0%, and 28 doctoral students and employees account for 14.0%. The specific demographic characteristics are shown in Table 10.

Table 10. Distribution of demographic characteristics of 200 learners in the entire sample

Feature category	Grouping	Number of people	Percentage
Years old	18–25	88	44.0%
	26–30	62	31.0%
	31–40	36	18.0%
	Over 40	14	7.0%
Professional Background	Computer science	72	36.0%
	Engineering	50	25.0%
	Business Studies	38	19.0%
	Humanities & social sciences	28	14.0%
	Other	12	6.0%
Online learning experience	Practicedness	144	72.0%
	Inexperience	56	28.0%
Sexual distinction	The male sex	112	56.0%
	The female sex	88	44.0%
Educational Level	Undergraduate course	98	49.0%
	Master degree candidate	74	37.0%
	PhD candidates and working professionals	28	14.0%

Evaluation results from 14 learners are selected and presented in Table 11.

Table 11. Evaluation results of MOOC teaching system learning efficiency based on Kirkpatrick's model and AHP fuzzy comprehensive evaluation method

Evaluation object	Reaction layer rating	Learning level rating	Behavioral level rating	Result level rating	Comprehensive learning efficiency rating
Student 1	82	78	85	80	81.2
Student 2	75	72	70	68	71.3
Student 3	90	88	92	85	88.7
Student 4	65	60	62	58	61.2
Student 5	80	75	82	78	78.8
Student 6	70	68	72	65	68.8
Student 7	88	85	90	82	86.3
Student 8	60	55	58	52	56.3
Student 9	78	73	80	75	76.5
Student 10	72	70	75	68	71.2
Student 11	85	82	88	80	83.8
Student 12	58	52	55	50	53.8
Student 13	76	71	78	73	74.5
Student 14	68	65	70	62	66.3

The descriptive statistics for the comprehensive learning efficiency scores of all 200 learners are presented in Table 12.

Table 12. Descriptive statistics of learning efficiency scores for 200 learners in the whole sample

Evaluation Indicator	Sample capacity	mean value	Standard deviation	Least value	Median	Crest value
Reaction layer scoring	200	74.6	10.2	52.0	75.0	93.0
Learning Layer Score	200	71.3	11.5	48.0	72.0	91.0
Behavior Layer Score	200	73.9	10.8	50.0	74.5	94.0
Result Layer Score	200	69.8	12.1	45.0	70.0	89.0
Comprehensive Learning Efficiency Score	200	72.5	10.5	51.5	73.0	91.0

The distribution patterns of scores at different levels and comprehensive scores were tested by skewness and kurtosis coefficients, and the results showed that the data was close to a normal distribution. The entire sample of 200 learners were divided into five levels of evaluation based on their comprehensive learning efficiency rating. Excellent levels correspond to scores of 90 or above, good levels correspond to scores of 80 to 89, medium levels correspond to scores of 70 to 79, poor levels correspond to scores of 60 to 69, and very poor levels correspond to scores below 60. The number and proportion of personnel in each level are shown in Table 13.

Table 13. Distribution of comprehensive learning efficiency levels of 200 learners in the whole sample

Order of evaluation	Fractional Interval	Number of people	Percentage
Outstanding	90 ≤ grade ≤ 100	18	9.0%
Good	80 ≤ grade ≤ 89	46	23.0%
Secondary	70 ≤ grade ≤ 79	72	36.0%
Range	60 ≤ grade ≤ 69	42	21.0%
Poor	grade < 60	22	11.0%

To validate the superiority of the research method, this paper further designs a comparative experiment. The proposed MOOC learning efficiency evaluation method—based on the Kirkpatrick's Model and AHP fuzzy comprehensive evaluation—serves as the experimental group. Methods from Huaren *et al*. [1–3] are selected for comparison, designated as Comparison Method 1, Comparison Method 2, and Comparison Method 3, respectively. The method described in Huaren and Chunyan [1] is a hierarchical evaluation method based on classical set theory, which strictly follows its steps to classify the same original indicator scores into clear levels through a fixed threshold set by it, and then conducts synthesis. The method Hossain *et al.* [2] adopts the traditional weighted average algorithm, using the same indicator weights as in this paper, to linearly weight and sum the standardized original scores to obtain the final evaluation score. The method described in Luki [3] is a decision-making approach based on simple rule-based reasoning, which replicates the if then rule library and inputs experimental data into its inference engine to obtain evaluation results. All comparison methods used the same raw metric data of 200 learners as input in this experiment. Under identical experimental conditions, all four methods assess MOOC learning efficiency, with evaluation accuracy results shown in Fig. 3.

The radar chart in Fig. 3 demonstrates the multidimensional advantages of the proposed evaluation method (experimental group).

The quantitative evaluation metrics employed for comparison include Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and classification accuracy. These metrics are calculated using experts 'comprehensive assessments as reference standards. Classification accuracy is defined as the proportion of samples whose final evaluation results match the experts' judgments, calculated as the number of consistent samples relative to the total sample size. The quantitative comparison results of the three comparison methods with the proposed method across all 200 learner datasets are presented in Table 14.

The data in Table 14 shows that the proposed method achieves an RMSE of 0.45, MAE of 0.35, and accuracy rate of 89.5%, all of which are the best among all comparison

methods. This indicates that the proposed method demonstrates higher consistency with expert evaluation results and smaller assessment errors compared to the three existing methods under comparison.

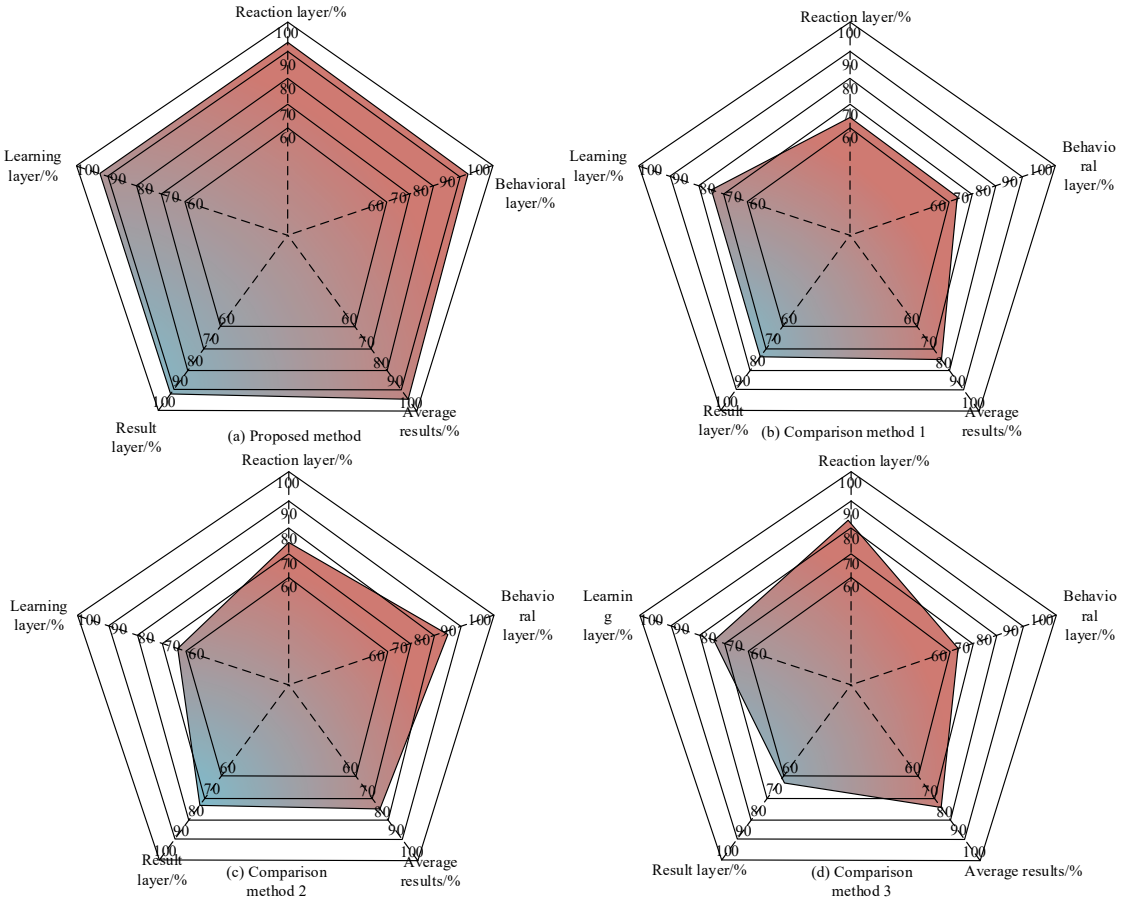

Fig. 3. Accuracy comparison of different evaluation methods for MOOC learning efficiency.

Table 14. Quantitative comparison results of different methods

Method	RMSE	MAE	Accuracy rate
Huaren and Chunyan [1] method	0.85	0.68	74.5%
Hossain *et al.* [2] method	0.62	0.50	81.0%
Luki [3] method	0.73	0.59	78.5%
Method of this paper	0.45	0.35	89.5%

V. Discussion

A. Interpretation of Test Results

The proposed method for evaluating MOOC teaching system learning efficiency, based on the Kirkpatrick's model and AHP fuzzy comprehensive evaluation, demonstrates significant scientific rigor and effectiveness. The Kirkpatrick's Model establishes a comprehensive framework spanning four progressive levels from reaction to outcome, covering key aspects of learning efficiency: the reaction level focuses on immediate perceptions; the learning level centers on knowledge, skills, and attitudes; the behavioral level examines application and competence; and the outcome level measures actual impact. This multi-level approach ensures a thorough and objective assessment.

The AHP method provides rigorous quantitative logic for the evaluation, determining indicator weights through expert comparisons and ensuring reasonable allocation through validation. The fuzzy comprehensive evaluation method effectively addresses ambiguity and uncertainty, converting qualitative aspects into quantitative measures. By integrating the strengths of these three approaches, the method ensures that evaluation results accurately reflect learning efficiency.

The AHP method, through expert comparison and calculation, ensures reasonable indicator weights, injecting scientific objectivity. The fuzzy comprehensive evaluation method handles fuzziness and uncertainty. The combination of these three elements accurately captures characteristics, providing reliable results under varying student numbers and verifying outstanding performance.

Compared to the comparison methods, the experimental group exhibits higher coordination and comprehensiveness across all four evaluation accuracy dimensions. Its radar chart coverage is larger and more balanced, attributed to the systematic framework of the Kirkpatrick's model, which captures features comprehensively and avoids limitations. Simultaneously, the integration of AHP and fuzzy comprehensive evaluation enhances the scientific rigor and precision of the assessment. The former ensures rational weight allocation for indicators, while the latter handles ambiguity and converts qualitative factors into quantitative measures. This multi-method synergy enables the experimental group to achieve superior assessment accuracy, validating its superiority.

B. Model Data Validation

This paper further validates the reliability of the evaluation model. The confidence interval is calculated based on the sampling distribution of the evaluation results. For each evaluation level, the confidence interval for its comprehensive evaluation value is defined as follows. If the mean comprehensive evaluation value of a certain level is U

and the standard error is *S*, then at the confidence level, the upper and lower limits of its confidence interval are *U* plus or minus *T* multiplied by *S*. *T* is the critical value of the t-distribution with *n* degrees of freedom minus 1 at the specified confidence level. This article adopts a 95% confidence level. The confidence level results are shown in Fig. 4.

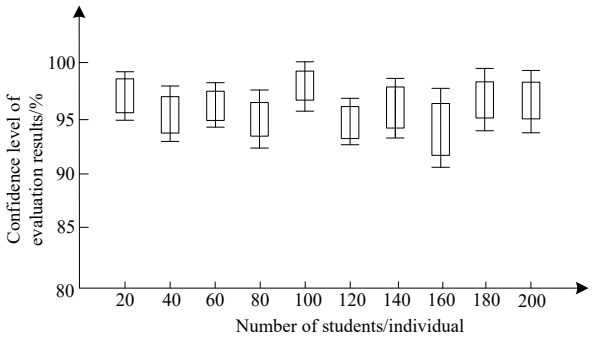

Fig. 4. Confidence results for MOOC teaching system learning efficiency assessment using the Kirkpatrick's model and AHP fuzzy comprehensive evaluation method.

As shown in Fig. 4, the learning efficiency assessment results for the MOOC teaching system using the Kirkpatrick's model and AHP fuzzy comprehensive evaluation method exhibit high confidence. They remain stable across varying student numbers, demonstrating strong reliability and stability. The Kirkpatrick's Model establishes a comprehensive framework covering key dimensions, avoiding one-sidedness and providing a thorough theoretical foundation.

To confirm the rigor of the above results, statistical tests were conducted on Table 14 data. The statistical verification of performance differences adopts paired sample t-test. The error in the evaluation results obtained by comparing this method with each comparative method on the same sample is considered as paired observations. Specifically, tests will be conducted on three types of indicators: RMSE, MAE, and accuracy. The null hypothesis of the test is that there is no difference in the mean performance indicators between the two methods, and the alternative hypothesis is that the mean performance indicators of our method are better than those of the comparison method. For the error metrics RMSE and MAE, the superiority is demonstrated by lower mean values; For accuracy metrics, superiority manifests as a higher mean. Therefore, a one-sided test is used. The statistical test results provide quantitative evidence to support the conclusion of significant superiority. All *p*-values of the tests were significantly lower than the significance level of 0.01, indicating rejection of the null hypothesis that our method outperforms the three compared methods in RMSE, MAE, and accuracy. Combined with the high confidence demonstrated by the confidence interval, this statistical validation result provides sufficient quantitative evidence for the conclusion that our method has excellent performance.

This article actually calculated the consistency ratio CR values for all levels. Among them, the criterion layer has a judgment matrix CR value of 0.03 for the target layer. The CR values of each judgment matrix in the indicator layer are as follows: the CR value of the corresponding matrix in the reaction layer is 0.08, the CR value of the corresponding matrix in the learning layer is 0.05, the CR value of the corresponding matrix in the behavior layer is 0.07, and the CR value of the corresponding matrix in the result layer is 0.04. The overall synthesis weight consistency ratio CR value under the overall target layer is 0.05. All CR values are significantly below the threshold of 0.1, indicating that experts have a high degree of consistency in overall and local judgments, and the weight allocation results are reliable.

C. Limitation Analysis

The experimental period is set at twelve weeks, which is a mid-term observation window. For outcome level indicators such as personal performance improvement and organizational efficiency growth that require longer periods of time to manifest, the evaluation may not be sufficient. At the methodological level, the comment set used in fuzzy comprehensive evaluation is divided into five levels: excellent, good, moderate, poor, and very poor. The classification criteria are relatively fixed, and the portrayal of subtle differences in learners' performance may not be precise enough, which limits the inference of causal relationships between different levels of results.

D. Impact Analysis

On the premise of maintaining the main relationship of the judgment matrix, introduce small perturbations to make the CR value close to 0.1, recalculate the weights, and evaluate the same batch of learner samples. The results showed that when the CR value increased from 0.05 to 0.10, the average change in the final comprehensive evaluation value was 0.02, which was much smaller than the confidence interval width reported in this article. This quantitative relationship indicates that in this research system, the actual calculated low CR value ensures the stability of weight allocation, and the confidence interval of the final evaluation result calculated from this can effectively cover the possible impact of slight inconsistencies in expert judgment. The evaluation result is robust.

VI. Conclusion

The main content of this study is to propose and validate a learning efficiency evaluation method for MOOC teaching systems based on the Kirkpatrick model and AHP fuzzy comprehensive evaluation method. The study first constructed an evaluation index system covering four dimensions: response, learning, behavior, and outcome. The Analytic Hierarchy Process was used to calculate the weight vectors of each level of indicators, and consistency checks were conducted to ensure the rationality of weight allocation. Subsequently, the fuzzy comprehensive evaluation method was used to construct a fuzzy relationship matrix by calculating membership degrees, and the weight vectors were combined with the fuzzy matrix for deblurring processing, ultimately achieving a comprehensive quantitative evaluation of learning efficiency.

Two hundred learners were selected for the experiment, and the proposed method was compared with three existing methods to verify its superior performance in root mean square error, mean absolute error, and classification accuracy metrics. The key contribution of this study is the effective integration of the systematic structure of the Kirkpatrick

model and the quantitative analysis advantages of AHP fuzzy comprehensive evaluation, establishing a multi-level and coherent evaluation framework. This method overcomes the limitations of traditional single dimensional scoring or simple weighted average methods, and can handle the inherent fuzziness and uncertainty in learning efficiency evaluation. It unifies scattered, accurate, or fuzzy feedback into a comprehensive diagnosis of overall teaching effectiveness, improving the reliability and scientificity of learning efficiency evaluation in MOOC teaching systems, and providing a scientific basis for optimizing MOOC teaching systems.

In view of the limitations of the method proposed in this article, future research can expand the sample size and diversity of the study, include more learners from different regions, cultural backgrounds, and occupational stages, and extend the time span of the tracking study to more comprehensively test the generalization ability and long-term predictive validity of this evaluation method. In terms of methodology, we can try to refine the rating of fuzzy evaluation comments or develop more flexible custom comment sets to improve the resolution of evaluation results.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization and data curation were performed by Y. Wu and Y. Lin; Writing—Original draft was completed by Y. Wu; Writing—Review & Editing, and Supervision were provided by N. G. Weng and B. Y. Leau. All authors approved the final version of this paper.

References

[1] Z. Huaren and H. Chunyan, "Combat effectiveness evaluation of synthetic brigade equipment based on combined grey entropy weight method," *Comput. Simul*., vol. 40, no. 4, pp. 26–30, 2023.

[2] M. A. Hossain, M. S. A. Mazumder, and M. H. Bari, "Impact assessment of machine learning algorithms on resource efficiency and management in urban developments," *Global Mainstream J*., vol. 1, no. 2, pp. 1–9, 2024.

[3] R. Luki, "Evaluation of the efficiency of banks in Serbia using the MABAC method," *Bankarstvo*, vol. 51, no. 2, pp. 10–60, 2022.

[4] N. Zbaltan, "Applying machine learning to audit data: Enhancing fraud detection, risk assessment and audit efficiency," *EDPACS*, vol. 69, no. 9, pp. 70–86, 2024.

[5] Y. Tong and Z. Zhan, "An evaluation model based on procedural behaviors for predicting MOOC learning performance: Students' online learning behavior analytics and algorithms construction," *Interact. Technol. Smart Educ*., vol. 20, no. 3, pp. 291–312, 2023.

[6] H. Yujie, X. Zhijie, and Y. Dingyu, "CDES: Data-driven efficiency evaluation methodology for cloud database," *Jisuanji Kexue*, vol. 51, no. 6, pp. 111–117, 2024.

[7] K. Kono, T. Matsunaga, and H. Okuda, "Parallel implementation and efficiency evaluation of multifrontal method," in *Proc. Comput. Mech. Conf.*, 2022, vol. 2022, no. 35, pp. 21–28.

[8] Ablyalimov, "Efficiency evaluation of 3VL80S electric locomotives by the method of averaging calculated of values," *Universum: Tech. Sci.*, vol. 98, pp. 5–11, 2022.

[9] K. Krupavathi, M. R. Babu, and A. Mani, "Assessment of water use efficiency for sustainable water management by integration of remote sensing and machine learning," *Adv. Geogr. Environ. Sci.*, vol. 1, pp. 61–78, 2024.

[10] H. A. Phan, A. N. Thi, and N. P. Dang, "Magnetic bead conjugated lung tumor cell binding efficiency assessment based on deep-learning approach," in *Proc. 1st Int. Conf. Health Sci. Technol. (ICHST)*, 2023, pp. 1–6.

[11] M. Ekmekcioglu, E. E. Baakn, and M. Zger, "Developing meta-heuristic optimization based ensemble machine learning algorithms for hydraulic efficiency assessment of storm water grate inlets," *Urban Water J*., vol. 19, no. 10, pp. 1093–1108, 2022.

[12] X. Chen, "Human-machine course evaluation with an ALBERT-based position embedding feature enhancement network," *Int. J. Knowl. Manag*., vol. 21, no. 21, pp. 1–22, 2025.

[13] G. Yang, "Enhanced IVIFN–ExpTODIM–MABAC technique for multi-attribute group decision-making and applications to college English teaching quality evaluation under interval-valued intuitionistic fuzzy sets," *Int. J. Fuzzy Syst*., vol. 27, no. 6, pp. 1846–1858, 2025.

[14] Y. Jin, "The user satisfaction evaluation of MOOC teaching platform based on multidimensional association rules," *Int. J. Continuing Eng. Educ. Life-Long Learn*., vol. 34, no. 1, pp. 391–404, 2024.

[15] J. Weng, Y. Hu, and H. Huang, "LPSA: Online educational video evaluation considering cognitive and teaching style," *Comput. Appl. Eng. Educ.*, vol. 32, no. 1, 22662, 2024.

[16] B. Qiao and Y. Mi, "Signal processing and development process based on," *J. Inf. Process. Syst*., vol. 20, no. 1, pp. 105–115, 2024.

[17] R. Pireddu and F. Truyen, "Enhancing reusability practices in massive online open courses," in *Proc. Eur. Conf. e-Learning*, 2024, vol. 23, no. 1, pp. 307–312.

[18] J. Zhu, "Application of MOOC mixed teaching model in college physical education teaching—Taking track and field teaching as an example," *Int. J. Embedded Syst.*, vol. 17, no. 1–2, pp. 126–137, 2024.

[19] Y. Guo, "A MOOC online teaching quality evaluation method based on fuzzy algorithm," *Int. J. Continuing Eng. Educ. Life-Long Learn*., vol. 1, no. 1, pp. 275–285, 2024.

[20] J. Rodrigues-Silva and N. Alsina, "Enhancing teachers' professional identity in a reflective learning MOOC," *Int. J. Emerg. Technol. Learn.*, vol. 19, no. 1, pp. 23–38, 2024.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).