

Optimization of Teaching Evaluation Systems Based on Combined Weighting Methods

Jun Lei¹, Jiangyue Lei², and Xichen Li³

¹ Associate Professor, School of Education Science, Minzu Normal University of Xingyi, Xingyi 562400 Guizhou, China

² Staff Journalist and Editor, News Center, China Media Group, Beijing, 100038, China

³ M.A, Ordos Education and Sports Bureau, General Support Center, Ordos, Inner Mongolia, 017010, E-mail: guozr79@163.com (corresponding author).

Project Management

Received November 13, 2025; revised January 1, 2026; accepted March 13, 2026

Available online March 13, 2026

Abstract: As educational reforms advance, optimizing teaching evaluation systems is crucial for improving educational quality. However, traditional evaluation systems typically rely on fixed weighting methods, which fail to adapt to the dynamic changes in evaluation indicators, leading to a lack of precision in results. To address this, this paper proposes an optimization method based on a combined weighting approach integrating the entropy weighting method with state variable weighting vectors. This method introduces state variable weighting theory to dynamically correct the initial entropy weights derived from data dispersion, constructing a new evaluation model using the Hadamard product. Experimental verification demonstrates that, compared to traditional methods, the proposed approach exhibits higher distinctiveness and stability when processing data across multiple disciplines and educational stages, effectively capturing dynamic characteristics in the teaching process. The evaluation system constructed in this study not only overcomes the limitations of static weighting but also provides a scientific reference for educational administrators to optimize resource allocation and enhance teaching quality.

Keywords: Combination weighting, optimization of teaching evaluation systems, entropy weighting method, state variable weighting theory.

Copyright © Journal of Engineering, Project, and Production Management (EPPM-Journal).

DOI 10.32738/IEPPM-2025-270

1. Introduction

In today's educational field, teaching evaluation serves as a crucial tool for measuring teaching quality and effectiveness, and its scientific rigor and rationality have garnered significant attention (Brzozka, 2025). Traditional teaching evaluation systems often use fixed weighting methods to integrate various evaluation indicators, which cannot deeply analyze the intrinsic information of data and adapt to the dynamic changes of indicators (Xhelaj and Burlando, 2024). In teaching evaluations across different disciplines, teaching stages, and student groups, the importance of various evaluation indicators may undergo significant changes, but traditional fixed-weighting evaluation systems cannot promptly and accurately reflect these changes (Bai et al., 2024). Scientific and reasonable teaching evaluation can provide teachers with accurate feedback on their teaching, helping them identify strengths and weaknesses in the teaching process, thereby enabling them to adjust teaching strategies and improve teaching methods in a targeted manner, ultimately enhancing teaching quality (Luan et al., 2024). Therefore, research on scientific, reasonable, and accurate teaching evaluation methods holds significance for improving teaching quality, optimizing the allocation of educational resources, and promoting educational equity (Fangyang, 2023).

Currently, teaching evaluation research primarily encompasses the construction of evaluation indicator systems, indicator weighting, and the optimization of evaluation models (Li, 2023). In terms of constructing evaluation indicators, the focus is on comprehensively assessing teaching quality from multiple dimensions, including teaching content, teaching methods, teacher quality, and student engagement (Tian et al., 2023). In terms of weighting methods, quantitative analysis techniques such as principal component analysis (Tang, 2024) and factor analysis (Tian et al., 2023) have been gradually introduced to reduce the influence of subjective factors. However, these methods still have certain limitations in handling nonlinear relationships and dynamic changes among indicators. Evaluation model research methods primarily include fuzzy decision-making, material element analysis, neural networks, and deep learning algorithms (Tang, 2024). Li (2023) emphasized the

importance of combining with China’s educational context, focusing on indicators with local characteristics such as teacher ethics and the conversion of educational and teaching achievements. Li et al. (2024) made significant improvements to the limitations of traditional entropy weighting methods in fuzzy data scenarios and applied them to the optimization of teaching evaluation systems. Li and Wang (2020) proposed segmented adjustments to weights based on indicator values, improving response speed by 30%, but the setting of transition points relies on experience. Qi et al. (2022) designed a Logistic variable weight function to smooth weight changes, achieving an anomaly detection accuracy of 85% in medical school evaluations. Li and Chen (2017) construct an Analytic Hierarchy Process (AHP) and Fuzzy Comprehensive Evaluation (FCE) model to assess engineering education quality, but it suffers from weight biases caused by expert preferences. Saadatdoost et al. (2023) construct a Quality-learning (Q-learning) weight adjustment model, but it requires training with millions of data points. Although existing research has made progress in weight determination, significant research gaps remain. First, subjective weighting methods like AHP rely heavily on expert experience, making them prone to bias. Second, traditional objective weighting methods (such as the single entropy weighting method), while utilizing data information, are sluggish in responding to sudden changes in indicator states. For instance, when a critical indicator (e.g., teaching effectiveness) shows an extremely low value, traditional methods fail to automatically “penalize” its weight, causing the comprehensive score to mask serious deficiencies. Therefore, there is an urgent need for a “penalty-incentive” mechanism that can both reflect data information content and dynamically adjust weights based on the real-time state of indicators. Through the analysis of the above literature, the following shortcomings are summarized: (1) Most existing teaching evaluation systems use static weighting methods and cannot dynamically adjust indicator weights based on actual conditions during the teaching process (Wang and Zhang, 2024). (2) Some studies fail to fully consider the complex relationships between evaluation indicators when constructing teaching evaluation systems (Wu, 2024). (3) The grading of indicators is not sufficiently detailed (Zhao and Yan, 2024).

In response to the aforementioned issues, this paper proposes an optimized teaching evaluation system based on a combination of entropy weighting and state variable weighting methods. The specific contributions of this paper are as follows: (1) Refining the grading of teaching quality indicators to comprehensively and accurately assess all elements of the teaching process, thereby constructing an optimized teaching evaluation system. (2) Introducing the combination of entropy weighting and state variable weighting methods to design a teaching evaluation method based on this approach. (3) Through actual experimental settings, the effectiveness of the method proposed in this study and traditional methods in teaching evaluation is compared, providing empirical support for the optimization of the teaching evaluation system.

2. Construction of Teaching Evaluation System

2.1. Teaching Evaluation System

The teaching evaluation system is a multidimensional, multi-level structure designed to comprehensively and objectively evaluate all aspects of teaching activities (Ouyang, 2025). The specific structure of the system is shown in Fig. 1.

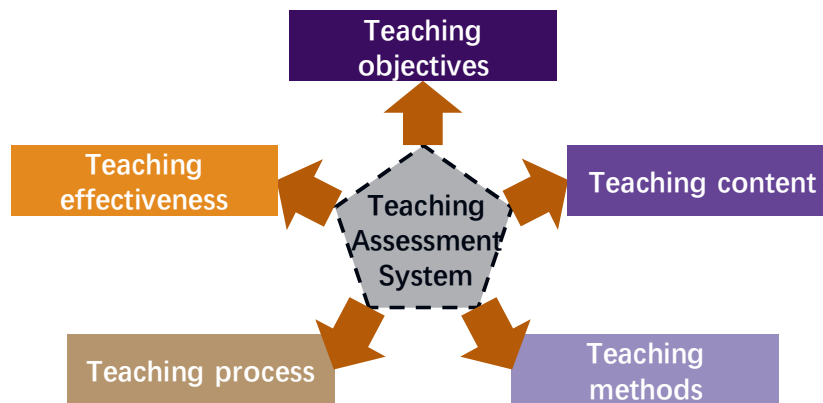


Fig. 1. Teaching evaluation system diagram

2.1.1. Teaching objectives dimension

Teaching objectives are the starting point and ultimate goal of teaching activities. It is important to assess whether teaching objectives are clear, specific, and aligned with curriculum standards and student’s actual needs. Additionally, it is essential to evaluate whether teaching objectives encompass multiple dimensions such as knowledge, skills, and attitudes, and whether they are feasible and measurable (Guo et al., 2024).

2.1.2. Teaching content dimension

The selection and organization of teaching content directly affect teaching effectiveness. It is necessary to assess whether the content is accurate, complete, and systematic, whether it is up-to-date, and whether it reflects the core knowledge and cutting-edge developments of the discipline. At the same time, it is also necessary to consider whether the difficulty level of the teaching content is appropriate for the student’s learning level.

2.1.3. Teaching methodology dimension

The diversity and flexibility of teaching methods are key to improving teaching quality. Evaluate whether teachers select appropriate teaching methods based on teaching content and student characteristics, such as lectures, discussions, and inquiry-based learning, and whether they can flexibly use modern educational technology to assist teaching, such as multimedia teaching and online teaching platforms.

2.1.4. Teaching process dimension

The teaching process includes lesson preparation, classroom instruction, and post-class tutoring. The evaluation assesses whether teachers have adequately prepared their lessons, whether the teaching process is complete and well-structured, whether the teaching pace is reasonable, whether teachers can pay attention to student’s classroom responses and adjust their teaching strategies promptly, and whether post-class tutoring is timely and effective.

2.1.5. Teaching effectiveness dimension

Teaching effectiveness is the ultimate manifestation of teaching activities and is primarily measured by student’s academic performance, improvement in abilities, and changes in learning attitudes. For example, whether students have improved in terms of knowledge mastery, thinking abilities, innovation abilities, and collaboration abilities, and whether their interest and enthusiasm for learning have increased.

The teaching quality indicator assessment matrix can be obtained from the teaching evaluation indicator system, as shown in Eq. (1).

$$E = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}_{n \times m} \tag{1}$$

Among them, x_{nm} represents the measured value of the m power quality indicator at time n.

2.2. Teaching Quality Indicator Grading

In order to more accurately assess teaching quality, it is necessary to classify various teaching quality indicators. In accordance with relevant national standards, the achievement levels of teaching assessment indicators are divided into five grades: Excellent I, Good II, Average III, Pass IV, and Fail V (Shao et al., 2024). By clearly defining the specific scoring criteria for each grade, Table 1 provides the quantitative standards for different indicators at each grade level.

Table 1. Grading of teaching evaluation indicators

Indicator	V	IV	III	II	I
Teaching objectives	≤1	≤2	≤3	≤4	>4
Teaching content	≤0.5	≤1.0	≤1.5	≤2.0	>2.0
Teaching methods	≤1.5	≤3.0	≤4.5	≤7.0	>7.0
Teaching process	≤0.5	≤1.0	≤1.5	≤2.0	>2.0
Teaching effectiveness	≤2	≤3	≤4	≤5	>5

For the teaching objectives dimension, assessment indicators are categorized into five levels. Level I indicators with values >4 primarily indicate a high degree of achievement of teaching objectives, meaning that students not only master the knowledge and skills specified in the teaching objectives but also exhibit significant positive changes in their attitudes and emotions. They can flexibly apply the knowledge and skills learned to real-life situations and perform exceptionally well in related extension learning or practical activities. Level II indicator values ≤4 primarily indicate a high level of achievement, meaning that students have achieved the teaching objectives to a significant extent. Level III indicator values ≤3 primarily indicate basic achievement, meaning that students have basically mastered the core knowledge and skills specified in the teaching objectives, have shown some positive changes in their emotional attitudes, but have certain limitations in the transfer and application of knowledge, and have a general level of mastery of expanded content. Level IV indicator values ≤2 primarily indicate partial achievement. Level V indicator values ≤1 primarily indicate failure to achieve the objectives. The grading principles for instructional objective dimension indicators are shown in Fig. 2.

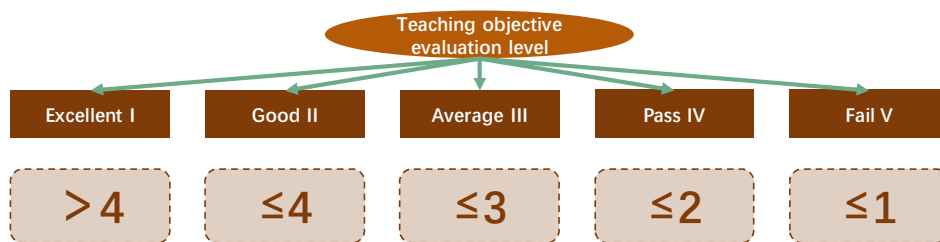


Fig. 2. Graded description of teaching objective dimension indicators

3. Entropy-State Weight Combination Weighting Method

3.1. Entropy Weighting Method

In order to achieve a grading assessment of teaching quality, the entropy weight method (Zeng et al., 2024) was used to conduct a comprehensive analysis of the teaching quality assessment grades. Through the teaching quality assessment grades, a comprehensive assessment initial matrix S was established, as presented in Eq. (2).

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1m} \\ \vdots & \cdots & \vdots \\ s_{n1} & \cdots & s_{nm} \end{bmatrix}_{n \times m} \quad (2)$$

In the equation, s_{mm} represents the teaching quality assessment grade.

To ensure that the evaluation process is based on the same measurement standards, the data needs to be standardized using Eq. (3).

$$z_{ij} = \frac{s_{ij} - \min_j \{s_{ij}\}}{\max_j \{s_{ij}\} - \min_j \{s_{ij}\}} \quad (3)$$

Consequently, the standardized matrix can be expressed as Eq. (4).

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{110} \\ \vdots & \cdots & \vdots \\ z_{n1} & \cdots & z_{n10} \end{bmatrix}_{n \times 10} \quad (4)$$

Among them, z_{ij} represents the standardized data of the j indicator assessment level at the i moment.

Since the data is discrete, this paper uses the probability of data frequency as input data for entropy calculation. Specifically, the difference between the maximum and minimum values of a set of data is used as the interval, which is divided into k equal parts. The probability of data existing in each interval is used as input for entropy calculation. The specific calculation is performed according to Eq. (5).

$$p_{ij} = \frac{t_{ij}}{n} \quad (5)$$

Among them, p_{ij} represents the frequency probability of data points falling into the i th interval under the j th indicator, and n_{ij} represents the number of n data points at the j th indicator level that falls into the i th interval. The constraint condition is as follows in Eq. (6).

$$1 \leq j \leq m, 1 \leq i \leq k, 0 \leq t \leq n, \sum_{i=1}^k p_{ij} = 1 \quad (6)$$

Based on the above, the probability input matrix is derived as shown in Eq. (7).

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \cdots & \vdots \\ p_{k1} & \cdots & p_{kn} \end{bmatrix}_{k \times m} \quad (7)$$

Based on the probability matrix derived above, to translate data dispersion into measurable weights, we introduce the concept of Information Entropy. In information theory, entropy measures the degree of uncertainty or disorder. Specifically, within the context of teaching evaluation, if the data distribution of a certain indicator is highly concentrated (high entropy), it implies slight variation across samples, meaning it cannot effectively distinguish teaching quality. Thus, a lower weight should be assigned. Conversely, greater data dispersion (low entropy) indicates that the indicator carries a higher amount of information, warranting a higher weight.

After obtaining the probability matrix, to quantify the amount of information contained in each indicator, we need to calculate its information entropy. A smaller information entropy implies a greater degree of data dispersion, and thus, a higher weight should be assigned. Under the same indicator, information entropy is used to quantify the degree of uncertainty in the data, reflecting the uncertainty or disorder of the probability distribution of the data under that indicator, which is calculated using Eq. (8).

$$I_{ij} = \ln \left(\frac{1}{p_{ij}} \right) = -\ln (p_{ij}) \quad (8)$$

Among them, I_{ij} represents the degree of uncertainty, which is complementary to the amount of information that can be provided, p_{ij} represents the probability that the data point falls within the i - th interval. As shown in Eq. (9).

$$E_j = \frac{\sum_{i=1}^k (p_{ij} \times I_{ij})}{\ln (k)} \quad (9)$$

Among them, E_j represents the entropy value of the j th indicator. When $p_{ij} \rightarrow 0$, since $\lim_{p_{ij} \rightarrow 0} \left(\frac{p_{ij}}{-\ln(p_{ij})} \right) = 0$ it follows that $\lim_{p_{ij} \rightarrow 0} (p_{ij} \times I_{ij}) = 0$.

Since the entropy value E_j represents a measure of uncertainty, it reflects the amount of information required to eliminate that uncertainty. Its value is inversely related to the amount of information needed: the higher the entropy, the less information is required. Conversely, the smaller the entropy value, the more information is needed. The value of entropy is complementary to the expected amount of information it can provide. Therefore, the expected information amount D_j is given in Eq. (10).

$$w_j = \frac{D_j}{\sum_{j=1}^m D_j} \quad (10)$$

In Eq. (10), when $E_j = 1$, D_j becomes 0, resulting in a normalized indicator weight of 0, which means the indicator does not effectively participate in the subsequent comprehensive evaluation. Therefore, an equidistant translation of the expected information amount is adopted to avoid the occurrence of multiplication or zero weight, as illustrated in Eq. (11).

$$w_j = \frac{D_j + \alpha}{\sum_{j=1}^m (D_j + \alpha)} \quad (11)$$

Where α represents the zero-point translation amount. Considering the differences and correlations among the indicators, this paper uses the Pearson correlation coefficient as a supplementary measure of correlation, as follows in Eq. (12).

$$\xi_{ij} = \frac{\sum_{t=1}^n (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2} \sqrt{\sum_{t=1}^n (y_t - \bar{y})^2}} \quad (12)$$

Where ξ_{ij} represents the correlation coefficient between indicator i and indicator j , x_t and y_t represent the t th standardized data for indicators i and j . Correspondingly, \bar{x} and \bar{y} represent the mean values of the standardized data for indicators i and j , respectively.

The conflict intensity between indicators is obtained from the Pearson correlation coefficient in Eq. (13).

$$I_j = \frac{\sum_{i=1}^{m-1} (1 - \xi_{ij})}{m-1} \quad (13)$$

Where I_j represents the conflict intensity of the j indicator. The comprehensive indicator weight value can then be determined by Eq. (14).

$$w_j = \frac{I_j (D_j + \alpha)}{\sum_{j=1}^m (I_j (D_j + \alpha))} \quad (14)$$

Based on the above principles, the entropy weight method flowchart is shown in Fig. 3.

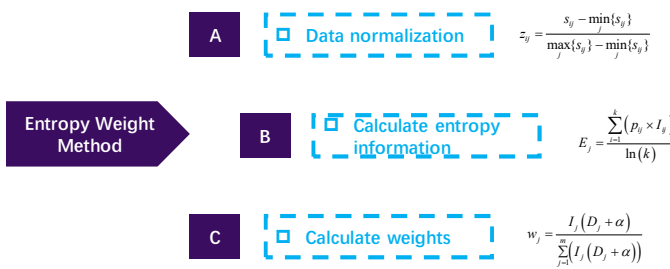


Fig. 3. Entropy weighting method

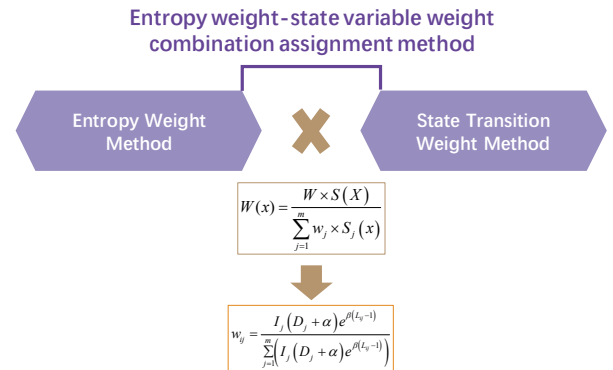


Fig. 4. Principle of the combination weighting method

3.2. State-Based Weighting Method

In order to respond to the real-time and dynamic nature of teaching evaluation, this paper adopts the state variable weight method based on variable weight theory to correct the weights. Relying solely on the entropy weighting method fails to reflect the impact of the indicator value's own state on the overall evaluation. Therefore, we introduce the state variable vector $S(X)$ to achieve dynamic weight correction by constructing a state function. According to variable weight theory (Ma et al., 2023), state variable weight rules are used to construct mappings, as expressed in Eq. (15).

$$S : [0,1]^m \rightarrow [0,1]^m \quad X \mapsto \mathbf{S}(X) = (S_1(x), S_2(x), \dots, S_m(x)) \quad (15)$$

Where $S(X)$ is an m -dimensional state variable weight vector. When addressing the relationship between the state variable and the single indicator level, to comply with the axiomatic definition of state variable weights, it is necessary to treat the data points as variables to establish the connection. By taking the indicator data as variables and applying grading constraints according to the level thresholds, the state variable weight factor can be obtained, as defined in Eq. (16).

$$S(x) = e^{\beta(L(x)-1)} \quad (16)$$

In this equation, $S(x)$ is the state variable weight factor for indicator data x , β is the gain | coefficient of the indicator variable weight factor. β is used to adjust the strength of the state variable weight for each constant weight point within the calculation period, to optimize the dynamic response characteristics of the weights at different time points. Through regression analysis, the optimal range of β values were determined, and $\beta = 0.85$ was finally selected. With this value, the dynamic adjustment of the weights maintains sensitivity to changes in state while avoiding excessive fluctuations, thus ensuring the stability and rationality of the evaluation results.

The constraint function is as in Eq. (17).

$$L(x) = \begin{cases} L_i \times \frac{(x - x_i)}{(x_{i-1} - x_i)} + \\ L_{i+1} \times \frac{(x - x_{i-1})}{(x_i - x_{i-1})} & x \in [x_{i-1}, x_i) \\ L_5 & x \in [x_4, +\infty) \end{cases} \quad (17)$$

In this Eq.(17): x represents the instantaneous indicator data at a 1 -minute interval, L_i is the level coefficient under the i th level, x_i is the threshold value for the i th level, where $i \in [1,5]$. Specifically, $x_0 = 0$.

3.3. Combination Weighting Method

The Hadamard product (element-wise matrix multiplication) was selected over linear weighted combinations in this study primarily due to its unique “multiplier effect.” While linear combinations (addition) tend to smooth out extreme states of indicators, the Hadamard product allows the state variable to act as a “gain coefficient” directly applied to the base entropy weight. This means that when an indicator is in a critical state (high state weight), its final weight is non-linearly amplified. This ensures the evaluation system remains highly sensitive to abnormal or exceptional states, aligning with the "penalty-incentive" logic inherent in variable weight theory.

In order to combine the advantages of the entropy weight method and the state variable weight method, a combined weighting approach is adopted. The specific principle is shown in Fig. 4. For the constant weight variable $W = (w_1, w_2, \dots, w_m)$. This paper uses the Hadamard product. As shown in Eq. (18).

$$W(x) = \frac{W \times S(X)}{\sum_{j=1}^m w_j \times S_j(x)} \quad (18)$$

At this point, satisfying the three axioms, when extended to multiple sets of data points, the variable weights obtained are as in Eq. (19).

$$w_{ij} = \frac{I_j (D_j + \alpha) e^{\beta(L_{ij}-1)}}{\sum_{j=1}^m (I_j (D_j + \alpha) e^{\beta(L_{ij}-1)})} \quad (19)$$

Where w_{ij} represents the weight of the j th indicator at time i . As shown in Eq. (20).

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1m} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nm} \end{bmatrix}_{n \times m} \quad (20)$$

Based on the above steps and methods, the weights of different data points under the same indicator are no longer the same. Instead, they are multiplied by their own grade weighting coefficient on the basis of the original entropy weighting method, which supplements the dynamic weighting of the indicator data, enhances the weighting resolution and dynamic

response of each indicator, and enables the weighting results to more accurately and quickly reflect the differences between indicators.

4. Optimization Process of Teaching Evaluation System Based on Combination Weighting Method

Combining the entropy weight method and the state variable weight method, this section proposes an optimization method for the teaching evaluation system based on the combination weighting method. The specific approach is shown in Fig. 5.

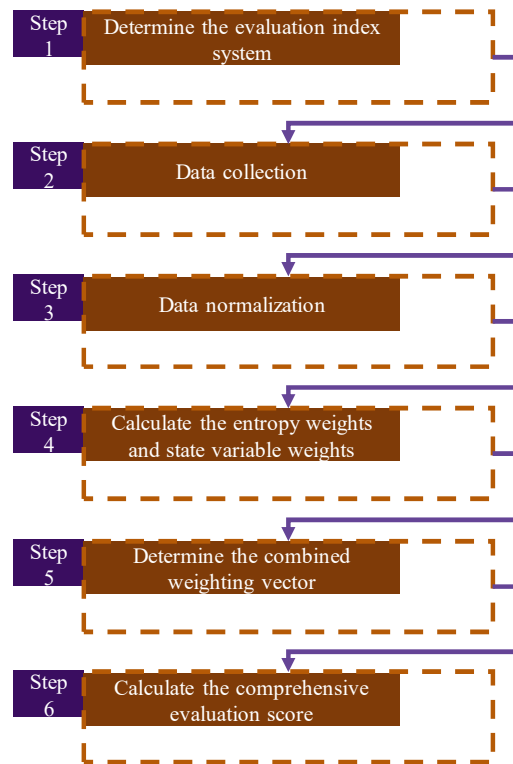


Fig. 5. Optimization process of the teaching evaluation system

Step 1: Determine the evaluation indicator system.

Based on the objectives and requirements of teaching evaluation, establish an evaluation indicator system that includes dimensions such as teaching objectives, content, methods, process, and effectiveness. Clarify the specific indicators under each dimension, their meanings, and measurement methods, as shown in Fig. 6.

Step 2: Data Collection.

Data for evaluation indicators are collected through various methods, as shown in Fig. 7. The achievement of teaching objectives can be comprehensively evaluated through student’s course exam scores, assignment completion, classroom performance, etc. The novelty of teaching content can be assessed by analyzing the syllabus, textbook content, and supplementary materials provided by teachers. The effectiveness of teaching methods can be evaluated by collecting data through student questionnaires, classroom observations, and other methods.

Step 3: Data standardization processing.

The collected raw data is processed using the data standardization method in the entropy weight method, converting data with different units and different value ranges into comparable standardized data for subsequent calculations.

Step 4: Calculate entropy weights and state change weights separately.

According to the entropy weight method calculation steps, calculate the entropy weights of each evaluation indicator grade. At the same time, according to the actual state weight function, calculate the state weights of each evaluation indicator.

Step 5: Determine the combination weighting vector.

According to the pre-set adjustment coefficient α , calculate the combination weighting vector using the combination weighting equation.

Step 6: Calculate the comprehensive assessment score.

Weighted sum of standardized assessment indicator data and combination weighting vector (Fig. 8),

The comprehensive score Z for teaching evaluation is obtained, and the calculation equation is given by Eq. (21).

$$Z = \sum \omega_{ji}y_{ij} \tag{21}$$

where y_{ij} is the normalized evaluation indicator data.

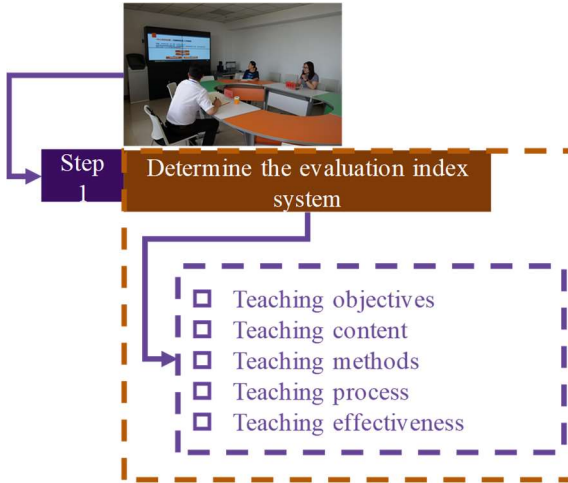


Fig. 6. Steps in constructing a teaching evaluation system

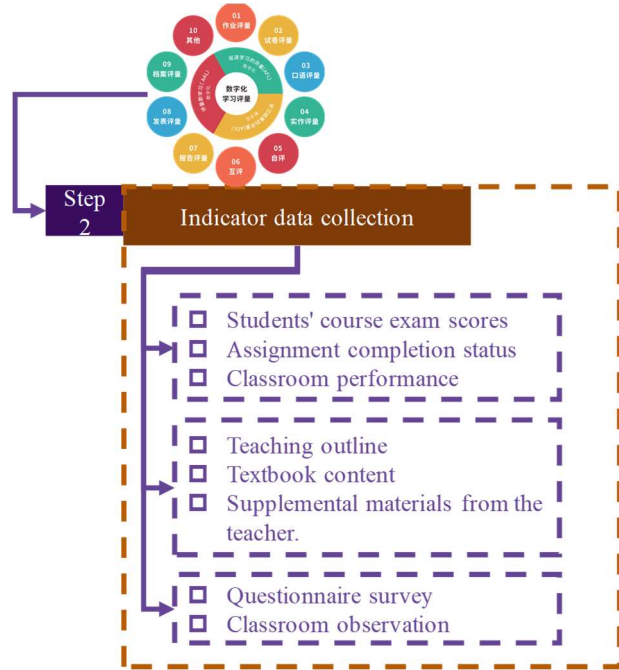


Fig. 7. Indicator data collection

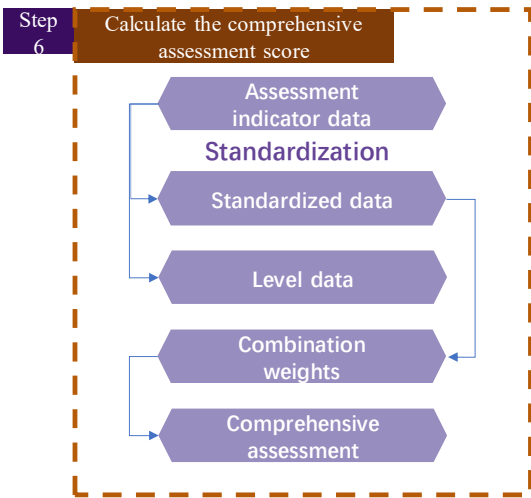


Fig. 8. Comprehensive assessment steps

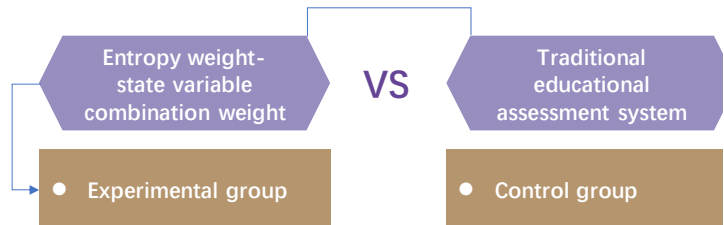


Fig. 9. Grouping situation

5. Experimental Verification

5.1. Experimental Setup

5.1.1. Selection of experimental subjects

A total of 1,441 test samples were selected from multiple schools, different grade levels (junior high school, high school, and university), and various disciplines (humanities, sciences, and engineering). These samples encompassed a wide range of teaching scenarios to ensure the results are more applicable and representative.

5.1.2. Group design

The experimental subjects were divided into an experimental group and a control group (Wang and Yang, 2023). The experimental group was evaluated using a teaching evaluation system based on a combination of entropy weight and state change weighting methods. The control group was evaluated using a traditional teaching evaluation system, as shown in Fig. 9.

5.1.3. Data collection and processing

During the experiment, the same data collection methods and standards were applied to each evaluation indicator to ensure the comparability of the two sets of data. The collected data included teacher teaching behavior data, such as lesson preparation records, classroom teaching videos, student learning data, such as academic performance, learning attitude survey results, and teaching environment data, such as teaching facility usage. (Abdrabo et al., 2024). Considering the differences in dimensions and evaluation standards of raw data across different educational stages (middle school, high school, university). For instance, differing maximum scores for exams, we conducted independent standardization for each subgroup before inputting data into the model. Using the range standardization method in Eq. (3), all indicator values were mapped to the $[0, 1]$ interval. This processing eliminated the incomparability of absolute values across different educational stages, ensuring the model focuses on the relative performance position of samples within their respective groups, thereby achieving a unified evaluation across stages. The collected data were processed according to the corresponding evaluation system to obtain the evaluation results for the experimental group and the control group, as shown in Fig. 10.

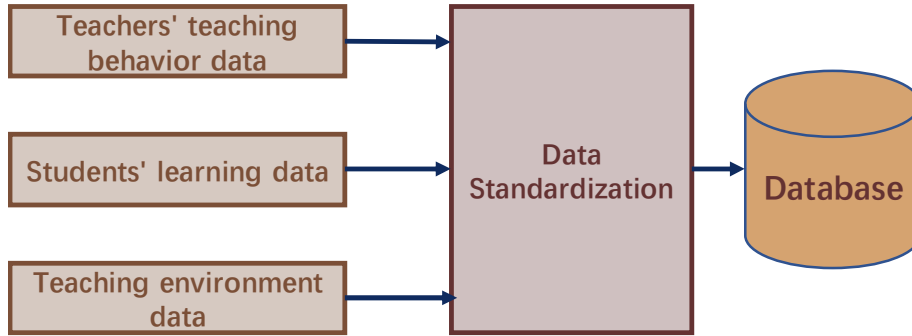


Fig. 10. Data collection and processing

5.2. Results Analysis

In order to analyze the accuracy and effectiveness of the optimization method of the teaching evaluation system based on the combination weighting method, this section conducts a verification analysis from the aspects of evaluation result accuracy, stability, and reliability, as well as the impact of different indicator weights on evaluation results.

5.2.1. Comparison of assessment accuracy

In order to analyze the accuracy of the evaluation results, this paper presents the comprehensive evaluation score distributions for the experimental group and the control group, as shown in Fig. 11 and Fig. 12.

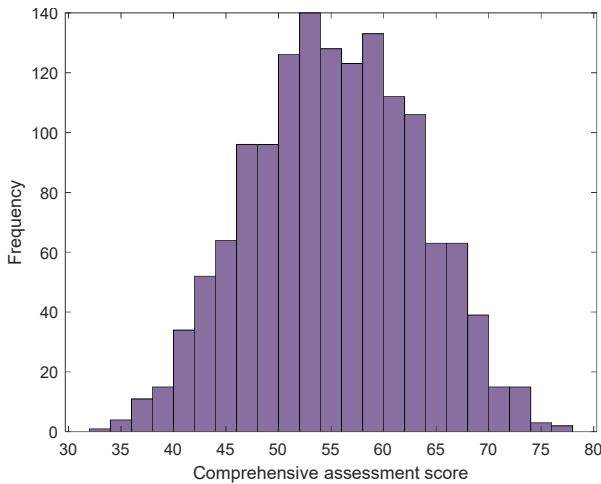


Fig. 11. Distribution of comprehensive assessment scores for teaching quality in the experimental group

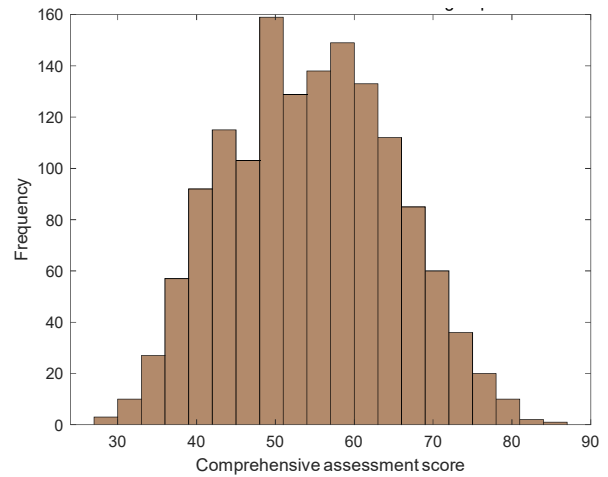


Fig. 12. Distribution of comprehensive assessment scores for teaching quality in the control group

As shown in Fig. 11, the score distribution exhibits characteristics of a normal distribution, indicating that the entropy-weight-state-variable-weight combination weighting method can reasonably distinguish between samples with different teaching quality. From Fig. 12, the score distribution is relatively concentrated, and the distinction is not as clear as in the experimental group. This is because the traditional evaluation system based on expert experience-weighted scores has relatively fixed weights, which cannot effectively adapt to changes in the importance of different indicators across various teaching scenarios.

5.2.2. Analysis of stability and reliability of assessment results

To analyze the stability and reliability of the evaluation method proposed in this paper, this section analyzes the fluctuations in evaluation scores for the experimental group and the control group in multiple rounds, yielding the results shown in Fig. 13 and Fig. 14. Specifically, the horizontal axis in these figures represents the sequence of assessment rounds (from Round 1 to Round 5), while the vertical axis denotes the normalized comprehensive evaluation score.

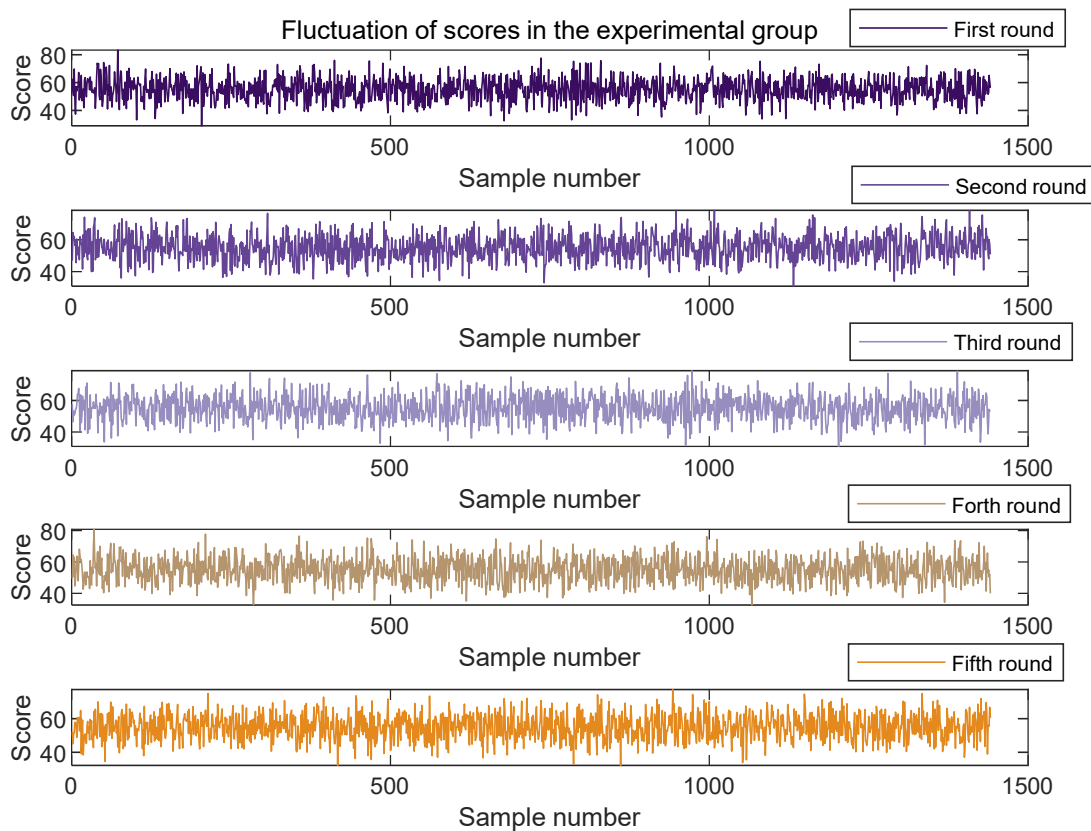


Fig. 13. Fluctuations in scores from multiple rounds of assessment in the experimental group

As can be seen from Fig. 13, the fluctuation range of the scores in each round of evaluation is relatively small, indicating that the evaluation system based on the entropy weight-state variable weight combination weighting method has high stability. This is because the method can adaptively adjust the weights according to the actual state and information content of the indicators, thereby reducing the impact of external factors on the evaluation results.

As seen in Fig. 14, the assessment scores of the control group fluctuated significantly, especially between samples from different types of schools and different grades. This is because the traditional assessment system lacks a dynamic adjustment mechanism and cannot adapt well to different teaching contexts, which affects the stability and reliability of the assessment results. To quantitatively assess the improvement in stability, we calculated the variance of scores across multiple evaluation rounds. The data shows that the score variance for the experimental group is 12.53, whereas the control group's variance is 45.21. Compared to the control group, the fluctuation amplitude in the experimental group was reduced by approximately 72.3%. Furthermore, the average Pearson correlation coefficient between scoring rounds in the experimental group reached 0.95, significantly higher than the 0.82 observed in the control group. These numerical results further demonstrate the advantage of the combined weighting method in maintaining the consistency of evaluation results.

5.2.3. Impact of different indicator weights on assessment results

To conduct a comparative analysis of the effectiveness of different weighting methods, this paper compares the evaluation results obtained using various weighting methods. The specific results are shown in Fig. 15. Histograms are used to visualize the distribution of evaluation results obtained using the three weighting methods, enabling an intuitive comparison of their differences. As shown in Fig. 15, the evaluation system based on the entropy-weighting and state-variable weighting combination method has a wider distribution, indicating that the scores are more dispersed. Compared with the other two methods, the entropy-weighting and state-variable weighting combination method is more effective in capturing the dynamic changes and real-time nature of teaching evaluation indicators.

6. Conclusion

This paper optimizes the teaching evaluation system by introducing an entropy-weighted and state-weighted combination weighting method, thereby constructing a more scientific and reasonable teaching evaluation system. First, research on the construction of the teaching evaluation system was conducted from five dimensions: teaching objectives, content, methods, process, and outcomes. Combining entropy weighting and state variable weighting theories, research on evaluation indicator

weighting methods was conducted. The entropy weight-state variable weight combination weighting method was applied to the teaching evaluation system, and multiple experimental verifications were conducted. The results indicate that this method has significant advantages in improving the accuracy, stability, and reliability of teaching evaluations. This study not only optimizes the weighting algorithm theoretically but also holds significant practical implications. Educational administrators can utilize this system to timely identify abnormal deficiencies in the teaching process, thereby conducting more precise teacher training and resource allocation. However, this study has limitations, particularly that the state variable weight parameter β is currently determined primarily based on regression analysis. Future work will focus on incorporating machine learning algorithms to achieve adaptive dynamic tuning of the parameter β by learning from massive historical evaluation data, thereby further enhancing the intelligence level of the evaluation system.

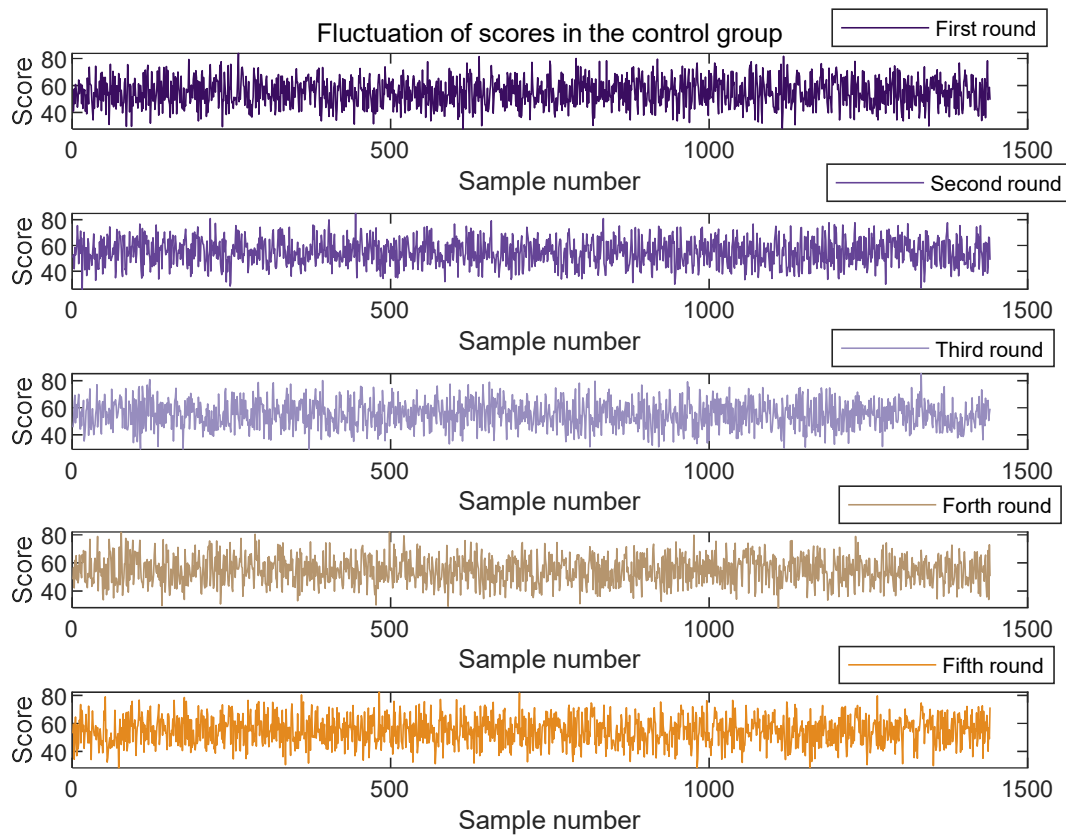


Fig. 14. Fluctuations in scores from multiple rounds of assessment in the control group

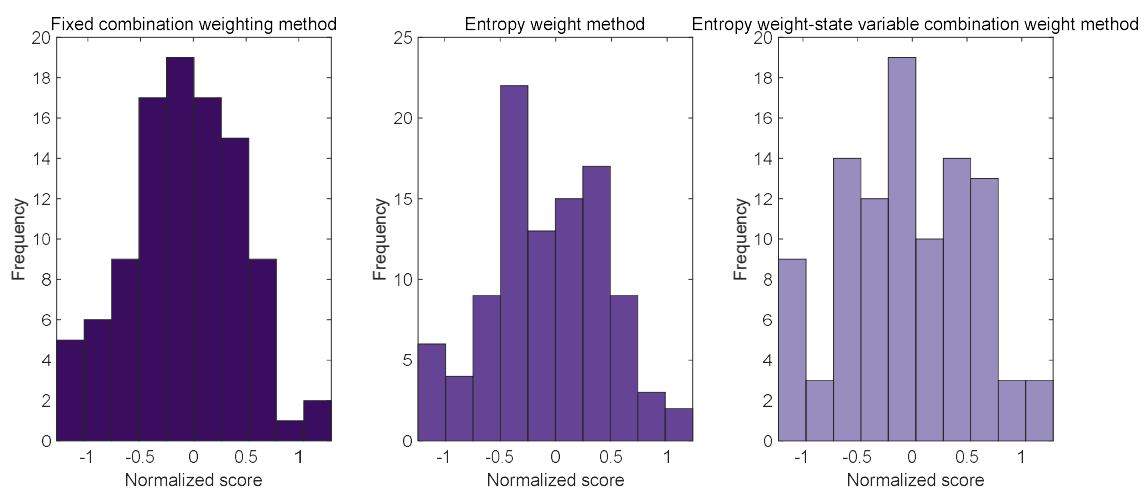


Fig. 15. Impact of different indicator weighting methods on assessment results

Author Contributions

Jun Lei contributed to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Jun Lei and

Jiangyue Lei contributed to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, and supervision. Xichen Li contributed to project administration and funding acquisition.

Funding

This research received the Guizhou Provincial Innovation Team on “Ecological Education System under the Macro Ecology Strategy” (Host Institution: Minzu Normal University of Xingyi).

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors confirm that no AI tools were used in the preparation of this manuscript.

References

- Abdrabo, K. I., Mabrouk, M., and Marzouk, A. (2024). Urban Competitiveness Assessment Using the Integration of the Global Power City Index and Multicriteria Decision-Making Methods. *Journal of Urban Planning and Development*, 150(3). <https://doi.org/10.1061/JUPDDM.UPENG-4789>
- Bai, W., Geng, G., and Fu, X. (2024). Research on the Optimization Method of Project-Based Learning Design for Chinese Teaching Based on Interference-Tolerant Fast Convergence Zeroing Neural Network. *International Journal of Computational Intelligence Systems*, 17(1), 1–14. <https://doi.org/10.1007/s44196-024-00532-6>
- Brzozka, B. (2025). Machine Learning Algorithms in Predicting College Students' Grades: A Review. *Journal of Applied Automation Technologies*, 3, 1–12. <https://doi.org/10.64972/jaat.2025v3.1>
- Fangyang, Z. (2023). Path Optimization of Physical Education Teaching Quality Management Integrated with Modern Network Technology. *International Journal of Science and Engineering Applications*. <https://doi.org/10.7753/ijsea1208.1016>
- Guo, Z., Ramli, M. F., and Zhang, W. (2024). Design and optimization of an open personalized human-computer interaction system for New Year Painting based on the learner's model. *Systems and Soft Computing*, 6. <https://doi.org/10.1016/j.sasc.2023.200070>
- Li, D., and Wang, F. (2020). State-based weighting theory and its application in intelligent control. *Acta Chimica Sinica*, 46(8), 1563–1574.
- Li, G. (2023). Research on the Teaching Optimization Path in Children's Creative Art Curriculum Under the Perspective of “Sanquan Education”. *Journal of Contemporary Educational Research*, 7(12), 136–141.
- Li, X. (2023). Optimization of the College Basketball Teaching Mode Based on the Applied Explainable Association Rule Algorithm and Cluster Analysis in Mobile Computing Environments. *Applied Artificial Intelligence*, 37(1), 23. <https://doi.org/10.1080/08839514.2023.2214768>
- Li, X., and Chen, Y. (2017). Applying AHP and Fuzzy Comprehensive Evaluation for Engineering Education Quality Assessment. *International Journal of Engineering Education*, 33(2A), 678–687. <https://doi.org/10.13140/RG.2.2.36235.39206>
- Li, Y., Cai, Q., and Wei, G. (2024). Grey relational analysis method for multi-attribute group decision making based on cumulative prospect theory under single value neutrosophic sets. *Journal of Intelligent and Fuzzy Systems*, 46(1), 15. <https://doi.org/10.3233/JIFS-231630>
- Luan, Y., Tang, X., and Ren, Y. (2024). Performance optimization of BP-DNN prediction model of suction caisson uplift bearing capacity employing modified Co-teaching method. *Computers and Geotechnics*, 176. <https://doi.org/10.1016/j.compgeo.2024.106756>
- Ma, H., Wang, J., and Wang, C. (2023). Modularized Fuzzy Evaluation Method for Temporary Support Devices in Coal Mine Tunnel Excavation Shields. *Journal of Xi'an University of Science and Technology*, 43(3), 576–585.
- Ouyang, R. (2025). Construction of online classroom instructional quality assessment system of university music based on BP neural network. *Scientific Reports*. <https://doi.org/10.1038/s41598-025-98556-1>
- Qi, X., Thomas, C., Min, L., Ismaila, T. S., Yun, D., and Ching, S. C. (2022). A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Computers and Education*, 189, 104582.
- Saadatdoost, R., Hiang Sim, A. T., and Yadegaridehkordi, E. (2023). Exploring the MOOC as a Community of Practice: developing a framework for knowledge construction. *Journal of E-Learning and Knowledge Society*, 19(1). <https://doi.org/10.20368/1971-8829/1135621>
- Shao, Z., Lin, Y., Chen, Y., and Chen, F. (2024). Two-stage robust optimization experimental case teaching of regional integrated energy systems. *Experimental Technology and Management*, 41(8), 175–181. <https://doi.org/10.16791/j.cnki.sjg.2024.08.024>
- Tang, D. (2024). Course Evaluation of Advanced Structural Dynamics Based on Improved SAPSO and FAHP. *Buildings*, 15. <https://doi.org/10.3390/buildings15010072>
- Tian, X., Chen, X., and Feng, L. (2023). Quality improvement path and countermeasures for future-oriented film and animation teaching: based on fuzzy comprehensive evaluation method. *Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology*, 44(2), 2981–2997.
- Wang, L., and Yang, X. (2023). Fuzzy comprehensive evaluation model in the evaluation of English teaching quality. *Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology*, 45(6), 10529–10543.
- Wang, S., and Zhang, S. (2024). Optimization of Higher Education Teaching Method System Based on Edge Intelligence. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-3007>

- Wu, Y. (2024). Research on optimization of teaching mental health monitoring and evaluation system based on intelligent algorithm. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1827>
- Xhelaj, A., and Burlando, M. (2024). Application of the teaching-learning-based optimization algorithm to an analytical model of thunderstorm outflows to analyze the variability of the downburst kinematic and geometric parameters. *Natural Hazards and Earth System Sciences*, 24(5), 23. <https://doi.org/10.5194/nhess-24-1657-2024>
- Zeng, Y., Jia, H., Yang, J., Wang, W., Zhong, H., and Han, J. (2024). A comprehensive contribution assessment method for virtual power plants based on fuzzy analytic hierarchy process, entropy weight method, and approximation of ideal solution ranking method. *Modern Electricity*, 41(1), 144–151.
- Zhao, Y., and Yan, Y. (2024). Research on the Optimization of English Translation Teaching Mode in Colleges and Universities Driven by Knowledge Mapping. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-2897>



Jun Lei, Associate Professor at the School of Education Science, Minzu Normal University of Xingyi, Guizhou. She received her Ph.D. in Education from Beijing Normal University, a Master of Education (M.Ed.) in English Education from East China Normal University, and a Master of Education (M.Ed.) from Beijing Normal University. Her research focuses on teacher education, professional development, educational human resources and teaching assessment. She has a wealth of teaching and research experience in the field of K-12 teaching. Integrating dual perspectives from K-12 basic education and higher education research, she effectively bridges the perspectives of basic and higher education. She strives to foster the synergy between classroom practices and educational reforms, aiming to establish an integrated educational research framework that combines theoretical rigor with practical applicability.



Yuejiang Lei, M.A. She received an undergraduate degree in Foreign Languages and Literature from the School of Foreign Studies at the University of Science and Technology Beijing and an M.A. in International Journalism and Communication from Beijing Foreign Studies University. Currently, she serves as an international news journalist in the news center of China Media Group. Her research focuses on English Language and Literature, Literary Ethics, and Intercultural Communication.



Xichen Li, M.A., is an official in a municipal Education and Sports Bureau. Her professional trajectory integrates frontline teaching, advanced study, and educational governance. Her primary research focuses on teacher professional development and evidence-based school improvement. She is particularly interested in designing actionable frameworks for teacher growth and formulating context-sensitive strategies for school enhancement, aiming to translate academic research into sustainable policy and practice. Her work bridges the gap between institutional policy formulation and classroom-level implementation, contributing to the development of systematic, research-informed approaches in public education.