WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2024.149882
ARCHIVES OF CIVIL ENGINEERING
Vol. LXX ISSUE 2 2024

pp. 579 –595

Research paper

Evaluation of flipped classroom teaching quality for civil engineering courses

Weipei Xue[1], Wei Jing[2]

Abstract: The impact of civil engineering course education on civil engineers is profound and crucial. Due to the hierarchical and ambiguous nature of quality assessment for flipped classroom teaching, there is an urgent demand for a rational and effective approach to conduct such assessments. This would enable the targeted formulation of instructional improvement methods based on assessment outcomes, ultimately elevating the quality of pedagogy. This study combines the analytic hierarchy process and fuzzy evaluation method. The fuzzy evaluation method is utilized to identify four primary evaluation factor sets, fourteen secondary judgment factor sets, and five evaluation outcome sets, with subsequent quantification of the assessment results. The analytic hierarchy process is employed to ascertain the weight coefficients of the evaluation factors. The comprehensive assessment model for flipped classroom teaching quality is established. The assessment results indicate that the overall quality of flipped classroom teaching in the civil engineering major at Anhui University of Science and Technology, conducted through the platform of Superstar Learning Hub, falls within the 'Good' category. The fuzzy comprehensive evaluation score for extracurricular learning quality is the lowest, and the weight proportion attributed to flipped classroom infrastructure is the highest. Consequently, several targeted improvement measures are proposed to enhance the quality of flipped classroom teaching.

Keywords: internet-enhanced education, analytic hierarchy process, fuzzy evaluation method, flipped classroom, quality assessment, education of civil engineers

---

[1] Assoc. Prof., PhD., Eng., Anhui University of Science and Technology, School of Civil Engineering and Architecture, Taifeng Street 168, 232001 Huainan, China, e-mail: xueweipei@aust.edu.cn, ORCID: 0000-0002-2551-8143

[2] Assoc. Prof., PhD., Anhui University of Science and Technology, School of Civil Engineering and Architecture, Taifeng Street 168, 232001 Huainan, China, e-mail: wjing@aust.edu.cn, ORCID: 0000-0002-0256-7655

1. Introduction

Civil engineering courses provide civil engineers with the necessary knowledge and skills to successfully participate in various civil engineering projects and make positive contributions to society and the environment [1]. These courses also help them continuously update and enhance their professional abilities to adapt to the ever-changing field of engineering. Therefore, civil engineering programs at higher education institutions are critically important for nurturing civil engineers with the requisite knowledge, skills, and professional competence [2].

The concept of Internet-enhanced education has been formally introduced against the backdrop of the deep integration between the internet and education. This concept amalgamates the internet with conventional instructional methodologies, devising novel modes of educational interaction that underscore enhanced intelligence and personalization. However, the assimilation and absorption of this new mode of teaching require the passage of time for refinement, as all developments exhibit a dual nature [3]. The accurate, effective, and rational evaluation of instructional quality within the framework of this new teaching approach, followed by comprehensive synthesis and summarization, is crucial.

The flipped classroom model is gaining popularity in higher education due to its use of internet-enhanced education [4], which is expected to improve learning outcomes and satisfaction while being cost-effective [5, 6]. However, it requires course redesign and some students struggle with self-regulation and time management. The model is controversial as it allows for more efficient use of in-class time but students' in-class performance may suffer if they fail to invest time in independent study [7, 8]. Furthermore, students may feel disoriented due to its novelty. Therefore, it is important to assess the model's instructional effectiveness.

Evaluating instructional quality in higher education, particularly in the context of flipped classroom teaching, is widely recognized as crucial for teaching quality management. However, it is challenging due to multiple layers and objectives, and the influence of evaluators' knowledge, cognitive abilities, and personal preferences, making it difficult to eliminate human biases [9]. Additionally, evaluation criteria often include qualitative descriptions with fuzzy attributes, adding complexity to the assessment process. While some institutions have tried to quantify these criteria, the scientific foundation for many indicators is often lacking, making the assessment of flipped classroom teaching quality a complex challenge.

In response to the hierarchical and fuzzy characteristics of flipped classroom teaching quality assessment, this study proposes a comprehensive assessment model that combines analytic hierarchy process and fuzzy evaluation method. This integration ensures a more scientific design and implementation of the instructional evaluation system. Additionally, building upon the presentation and discussion of flipped classroom teaching practices, this study offers constructive recommendations based on the assessment results. The aim is to facilitate the enhancement of flipped classroom teaching quality.

2. Flipped classroom based on Superstar Learning Hub

Superstar Learning Hub is a mobile-friendly learning platform that offers a wide range of books and course resources for smartphones and tablet computers [10]. Students can access courses online, download materials, and participate in group discussions. Educators can use the platform to create courses, ask questions, assign homework, and track students' engagement and performance with analytics. It effectively extends traditional classroom teaching by providing opportunities for pre- and post-class knowledge expansion, making it an essential tool for building flipped classroom environments [10]. Taking 'Civil Engineering Testing Techniques' as an example, the construction of a flipped classroom using Superstar Learning Hub involves the measures depicted in Figure 1.

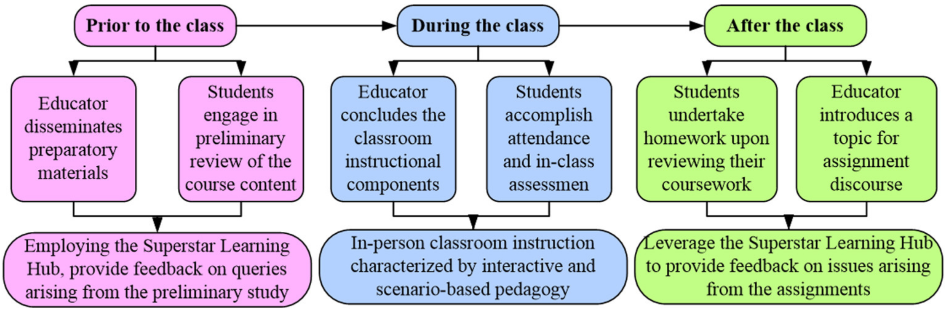

Fig. 1. Flipped classroom construction based on superstar learning hub

3. An comprehensive assessment model

To address the inherent hierarchical and fuzzy nature of assessing flipped classroom teaching quality, this study combines the analytic hierarchy process and fuzzy evaluation method. This integration results in a comprehensive evaluation model for flipped classroom teaching quality, harnessing the strengths of both methods. The analytic hierarchy process breaks down the problem into multiple levels, improving understanding and problem-solving [11]. Meanwhile, fuzzy comprehensive evaluation deals with ambiguous or uncertain data. This combination allows the transformation of fuzzy data into comparable values while hierarchically assessing the significance of various factors [12]. This approach enhances the accuracy, effectiveness, and rationality of assessing the educational outcomes of flipped classroom instruction.

3.1. Analytic hierarchy process

The analytic hierarchy process is a multi-objective decision analysis methodology that combines both quantitative and qualitative aspects. This approach proves highly practical in cases where the objective structure is intricate and essential data is lacking [13]. It is particularly suitable for addressing problems that are fuzzy and difficult to quantify. The primary steps involved in this process are outlined as follows.

3.1.1. Modeling of hierarchical structures

Decomposing a complex decision problem involves breaking it down into individual factors, hierarchicalizing them based on their attributes and relationships. Factors in the higher level exert dominance over factors in the lower level, while factors in the lower level exert influence over factors in the higher-level associations. The uppermost level is the target layer, representing the decision goal of a problem. The intermediate layer consists of indicators and other requisites for achieving the objectives, often referred to as the indicator layer. The lowest layer is typically termed the alternative layer or solution layer, encompassing various alternative options and measures aligned with the objectives, thus also known as the measure layer [14].

3.1.2. Establishing pairwise comparison matrix

For the factors $X = [x_1, x_2, \ldots, x_n]$ in the measure layer B that are subordinates to a factor in the indicator layer A, the pairwise comparison matrix A is constructed as follows.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{3.1}$$

where a_{ij} represents the importance ratio between x_i and x_j and the importance ratio between x_j and x_i is $a_{ji} = 1/a_{ij}$. When i equals j, $a_{ij} = a_{ij} = 1$. The values of a_{ij} can be determined using 1–9 ratio scale method for importance levels [15], as outlined in Table 1.

Table 1. 1–9 ratio scale method for importance levels

Significance of Importance	a_{ij}	a_{ji}
x_i and x_j are equal	1	1
x_i slight more important than x_j	3	1/3
x_i moderately more important than x_j	5	1/5
x_i significantly more important than x_j	7	1/7
x_i absolutely more important than x_j	9	1/9
The intermediate value between adjacent levels of importance	2, 4, 6, 8	1/2, 1/4, 1/6, 1/8,

3.1.3. Hierarchical single ordering and consistency checking

Hierarchical single ranking involves calculating the largest eigenvalue $\lambda_{\max}$ of matrix A, followed by normalization to derive the corresponding eigenvector w_i (with its elements summing to 1). This normalized eigenvector w_i can then serve as the weight coefficients for ranking the importance of factors in the current layer with respect to a specific factor in the upper layer.

Utilizing the root method to obtain the matrix weight vector involves calculating the initial weight vector w_i'. By undergoing a normalization process, the normalized eigenvector w_i is derived. As the condition $a_{ij}a_{jk} = a_{ik}$ is satisfied and $\forall i, j, k = 1, 2, \ldots, n$, a positive reciprocal matrix is considered a consistent matrix. Deviations from consistency are acceptable only within a certain range. Therefore, a consistency check is necessary.

Conduct the consistency check for matrix A according to Eq. (3.2).

$$CR = \frac{CI}{RI} \tag{3.2}$$

where *CI* represents the random consistency ratio. When *CI* is less than 0.1, the consistency check is considered satisfactory, and the weight coefficient distribution is deemed rational. Otherwise, adjustments to matrix element values are necessary, leading to the reassignment of weight coefficient values. *CI* represents the consistency index, and *RI* signifies the average random consistency indicator.

According to the theorem, it is known that the greater the matrix order n, the more severe the inconsistency in matrix A, in relation to the maximum eigenvalue $\lambda_{\max}$. Consequently, compute the consistency index *CI* in accordance with Eq. (3.3).

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{3.3}$$

The corresponding relationship between n and *RI* values is obtained through extensive computation, as depicted in Table 2.

Table 2. Correspondence of n and *RI* values

n	10	9	8	7	6	5	4	3	2	1
RI	1.49	1.45	1.41	1.32	1.24	1.12	0.90	0.58	0	0

3.1.4. Hierarchical total ordering and consistency checking

Hierarchical total ranking involves establishing the importance ranking weight values of all factors within a certain layer with respect to the overall objective, progressing in sequence from the uppermost layer to the lowest. The highest-level overall objective is denoted as Z. Within the indicator layer A, there exist m factors $A_1, A_2, \ldots, A_m$. Their respective rankings concerning the overarching objective Z are represented as $a_1, a_2, \ldots, a_m$. In the measure layer B, there are n factors pertaining to each factor in the indicator layer A_j. These single-layer rankings are denoted as $b_{1j}, b_{2j}, \ldots, b_{nj}$ $(j = 1, 2, 3, \ldots, m)$.

The weight of the i factor in the action layer concerning the overall objective is represented by $\sum_{j=1}^{m} a_j b_{ij}$, the calculation of overall ranking consistency follows Eq. (3.4).

$$CR = \frac{\sum_{j=1}^{m} CI(j) a_j}{\sum_{j=1}^{m} RI(j) a_j} \tag{3.4}$$

where $CI(j)$, $j = 1, \ldots, m$ denotes the consistency index for single ranking, and $RI(j)$, $j = 1, \ldots, m$ represents the random consistency index. Similarly, only when CR is determined to be less than 0.1 can the hierarchical total ranking pass the consistency check. Consequently, the standardized overall eigenvector $W = [w_1, w_2, \ldots, w_n]$ can be obtained.

3.2. Fuzzy evaluation method

The fuzzy evaluation methodology employs the foundational principles of fuzzy mathematics to quantitatively assess ambiguous and uncertain phenomena in the real world, thus facilitating objective, accurate, and realistic evaluations [16]. The primary steps in applying the fuzzy evaluation method are as follows.

3.2.1. Clarifying the factor set

The factor set is a collection of primary factors that can be used to evaluate objects, represented as $U = [u_1, u_2, \ldots, u_m]$, where m is the number of evaluative factors and u_i is the i evaluative factor. Depending on the specific context, evaluative factors can be categorized into different attributes, including the first-level evaluative factor set, subsidiary second-level evaluative factor sets, and even third-level evaluative factor sets. These factors typically exhibit varying degrees of fuzziness.

3.2.2. Establishing comprehensive set of judgments

The judgment set is a collection of potential judgment outcomes for evaluative objects, denoted as $V = [v_1, v_2, \ldots, v_n]$, where n represents the number of judgment outcomes, and v_j stands for the j judgment outcome.

3.2.3. Determining factor weight vectors

Due to the varying importance of each factor, which is reflected in different weights, let the weights of each factor u_i be denoted as w_i. Consequently, the fuzzy set of weight collection for each factor can be represented as $W = [w_1, w_2, \ldots, w_n]$. This study employs the analytic hierarchy process to compute the factor weight vector.

3.2.4. Single-factor fuzzy evaluation obtaining the judgment matrix

Single-factor fuzzy evaluation begins by establishing the degree of membership of evaluative objects to the judgment set V, based on a single factor. Let r_{ij} represent the degree of membership of the i element in $U = [u_1, u_2, \ldots, u_m]$ to the first element in the judgment set $V = [v_1, v_2, \ldots, v_n]$. Consequently, the result of single-factor judgment for the i element can be denoted as $R_i = [r_{i1}, r_{i2}, \ldots, r_{im}]$. When m sets of single-factor judgments $R_1, R_2, \ldots, R_m$ are arranged as rows, the resulting matrix R constitutes the fuzzy comprehensive judgment matrix.

3.2.5. Multi-indicator synthesis judgment

For the given factor weight vector W and matrix R, the fuzzy vector W is transformed from the factor set U to the judgment set V using a weighted average type fuzzy operator $M(\cdot, \oplus)$ [17]. This results in the fuzzy vector B, as illustrated in Eq. (3.5).

$$B = W \cdot R = (w_1, w_2, \cdots w_n) \cdot \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \tag{3.5}$$

where B represents the degrees of membership of various factors within the judgment set. According to the principle of maximum membership, the judgment result is determined by selecting the maximum value.

3.2.6. Synthesizing the results of judging

Based on the degrees of membership of various factors within the judgment set and the scores of each indicator, the final fuzzy comprehensive evaluation result is obtained.

$$Q = B \cdot S^T \tag{3.6}$$

where S represents the score vector corresponding to the assessment set, categorized into levels such as Poor, Weak, Medium, Good, and Excellent, with corresponding scores of 30, 50, 60, 80, and 100. Therefore, the $S = [30\ 50\ 60\ 80\ 100]$.

From the aforementioned description, it is evident that the combination of the analytic hierarchy process (AHP) and fuzzy evaluation method yields the fuzzy hierarchical comprehensive evaluation model for flipped classroom teaching quality. This is depicted in Fig. 2.

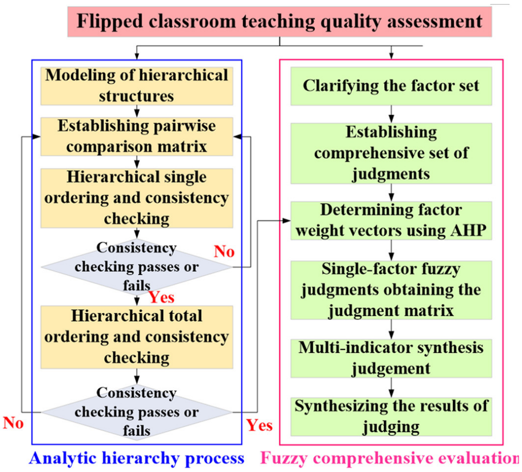

Fig. 2. Fuzzy hierarchical comprehensive evaluation model

4. Case study of flipped classroom teaching quality evaluation

4.1. Figures defining evaluation indicators and formulating the set of evaluative factors

The initial step in creating a comprehensive evaluation model for flipped classroom teaching quality is defining the evaluation indicator system and its sub-indicators. Success in teaching quality evaluation depends on the quality of this indicator system. To establish this system, interviews and questionnaires were conducted with 4 flipped classroom instructors and 208 students at Anhui University of Science and Technology, specifically in the civil engineering major, using the Superstar Learning Hub. Interview records were analyzed and consolidated, resulting in the identification of the primary factors influencing flipped classroom teaching quality. These primary factors are as follows: classroom teaching quality, extracurricular learning quality, flipped classroom infrastructure, and participant behavioral performance. Using the Likert scale, secondary evaluation factors were further determined under each primary factor along with their corresponding relationships [18], as presented in Table 3.

Classroom teaching is a central aspect of the flipped classroom process, with classroom teaching quality as the primary evaluation factor. Within this factor, four secondary evaluation criteria are identified: Classroom Teaching Methods: This evaluates the innovative teaching activities carried out by instructors in the classroom, focusing on student engagement. Classroom Discussion Atmosphere: This assesses the level of student participation and enthusiasm during in-person teaching interactions in the classroom. Knowledge Absorption in the Classroom: This measures students' understanding and acceptance of lecture content and their ability to apply it effectively. Classroom Performance Assessment: This encompasses various assessment methods, including in-class quizzes, attendance records, and written exams.

Table 3. Evaluation indicators for flipped classroom teaching quality

Factors	Primary judgment factors	Secondary judgment factors
U Flipped classroom teaching quality	u_1 Classroom teaching quality	u_{11} Classroom teaching methods
		u_{12} Classroom discussion atmosphere
		u_{13} Knowledge absorption in the classroom
		u_{14} Classroom performance assessment
	u_2 Extracurricular learning quality	u_{21} Extracurricular learning mode
		u_{22} Extracurricular teacher-student interaction
		u_{23} Self-disciplined extracurricular learning
		u_{24} Extracurricular performance assessment
	u_3 Flipped classroom infrastructure	u_{31} Teacher-side hardware facilities
		u_{32} Student-side hardware facilities
		u_{33} Flip platform operation
	u_4 Participant behavioral performance	u_{41} Teacher teaching behaviors
		u_{42} Student learning behaviors
		u_{43} Platform operator engagement behaviors

Extracurricular learning is equally important in the flipped classroom teaching process, with Extracurricular Learning quality as the primary evaluation factor. Within this factor, four secondary criteria are established: Extracurricular Learning Mode: This evaluates the methods students employ for independent pre-study of course content outside the classroom. Extracurricular Teacher-Student Interaction: This assesses the interaction and communication between teachers and students on the Superstar Learning platform outside the physical classroom. Self-Disciplined Extracurricular Learning: This measures students' ability to efficiently manage their extracurricular time and complete their learning tasks with quality and quantity. Extracurricular Performance Assessment: This includes the recording of task completion on the Superstar Learning platform.

Teaching with internet technology in flipped classrooms requires the use of hardware equipment, including devices like computers, tablets, mobile phones, and other electronic devices. This applies to both teacher and student hardware facilities. Additionally, in the flipped classroom instruction process, the use of interactive platforms like the Superstar Learning Hub involves aspects such as interface design, user operations, feedback mechanisms, server capacity, and other critical factors that influence the teaching experience.

In flipped classroom teaching, the focus is on the participation and behaviors of three key parties: teachers, students, and platform operators. Teachers: They should be well-prepared, display enthusiasm during lessons, employ clear teaching strategies, have proficiency in using the platform, and be patient in addressing student inquiries. Students: They are expected to engage in independent learning outside of class, actively participate in discussions, and approach learning as an interest-driven pursuit. Platform Operators: Their responsibilities include continually improving the interactive interface of the Superstar Learning Hub, promptly addressing any potential bugs, and swiftly resolving software-related issues encountered by teachers and students during usage.

4.2. Establishing the fuzzy comprehensive evaluation set

In this study, the evaluation results for flipped classroom teaching quality are categorized into five levels: Poor, Weak, Medium, Good, and Excellent. This classification is represented as the judgment set $V = [\text{ Poor } v_1,\ \text{Weak } v_2,\ \text{Medium } v_3,\ \text{Good } v_4,\ \text{Excellent } v_5]$. Furthermore, a correspondence between the levels of fuzzy comprehensive evaluation and the corresponding score ranges is provided, as illustrated in Table 4.

Table 4. Correspondence between score values and fuzzy comprehensive evaluation levels

Score range	< 50	50–59	60–79	80–89	90–100
Grade	Poor	Weak	Medium	Good	Excellent

4.3. Obtaining the judgment matrix and determining indicator weights

In assessing flipped classroom teaching quality, indicator weights vary and are determined using the analytic hierarchy process (AHP). This method minimizes subjectivity through rigorous mathematical analysis. The weights' rationality is evaluated by ensuring consistency in the judgment matrix, aiming to align them with objective reality for quantitative representation, thereby enhancing the reliability, accuracy, and objectivity of fuzzy comprehensive evaluation. Initially, a 1–9 scale scores primary indicators to create pairwise comparison matrices. Consistency is assessed, and the primary indicator factor weight vector is obtained (Table 5). The same approach is then used to calculate the weights for secondary indicators corresponding to the four consistent criteria, presented in Tables 6, 7, 8, and 9. All these weights have passed the consistency check.

4.4. Fuzzy comprehensive evaluation matrix

The implementation of the flipped classroom teaching method in the field of civil engineering at Anhui University of Science and Technology has been introduced only in recent years. Among these implementations, the use of the Superstar Learning Hub platform has been the most prominent, with a duration of no more than four years. Other platforms for flipped classroom instruction have been in operation for approximately two year. Therefore, for this study, a questionnaire survey was exclusively conducted among instructors and students engaged in flipped classroom instruction through the Superstar Learning Hub. The survey utilized the fuzzy comprehensive evaluation framework established in Section 4.2 and yielded a total of 212 valid responses. The organized results are presented in Table 10.

Table 5. AHP calculation results for primary indicator

Indicators	u_1	u_2	u_3	u_4	Weight	$\lambda_{\max}$	4.196
u_1	1	3	1/3	1/4	0.126	*CI*	0.065
u_2	1/3	1	1/8	1/6	0.052	*RI*	0.900
u_3	3	8	1	3	0.521	*CR*	0.073
u_4	4	6	1/3	1	0.301	Consistency checking passed	

Table 6. AHP calculation results for secondary indicator (classroom teaching quality)

Indicators	u_{11}	u_{12}	u_{13}	u_{14}	Weight	$\lambda_{\max}$	4.241
u_{11}	1	4	2	3	0.455	*CI*	0.080
u_{12}	1/4	1	1/3	1/4	0.078	*RI*	0.900
u_{13}	1/2	3	1	3	0.299	*CR*	0.089
u_{14}	1/3	4	1/3	1	0.168	Consistency checking passed	

Table 7. AHP calculation results for secondary indicator (extracurricular learning quality)

Indicators	u_{21}	u_{22}	u_{23}	u_{24}	Weight	$\lambda_{\max}$	4.238
u_{21}	1	2	2	3	0.397	*CI*	0.079
u_{22}	1/2	1	2	4	0.301	*RI*	0.900
u_{23}	1/2	1/2	1	5	0.225	*CR*	0.088
u_{24}	1/3	1/4	1/5	1	0.077	Consistency checking passed	

Table 8. AHP calculation results for secondary indicator (flipped classroom infrastructure)

Indicators	u_{31}	u_{32}	u_{33}	$\lambda_{\max}$	3.009
u_{31}	1	1	1/3	*CI*	0.005
u_{32}	1	1	1/4	*RI*	0.580
u_{33}	3	4	1	*CR*	0.008
Weight	0.192	0.174	0.634	Consistency checking passed	

Table 9. AHP calculation results for secondary indicator (participant behavioral performance)

Indicators	u_{41}	u_{42}	u_{43}	$\lambda_{\max}$	3.074
u_{41}	1	1/3	3	*CI*	0.037
u_{42}	3	1	4	*RI*	0.580
u_{43}	1/3	1/4	1	*CR*	0.063
Weight	0.268	0.614	0.118	Consistency checking passed	

Table 10. Flipped classroom teaching quality scores

Judgment factors		Fuzzy comprehensive evaluation set				
		Poor	Weak	Medium	Good	Excellent
u_1	u_{11}	4	8	102	85	13
	u_{12}	7	23	145	26	11
	u_{13}	3	10	40	113	46
	u_{14}	3	7	46	119	37
u_2	u_{21}	1	4	184	19	4
	u_{22}	48	34	86	40	4
	u_{23}	11	41	137	15	8
	u_{24}	8	20	54	88	42
u_3	u_{31}	0	1	25	53	133
	u_{32}	1	4	16	35	156
	u_{33}	2	1	32	92	85
u_4	u_{41}	3	10	54	71	74
	u_{42}	2	7	42	104	57
	u_{43}	14	26	67	84	21

The scores obtained for various indicators of flipped classroom teaching quality in Table 10 are normalized to establish four sub-level evaluation matrices using the method of normalization. These matrices are presented below.

$$R_1 = \begin{bmatrix} 0.019 & 0.038 & 0.481 & 0.401 & 0.061 \\ 0.033 & 0.108 & 0.684 & 0.123 & 0.052 \\ 0.014 & 0.047 & 0.189 & 0.533 & 0.217 \\ 0.014 & 0.033 & 0.217 & 0.561 & 0.175 \end{bmatrix} \tag{4.1}$$

$$R_2 = \begin{bmatrix} 0.005 & 0.019 & 0.868 & 0.090 & 0.019 \\ 0.226 & 0.160 & 0.406 & 0.189 & 0.019 \\ 0.052 & 0.193 & 0.646 & 0.071 & 0.038 \\ 0.038 & 0.094 & 0.255 & 0.415 & 0.198 \end{bmatrix} \tag{4.2}$$

$$R_3 = \begin{bmatrix} 0.000 & 0.005 & 0.118 & 0.250 & 0.627 \\ 0.005 & 0.019 & 0.075 & 0.165 & 0.736 \\ 0.009 & 0.005 & 0.151 & 0.434 & 0.401 \end{bmatrix} \tag{4.3}$$

$$R_4 = \begin{bmatrix} 0.014 & 0.047 & 0.255 & 0.335 & 0.349 \\ 0.009 & 0.033 & 0.198 & 0.491 & 0.269 \\ 0.066 & 0.123 & 0.316 & 0.396 & 0.099 \end{bmatrix} \tag{4.4}$$

4.5. Computation of fuzzy comprehensive evaluation results

The comprehensive evaluation results are obtained from Eq. (3.5), where w_i represents the vector of sub-level indicator weights. According to the characteristics of the operator, the fuzzy transformation composition operator $M(\cdot, \oplus)$ is employed [19]. The integrated evaluation results for primary indicators u_1, u_2, u_3, and u_4 are as follows.

$$\begin{aligned} B_1 &= w_1 \cdot R_1 = \\ &= [0.455\ 0.078\ 0.299\ 0.168] \cdot \begin{bmatrix} 0.019 & 0.038 & 0.481 & 0.401 & 0.061 \\ 0.033 & 0.108 & 0.684 & 0.123 & 0.052 \\ 0.014 & 0.047 & 0.189 & 0.533 & 0.217 \\ 0.014 & 0.033 & 0.217 & 0.561 & 0.175 \end{bmatrix} = \\ &= [0.018\ 0.045\ 0.365\ 0.446\ 0.126] \end{aligned} \tag{4.5}$$

$$\begin{aligned} B_2 &= w_2 \cdot R_2 = \\ &= [0.397\ 0.301\ 0.225\ 0.077] \cdot \begin{bmatrix} 0.005 & 0.019 & 0.868 & 0.090 & 0.019 \\ 0.226 & 0.160 & 0.406 & 0.189 & 0.019 \\ 0.052 & 0.193 & 0.646 & 0.071 & 0.038 \\ 0.038 & 0.094 & 0.255 & 0.415 & 0.198 \end{bmatrix} = \\ &= [0.085\ 0.107\ 0.632\ 0.140\ 0.037] \end{aligned} \tag{4.6}$$

$$\begin{aligned} B_3 &= w_3 \cdot R_3 = \\ &= [0.192\ 0.174\ 0.634] \cdot \begin{bmatrix} 0.000 & 0.005 & 0.118 & 0.250 & 0.627 \\ 0.005 & 0.019 & 0.075 & 0.165 & 0.736 \\ 0.009 & 0.005 & 0.151 & 0.434 & 0.401 \end{bmatrix} = \\ &= [0.007\ 0.007\ 0.131\ 0.352\ 0.503] \end{aligned} \tag{4.7}$$

$$\begin{aligned} B_4 &= w_4 \cdot R_4 = \\ &= [0.268\ 0.614\ 0.118] \cdot \begin{bmatrix} 0.014 & 0.047 & 0.255 & 0.335 & 0.349 \\ 0.009 & 0.033 & 0.198 & 0.491 & 0.269 \\ 0.066 & 0.123 & 0.316 & 0.396 & 0.099 \end{bmatrix} = \\ &= [0.017\ 0.047\ 0.227\ 0.438\ 0.270] \end{aligned} \tag{4.8}$$

Finally, a comprehensive assessment is conducted based on the weights of the primary indicators, as demonstrated below.

$$\begin{aligned} B &= W \cdot R = \\ &= [0.126\ 0.052\ 0.521\ 0.301] \cdot \begin{bmatrix} 0.018 & 0.045 & 0.365 & 0.446 & 0.126 \\ 0.085 & 0.107 & 0.632 & 0.140 & 0.037 \\ 0.007 & 0.007 & 0.131 & 0.352 & 0.503 \\ 0.017 & 0.047 & 0.227 & 0.438 & 0.270 \end{bmatrix} = \\ &= [0.015\ 0.029\ 0.216\ 0.379\ 0.361] \end{aligned} \tag{4.9}$$

From the computed results, it can be observed that in this study, the assessment of flipped classroom teaching quality conducted through the Superstar Learning Hub platform indicates a ‘Good’ rating of 37.9%. Applying the principle of maximum membership degree, the evaluation of flipped classroom teaching quality for the civil engineering major at Anhui University of Science and Technology should be classified as ‘Good’.

4.6. Computation of fuzzy comprehensive evaluation scores

To further quantify the fuzzy evaluation results, and in accordance with the grading values specified in Section 4.2, utilizing Eq.(3.6) not only yields the fuzzy comprehensive evaluation score, but also provides the scores for various secondary factors under the primary judgment factor. This is demonstrated below.

$$Q = B \cdot S^T = [0.015\ 0.029\ 0.216\ 0.379\ 0.361] \cdot [3\ 50\ 60\ 80\ 100]^T = 81.259 \tag{4.10}$$

$$Q_1 = B_1 \cdot S^T = [0.018\ 0.045\ 0.365\ 0.446\ 0.126] \cdot [30\ 50\ 60\ 80\ 100]^T = 72.973 \tag{4.11}$$

$$Q_2 = B_2 \cdot S^T = [0.085\ 0.107\ 0.632\ 0.140\ 0.037] \cdot [30\ 50\ 60\ 80\ 100]^T = 60.678 \tag{4.12}$$

$$Q_3 = B_3 \cdot S^T = [0.007\ 0.007\ 0.131\ 0.352\ 0.503] \cdot [30\ 50\ 60\ 80\ 100]^T = 86.869 \tag{4.13}$$

$$Q_4 = B_4 \cdot S^T = [0.017\ 0.047\ 0.227\ 0.438\ 0.270] \cdot [30\ 50\ 60\ 80\ 100]^T = 78.572 \tag{4.14}$$

4.7. Results of case study

This study employed both the analytic hierarchy process and fuzzy evaluation method to assess flipped classroom teaching quality, resulting in an overall assessment falling into the ’Good’ category. Among the primary judgment factors, hardware facilities for flipped classrooms achieved a ’Good’ level, while classroom teaching quality, extracurricular learning quality, and participant behavioral performance were rated as ’Medium’. In terms of scores, extracurricular learning quality had the lowest score and was close to a lower rating threshold, suggesting the need for future attention in this aspect of flipped classroom teaching. These evaluation results align with the current state of flipped classroom teaching using the Superstar Learning Hub platform in civil engineering at Anhui University of Science and Technology. While the current rating is ’Good’, the scores are at the lower end of the ’good’ range in the fuzzy evaluation. Therefore, it is advisable to focus more on extracurricular learning quality and classroom teaching quality in the future.

This study tackles the hierarchical and fuzzy nature of assessing flipped classroom teaching quality using a comprehensive evaluation model, and the results have the following characteristics [20]. Firstly, the evaluation conclusions are highly reliable. Both the analytic hierarchy process and fuzzy evaluation method are effective in handling imprecise and ambiguous issues. Among these methods, the fuzzy assessment approach excels in comprehensive judgment. When combined with indicator weights determined through the analytic hierarchy process, the fuzzy assessment gains greater scientific validity. Secondly, the evaluation conclusions are versatile. The results for flipped classroom teaching quality can be used to establish best practices and recognize achievements. Single-indicator evaluations can help identify areas of strength and areas that need improvement, enabling targeted suggestions to enhance teaching methods and improve teaching quality.

5. Discussion and implications of the study

This study found that extracurricular learning performance is significantly affected by the mode of learning, teacher-student interaction, and self-disciplined learning. The lower performance in extracurricular learning may be due to challenges in adapting to the flipped classroom model. Difficulties in adjusting study habits, particularly when the flipped classroom is new and lacks experience, negatively impact student satisfaction, as students have to rely on their own experimentation and exploration.

While many survey respondents excelled in extracurricular performance assessments, subtle signs suggest that some students in the flipped classroom exhibit unfavorable study habits, such as procrastination and inattentiveness during instructional videos. These habits may stem from the novelty of the flipped classroom model or existing learning issues. What's concerning is that these students still achieve high assessment scores, which could undermine their motivation for academic improvement. Therefore, there's a need for a fundamental shift in students' attitudes toward extracurricular learning. They should recognize that it's a lifelong endeavor, particularly in self-directed learning, which not only benefits their current academics but also significantly influences future career prospects through proactive education.

In assessing the flipped classroom, hardware facilities take the top spot in importance, followed by the operational efficiency of the online platform, which holds a significant coefficient of 0.634. This highlights the critical role of selecting the right online instructional platform in flipped classroom pedagogy. In this study, using both the desktop version and the mobile app of Superstar Learning platform enhances convenience for both educators and students. However, for optimal results, the platform should offer a user-friendly interface, enabling easy input of complex content like formulas and diagrams, along with real-time audio interaction. The platform's video and audio resources must also be of the highest quality.

Analyzing indicator importance in flipped classroom assessment identifies areas needing priority attention for pedagogical improvement. To enhance instructional quality, focus on both low-scoring and high-weight indicators. This strategic approach identifies ways to improve pedagogical quality and effectiveness, providing a basis for measures to elevate flipped classroom instruction comprehensively.

6. Conclusions

The flipped classroom teaching method has emerged as a significant trend in contemporary higher education. In the field of civil engineering, an increasing number of institutions are adopting this pedagogical approach. The quality of education plays a pivotal role in determining the extent to which civil engineering students grasp and absorb knowledge. Therefore, the quality of flipped classroom instruction directly impacts the cultivation of civil engineers.

This study introduces a comprehensive evaluation model for assessing the quality of flipped classroom instruction using Superstar Learning Hub. The model, which combines the analytic hierarchy process and fuzzy evaluation method, addresses hierarchy and ambiguity challenges. It includes four primary evaluation factors: classroom teaching quality, extracurricular learning quality, flipped classroom infrastructure, and participant behavioral performance, each with secondary clusters totaling fourteen aspects. The evaluation for the quality of flipped classroom instruction in civil engineering at Anhui University of Science and Technology is classified as 'Good' using the principle of maximum membership degree within a fuzzy framework. The study also proposes measures and recommendations to enhance the practice of flipped classroom instruction based on fuzzy comprehensive assessment scores and factor weights.

Acknowledgements

The underlying project was funded by the Anhui Provincial Higher Education Quality Engineering Provincial Teaching Research Project (2021jyxm0375, 2019jyxm1126, 2022jyxm400); the Research Project of Education and Teaching Reform of Anhui University of Science and Technology (2021xjjy13, 2018xjjy058); the Anhui Provincial Quality Engineering Online and Offline Mixed Curriculum Project (2021xsxxkc070) and the Cooperative Education Project of Industry-University Cooperation of the Ministry of Education of China (202101353001, 202102234039).

References

[1] K. Gavin, "Case study of a project-based learning course in civil engineering design", *European Journal of Engineering Education*, vol. 36, no. 6, pp. 547–558, 2011, doi: 10.1080/03043797.2011.624173.

[2] H. Zobel, A. Pawlak, M. Pawlik, P. Żółtowski, R. Czubacki, and T. Al-Khafaji, "Hyperloop-Civil engineering point of view according to Polish experience", *Archives of Civil Engineering*, vol. 68, no. 1, pp. 5–26, 2022, doi: 10.24425/ace.2022.140153.

[3] J. Szelka, "Information and decision modelling in civil engineering", *Archives of Civil Engineering*, vol. 68, no. 4, pp. 97–110, 2022, doi: 10.24425/ace.2022.143028.

[4] Y. Doğan, V. Batdı, and M.D. Yaşar, "Effectiveness of flipped classroom practices in teaching of science: a mixed research synthesis", *Research in Science & Technological Education*, vol. 41, no. 1, pp. 393–421, 2023, doi: 10.1080/02635143.2021.1909553.

[5] K. Missildine, R. Fountain, L. Summers, and K. Gosselin, "Flipping the classroom to improve student performance and satisfaction", *Journal of Nursing Education*, vol. 52, no. 10, pp. 597–599, 2013, doi: 10.3928/01484834-20130919-03.

[6] J. O'Flaherty and C. Phillips, "The use of flipped classrooms in higher education: A scoping review", *The Internet and Higher Education*, vol. 25, pp. 85–95, 2015, doi: 10.1016/j.iheduc.2015.02.002.

[7] M.C. Schlairet, R. Green, and M.J. Benton, "The flipped classroom: strategies for an undergraduate nursing course", *Nurse Educator*, vol. 39, no. 6, pp. 321–325, 2014, doi: 10.1097/NNE.0000000000000096.

[8] H.M. Bingen, S.A. Steindal, R. Krumsvik, and B. Tveit, "Nursing students studying physiology within a flipped classroom, self-regulation and off-campus activities", *Nurse Education in Practice*, vol. 35, pp. 55–62, 2019, doi: 10.1016/j.nepr.2019.01.004.

[9] S. Bhat, R. Raju, S. Bhat, and R. D'Souza, "Redefining quality in engineering education through the flipped classroom model", *Procedia Computer Science*, vol. 172, pp. 906–914, 2020, doi: 10.1016/j.procs.2020.05.131.

[10] W.P. Xue, Z.S. Yao, and X.J. Wang, "Teaching exploration of civil engineering testing technology course in the context of new engineering", *Journal of Changchun Normal University*, vol. 40, no. 6, pp. 148–150, 2021.

[11] R.W. Saaty, “The analytic hierarchy process-what it is and how it is used”, *Mathematical Modelling*, vol. 9, no. 3-5, pp. 161–176, 1987, doi: 10.1016/0270-0255(87)90473-8.

[12] W.Q. Shao, “Evaluation of international port city based on fuzzy comprehensive evaluation”, *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 6, pp. 7027–7032, 2020, doi: 10.3233/JIFS-179780.

[13] A.F. Shapiro and M.C. Koissi, “Fuzzy logic modifications of the analytic hierarchy process”, *Insurance: Mathematics and Economics*, vol. 75, pp. 189–202, 2017, doi: 10.1016/j.insmatheco.2017.05.003.

[14] S. Sipahi and M. Timor, “The analytic hierarchy process and analytic network process:an overview of applications”, *Management Decision*, vol. 48, no. 5, pp. 775-808, 2010, doi: 10.1108/00251741011043920.

[15] A. Özdağoğlu and G. Özdağoğlu, “Comparison of AHP and fuzzy AHP for the multi-criteria decision making processes with linguistic evaluations”, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 6, no. 11, pp. 65–85, 2007.

[16] D. Tadic, A. T. Gumus, S. Arsovski, A. Aleksic, and M. Stefanovic, “An evaluation of quality goals by using fuzzy AHP and fuzzy TOPSIS methodology”, *Journal of Intelligent & Fuzzy Systems*, vol. 25, no. 3, pp. 547–556, 2013, doi: 10.3233/IFS-120659.

[17] C. Wang, R.P. Agarwal, and D. O’Regan, “Calculus of fuzzy vector-valued functions and almost periodic fuzzy vector-valued functions on time scales”, *Fuzzy Sets and Systems*, vol. 375, pp. 1–52, 2019, doi: 10.1016/j.fss.2018.12.008.

[18] A.T. Jebb, V. Ng, and L. Tay, “A review of key Likert scale development advances: 1995–2019”, *Frontiers in Psychology*, vol. 12, art. no. 637547, 2021, doi: 10.3389/fpsyg.2021.637547.

[19] P. Sussner, “Lattice fuzzy transforms from the perspective of mathematical morphology”, *Fuzzy Sets and Systems*, vol. 288, pp. 115–128, 2016, doi: 10.1016/j.fss.2015.09.018.

[20] Y. Owusu-Agyeman, O. Larbi-Siaw, B. Brenya, and A. Anyidoho, “An embedded fuzzy analytic hierarchy process for evaluating lecturers’ conceptions of teaching and learning”, *Studies in Educational Evaluation*, vol. 55, pp. 46–57, 2017, doi: 10.1016/j.stueduc.2017.07.001.