International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

Developing a Classroom Teaching Quality Evaluation Index System in the Context of Digital Transformation: Based on Delphi and AHP

[1], Hao Rongxia, [2]Amando C. Yutuc,

[1], Graduate School, Angeles University Foundation, Philippines; NIC, Shandong Second Medical University, China
hao.rongxia@auf.edu.ph
[2] Graduate School, Angeles University Foundation, Philippines; DepEd of Philippines
Amando.yutuc@auf.edu.ph

Article History
Receive Date: 2025/3/10
Revise Date: 2025/5/22
Accept Date: 2025/5/26
Publish Date: 2025/5/26

Abstract

Amid the wave of digital transformation, classroom teaching in universities is experiencing innovations in educational concepts, teaching methods, and learning styles. To adapt to these changes, the traditional evaluation index system for classroom teaching quality urgently needs updating. This study aimed to construct an evaluation index system for classroom teaching quality that meets the requirements of digital transformation, thereby enhancing the scientific and effective nature of evaluations. By comprehensively utilizing literature review, policy analysis, and specialized group discussions, a preliminary evaluation index questionnaire was drafted. The Delphi method was then employed, inviting experts in the field to evaluate and revise the questionnaire items. After two rounds of expert consultations, an evaluation system including 6 primary indicators and 32 secondary indicators was established. The Analytic Hierarchy Process (AHP) was used to determine the weight of each indicator, ensuring the validity and effectiveness of the evaluation system. This system had received high recognition from the experts, indicating a high level of consensus, and demonstrating excellent scientific reliability. The findings from this study not only shed light on the practice of teaching quality evaluation but also paved the way for reforming classroom teaching at universities.

Keywords: *Digital transformation; Classroom teaching quality; Evaluation indicators; Delphi method; AHP*

Introduction

As the core pathway for talent cultivation, the quality of classroom teaching directly impacts educational outcomes. Effective classroom teaching evaluation can not only timely identify and resolve issues in the teaching process, but also be crucial for enhancing teaching quality. Classroom teaching quality evaluation involves making value judgments on the teaching process and its outcomes based on educational goals and relevant standards, using teaching evaluation theories and technical means. This process includes a comprehensive consideration of multiple dimensions such as teaching attitude, content, methods, and effects. By providing feedback on teaching information, it guides teachers and students to optimize subsequent teaching activities (Chen, 1999). Luo and Guo (2020) further pointed out that the evaluation process

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

should cover multiple aspects of teaching to ensure comprehensiveness.

In 2020, the Central Committee of the Communist Party of China and the State Council issued the "Overall Plan for Deepening the Reform of Education Evaluation in the New Era," which emphasized the pivotal role of education evaluation(Zhang,2024) Ashfaq, R., & Nabi, Z. (2022). It called for attention to students' overall development, emphasis on the teaching process, encouragement of teaching innovation, advocacy for personalized teaching, strengthening of teachers' professional development, and improvement of outcome evaluation, process evaluation, value-added evaluation, comprehensive evaluation, and innovative evaluation tools. The plan also highlighted the full utilization of information technology to enhance the scientific, professional, and objective nature of educational evaluation. In 2022, the Ministry of Education launched the "Education Digitalization Strategy Action," marking the deep application of digital transformation in the field of education. Digital transformation has not only changed traditional teaching methods and means, optimized the sharing and allocation of educational resources, but also provided new perspectives and opportunities for the innovation of classroom teaching quality evaluation systems (Zhu, Guo, & Yi, 2023). These requirements reflect higher expectations and more refined demands for the evaluation of classroom teaching quality in colleges and universities, mainly manifested in: modern classroom teaching design not only poses new and higher requirements for teachers' lectures but also emphasizes a student-centered approach; the application of advanced evaluation methods and means in classroom teaching evaluation; the need to combine condition evaluation, process evaluation, and outcome evaluation in classroom teaching quality evaluation, focusing on the entire teaching process; and a greater emphasis on evaluating teachers' classroom teaching philosophies and improving students' comprehensive qualities (Liu & Xu, 2023). Educational evaluators must establish new concepts of classroom teaching evaluation and construct a scientific and objective indicator system for evaluating classroom teaching quality (Cao & Wang, 2023).

The evaluation index system for classroom teaching quality plays a significant guiding role in teaching practice, directing teaching towards efficiency and effectiveness. Therefore, in the context of digital transformation, it is crucial to deeply reflect on the existing classroom teaching quality evaluation index system and construct an evaluation system that meets the needs of the new era to improve educational teaching quality (Zhu et al., 2023). However, there are few data on the evaluation index systems for classroom teaching quality at universities. Guided by the fourth-generation education evaluation theory, this study adopts empirical research methods to construct a scientific and effective classroom teaching quality evaluation index system that meets the requirements of digital transformation. This system aims to provide a solid scientific basis for improving teaching methods and enhancing teaching effectiveness, while also offering practical references and theoretical guidance for evaluating teaching quality in colleges and universities.

2. Methods

2.1 The Questionnaire Item Pool

A comprehensive review of literature and policy was conducted in Mainland China and other countries or regions. The goals of higher education and the meanings of classroom teaching quality evaluation were identified and clarified. A questionnaire item pool for classroom teaching quality evaluation specifically in the context of digital transformation was then formed. Through in-depth interviews with individual experts and multiple discussions within the focused groups, a preliminary classroom teaching quality evaluation index system based on the questionnaire item pool was established. This system includes 6 primary indicators and 36 secondary indicators, as presented in Tables 2 and 3.

2.2 Indicators Refined by the Delphi Method

2.2.1 Participants

To ensure the validity and reliability of each indicator, the Delphi method was adopted. It's recommended to have 15-30 experts for the Delphi method (Hsu & Sandford, 2007); Kenawy, M., Elkhweet, S., Elatrebi, H., & Elwakil, F. (2024). By means of a purposive sampling, 20 experts with over 10 years of experience in the relevant fields from Chinese universities and research institutions were invited to participate in this survey. Their expertise includes educational management, front-line teaching, ideological and political education, psychology, and statistics.

2.2.2 The Consultation Form

The consultation form was composed of an introduction to the research context, instructions for completion, and guidelines on relevant knowledge, followed by the main survey content. The main survey content included sections on the experts' demographics, an evaluation indicator judgment form, and a survey on expert consultation reliability. The evaluation indicator judgment form employed a 5-point Likert scale, prompting experts to rate

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

the importance of each indicator and provide feedback in a "modification comments" section.

2.2.3 Data Collection

The evaluation index system was constructed and refined through 2 rounds of opinion collection. In the first round, the consultation form above was sent to the invited participants through the Questionnaire Star Platform. The participants completed the consultation form on their mobile phones or computers and submitted their responses and feedback to the platform from which the data were easily collected. The consultation form was then revised based on an analysis of the data collected from the first round. In the second round, the revised consultation form was distributed to the same group of participants along with the statistical data and responses to the controversial issues from the first round in a similar manner.

2.3 Indicator Weights Determined by the AHP

The AHP is a systematic analysis method that combines qualitative and quantitative approaches, suitable for addressing complex decision-making problems (Shen et al., 2019). The basic steps of this study included establishing a hierarchical structure model, constructing pairwise comparison matrices, calculating weight vectors and consistency testing, and computing combined weight vectors.

2.4 Statistical Analysis

Data entry and analysis were conducted using Excel 2019 (Microsoft Corporation), SPSS 25.0 (IBM Corporation), and Yaahp 10.3 (Decision Tree Software). Descriptive statistics were used to represent the demographics of the Experts. The reliability of the Delphi expert consultation was measured through the experts' activity coefficient, degree of authority, degree of opinion concentration, and degree of coordination. The statistical significance was set at $P < 0.05$. Indicators with a mean of > 3, a standard deviation of > 1, and a coefficient of variation of > 0.25 were retained (Shen et al., 2019). The experts' coordination coefficient W was deemed to indicate consensus when it exceeded 0.4, after which consultation would cease (Zeng & Cheng, 2016).

Additionally, based on the mean importance scores of each indicator from the final round of consultation, Satty's scale of importance levels was assigned to compare indicators within the same level and construct pairwise judgment matrices. The product-sum method was used to calculate the maximum eigenvector and indicator weights of the pairwise comparison matrix, while the product method was used to compute the comprehensive weights of the corresponding indicators, resulting in the overall weight coefficients. For judgment matrices with an order of $n \geq 3$, the consistency test standard requires that the consistency ratio (CR) be less than 0.1. If the CR exceeds this threshold, the scale values of the elements must be adjusted, and the matrix must be reconstructed until the consistency requirements are satisfied (Zhang et al., 2023). Satty's scale values are shown in Table 1 (Gao, 2022).

Table 1 Satty's Relative Importance Scale

Satty's Scale Value	Difference in Mean Importance Scores	Satty's Scale Value	Difference in Mean Importance Scores	Significance
1	X-Y=0.0	1		Equally Important
2	0.00<X-Y≦0.25	1/2	-0.25<X-Y≦0.00	
3	0.25<X-Y≦0.50	1/3	-0.50<X-Y≦-0.25	Slightly Important
4	0.50<X-Y≦0.75	1/4	-0.75<X-Y≦-0.50	
5	0.75<X-Y≦1.00	1/5	-1.00<X-Y≦-0.75	Moderately Important
6	1.00<X-Y≦1.25	1/6	-1.25<X-Y≦-1.00	
7	1.25<X-Y≦1.50	1/7	-1.50<X-Y≦-1.25	Strongly Important
8	1.50<X-Y≦1.75	1/8	-1.75<X-Y≦-1.50	
9	X-Y>1.75	1/9	X-Y<11.75	Extremely Important

Note: X and Y represent the mean importance scores of two different indicators at the same level.

2.5 Ethics

This study followed the principles of informed consent, confidentiality, fairness, and non-maleficence. Data were used exclusively for statistical analysis, ensuring the confidentiality of personal information. The research received approval from the Ethics Committee of Angeles University Foundation.

3. Results

3.1 Demographics of the participants

The 20 consulting experts recruited for this study possessed extensive experience in education. Regarding teaching experience, 80% of the experts had over 30 years of experience, 15%, between 21 to 30 years, and 5%, between 10 to 20 years. In terms of professional titles, 85% of the experts were professors, while 15% were

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

associate professors. Educational background shows that 45% of the experts were doctorate degree holders, 30%, master's degree holders, and 25%, bachelor's degree holders. The participants' expertise included include educational management (40%), front-line education and teaching (40%), ideological and political education (5%), psychology (5%), and statistics (10%).

3.2 Expert Activity Coefficient

The expert activity coefficient was measured by the response rate of the consultation questionnaires, reflecting the experts' attention and engagement in the research. In the first round, 20 questionnaires were distributed and 19 were returned, with a response rate of 95%. In the second round, 19 questionnaires were distributed and all 19 were returned, achieving a response rate of 100%. This indicated the high attention and active response from the experts involved in the study.

3.3 Expert Authority Degree

The degree of expert authority was measured by the authority coefficient (Cr), which is the arithmetic mean of the judgment coefficient (Ca) and the familiarity coefficient (Cs)(Li, et al.,2024). A Cr greater than 0.7 is considered statistically significant, with higher Cr values indicating greater persuasiveness of the expert group (Shen, Ou & Ou, 2021). In this study, the Cs, Ca and Cr were 0.92, 0.94 and 0.93, respectively, indicating that the expert group possessed a high authority and persuasiveness.

3.4 Expert Coordination Degree

The degree of expert coordination was evaluated using the coefficient of variation (CV) and the coordination coefficient (W). The CV measures the coordination of experts on individual indicators, with a smaller CV indicating greater coordination. The W measures the consistency of experts' weight assessments of indicators, with a larger W indicating improved coordination. When the P-value corresponding to the W-value is below 0.05, the experts' ratings are considered to be consistent (Shen, Ou & Ou, 2021).

The results of the first round of consultation showed that the CV of the primary indicators ranged from 0.000 to 0.1133, with a W-value of 0.474 and a P-value of less than 0.05, meeting the consistency requirement. The CV of the secondary indicators was between 0.0464 and 0.2982, with a W-value of 0.297, which was not in compliance with the requirement. Hence, a subsequent round of questionnaires was conducted for the secondary indicators. After optimization, the CV of the second round of questionnaires ranged from 0.000 to 0.1147, with an improved W-value of 0.403 and a P-value of less than 0.05, suggesting significantly improved coordination among experts, making the results reliable. The third round of questionnaires was considered unnecessary.

3.5 Evaluation of Indicator Importance

3.5.1 Results of the First Round of Expert Consultation

The results of the first round of expert consultation on primary and secondary indicators are presented in Tables 2 and 3.

Table 2 The results of the first round of expert consultation on primary indicators

Primary Indicator	SD	Mean	CV
A Teaching Belief	0.000	5.00	0.0000
B Teaching Attitude	0.000	5.00	0.0000
C Teaching Content	0.000	5.00	0.0000
D Teaching Skills	0.000	5.00	0.0000
E Teaching Resources	0.513	4.53	0.1133
F Teaching Effectiveness	0.000	5.00	0.0000

From Table 2, it can be seen that no indicators met the criteria for deletion. Therefore, all primary indicators were retained.

Table 3 The results of the first round of expert consultation on secondary indicators

Secondary Indicator	SD	Mean	CV
A1 Alignment of Teaching Objectives with Student Development Goals	0.315	4.89	0.0644
A2 Integration of Student-Centered Philosophy	0.535	4.79	0.1118
A3 Emphasis on Critical Thinking and Lifelong Learning Skills	0.229	4.95	0.0464
A4 Effective Use of Digital Thinking to Enhance Classroom Teaching	0.749	4.32	0.1736
A5 Effective Integration of Ideological and Political Education	0.582	4.68	0.1243
B1 Enthusiasm for Work	0.315	4.89	0.0644
B2 Neat and Presentable Appearance	1.177	3.95	0.2982
B3 Care for Student Growth and Progress	0.562	4.74	0.1186
B4 Timely Responses to Student Questions	0.772	4.47	0.1726
B5 Accurate and Thorough Analysis of Student Learning Conditions	0.478	4.68	0.1020

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

B6 Emphasis on Teaching Reflection and Student Feedback	0.375	4.84	0.0774
C1 Adaptation to Students' Cognitive Levels and Learning Needs	0.315	4.89	0.0644
C2 Integration with the Latest Research Results and Practical Cases	0.507	4.58	0.1108
C3 Content Richness and Logical Coherence	0.315	4.89	0.0644
C4 Appropriate Handling of Key and Difficult Points	0.419	4.79	0.0875
C5 Challenging Assignments	1.026	3.95	0.2599
D1 Flexible and Diverse Teaching Methods Suitable for Content	0.229	4.95	0.0464
D2 Effective Use of Digital Technology for Personalized Instruction	0.612	4.53	0.1352
D3 Clear and Organized Language Expression	0.315	4.89	0.0644
D4 Reasonable Time Management	1.026	3.95	0.2599
D5 Effective Interaction between Teacher and Students	0.452	4.74	0.0955
D6 Timely and Accurate Student Evaluation	0.671	4.32	0.1555
D7 Implementation of Differentiated Teaching Strategies	1.015	3.84	0.2641
D8 Strong Classroom Management Skills	0.315	4.89	0.0644
E1 Rich and Diverse Teaching Resources	0.478	4.68	0.1020
E2 Intelligent and Accurate Learning Resource Push	0.478	4.68	0.1020
E3 Timely Update of Online Learning Resources	1.049	3.89	0.2692
E4 User-Friendly and Functional Teaching Platform Tools	0.612	4.53	0.1352
E5 Advanced and Intelligent Teaching Facilities	0.692	4.58	0.1512
F1 Harmonious and Pleasant Classroom Atmosphere	0.769	4.42	0.1738
F2 Improvement in Students' Knowledge Mastery and Application Skills	0.315	4.89	0.0644
F3 Enhancement of Students' Thinking and Innovation Abilities	0.315	4.89	0.0644
F4 Improvement in Students' Communication, Collaboration, and Teamwork Skills	0.419	4.79	0.0875
F5 Enhancement of Students' Self-Learning Abilities	0.229	4.95	0.0464
F6 Improvement in Students' Character and Social Responsibility	0.315	4.89	0.0644
F7 Overall Improvement for Students with Different Backgrounds	1.150	3.89	0.2952

Table 3 shows that based on their standard deviation and coefficient of variation, the indicators B2, C5, D4, D7, E3, and F7 met the criteria for deletion and were therefore removed. Additionally, one expert suggested adding a secondary indicator "Content Depth and Challenge" under the primary indicator "C: Teaching Content," and two experts proposed adding "Reasonable Teaching Design and Proper Organization" under the primary indicator "D: Teaching Skills." Another expert recommended moving "F1: Harmonious and Pleasant Classroom Atmosphere" from the primary indicator "F: Teaching Effect" to "D: Teaching Skills," and modifying F1 to "Good at Activating Classroom Atmosphere." These suggestions were adopted after discussion within the research team.

One expert suggested adding "High Teacher Engagement" under the primary indicator "B: Teaching Attitude," However, when talking to this expert, "High Teacher Engagement" was considered too vague in meaning and, to some extent, already reflected in other indicators like "Student Analysis" and "Teaching Design," and thus, was not adopted. Additionally, experts suggested modifications to the descriptions of four indicators: changing A1 to "Alignment of Teaching Goals with Talent Cultivation Goals," A5 to "Integrating Ideological and Political Education Effectively," D2 to "Using Digital Technology for Differentiated Instruction," and D5 to "Active and Effective Teacher-Student and Student-Student Interaction." These changes were adopted after discussion within the research team.

3.5.2 Results of the Second Round of Expert Consultation

After the first round of indicator optimization, the number of secondary indicators was adjusted to 32. The refined indicators were redistributed to the experts. Table 4

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

presents the results of the second round of expert consultation on secondary indicators.

Table 4 The results of the second round of expert consultation on secondary indicators

Secondary Indicator	SD	Mean	CV
A1 Alignment of Teaching Goals with Talent Development Goals	0.000	5.00	0.0000
A2 Integration of Student-Centered Philosophy	0.507	4.42	0.1147
A3 Emphasis on Critical Thinking and Lifelong Learning Skills	0.000	5.00	0.0000
A4 Effective Use of Digital Thinking to Enhance Classroom Teaching	0.496	4.37	0.1134
A5 Integration of Moral Education and Effective Incorporation of Ideological and Political Education	0.000	5.00	0.0000
B1 Enthusiasm for Work	0.000	5.00	0.0000
B2 Care for Student Growth and Progress	0.513	4.53	0.1133
B3 Timely Responses to Student Questions	0.513	4.47	0.1147
B4 Accurate and Thorough Analysis of Student Learning Conditions	0.513	4.47	0.1147
B5 Emphasis on Teaching Reflection and Student Feedback	0.000	5.00	0.0000
C1 Adaptation to Students' Cognitive Levels and Learning Needs	0.000	5.00	0.0000
C2 Integration with the Latest Research Results and Practical Cases	0.513	4.53	0.1133
C3 Rich Content and Logical Coherence	0.000	5.00	0.0000
C4 Appropriate Handling of Key and Difficult Points	0.51	4.5	0.113
	3	3	3
C5 Depth of Content and Challenge of Problems	0.513	4.47	0.1147
D1 Flexible and Diverse Teaching Methods Suitable for Content	0.000	5.00	0.0000
D2 Effective Use of Digital Technology for Personalized Instruction	0.496	4.37	0.1134
D3 Clear and Organized Language Expression	0.000	5.00	0.0000
D4 Reasonable Teaching Design and Organization	0.000	5.00	0.0000
D5 Active and Effective Teacher-Student and Student-Student Interaction	0.507	4.42	0.1147
D6 Timely and Accurate Student Evaluation	0.478	4.32	0.1107
D7 Strong Classroom Management Skills	0.000	5.00	0.0000
D8 Good at Activating Classroom Atmosphere	0.478	4.32	0.1107
E1 Rich and Diverse Teaching Resources	0.513	4.53	0.1133
E2 Intelligent and Accurate Learning Resource Push	0.513	4.53	0.1133
E3 User-Friendly and Functional Teaching Platform Tools	0.513	4.53	0.1133
E4 Advanced and Intelligent Teaching Facilities	0.496	4.37	0.1134
F1 Improvement in Students' Knowledge Mastery and Application Skills	0.000	5.00	0.0000
F2 Enhancement of Students' Thinking and Innovation Abilities	0.000	5.00	0.0000
F3 Improvement in Students' Communication, Collaboration, and Teamwork Skills	0.000	5.00	0.0000

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

F4 Enhancement of Students' Self-Learning Abilities	0.0000	5.000	0.00000
F5 Improvement in Students' Character and Social Responsibility	0.0000	5.000	0.00000

From the data in Table 4, it can be observed that all secondary indicators had an average value greater than 3 and a coefficient of variation less than 0.25, which met the requirements for keeping. The experts did not provide any further comments, so all 32 secondary indicators were retained.

3.6 Indicator Weight and Consistency Evaluation

This study used Yaahp 10.3 data processing software to determine the weights of the 6 primary indicators and the weights of the secondary indicators under each primary indicator. The results of the indicator weights and consistency calculations are shown in Table 5.

Table 5 Indicator Weights and Consistency Calculation Results

Indicator	Wi	λmax	CI	CR	Results
A	0.1875	6.0000	0.00000	0.0000	Pass
B	0.1875				
C	0.1875				
D	0.1875				
E	0.0625				
F	0.1875				
A1	0.2840	5.0593	0.01484	0.0132	Pass
A2	0.0843				
A3	0.2840				
A4	0.0636				
A5	0.2840				
B1	0.3495	5.0265	0.00662	0.0059	Pass
B2	0.1387				
B3	0.0811				
B4	0.0811				
B5	0.3495				
C1	0.3399	5.0265	0.00662	0.0059	Pass
C2	0.1235				
C3	0.3399				
C4	0.1235				
C5	0.0732				
D1	0.1975	8.1715	0.02449	0.0174	Pass
D2	0.0574				
D3	0.1975				
D4	0.1975				
D5	0.0670				
D6	0.0429				
D7	0.1975				
D8	0.0429				
E1	0.2857	4.0000	0.00000	0.0000	Pass
E2	0.2857				
E3	0.2857				
E4	0.1429				
F1	0.2000	5.0000	0.00000	0.0000	Pass
F2	0.2000				
F3	0.2000				
F4	0.2000				
F5	0.2000				

From Table 5, it can be seen that the consistency ratio (CR) for each dimension was less than 0.1, indicating that the consistency check has passed.

Based on the calculation results of the judgment matrices and weights for each indicator, the summarized weights between the hierarchical levels are shown in Table 6.

Table 6 Summary of Indicator Weights

Target Layer	Primary Indicators	Relative Weight	Secondary Indicators	Relative Weight	Overall Weight
S	A	0.1875	A1	0.2840	0.0533
			A2	0.0843	0.0158
			A3	0.2840	0.0533
			A4	0.0636	0.0119
			A5	0.284	0.0533
	B	0.1875	B1	0.3495	0.0655
			B2	0.1387	0.0260
			B3	0.0811	0.0152
			B4	0.0811	0.0152
			B5	0.3495	0.0655
	C	0.1875	C1	0.3399	0.0637
			C2	0.1235	0.0232
			C3	0.3399	0.0637
			C4	0.1235	0.0232
			C5	0.0732	0.0137
	D	0.1875	D1	0.1975	0.0370
			D2	0.0574	0.0108
			D3	0.1975	0.0370
			D4	0.1975	0.0370
			D5	0.0670	0.0126
			D6	0.0429	0.0080

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

			D7	0.1975	0.0370
			D8	0.0429	0.0080
	E	0.0625	E1	0.2857	0.0179
			E2	0.2857	0.0179
			E3	0.2857	0.0179
			E4	0.1429	0.0089
	F	0.1875	F1	0.2000	0.0375
			F2	0.2000	0.0375
			F3	0.2000	0.0375
			F4	0.2000	0.0375
			F5	0.2000	0.0375

From Table 6, it can be seen that the primary indicators A Teaching Belief, B Teaching Attitude, C Teaching Content, D Teaching Skills, and F Teaching Effectiveness all had the same weight of 0.1875, while the weight for E Teaching Resources was the lowest at 0.0625. This suggests that within the current evaluation system, although necessary, the teaching resources may be less important in terms of weight allocation compared to other indicators that are more directly related to the teaching process and student development.

Among the secondary indicators, those related to digital technology had lower weights. This may be due to concerns about the varying degrees of technology adoption, differences in teachers' technological literacy, or apprehensions about over-reliance on technology.

4. Discussion

4.1 Reliability and validity of this study

The success of the Delphi method largely depends on the selection of experts (Brown, 1987). In this study, the 20 carefully selected experts included 85% with professor titles and 80% with over 30 years of teaching experience. The high authority coefficient of 0.93, well above the 0.7 standard, indicates that the experts have an excellent familiarity with the research and their evaluations are based on solid theoretical knowledge and extensive practical experience, ensuring high reliability of the assessment results. The response rates in the two rounds of expert consultation were 95% and 100%, respectively. According to the conventional standards, a Delphi method response rate of 70% is considered very good (Shen, Ou & Ou, 2021). This result reflects the experts' high attention to and active participation in this study.

During the expert consultation process, the coefficient of variation (CV) for the primary indicators ranged from 0.000 to 0.1133. The Kendall's W coefficient was 0.474. For the secondary indicators, the CVs in the two rounds of consultation were from 0.0464 to 0.2982 and from 0.000 to 0.1147, respectively. In the two rounds, the Kendall's W coefficients for the secondary indicators were 0.297 and 0.403, respectively, with P-values less than 0.05. The significant improvement in opinion concentration in the second round, falling within an ideal range, indicates a good coordination among experts, making the assessment results reliable.

The AHP was used to determine indicator weights, a scientifically robust decision-making tool. This study utilized Yaahp 10.3 software to construct judgment matrices based on the average importance scores of the indicators from the second round of consultations and conducted weight calculations and consistency checks. The product-sum method was used to calculate the comprehensive weights of each indicator, clarifying their relative importance in the evaluation system. The CR for both primary and secondary indicators were less than 0.10, successfully passing the consistency check, which validates the rationality of the weight distribution.

4.2 Novelty of the Indicator System

Firstly, the final indicator system shows a clear distinction from traditional classroom evaluation systems. The evaluation system developed in this study significantly reflects the core principles given by the fourth-generation of the educational evaluation theory: multidimensionality, process-oriented and developmental evaluation, as well as reflection and self-regulation within the evaluation process (Guba & Lincoln, 1989). This system emphasizes the use of digital technology in teaching, which represents not only an update in tools, but also a revolution in teaching methods. For instance, the indicator D2 "Effective Use of Digital Technology for Differentiated Instruction" reflects the design of personalized learning paths, while indicators in category E emphasize the flexible use of teaching resources and platforms. This necessitates rethinking how technology can promote active and deep learning among students. Traditional systems focus more on teaching content, methods, and basic teacher skills, with less consideration for use of technology.

Secondly, the system in this study places a stronger emphasis on student-centered teaching philosophies. For example, indicator A2 "Integration of Student-Centered Concepts" highlights the importance of considering students' needs and cognitive levels in teaching, encouraging their involvement in teaching design and evaluation to foster self-development and realization. In contrast, traditional systems often evaluate teaching performance and outcomes from the teacher’s perspective.

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

Thirdly, the new indicator system also highlights personalized teaching and classroom interaction. Indicators such as B4 "Accurate Student Situation Analysis," D5 "Active and Effective Teacher-Student and Student-Student Interaction," and D6 "Timely and Accurate Student Evaluation" reflect the multidimensional and process-oriented nature of teaching interactions. This system emphasizes the quality and effectiveness of interactions and how they contribute to students' cognitive and emotional development, while traditional systems tend to uniform teaching models and focus more on the teacher's teaching expression.

Finally, the new indicators are more comprehensive, including aspects, such as the F category, which not only focuses on knowledge transfer, but also on cultivating various student abilities. This aligns with the fourth-generation evaluation theory's emphasis on promoting students' holistic development, including cognitive, emotional, and social aspects. Traditional systems, on the other hand, primarily focus on students' knowledge acquisition and lack comprehensive evaluation of student capabilities.

4.3 Continuity of the Indicator System

Despite the opportunities and challenges brought by digital transformation, the continuity and connection between the traditional and the new indicator systems remain strong, as dictated by the essence of education and teaching principles. Regardless of technological advancements, the fundamental goal of education—promoting students' holistic development and lifelong learning—remains unchanged. Both the traditional and the new systems emphasize the achievement of teaching objectives and the importance of talent cultivation, that is, the long-term goals and social responsibilities of education.

Teacher professionalism, teaching attitudes, and teaching skills remain the crucial factors in determining teaching quality. Nevertheless, under the new evaluation system, the teacher's responsibilities are more extensive, encompassing not only the dissemination of knowledge but also guidance, design, and reflection. While still focusing on the quality of teaching content, the new system places greater emphasis on timeliness, relevance, and innovation. The curriculum should incorporate the most recent research findings and practical examples to foster students' critical thinking and creativity. Student learning outcomes are a critical basis for evaluation. The new system builds upon traditional approaches to provide a more comprehensive assessment of student outcomes, including knowledge acquisition, thinking skills, communication abilities, self-directed learning capabilities, and social responsibility. Although both the traditional and the new systems advocate for continuous improvement in teaching based on evaluation results, the new system also emphasizes teaching reflection and student feedback, requiring a higher level of openness and interaction in the evaluation process.

4.4 Application Positioning of the Indicator System

The new evaluation indicator system, while building upon the traditional frameworks, places a stronger emphasis on the requirements of digital transformation. This shift represents not only an increased focus on technology application but also a profound reflection and update of educational essence and teaching philosophies. In the context of educational digital transformation, the emphasis should not be put merely on the adoption or fascination with information technology. The focus should be on how classroom teaching differs from traditional methods and how educational concepts evolve within this broader context, instead of on the digital technology tools themselves (yang,2023).

As educators, it is crucial to correctly understand educational digitalization and to appropriately use digital technologies. This involves updating educational philosophies and redefining the roles of teachers and students. Only by doing so can we effectively leverage modern digital technologies to enhance classroom teaching quality. For educational administers, it is essential to recognize that evaluation indicators play a guiding role in education and teaching. When a new time comes, it's important to think carefully and research the changes instead of just following the latest trends. The ultimate goal of evaluating classroom teaching quality should always be the effectiveness of teaching, ensuring that teaching remains aligned with its fundamental purpose.

5 Conclusions

This study has developed an evaluation indicator system for classroom teaching quality in line with the fourth-generation educational evaluation theory and the context of digital transformation. Through literature review, policy analysis, in-depth interviews, as well as by incorporating domestic and international research findings, a comprehensive and scientifically validated indicator system has been developed. Delphi and AHP methods were used to ensure the scientific and practical applicability of the evaluation indicators.

Currently, there are few studies focusing on classroom teaching evaluation systems under the digital transformation context. The indicator system developed in

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

this study represents a significant supplement and innovation to existing university classroom teaching evaluation frameworks. The results provide robust guidance and support for teaching evaluation practices and contribute to the enhancement of classroom teaching quality in higher education institutions.

However, limitations in this study should be addressed. For instance, it would be feasible to increase the number of experts and cover more subject areas. Due to constraints like time and cost, this study primarily focused on constructing the evaluation indicator system without extensive empirical data analysis. Future research could collect the empirical data related to this indicator system, whereby further refining and validating the indicators' effectiveness. Additionally, the validity and reliability of the indicator system will be evaluated to develop a mature system tested by practice.

Ethical Approval Declaration

"All procedures involving human participants in this study were conducted in accordance with the ethical standards set by applicable research guidelines and the principles of the 1964 Declaration of Helsinki and its subsequent amendments. Ethical approval was secured before the commencement of data collection."

Funding: -

This study did not receive any external funding.

Data availability:-

The datasets generated and analysed during the current study will be available from the author upon reasonable request.

Consent for publication:-

I hereby provide consent for the publication of the manuscript detailed above.

Competing interests:-

The authors declare no competing interests

References

[1] Ashfaq, R., & Nabi, Z. (2022). Media Literacy and Learning: Conceptual Contribution in the Field of Media Education. International Journal of Instructional Technology and Educational Studies, 3(4), 1-11. doi: 10.21608/ihites.2021.107738.1082

[2] Brown, B. (1987). Delphi process: A methodology using for the elicitation of opinions of experts. and Corporation, 9, 39.

[3] Cao, P. J., & Wang, A. X. (2023). How new generation digital technology empowers educational evaluation reform. People's Education, (20), 30-32.

[4] Chen, Y. (1999). Educational Evaluation. Beijing: People's Education Press.

[5] Gao, Q. (2022). Construction of quality-sensitive indicators for nutritional care in gastric cancer patients [Master's thesis, Lanzhou University]. https://doi.org/10.27204/d.cnki.glzhu.2022.002555

[6] Guba, E. G., & Lincoln, Y. S. (1989). Fourth-generation evaluation. Sage Publications.

[7] Hsu, C., & Sander, E. S. (2007). The Delphi technique: Making sense of consensus. Practical Assessment, Research & Evaluation, 12(10), 1-8. https://pareonline.net/getvn.asp?v=12&n=10

[8] Kenawy, M., Elkhweet, S., Elatrebi, H., & Elwakil, F. (2024). Challenges of creativity education in the age of digital transformation through the lens of public high school principals and teachers in Egypt. International Journal of Instructional Technology and Educational Studies, 5(1), 46-54. doi: 10.21608/ihites.2025.268461.1175

[9] Li, X., Li, R., Lin, A., et al. (2024, March 7). Construction of a postoperative delirium nursing care plan for patients with Stanford type A aortic dissection: A Delphi study (Version 1) [Preprint]. Research Square. https://doi.org/10.21203/rs.3.rs-3964275/v1

[10] Liu, Z., & Xu, Y. (2023). Practical research on the teaching quality evaluation system of curriculum in the information age. Chinese Science and Technology Papers, 18(2), 234.

[11] Luo, Z., & Guo, C. (2020). Retrospective and prospect of classroom teaching evaluation standards in the 70 years since the founding of New China. China Education, (1), 55-61.

[12] Shen, L., Yang, J., & Jin, X. (2019). Based on the Delphi method and analytic hierarchy process to construct the evaluation index system of nursing simulation teaching quality. Nurse Education Today, 79, 67-73.

[13] Shen, Q. Y., Ouyang, H., & Ouyang, Y. L. (2021). Construction of an evaluation index system for the achievement of industry-education integration goals. Higher Education Exploration, 12, 104-109.

[14] Yang, Z. (2023). The connotation, stages, and implementation path of digital development of higher education. China Higher Education, (5), 16-20.

[15] Zeng, Z., & Cheng, X. (2016). Analysis of problems in the application of the Delphi method research—Based on 38 CSSCI (2014-2015) source journals. Library Information Work, (16), 116-120. https://doi.org/10.13266/j.issn.0252-3116.2016.16.014

[16] Zhang, H., Li, K., Wang, H., Hu, Y., Liu, Y., Liu, Y., ... & Chen, W. (2023). Construction of a clinical online teaching quality evaluation system based on the CIPP model and analytic hierarchy process. Journal of Bengbu Medical College, (11), 1607-1612. https://doi.org/10.13898/j.cnki.issn.1000-2200.2023.11.029

[17] Zhang, L. (2024). Investigation on student evaluation of higher vocational evaluation in the context of digital transformation in education. Frontiers in

International Journal of Instructional Technology and Educational Studies (IJITES)
ISSN (Print): 2682-3918 - ISSN (online): 2682-3926
Volume6 / Issue3, August, 2025
DOI: 10.21608/ihites.2025.311799.1203

Educational Research, 7(7), 10-15. https://doi.org/10.25236/FER.2024.070702

[18] Zhu, H., Huang, Y., Li, Q., Cao, L., Wang, H., & Wang, D. (2023). Revision and improvement of classroom teaching evaluation index system in medical colleges: A case study of Hainan Medical College. Chinese Higher Medical Education, (1), 39-40.

[19] Zhu, N., Guo, L., & Yi, J. (2023). The dilemma and breakthrough of college classroom teaching quality evaluation in the era of artificial intelligence. China Agricultural Education, (4), 49-56.