Available Online: 11 August 2026 DOI: 10.54254/2977-5701/2026.36076

Construction of a fuzzy comprehensive evaluation system for the effectiveness of tiered instruction in higher vocational e-commerce programs

Zhiying Wang

Suining Energy Vocational College, Suining, China

zhwa457@outlook.com

Abstract. The intelligent business environment is driving instructional practices in higher vocational e-commerce programs from a uniform teaching schedule toward differentiated competency development. Conventional evaluation methods are no longer capable of accurately reflecting students' developmental trajectories in tasks such as e-commerce operations, digital marketing, customer service, and data analytics. To address the need for evaluating the effectiveness of tiered instruction, this study develops a fuzzy comprehensive evaluation system centered on competency growth and supported by evidence generated from intelligent learning environments. The proposed framework integrates occupational competency standards, course task data, learning platform behavioral records, and teacher evaluation information to establish a tiered evaluation index system. It further incorporates a hybrid weighting approach combining subjective and objective methods, a contextualized membership degree calculation model, and a comprehensive evaluation model for instructional effectiveness. The system effectively addresses the challenges of ambiguous evaluation boundaries, heterogeneous evidence sources, and substantial differences across instructional tiers. It provides a practical framework for improving teaching quality, supporting differentiated student development, and implementing intelligent assessment in higher vocational e-commerce education.

Keywords: higher vocational e-commerce program, tiered instruction, intelligent assessment, fuzzy comprehensive evaluation

1. Introduction

Against the backdrop of the digital economy, competency requirements for e-commerce professionals have expanded beyond proficiency in operating a single platform to encompass integrated capabilities, including content planning, product operations, customer communication, data analytics, and the application of intelligent technologies. Students enrolled in higher vocational e-commerce programs exhibit considerable differences in their academic foundations, learning pace, career interests, and ability to transfer knowledge across tasks. Consequently, tiered instruction has become an important instructional strategy for enhancing curriculum adaptability and meeting diverse learning needs. Existing evaluation approaches primarily rely on final examination scores, project performance, and teachers' subjective judgments. Although these methods capture certain learning outcomes, they are inadequate for reflecting competency development among students with different starting points and fail to fully exploit the value of learning behavior data recorded by intelligent learning platforms. As teaching platforms, practical training systems, and e-commerce simulation environments continue to accumulate large volumes of process-oriented learning data, the focus of educational evaluation has gradually shifted from static academic performance to a comprehensive assessment of learning behaviors, task performance, and occupational competencies. Transforming these heterogeneous data sources into evaluation results that are interpretable, comparable, and actionable for instructional improvement has therefore become a critical challenge in enhancing the quality of tiered instruction in higher vocational e-commerce programs.

2. Analytical framework for evaluating the effectiveness of tiered instruction in higher vocational e-commerce programs

2.1. Competency growth orientation of tiered instructional objectives

Evaluation of tiered instruction in higher vocational e-commerce programs should place competency growth at its core. For students at the foundation level, evaluation should emphasize the stability with which they complete fundamental job tasks, focusing primarily on their ability to process product information, provide customer service, understand platform regulations, and interpret basic operational data. For students at the development level, the emphasis should shift to integrated task execution, assessing competencies such as product listing, promotional campaign implementation, customer feedback management, and interpretation of operational data [1]. For students at the advanced level, evaluation should focus on comprehensive project planning and the transferability of intelligent technologies, examining capabilities in traffic analysis, content optimization, marketing strategy adjustment, and collaboration with intelligent customer service systems. The fundamental value of tiered evaluation lies in identifying students' competency development relative to their individual starting points, thereby shifting the focus of instructional evaluation from one-time performance outcomes to continuous learning progress.

2.2. Alignment between e-commerce occupational competencies and student differences

Occupational competencies in the e-commerce industry are characterized by well-defined task sequences, strong dependence on digital tools, and rapid performance feedback. Differences among students are primarily reflected in task comprehension, operational efficiency, data awareness, and problem-solving capabilities. Accordingly, the evaluation system should decompose occupational competencies into observable task components while incorporating learner differences across instructional tiers into the evaluation criteria. For foundation-level students, learning effectiveness may be assessed through task standardization and the speed of error correction. For development-level students, evaluation should emphasize task quality and the accuracy of data interpretation. For advanced-level students, greater attention should be given to solution optimization capabilities and the depth of intelligent technology application [2].

2.3. A multidimensional integration path for evidence of teaching effectiveness

Intelligent teaching environments provide more comprehensive sources of evidence for instructional evaluation. Learning management platforms can record students' learning duration, resource utilization, quiz revisions, and participation in discussion activities. E-commerce practical training systems can capture operational data related to product publishing, order processing, customer service responses, promotional campaign configuration, and data reporting. Teacher evaluations can supplement these records by assessing students' task comprehension, collaborative attitudes, and classroom performance, while students' self-assessments provide insights into learning difficulties and strategy adjustments [3]. The evaluation framework should transform these heterogeneous sources of evidence into comparable indicators and employ fuzzy comprehensive evaluation to address the ambiguity inherent in indicator boundaries. In doing so, the resulting evaluation not only preserves the contextual characteristics of the instructional process but also provides robust analytical capability for model-based assessment.

3. Design of the evaluation index system for the effectiveness of tiered instruction

3.1. Integration of indicator sources

The evaluation indicators are derived from four primary sources. The first source consists of occupational competency requirements for e-commerce positions, including job responsibilities in online store operations, live-streaming e-commerce, customer service, digital marketing, and data analytics. The second source comprises curriculum standards and instructional objectives, from which the core learning tasks and assessment requirements are identified. The third source is learning records generated by intelligent platforms, covering resource access, assignment submission, quiz revisions, practical training activities, and interactive feedback. The fourth source includes teacher observations and project evaluations, which complement platform-generated data by capturing aspects that are difficult to quantify automatically, such as strategic decision-making, communication skills, and problem-solving ability. The integration of these diverse sources forms a comprehensive pool of evaluation indicators, thereby reducing the limitations associated with relying solely on examination scores.

3.2. Establishment of core evaluation dimensions

Considering the instructional characteristics of higher vocational e-commerce education, the evaluation index system is structured around five core dimensions: fundamental occupational competencies, completion of operational tasks, application of data analytics, collaboration with intelligent tools, and learning competency growth. Fundamental occupational competencies assess whether students possess the essential skills required to perform basic e-commerce job responsibilities [4]. Completion of operational tasks evaluates the quality of students' performance in activities such as product management, promotional campaign execution, customer service, and order processing. The application of data analytics measures students' ability to interpret traffic metrics, conversion rates, average order value, and customer feedback. Collaboration with intelligent tools reflects students' proficiency in utilizing intelligent customer service systems, content generation technologies, data dashboards, and recommendation tools. Learning competency growth evaluates the extent of students' progress relative to their initial competency levels.

3.3. Refinement of observable indicators

Each of the five core dimensions is further operationalized through a set of observable indicators. Fundamental occupational competencies include understanding platform regulations, standardization of product information entry, appropriateness of customer service communication, and proficiency in order processing procedures. Completion of operational tasks encompasses the quality of product listing, accuracy of promotional campaign configuration, effectiveness in handling after-sales issues, and stability of project delivery [5]. Application of data analytics includes the interpretation of key performance indicators, identification of causes of abnormal data, report analysis, and formulation of operational recommendations. Collaboration with intelligent tools covers the use of intelligent customer service systems, revision of AI-generated content, utilization of data dashboards, and verification of outputs produced by intelligent tools. Learning competency growth includes reductions in task revision frequency, adjustments to learning strategies, responsiveness to classroom feedback, and the transfer of acquired competencies to comprehensive tasks. Each indicator should be explicitly linked to specific instructional activities to ensure that the evaluation remains closely aligned with authentic classroom practice.

3.4. Validation and screening of evaluation indicators

The screening and validation of evaluation indicators are conducted through three complementary approaches: expert review, comparison with students' learning records, and alignment with course tasks. Expert review is employed to determine whether the indicators accurately reflect occupational competency requirements. Comparison with learning records assesses whether valid evidence for each indicator can be obtained from intelligent learning platforms. Alignment with course tasks evaluates whether the indicators can be continuously observed and measured throughout classroom instruction [6]. Indicators that are difficult to collect, conceptually redundant, or incapable of providing meaningful guidance for instructional improvement should be eliminated. The resulting evaluation index system not only reflects the competency objectives of higher vocational e-commerce education but also satisfies the computational requirements of the proposed evaluation model.

4. Construction of the fuzzy comprehensive evaluation model

4.1. Establishment of the evaluation grade set and linguistic rating set

Table 1. Evaluation grades and model score assignments

Grade Symbol	Evaluation Grade	Representative Score	Instructional Interpretation
v_1	Excellent	95	Able to independently complete comprehensive e-commerce tasks and explain task outcomes.
v_2	Good	85	Able to complete major tasks, with only limited guidance required in a few stages.
v_3	Fair	75	Able to accomplish basic tasks, but demonstrates insufficient adaptation in complex situations.
v_4	Pass	65	Able to complete fundamental tasks under instructional guidance.
v_5	Needs Improvement	55	Task performance remains unstable, and key operational skills require further training.

The effectiveness of tiered instruction is inherently characterized by uncertainty and fuzziness. A student may demonstrate a high level of operational proficiency in a particular task while exhibiting only moderate performance in data interpretation. Likewise, a student may obtain a relatively modest final score yet achieve substantial competency growth compared with their initial level. To accommodate such ambiguous evaluation scenarios, this study defines the evaluation grade set as "Excellent, Good, Fair, Pass, Needs Improvement".

For subsequent model computation, each evaluation grade is assigned a corresponding numerical value. These values are used solely for computational purposes and are not intended to replace teachers' professional judgments (Table 1).

The proposed evaluation grade set establishes a unified evaluation language, enabling students at the foundation, development, and advanced instructional levels to be assessed within the same evaluation framework while preserving the differentiated requirements of tiered instruction.

4.2. Determination of hybrid subjective–objective weights

Evaluation in higher vocational e-commerce programs should integrate both expert knowledge and evidence generated from intelligent learning platforms. Excessive reliance on expert judgment may amplify subjective bias, whereas dependence solely on platform-generated data may overlook task quality and professional competence. Therefore, this study adopts a hybrid subjective–objective weighting method that combines expert-assigned weights with the degree of data dispersion. Let the expert-assigned weight of the j-th indicator be a_j, the objective weight obtained using the entropy weighting method be e_j, the adjustment coefficient be α , and the integrated weight be w_j. The combined weighting formula is expressed as Equation (1):

$$w_j = \frac{a_j^{\alpha} e_j^{1-\alpha}}{\sum_{j=1}^{m} a_j^{\alpha} e_j^{1-\alpha}}\ ,\ 0 < \alpha < 1,\ \sum_{j=1}^{m} w_j = 1 \tag{1}$$

where m denotes the number of evaluation indicators. The subjective weight a_j reflects experts' judgments regarding the importance of occupational competencies, while the objective weight e_j represents the discriminative power of each indicator across the sample data. The adjustment coefficient α balances the relative influence of expert judgment and empirical evidence. This weighting approach is applied during the indicator weighting stage, ensuring that the resulting weights are aligned with professional training objectives while simultaneously capturing the actual variations revealed by intelligent learning platform data. The integrated weights of the first-level dimensions and their corresponding second-level observable indicators are presented in Table 2.

Table 2. Evaluation indicators and integrated weights for the effectiveness of tiered instruction

First-Level Dimension	Second-Level Observable Indicator	Symbol	Integrated Weight
Fundamental Occupational Competencies	Understanding of Platform Regulations	u_{11}	0.052
Fundamental Occupational Competencies	Standardization of Product Information Entry	u_{12}	0.061
Fundamental Occupational Competencies	Appropriateness of Customer Service Communication	u_{13}	0.047
Completion of Operational Tasks	Quality of Product Listing	u_{21}	0.073
Completion of Operational Tasks	Accuracy of Promotional Campaign Configuration	u_{22}	0.068
Completion of Operational Tasks	Handling of After-sales Issues	u_{23}	0.059
Application of Data Analytics	Interpretation of Key Performance Indicators	u_{31}	0.071
Application of Data Analytics	Identification of Causes of Abnormal Data	u_{32}	0.084
Application of Data Analytics	Formulation of Operational Recommendations	u_{33}	0.079

Table 2. Continued

Collaboration with Intelligent Tools	Utilization of Intelligent Customer Service Systems	u_{41}	0.064
Collaboration with Intelligent Tools	Revision of AI-generated Content	u_{42}	0.067
Collaboration with Intelligent Tools	Utilization of Data Dashboards	u_{43}	0.074
Learning Competency Growth	Reduction in Task Revision Frequency	u_{51}	0.058
Learning Competency Growth	Responsiveness to Classroom Feedback	u_{52}	0.069
Learning Competency Growth	Transfer of Competencies to Comprehensive Tasks	u_{53}	0.074

As shown in Table 2, identification of causes of abnormal data, formulation of operational recommendations, utilization of data dashboards, and transfer of competencies to comprehensive tasks receive relatively high weights. This finding indicates that evaluation in intelligent e-commerce education should place greater emphasis on students' ability to interpret data and effectively utilize intelligent technologies. Although fundamental operational skills remain indispensable, their instructional value should ultimately support students' successful completion of more comprehensive and higher-order professional tasks.

4.3. Contextualized design of the membership function

Evaluation criteria should not be identical across different instructional levels. For foundation-level students, the emphasis is placed on the completion of fundamental tasks and the correction of operational errors, whereas for advanced-level students, greater attention is given to handling complex tasks and transferring competencies through the use of intelligent technologies. To reconcile the use of a unified evaluation model with the differentiated requirements of tiered instruction, this study develops a contextualized membership function. Let x_j denote the standardized score of a student on the j-th evaluation indicator. Let l_{jk}, m_{jk}, and u_{jk} represent the lower bound, center value, and upper bound, respectively, of the k-th evaluation grade for that indicator. The membership degree is calculated according to Equation (2):

$$\mu_{jk}\left(x_j\right)=\begin{cases}0, & x_j \leq l_{jk} \\ \frac{x_j-l_{jk}}{m_{jk}-l_{jk}}, & l_{jk}<x_j \leq m_{jk} \\ \frac{u_{jk}-x_j}{u_{jk}-m_{jk}}, & m_{jk}<x_j<u_{jk} \\ 0, & x_j \geq u_{jk}\end{cases} \tag{2}$$

where $\mu_{jk}\left(x_j\right)$ denotes the degree to which a student's performance on the k-th indicator belongs to the j-th evaluation grade. The parameters l_{jk}, m_{jk}, and u_{jk} can be determined according to the task difficulty associated with different instructional levels. Equation (2) is employed to convert platform-generated scores, project evaluation results, and teacher observation records into fuzzy membership degrees, thereby avoiding the rigid assignment of borderline scores to a single evaluation category.

4.4. Construction of the tiered evaluation matrix

Suppose the evaluation indicator set for an individual student or a group of students within the same instructional level is denoted by $U=\{u_1,u_2,\ldots,u_m\}$, and the evaluation grade set is denoted by $V=\{v_1,v_2,\cdots,v_n\}$. Based on the membership function, the evaluation matrix can be constructed as shown in Equation (3):

$$R = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \mu_{12} & \mu_{22} & \cdots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m1} & \mu_{m2} & \cdots & \mu_{mn} \end{bmatrix} \tag{3}$$

In this matrix, each row represents the membership distribution of an observable indicator across the evaluation grades, while each column reflects the extent to which the student's overall performance approaches a particular evaluation grade. Separate evaluation matrices may be established for the foundation, development, and advanced instructional levels to reveal structural differences in learning performance across tiers.

4.5. Identification of comprehensive evaluation results

The comprehensive evaluation is obtained through the operation of the integrated weight vector and the evaluation matrix. Let the integrated weight vector be denoted by $W = (w_1, w_2, \ldots, w_m)$, the comprehensive evaluation vector by $B = (b_1, b_2, \ldots, b_n)$, and the representative score vector of the evaluation grades by C. The comprehensive evaluation is then computed according to Equation (4):

$$B = WR,\ F = BC^T \tag{4}$$

where B represents the comprehensive membership degrees of an individual student or a student group with respect to each evaluation grade, and F denotes the final comprehensive evaluation score. Equation (4) is used to generate the overall evaluation result, while the evaluation grade can be identified according to the maximum membership principle based on the elements of B. Compared with the use of a single aggregate score, this approach provides a more informative representation of students' performance across adjacent evaluation grades. For example, a student may obtain a comprehensive score close to the Good level while still exhibiting a relatively high membership degree in the Needs Improvement category. In such cases, instructors can further examine the corresponding evaluation indicators to identify specific weaknesses and provide targeted instructional interventions.

5. Validation of the fuzzy comprehensive evaluation system for tiered instruction

5.1. Sample selection and data collection

To examine the applicability of the proposed fuzzy comprehensive evaluation system in authentic instructional settings, this study selected two intact classes from the E-commerce program at a higher vocational college as the validation sample. All participants had completed the intermediate-stage learning of two core courses, Online Store Operations and E-commerce Data Analytics, which covered essential tasks including product information editing, basic online store management, customer inquiry handling, promotional campaign configuration, business data interpretation, and the application of intelligent support tools. Rather than selecting participants based on one-time examination scores, the sampling criteria emphasized whether students possessed complete learning records, had participated continuously in practical training activities, and had produced traceable project outcomes [6]. This approach minimizes the influence of incidental examination performance on the evaluation results and enables the validation process to more accurately reflect the actual implementation of tiered instruction.

Student grouping was determined according to results from entry-level diagnostic tasks, platform-based baseline assessments, teachers' classroom observations, and initial performance in practical training activities. Foundation-level students generally demonstrated an unstable understanding of platform regulations, relied heavily on instructional guidance to complete tasks, and exhibited relatively weak data interpretation skills. Development-level students were able to complete most fundamental tasks independently but still required teacher support when configuring promotional campaigns, identifying abnormal data patterns, and formulating operational recommendations. Advanced-level students demonstrated strong autonomous operational capabilities and adapted quickly to intelligent customer service systems, AI-assisted content generation tools, and data dashboards; however, their ability to optimize comprehensive business solutions still required further development through project-based learning. This tiered classification not only reflects differences in students' initial competency levels but also preserves flexibility for dynamic adjustment throughout the instructional process.

Data were collected throughout the pre-class, in-class, and post-class stages. Pre-class data included diagnostic assessments, baseline platform operation performance, and learning preparedness. In-class data consisted of classroom task submissions, operational logs from the practical training system, teachers' observation records, and records of group project collaboration. Post-class data included the number of project revisions, students' responsiveness to instructional feedback, and evidence of competency transfer in comprehensive tasks [7]. Records generated by the intelligent learning platform provided behavioral evidence, while teachers' observations supplemented information regarding learning strategies, communication quality, and task comprehension. Project evaluations captured students' overall performance in authentic occupational tasks. By integrating these heterogeneous data sources into the evaluation model, the proposed framework produces results that are both data-driven and pedagogically interpretable within the context of vocational education. The sample composition and data sources are summarized in Table 3.

Table 3. Sample composition and data sources

Student Tier	Number of Students	Mean Entry Diagnostic Score	Valid Platform Records	Practical Training Tasks	Teacher Observation Records
Foundation Level	28	58.4	2146	392	168
Development Level	31	67.9	2385	434	186
Advanced Level	24	76.2	1918	336	144
Total	83	67.1	6449	1162	498

As shown in Table 3, the three instructional tiers exhibit clear differences in their mean entry diagnostic scores, indicating that the tiered classification is supported by observable evidence. The substantial numbers of valid platform records and practical training tasks provide sufficient data for continuous analysis of students' learning processes. Although the number of teacher observation records is smaller than that of platform-generated records, their primary function is to explain the underlying learning conditions reflected in students' behavioral data. For example, platform records may indicate that a student repeatedly revised a product title, whereas teacher observations can further determine whether these revisions resulted from an improved awareness of operational standards or from an incomplete understanding of task requirements. By integrating these complementary sources of evidence, the evaluation model avoids equating simple operational frequency with learning quality while simultaneously reducing the subjectivity associated with teachers' personal judgments.

5.2. Consistency between evaluation results and competency growth

The evaluation of tiered instruction should not rely solely on students' final project performance but should also consider the extent of their competency development relative to their initial proficiency levels. In this study, the mean score of the entry diagnostic assessment is used to represent students' baseline competency at the beginning of the course, while the mean score of the final comprehensive project reflects their overall task performance upon course completion. The difference between these two scores is regarded as the observed measure of competency growth [8]. In contrast, the fuzzy comprehensive evaluation score is calculated by integrating indicator weights, the membership matrix, and the numerical assignments of evaluation grades. Comparing these two sets of results enables an assessment of whether the proposed model accurately captures students' learning progress.

Specifically, the entry diagnostic assessment included tasks such as identifying platform regulations, entering product information, responding to basic customer inquiries, and interpreting simple operational data. The final comprehensive project required students to complete a series of tasks within a simulated online store scenario, including product optimization, promotional campaign configuration, customer issue resolution, data diagnosis, and the preparation of operational recommendations. These two assessments represent successive stages of learning, reflecting students' progression from fundamental operational skills to comprehensive professional competencies. Rather than directly adopting the final project score, the fuzzy comprehensive evaluation model decomposes project performance into multiple observable indicators and combines these with platform-generated learning records and teacher evaluations to produce an integrated assessment. If the fuzzy evaluation scores are consistent with the direction of competency growth, the model can be considered effective in identifying the genuine learning gains achieved through tiered instruction. The comparison results are presented in Table 4.

Table 4. Comparison between fuzzy evaluation results and competency growth

Student Tier	Mean Entry Diagnostic Score	Mean Final Project Score	Mean Competency Growth	Fuzzy Evaluation Score	Grade with Maximum Membership
Foundation Level	58.4	72.6	14.2	74.8	Fair
Development Level	67.9	81.5	13.6	82.3	Good
Advanced Level	76.2	88.7	12.5	87.6	Good
Overall Sample	67.1	80.9	13.8	81.2	Good

As shown in Table 4, foundation-level students began the course with relatively low baseline competency but achieved the greatest average competency growth, indicating that tiered instruction provides particularly strong support for students with weaker academic foundations. At the beginning of the course, these students commonly experienced difficulties in understanding platform regulations, accurately completing product attribute information, and maintaining appropriate communication styles in customer service interactions. Following systematic practice on fundamental tasks and continuous formative feedback, the stability of their project performance improved substantially. The fuzzy comprehensive evaluation model classified this group into the Fair category based on the maximum membership principle, suggesting that although these students demonstrated considerable progress, their overall competency had not yet reached the Good level. This assessment is consistent with their actual learning status.

Students in the development level achieved a fuzzy evaluation score of 82.3, with Good identified as the grade of maximum membership. These students already possessed a certain degree of operational competence at the beginning of the course, and their competency growth was primarily reflected in improved task integration. They performed consistently in product listing, promotional campaign configuration, and after-sales service, although they still exhibited insufficient depth in interpreting operational data. The model's classification of this group as Good indicates that they were capable of completing the core tasks required by the curriculum but still needed further improvement in identifying abnormal data patterns within complex business scenarios.

Students in the advanced level obtained both the highest mean final project score and the highest fuzzy evaluation score, although their average competency growth was lower than that of the foundation and development groups. This finding suggests that students with stronger initial competencies have relatively limited room for quantitative growth, and therefore instructional effectiveness should not be evaluated solely on the magnitude of competency improvement. Advanced-level students demonstrated clear strengths in utilizing intelligent customer service systems, operating data dashboards, and presenting comprehensive business solutions. Nevertheless, some students exhibited an excessive reliance on AI-generated outputs, characterized by the ability to adopt intelligent tool recommendations without sufficiently validating or adapting them to specific business contexts. The fuzzy comprehensive evaluation model classified this group as Good, indicating that although their overall competency level was high, they had not yet achieved consistently Excellent performance. These findings provide practical guidance for instructors to strengthen advanced-level students' competencies in solution diagnosis, strategic adjustment, and critical evaluation of outputs generated by intelligent technologies.

5.3. Discriminative analysis of evaluation results across student tiers

Whether an evaluation system can effectively support tiered instruction depends largely on its ability to distinguish differences in competency structures among students at different instructional levels. If the evaluation results across tiers are overly similar, the model will provide limited support for task allocation and instructional adjustment. Accordingly, this study compares students' performance across five core dimensions—fundamental occupational competencies, completion of operational tasks, application of data analytics, collaboration with intelligent tools, and learning competency growth—to examine the model's effectiveness in differentiating student performance.

Fundamental occupational competencies reflect students' mastery of platform regulations, product information standards, and customer service communication. Completion of operational tasks measures students' ability to translate fundamental operational skills into successful project outcomes. Application of data analytics emphasizes students' capacity to interpret store performance data, promotional campaign effectiveness, and customer feedback. Collaboration with intelligent tools assesses whether students can effectively utilize intelligent customer service systems, AI-assisted content generation tools, and data dashboards while critically evaluating the outputs generated by these technologies. Learning competency growth focuses on students' progress relative to their initial competency levels, emphasizing task revision, responsiveness to instructional feedback, and competency transfer. Together, these five dimensions represent different aspects of occupational competency in e-commerce and provide a comprehensive assessment of the effectiveness of tiered instruction. The evaluation results for each instructional level are presented in Table 5.

Table 5. Evaluation results across core dimensions for different student tiers

Core Dimension	Foundation Level (Mean)	Development Level (Mean)	Advanced Level (Mean)	Test Statistic for Between-Tier Differences
Fundamental Occupational Competencies	76.3	82.1	87.4	18.26
Completion of Operational Tasks	73.5	81.7	88.2	24.11
Application of Data Analytics	70.8	79.4	86.5	27.43
Collaboration with Intelligent Tools	72.6	80.2	89.1	31.08
Learning Competency Growth	80.4	82.7	84.3	6.52

As shown in Table 5, students at different instructional levels exhibit a clear progressive pattern across all five core dimensions. Among these dimensions, collaboration with intelligent tools yields the largest between-tier test statistic, indicating that it provides the strongest discriminative power in evaluating students' competencies in intelligent e-commerce tasks. Foundation-level students are generally able to utilize intelligent customer service systems and basic AI-assisted content generation tools with teacher guidance; however, they often lack the ability to critically evaluate AI-generated outputs and may directly incorporate content that is inconsistent with the product context into project tasks. Development-level students are capable of revising AI-generated content according to task requirements, although they still demonstrate inconsistencies when interpreting data dashboards and formulating operational recommendations. Advanced-level students are able to refine intelligent tool outputs by considering product positioning, customer profiles, and promotional objectives, thereby demonstrating stronger competency transfer and task adaptation.

The differences among instructional levels are also pronounced in the application of data analytics dimension. Foundation-level students generally remain at the stage of basic data interpretation and can identify key indicators such as page views, visitor numbers, and order volume, but they exhibit limited ability to explain the underlying causes of data variation. Development-level students can make preliminary judgments by relating operational data to promotional schedules, product pricing, and customer inquiries, although their conclusions are sometimes insufficiently supported by evidence. Advanced-level students are capable of integrating multiple performance indicators into comprehensive operational analyses and proposing relatively complete optimization strategies. These findings suggest that the application of data analytics has become a key dimension for differentiating students' competencies in the evaluation of tiered instruction within higher vocational e-commerce education.

The between-tier difference in learning competency growth is comparatively smaller, indicating that students across all instructional levels achieved measurable improvement during the learning process. Foundation-level students attained a mean score of 80.4 in this dimension, reflecting substantial developmental progress. Although development-level and advanced-level students entered the course with stronger initial competencies, they also demonstrated continued improvement in project revision, responsiveness to instructional feedback, and the transfer of competencies to comprehensive tasks. These findings suggest that tiered evaluation should not focus exclusively on differences among instructional levels but should also emphasize each student's developmental trajectory. For foundation-level students, instructional priorities should center on task standardization and confidence building. For development-level students, greater emphasis should be placed on data interpretation and task integration. For advanced-level students, instruction should focus on the critical evaluation of outputs generated by intelligent technologies and the optimization of project strategies.

5.4. Validation of the stability and applicability of the evaluation system

The stability of an evaluation system determines whether it can be adopted as a sustainable tool for long-term instructional practice. If the evaluation results fluctuate substantially across different task types or assessment periods, the model cannot provide a reliable basis for instructional improvement. Accordingly, this study evaluates the proposed system from five perspectives: test–retest reliability, internal consistency, agreement with teacher judgments, completeness of data sources, and stability of tier classification. The same evaluation index system was applied to both the midterm and final project assessments. Three instructors with expertise in e-commerce independently evaluated students' project performance, and their ratings were subsequently compared with the evaluation results generated by the proposed model. The completeness of data sources was verified by examining the effective coverage of records collected from the learning management platform, the practical training system, and teachers' observation records. The results of these validation tests are summarized in Table 6.

Test–retest reliability was employed to examine the consistency of the evaluation model across different instructional stages. The midterm project emphasized the integration of fundamental operational tasks, whereas the final project focused on comprehensive e-commerce operational planning. Although the two assessments differed in task complexity, they measured the same underlying competency structure. The resulting correlation coefficient of 0.81 indicates that the proposed model maintains a high degree of stability despite changes in task characteristics. Internal consistency was used to determine whether the evaluation indicators collectively measured the effectiveness of tiered instruction. The coefficient of 0.87 demonstrates a strong internal structure among the evaluation indicators. Agreement with teacher judgments was assessed to determine whether the model's evaluation outcomes were consistent with instructors' professional assessments. The resulting agreement coefficient of 0.76 indicates substantial consistency between the model and teachers' overall evaluations.

Table 6. Validation results for the stability of the evaluation system

Validation Item	Assessment Content	Result	Interpretation
Test–Retest Reliability	Correlation between model scores for the midterm and final projects	0.81	High consistency across evaluation stages
Internal Consistency	Consistency coefficient of the evaluation index system	0.87	Stable internal structure of the indicators
Agreement with Teacher Judgments	Agreement coefficient between model classifications and teachers' overall ratings	0.76	Model evaluations are highly consistent with professional judgments
Completeness of Data Sources	Coverage coefficient of valid evaluation records	0.92	Platform and practical training data adequately support evaluation
Stability of Tier Classification	Consistency coefficient of student classifications in repeated evaluations	0.79	Tier classifications remain relatively stable

As shown in Table 6, the proposed evaluation system demonstrates satisfactory stability and practical applicability in classroom settings. The high test–retest reliability indicates that the model produces consistent evaluations across different stages of instruction without being unduly influenced by the performance of a single project. The strong internal consistency confirms that the five core dimensions—fundamental occupational competencies, completion of operational tasks, application of data analytics, collaboration with intelligent tools, and learning competency growth—constitute a coherent evaluation framework. The substantial agreement with teacher judgments further indicates that, although the model incorporates intelligent learning data and fuzzy computation, its evaluation outcomes remain closely aligned with professional instructional expertise. The high completeness of data sources demonstrates that learning management platforms and practical training systems provide sufficient evidence to support the evaluation process without imposing substantial additional documentation burdens on instructors. Furthermore, the stability of tier classification suggests that the model consistently identifies students' instructional levels while remaining sensitive to those who demonstrate sufficient progress to transition across tiers.

From the perspective of instructional improvement, the proposed evaluation system enables teachers to make more targeted pedagogical adjustments. For foundation-level students who obtain relatively low scores in fundamental occupational competencies, instructors may strengthen instruction on platform regulations, standard procedures for product information entry, and customer service communication. Conversely, if these students demonstrate strong learning competency growth, they may be assigned selected tasks designed for the development level to prevent prolonged exposure to low-difficulty training. For development-level students who perform well in completion of operational tasks but exhibit weaker performance in application of data analytics, instructors may introduce additional activities involving business data interpretation, post-campaign performance analysis, and diagnosis of abnormal operational indicators. For advanced-level students who achieve high scores in collaboration with intelligent tools but demonstrate inconsistent performance in the transfer of competencies to comprehensive tasks, instructors may assign more complex store diagnostic projects and require explicit justification for the selection and modification of outputs generated by intelligent technologies.

The proposed evaluation framework also improves the quality of feedback provided to students. Conventional assessment methods typically report only an overall score and a final grade, making it difficult for students to identify their specific strengths and weaknesses. In contrast, the fuzzy comprehensive evaluation model decomposes the overall evaluation into individual dimensions and observable indicators, allowing students to clearly recognize their performance in areas such as e-commerce operations, data analytics, intelligent technology application, and competency growth. Consequently, teachers can replace generalized comments with task-oriented recommendations. For example, if a foundation-level student receives an overall Fair rating but demonstrates strong learning competency growth, the instructor may acknowledge the student's substantial progress while recommending further improvement in product title standardization and the accuracy of customer service communication. Similarly, if an advanced-level student achieves a high overall evaluation score but performs poorly in validating outputs generated by intelligent tools, the instructor may require the student to provide explicit human justification for adopting or modifying AI-generated recommendations in project reports. In this way, the evaluation results become directly integrated into the process of continuous instructional improvement rather than serving merely as a summative assessment outcome.

6. Conclusion

This study addresses the evaluation of tiered instruction in higher vocational e-commerce programs by developing a fuzzy comprehensive evaluation system centered on competency growth and supported by evidence collected from intelligent learning platforms. The proposed framework integrates occupational competencies, individual learner differences, learning process data, and the application of intelligent technologies into a unified evaluation model, thereby overcoming several limitations of conventional assessment approaches, including an excessive reliance on outcome-based evaluation, insufficient representation of differences across instructional tiers, and inadequate utilization of process-oriented learning evidence. The validation results demonstrate that the proposed evaluation system effectively identifies competency development among students at different instructional levels and provides actionable evidence for teachers to adjust task difficulty, refine instructional support strategies, and optimize curriculum design. By combining intelligent learning data with fuzzy comprehensive evaluation, the framework not only enhances the objectivity and interpretability of instructional assessment but also offers practical support for implementing differentiated instruction in higher vocational e-commerce education. Future research may further integrate authentic e-commerce operational data into the evaluation framework, thereby extending its applicability across multiple courses and interdisciplinary educational settings while enhancing its practical value in intelligent vocational education.

References

[1] Dong, W. Z., Tian, W., & Hou, L. (2026). Construction and evaluation of an engineering application-oriented talent cultivation model under the background of university–enterprise collaborative education. *Gaojiao Xuekan, 12*(2), 170–173. [In Chinese]

[2] Wang, F. L. (2025). Evaluation of competency levels of higher vocational teachers based on the analytic hierarchy process. *Shandong Shangye Zhiye Jishu Xueyuan Xuebao, 25*(6), 44–50. [In Chinese]

[3] Song, S. (2025). Evaluation of the effectiveness of university–enterprise collaborative education for the engineering cost major under the background of industry–education integration. *Chanye Yu Keji Luntan, 24*(11), 140–142. [In Chinese]

[4] Zhang, L. X., Qian, J., Xie, L. Y., et al. (2023). Evaluation of university–enterprise co-built training bases based on the analytic hierarchy process and fuzzy comprehensive evaluation method: A case study of the Shimadzu Joint Laboratory at Zhejiang Pharmaceutical University. *Zhongguo Xiandai Jiaoyu Zhuangbei*, (17), 163–166, 176. [In Chinese]

[5] Xu, L., Fang, Y., & Xu, Z. J. (2023). Evaluation of university campus landscapes based on the AHP–fuzzy comprehensive evaluation method: A case study of Jiangpu Campus of Nanjing Tech University. *Chengshi Jianzhu, 20*(15), 197–201. [In Chinese]

[6] Mi, N. (2023). Effectively constructing evaluation indicators for university–enterprise cooperation and industry–education integration. *Renli Ziyuan*, (10), 158–159. [In Chinese]

[7] Shao, L. (2022). Application of the fuzzy comprehensive evaluation method in evaluating the quality of university–enterprise cooperation in vocational colleges. *Zhiye Jiaoyu, 21*(7), 48–51. [In Chinese]

[8] Zhang, X. L. (2020). Performance evaluation of university–enterprise cooperation in higher vocational colleges based on the analytic network process (ANP). *Diannao Yu Dianxin*, (12), 22–27. [In Chinese]